ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Ebert, Cara; Flörchinger, Daniela; Frohnweiler, Sarah; Ihring, Stephanie; Rosadio Cayllahua, Karen Micaela

Research Report

Employment and income effects of skills development interventions: An impact evaluation of three employment promotion measures in Eastern Africa within GIZ's employment and skills for development program

RWI Projektberichte

Provided in Cooperation with:

RWI – Leibniz-Institut für Wirtschaftsforschung, Essen

Suggested Citation: Ebert, Cara; Flörchinger, Daniela; Frohnweiler, Sarah; Ihring, Stephanie; Rosadio Cayllahua, Karen Micaela (2021) : Employment and income effects of skills development interventions: An impact evaluation of three employment promotion measures in Eastern Africa within GIZ's employment and skills for development program, RWI Projektberichte, RWI - Leibniz-Institut für Wirtschaftsforschung, Essen

This Version is available at:
https://hdl.handle.net/10419/251877

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Project Report

RWI – Leibniz-Institut für Wirtschaftsforschung

Employment and income effects of skills development interventions

An impact evaluation of three employment promotion measures in Eastern Africa within GIZ's Employment and Skills for Development program

Project Report commissioned by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

June 2021

Impressum

Herausgeber:

RWI – Leibniz-Institut für Wirtschaftsforschung

Hohenzollernstraße 1–3 | 45128 Essen, Germany

Postanschrift:

Postfach 10 30 54 | 45030 Essen, Germany

Fon: +49 201–81 49-0 | E-Mail: rwi@rwi-essen.de
www.rwi-essen.de

Vorstand

Prof. Dr. Dr. h. c. Christoph M. Schmidt (Präsident)

Prof. Dr. Thomas K. Bauer (Vizepräsident)

Dr. Stefan Rumpf

© RWI 2021

Der Nachdruck, auch auszugsweise, ist nur mit Genehmigung des RWI gestattet.

RWI Projektbericht

Schriftleitung: Prof. Dr. Dr. h. c. Christoph M. Schmidt

Gestaltung: Daniela Schwindt, Magdalena Franke, Claudia Lohkamp

Employment and income effects of skills development interventions

An impact evaluation of three employment promotion measures in Eastern Africa within GIZ's Employment and Skills for Development program

Project Report commissioned by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Project Team

Dr. Cara Ebert, Daniela Flörchinger, Sarah Frohnweiler, Stephanie Ihring, and Karen Micaela Rosadio Cayllahua

Research Assistance

Dorothee Seybold and Veerangna Kohli

This research project is a collaborative effort comprehensively involving the GIZ-E4D teams of the GIZ headquarter and the country offices in Kenya and Uganda. The RWI team gratefully acknowledges their perpetual contributions without which this study could not have been realized.

Project Report

RWI – Leibniz-Institut für Wirtschaftsforschung

Employment and income effects of skills development interventions

An impact evaluation of three employment promotion measures in Eastern Africa within GIZ's Employment and Skills for Development program

Project Report commissioned by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

June 2021

rwi

Table of contents

Executive Summary **17**

1 Introduction **26**

2 Quantitative Evaluation of the KAM program **28**

2.1 The KAM program 28
2.2 Research questions 29
2.3 Research design 30
2.3.1 Estimation methods 30
2.3.2 Outcome indicators 32
2.4 Data collection and sample 35
2.4.1 Data collection through IPA 35
2.4.2 Timing of the data collection 36
2.4.3 Data specificities and modifications 38
2.5 Background characteristics of study participants 41
2.5.1 Regions and trades of study participants 41
2.5.2 Sociodemographic characteristics 43
2.5.3 Employment status and earnings prior to the KAM program 45
2.5.4 Restoring the comparability of treatment and control groups 47
2.6 Descriptive statistics on the KAM program 48
2.6.1 Outreach of the program 48
2.6.2 Internship placements 48
2.6.3 Participants' satisfaction with the KAM program 51
2.7 Descriptive statistics on participants' job search 58
2.7.1 Alternative employment promoting programs 58
2.7.2 Participation in interviews 58
2.7.3 Participant-reported evaluation of KAM on job search 59
2.8 Descriptive statistics on labor market outcomes 60
2.8.1 Changes in labor market characteristics between baseline and follow-up 60
2.8.2 Participant-reported evaluation of KAM on current jobs 63
2.9 Estimated program effects on labor market outcomes 64
2.9.1 Program impact on employment status 65
2.9.2 Program impact on income 66
2.9.3 Program impact on job search 68
2.9.4 Robustness checks 69
2.10 Program impact for sub-groups of study participants 70
2.10.1 Program effect heterogeneity by respondents' gender 71
2.10.2 Program effect heterogeneity by respondents' age 72
2.10.3 Program effect heterogeneity by respondents' work experience 73
2.11 Program impact on secondary outcomes 74
2.11.1 Employment and earnings aspirations 74
2.11.2 Family structure 75
2.11.3 Banking and savings 75

2.12 Summary and discussion ... 76

3 Qualitative evaluation of KAM's `sustainability program' ... 80

3.1 Introduction: Sustainability problematic of development cooperation projects ... 80

3.2 Overview of case background: KAM TVET project ... 80

3.3 Research questions and design ... 81

3.4 Theoretical background information on ownership and sustainability ... 83

3.5 Evaluation of overall KAM TVET program by companies ... 85

3.6 Suggestions for improvement of overall KAM TVET program by companies ... 85

3.7 Research question 1: Ownership of internship placements and funding switch ... 87

3.7.1 Companies' ownership of internship placements ... 87

3.7.1.1 Applied ownership definition for internship placements ... 87

3.7.1.2 Evaluation of companies' ownership of internship placements ... 87

3.7.1.3 Underlying reasons for companies' ownership of internship placements ... 89

3.7.1.4 Relevance of funding switch for companies' ownership of internships ... 90

3.7.2 Companies' ownership of funding switch ... 90

3.7.2.1 Applied ownership definition for funding switch ... 90

3.7.2.2 Evaluation of companies' ownership of funding switch ... 90

3.7.2.3 Reasons for mixed ownership of funding switch among companies ... 90

3.7.3 Comparison of ownership success factors to research findings on ownership ... 92

3.7.4 Research question 3 (part 1): Lessons learnt and potential solutions to enhance companies' ownership of the funding switch ... 93

3.8 Research question 2: Sustainability of internship placements after the switch to the sustainability program ... 95

3.8.1 Applied sustainability definition for internship placements ... 95

3.8.2 Sustainability of internship placements after funding switch ... 95

3.8.3 Sustained employment benefits for interns after funding switch ... 96

3.8.4 Reasons for companies' (non-)participation in the sustainability program ... 97

3.8.4.1 Companies' motives (not) to pay stipends to interns ... 97

3.8.4.2 Companies' satisfaction with (former) KAM interns ... 101

3.8.5 Research question 3 (part 2): Lessons learnt and potential solutions to enhance the sustainability of the internship placements in the future ... 102

3.8.5.1 Strategies to increase the companies' willingness to pay stipends in the future ... 103

3.8.5.2 Strategies to increase companies' satisfaction with KAM interns ... 106

3.9 Evaluation of ownership and sustainability at project level ... 108

3.9.1 Ownership at project level ... 108

3.9.2 Sustainability at project level ... 108

3.9.3 Comparison of sustainability success factors to research findings on sustainability ... 110

3.9.4 Lessons learnt and potential solutions to enhance sustainability at project level ... 111

4 Quantitative Evaluation of the RtW program ... 112

4.1 The Ready to Work Program ... 112

4.2 Research Questions 113
4.3 Research design 114
4.3.1 Estimation methods 114
4.3.2 COVID-19 and the timing of data collection 117
4.3.3 The pre-COVID-19 research design 118
4.3.4 Outcome indicators 118
4.4 Data collection 120
4.4.1 Sample size and response rates 120
4.4.2 Study participants by treatment group, schools, and trades 120
4.4.3 Timing of survey and program roll-out 122
4.5 Background characteristics of study participants 124
4.5.1 Sociodemographic characteristics 124
4.5.2 Employment status and earnings prior to the RtW program 127
4.5.3 Restoring the comparability of treatment and control groups 128
4.6 Descriptive statistics on the ReadyToWork program and UMA internship placements 128
4.6.1 Participants' satisfaction with the RtW training program 129
4.6.2 Descriptive statistics of the internship placements 131
4.6.3 Participation in alternative employment promoting programs 134
4.7 Descriptive statistics on participants' job search 135
4.7.1 Interviews 135
4.7.2 Length of job search 137
4.7.3 Strategies of job search 138
4.8 Descriptive statistics on labor market outcomes 140
4.8.1 Labor market characteristics in February 2020 140
4.8.2 Changes in labor market characteristics between baseline and follow-up 143
4.8.2.1 Changes in employment status 143
4.8.2.2 Changes in income 144
4.8.2.3 Changes in hourly wages and working hours 146
4.9 Estimated program effects on labor market outcomes 146
4.9.1 Program impact on employment status 147
4.9.2 Program impact on incomes and wages 148
4.9.3 Robustness checks 149
4.10 Program impact for sub-groups of RtW respondents 150
4.10.1 Program effect heterogeneity by respondents' age 150
4.10.2 Program effect heterogeneity by respondents' prior sustained work experience 151
4.10.3 Program effect heterogeneity by respondents' construction sector experience 152
4.10.4 Program effect heterogeneity by respondents' electrical or mechanical sector experience 152
4.10.5 Program effect heterogeneity by respondents' vocational training level 153
4.11 Program impact on secondary outcomes 154
4.11.1 Employment and earnings aspirations 154

4.11.2 Migration intentions 156
4.12 Summary and discussion 158
5 Quantitative Evaluation of the S4C program 160
5.1 The Skills for Construction program 160
5.2 Research questions 161
5.3 Research design 162
5.3.1 Estimation methods 162
5.3.2 COVID-19 and the timing of data collection 164
5.3.3 Outcome indicators 164
5.4 Data collection 166
5.4.1 Sample size and response rates 166
5.4.2 Study participants by treatment group, region of residence, and education 167
5.4.3 Timing of survey and program roll-out 169
5.5 Background characteristics of study participants 171
5.5.1 Sociodemographic characteristics 171
5.5.2 Employment status and earnings prior to the S4C program 174
5.5.3 Restoring the comparability of treatment and control groups 176
5.6 Descriptive statistics on S4C placements 176
5.6.1 Who was placed? 176
5.6.2 Who is still expecting to be placed? 179
5.6.3 When were beneficiaries placed in internships? 180
5.7 Participation in alternative employment promoting programs 182
5.8 Descriptive statistics on respondents' job search 184
5.8.1 Interviews 184
5.8.2 Length of job search 187
5.8.3 How did respondents search for a job? 188
5.9 Descriptive statistics on labor market outcomes 191
5.9.1 Labor market characteristics at endline 191
5.9.2 Changes in labor market characteristics between baseline and follow-up 197
5.10 Estimated program effects on labor market outcomes 203
5.10.1 Program impact on employment status 204
5.10.2 Program impact on incomes and wages 209
5.10.3 Robustness checks 214
5.11 Program impact for sub-groups of S4C respondents 216
5.11.1 Program effect heterogeneity by respondents' gender 217
5.11.2 Program effect heterogeneity by respondents' age 218
5.11.3 Program effect heterogeneity by respondents' vocational training level 219
5.11.4 Program effect heterogeneity by respondents' previous work experience 220
5.11.5 Program effect heterogeneity by respondents' construction sector experience 222
5.12 Program impact on secondary outcomes 223
5.12.1 Employment and earnings aspirations 223
5.12.2 Migration intentions 225

5.13 Summary and discussion 228

6 Labor Mobility and the Relevance of Regional Income Differentials among S4C Participants 231

6.1 Introduction 231

6.2 Descriptive evidence on migration intentions 231

6.2.1 Internal and international migration intentions and preparations 231

6.2.2 Destination choices 234

6.2.3 Internal migration for internships and jobs opportunities in the oil and gas sector 236

6.3 The importance of information about wage differentials 240

6.3.1 The information treatment 240

6.3.2 Estimation strategy and randomization 242

6.3.3 Results 244

6.3.3.1 Descriptive evidence on respondents' expected regional wages 244

6.3.3.2 Impacts of the information treatment 246

6.3.3.3 Impacts of the information treatment by S4C beneficiary status 249

6.4 Summary and conclusion 251

7 Value for Money 253

7.1 Value for Money as a concept 253

7.2 Why evaluate Value for Money? 255

7.3 Caveats of Value for Money 255

7.4 Conceptual framework of Value for Money: *The 4E approach* 256

7.5 Approaches to Value for Money 259

7.6 Tools to evaluate Value for Money 264

7.7 Cost-effectiveness analysis 267

7.7.1 Ingredients and procedure 267

7.7.2 Example calculation for the KAM and S4C program 272

7.8 Conclusion 275

8 Indirect and Induced Employment Effects 278

8.1 Concept 278

8.2 Literature and Implementation of Theory 278

8.3 The case of human capital interventions 281

8.3.1 Conceptualization of employment effects of human capital interventions 281

8.3.2 The calculation of employment effects of human capital interventions 283

8.4 E4D Application 284

8.5 Conclusion and Recommendation 286

9 Lessons Learned for Future Program Designs 287

9.1 Program design 289

9.1.1 Trainings 289

9.1.2 Placements 290

9.2 Target group 291

9.3 Recommendations: 291

10 Lessons Learned for Future Impact Evaluations ... 293

11 References ... 299

A Appendix ... 305

A1 Appendix of the quantitative KAM evaluation ... 305

A1.1 Additional descriptive statistics ... 305

A1.2 Results from alternative sample compositions ... 311

A1.3 Subsample estimations on income, wages, working hours, and job search ... 315

A2 Appendix of the qualitative KAM evaluation ... 320

A2.1 Overview of interviews ... 320

A2.2 Blueprints of consent forms for interviews with companies ... 321

A2.3 MAXQDA code system used for analysis of interviews ... 323

A2.4 Detailed description of factor model on ownership and sustainability ... 329

A2.4.1 Factors promoting ownership ... 329

A2.4.2 Factors promoting project sustainability ... 333

A2.5 Overview of data contradictions at each company by source of information and discrepancies in these numbers across sources ... 340

A2.6 Zooming in on companies' main reasons for (non-)participation in pilot of funding switch ... 353

A3 Appendix of the quantitative RtW evaluation ... 356

A3.1 Original RtW cohort differences-in-differences evaluation design ... 356

A3.2 Results from Double Machine Learning estimations ... 357

A3.3 Subsample estimations on income, wages, and working hours ... 359

A4 Appendix of the quantitative S4C evaluation ... 364

A4.1 Program impact for subsample of respondents who completed training ... 364

A4.2 Program heterogeneity with respect to endline in March 2020 ... 369

List of tables

Table 2.4.1 Total survey participation ... 35
Table 2.5.1 Sociodemographic characteristics of study participants, all TVETs ... 44
Table 2.5.2 Pre-intervention employment characteristics of study participants, all TVETs ... 46
Table 2.5.3 Pre-intervention employment characteristics of employed study participants, all TVETs ... 47
Table 2.6.1 Outreach of the KAM program ... 48
Table 2.7.1 Impact of KAM program on job interviews ... 59
Table 2.8.1 Impact of KAM program on current jobs ... 64
Table 3.4.1 Overview of factor model ... 83
Table 3.7.1 Comparison of ownership success factors to research findings on ownership ... 92
Table 3.8.1 Companies' reasons to pay stipends to interns ... 98
Table 3.8.2 Companies' reasons to not pay stipends to interns ... 100
Table 3.8.3 Companies' reasons for dissatisfaction with KAM interns ... 101
Table 3.8.4 Companies' reasons for satisfaction with KAM interns ... 102
Table 3.9.1 Comparison of sustainability success factors to research findings on sustainability ... 110
Table 4.4.1 Number of respondents by school type and treatment group ... 120
Table 4.5.1 Background characteristics of study participants ... 125
Table 4.5.2 Pre-intervention employment characteristics ... 127
Table 4.5.3 Background characteristics of study participants ... 128
Table 4.7.1 Average number of months of job search among those that searched at all ... 138
Table 5.4.1 Number of respondents by cohort and treatment group ... 166
Table 5.5.1 Sociodemographic background characteristics of study participants ... 172
Table 5.5.2 Employment status and earnings at registration ... 175
Table 5.8.1 Average duration of job search for those who searched at all ... 188
Table 6.3.1 Comparison of background characteristics of treatment and control group ... 243
Table 6.3.2 Information treatment effects on internal migration intentions ... 247
Table 6.3.3 Information treatment effects on destination choices measured by potential wage improvements ... 249
Table 6.3.4 Information treatment effects on other migration intentions outcomes ... 249
Table 6.3.5 Information treatment effects on internal migration intentions by S4C treatment group ... 250
Table 6.3.6 Information treatment effects on destination choices measured by potential wage improvements by S4C treatment group ... 251
Table 7.1.1 Value for Money Definitions ... 254
Table 7.4.1 The 4Es: concept and examples ... 258
Table 7.5.1 DFID's Value for Money enablers and the Value for Money cycle ... 260
Table 7.6.1 Tools for evaluating Value for Money ... 265
Table 7.7.1 Example calculation for inflation adjustment ... 274
Table 7.7.2 Sensitivity analyses of the cost-effectiveness ratio ... 275
Table 8.4.1 Employment multiplier estimates from the literature ... 285
Table 8.4.2 Calculation results of (economy-wide) employment effects ... 286
Table 9.1. Summary of the results of the quantitative evaluations of the KAM, RtW, and S4C programs ... 288

Table A1.1.1 Assumed thresholds of weekly working hours to define the working days ... 305
Table A1.1.2 Sociodemographic characteristics of study participants, mixed TVETs ... 308
Table A1.1.3 Pre-intervention employment characteristics of study participants, mixed TVETs ... 309
Table A1.1.4 Pre-intervention employment characteristics of employed study participants, mixed TVETs ... 309
Table A2.1.1 Referencing and overview of interviews ... 320
Table A2.5.1 Legend for Table A2.5.2 "Overview of data contradictions at each company by source of information and discrepancies in these numbers across sources" ... 340
Table A2.5.2 Overview of data contradictions at each company by source of information and discrepancies in these numbers across sources ... 341
Table A2.6.1 Overview of companies' main reasons to (not) participate in the new payment scheme ... 353

List of figures

Figure 2.4.1 Time gap between follow-ups and baseline by treatment and control group ... 36
Figure 2.4.2 Time gap between KAM start and baseline for treatment group ... 37
Figure 2.4.3 Time gap between placement end and survey wave for treatment group ... 37
Figure 2.4.4 Time gap between graduation and survey wave for control group ... 38
Figure 2.4.5 Number of observations by wave and study group, all TVETs ... 40
Figure 2.5.1 Number of study participants by region and by study group ... 42
Figure 2.5.2 Study participants by TVET course visited and study group ... 42
Figure 2.6.1 Placement share by gender and study group ... 49
Figure 2.6.2 Number of placement sections within internships by study group ... 49
Figure 2.6.3 Placement length by study group ... 50
Figure 2.6.4 Time gap between KAM start and placement start ... 50
Figure 2.6.5 Placement status by TVET course taken ... 51
Figure 2.6.6 Feedback of KAM participants on KAM program ... 52
Figure 2.6.7 Level of preparation for placement through the KAM training ... 52
Figure 2.6.8 Perceived employment opportunities after KAM program ... 53
Figure 2.6.9 Key goals at the time leaving the KAM program ... 54
Figure 2.6.10 Achievement of key goals ... 54
Figure 2.6.11 Indicators of satisfaction with placement (I) ... 55
Figure 2.6.12 Indicators of satisfaction with placement (II) ... 56
Figure 2.6.13 Reasons to recommend the KAM program ... 56
Figure 2.6.14 Reasons not to recommend the KAM program ... 57
Figure 2.6.15 Reasons for dropping out of the KAM program ... 57
Figure 2.7.1 Knowledge and usage of labor market matching services by study group ... 58
Figure 2.7.2 Participation in interviews at baseline by study group ... 59
Figure 2.8.1 Change in employment status by study group ... 61
Figure 2.8.2 Change in decent employment status by study group ... 61
Figure 2.8.3 Change in mean total income by study group ... 62
Figure 2.8.4 Change in mean hourly earnings among employed by study group ... 63
Figure 2.9.1 Estimated effects on employment indicator ... 66
Figure 2.9.2 Estimated effects on income and wages ... 67

Figure 2.9.3 Estimated effects on weekly working hours among employed 67
Figure 2.9.4 Estimated effects on job search 68
Figure 2.10.1 Estimated treatment effects on employment indicators, by gender 71
Figure 2.10.2 Estimated treatment effects on employment indicators, by age group..... 72
Figure 2.10.3 Estimated treatment effects on employment indicators, by experience 73
Figure 2.11.1 Estimated treatment effects on participants' aspirations 74
Figure 2.11.2 Estimated treatment effects on family structure 75
Figure 2.11.3 Estimated treatment effects on banking and savings behaviour 76
Figure 4.3.1 Map of Uganda with locations of study schools 115
Figure 4.4.1 Distribution of respondents across schools by treatment group 121
Figure 4.4.2 Distribution of respondents across courses by treatment group 122
Figure 4.4.3 Time between relevant program activities and date of survey 123
Figure 4.6.1 Beneficiaries' satisfaction with the RtW training, excluding the satisfaction with placement services and experiences 129
Figure 4.6.2 Most successful components if the RtW training 129
Figure 4.6.3 Beneficiaries' satisfaction with different aspects of the RtW training conditions 130
Figure 4.6.4 Percentage share of beneficiaries placed by treatment school and trades..... 131
Figure 4.6.5 Percentage share of beneficiaries placed by gender and age..... 132
Figure 4.6.6 Timing of UMA placement completion by school type..... 132
Figure 4.6.7 Timing of internships (facilitated through UMA and any others) that took place after graduation by treatment group and school type 133
Figure 4.6.8 Participation in employment promoting programs or activities (12.a) and the number of activities conditional on participation (12.b)..... 134
Figure 4.6.9 Type of employment promoting program or activity by treatment group 135
Figure 4.7.1 Number of interview invitations since 2019 by treatment group 136
Figure 4.7.2 Average percentage share of interview invitations that were attended by treatment group 136
Figure 4.7.3 Average success rates of attended interviews by treatment group 137
Figure 4.7.4 Number of months participants had searched for a job in the past year by treatment group and employment status..... 138
Figure 4.7.5 Most successful strategies of job search by treatment group..... 139
Figure 4.7.6 Main challenges of job search by treatment group 139
Figure 4.8.1 Employment status at the time of survey in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA 140
Figure 4.8.2 Number of jobs by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA 141
Figure 4.8.3 Monthly income and income among employed (in UGX) at the time of the survey in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA..... 142
Figure 4.8.4 Hourly wage and total hours worked per month among employed at the time of the survey in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA..... 143
Figure 4.8.5 Change in employment status between February 2019 and February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA..... 144
Figure 4.8.6 Mean monthly income and mean monthly income among employed before the RtW program in February 2019 and after the RtW program

in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA 145
Figure 4.8.7 Mean hourly wage and hours worked per month among employed before the RtW program in February 2019 and after the RtW program in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA 146
Figure 4.9.1 Estimated treatment effects on binary employment indicators..... 148
Figure 4.9.2 Estimated treatment effects on total income, income among employed and wages among employed. Income and wages were included as the natural logarithm of the respective income and wage measure 148
Figure 4.9.3 Estimated treatment effects on monthly working hours among employed respondents..... 149
Figure 4.10.1 Estimated treatment effects on employment by respondents' age 151
Figure 4.10.2 Estimated treatment effects on employment by respondents' prior work experience 151
Figure 4.10.3 Estimated treatment effects on employment by respondents' prior construction sector experience 152
Figure 4.10.4 Estimated treatment effects on employment by respondents' prior experience in electrical and mechanical work 153
Figure 4.10.5 Estimated treatment effects on employment by respondents' vocational training level 153
Figure 4.11.1 Response to the question in a) "How optimistic are you that you will have work in the short term, i.e., during the next 12 months?" and b) "How optimistic are you in finding work that satisfies your expectations in terms of tasks and salary in the long term, i.e., after 12 months from now?" by treatment group..... 154
Figure 4.11.2 Monthly income aspirations (in 1,000 UGX) for a job held a) in the next 12 months and b) after 12 months by treatment group..... 155
Figure 4.11.3 Estimated treatment effects on short-term (up to 12 months) and long-term (after 12 months) employment aspirations 156
Figure 4.11.4 Estimated treatment effects on short-term (up to 12 months) and long-term (after 12 months) income aspirations 156
Figure 4.11.5 International migration intentions by treatment group 157
Figure 4.11.6 Estimated treatment effects on migration intentions and plans to migrate in the next 12 months 157
Figure 5.4.1 Distribution of respondents in different treatment groups by level of education..... 167
Figure 5.4.2 Distribution of respondents in different training levels by level of education..... 168
Figure 5.4.3 Distribution of respondents by treatment group and region of residence 168
Figure 5.4.4 Distribution of respondents by training level and region of residence..... 169
Figure 5.4.5 Time between relevant program activities and date of the survey 170
Figure 5.6.1 Percentage share of beneficiaries placed by cohort and training site 177
Figure 5.6.2 S4C placements by training level..... 177
Figure 5.6.3 Percentage share of beneficiaries placed by level of education 178
Figure 5.6.4 Percentage share of beneficiaries placed by gender and age..... 178
Figure 5.6.5 Percentage share of beneficiaries who are still expecting to receive a placement by cohort and training site 179

Figure 5.6.6 Percentage share of beneficiaries who are still expecting to receive a placement by level of education ... 180
Figure 5.6.7 Percentage share of beneficiaries who are still expecting to receive a placement by gender and age ... 180
Figure 5.6.8 Timing of S4C internship placement ... 181
Figure 5.6.9 Timing of any internship ... 181
Figure 5.7.1 Participation in any employment promoting program among pure control group ... 182
Figure 5.7.2 Type of employment promoting program attended by control group ... 182
Figure 5.7.3 Types of skills trainings attended by control group ... 183
Figure 5.7.4 Duration of trainings attended by control group ... 183
Figure 5.7.5 Success of matching services ... 184
Figure 5.8.1 Invitation to job interviews ... 185
Figure 5.8.2 Number of interview invitations ... 185
Figure 5.8.3 Percentage share of attended interviews ... 186
Figure 5.8.4 Percentage share of attended interviews that led to a job ... 186
Figure 5.8.5 Length of job search by treatment group and employment status in November/December ... 187
Figure 5.8.6 Length of job search by training level and employment status in November/December ... 187
Figure 5.8.7 Methods of job search ... 188
Figure 5.8.8 Most successful methods of job search ... 189
Figure 5.8.9 Internet access for job search ... 189
Figure 5.8.10 Challenges in job search ... 190
Figure 5.8.11 Impacts of the COVID-19 pandemic on job search ... 191
Figure 5.9.1 Employment outcomes by treatment group ... 191
Figure 5.9.2 Employment outcomes by training level ... 192
Figure 5.9.3 Number of jobs by treatment group ... 192
Figure 5.9.4 Number of jobs by training level ... 193
Figure 5.9.5 Total monthly income by treatment group ... 193
Figure 5.9.6 Total monthly income by training level ... 194
Figure 5.9.7 Total monthly income among employed by treatment group ... 194
Figure 5.9.8 Total monthly income among employed by training level ... 195
Figure 5.9.9 Hourly wage among employed by treatment group ... 195
Figure 5.9.10 Hourly wage among employed by training level ... 196
Figure 5.9.11 Working hours among employed by treatment group ... 196
Figure 5.9.12 Working hours among employed by training level ... 197
Figure 5.9.13 Changes in employment status by treatment group ... 198
Figure 5.9.14 Changes in employment status by training level ... 198
Figure 5.9.15 Changes in income by treatment group ... 199
Figure 5.9.16 Changes in income by training level ... 199
Figure 5.9.17 Changes in total income among employed by treatment group ... 200
Figure 5.9.18 Changes in total income among employed by training level ... 200
Figure 5.9.19 Changes in hourly wage among employed by treatment group ... 201
Figure 5.9.20 Changes in hourly wage among employed by training level ... 201
Figure 5.9.21 Change in working hours among employed by treatment group ... 202
Figure 5.9.22 Change in working hours among employed by training level ... 202
Figure 5.9.23 Organization of work during the pandemic ... 203
Figure 5.10.1 Estimated treatment effects of S4C training and placement on binary employment indicators ... 205
Figure 5.10.2 Level 1 test scores and probability of level 2 participation ... 205

Figure 5.10.3 S4C level 1 test scores and employment outcomes in November/December ... 206
Figure 5.10.4 S4C level 1 test scores and employment outcomes in February/March ... 207
Figure 5.10.5 Estimated treatment effects of S4C level 2 training on binary outcome indicators ... 208
Figure 5.10.6 Estimated treatment effects of S4C training and placement on total income, income among employed and wages among employed ... 209
Figure 5.10.7 Estimated treatment effects of S4C training and placement on hours worked among employed ... 210
Figure 5.10.8 S4C level 1 test score, incomes, and wages in November ... 211
Figure 5.10.9 Level 1 test score, incomes, and wages in March ... 212
Figure 5.10.10 Estimated treatment effects of S4C level 2 training on incomes and wages ... 213
Figure 5.10.11 Level 1 test scores and monthly working hours ... 213
Figure 5.10.12 Estimated treatment effects of S4C level 2 training on monthly working hours ... 214
Figure 5.10.13 Estimated treatment effects of S4C training and placement on binary outcomes for alternative November endline ... 214
Figure 5.10.14 Estimated treatment effects of S4C level 2 training on binary indicators for alternative November endline ... 215
Figure 5.10.15 Estimated treatment effects of S4C training and placement on incomes and wages for alterna*tive November endline* ... 215
Figure 5.10.16 Estimated treatment effects of S4C level 2 training on incomes and wages for alternative November endline ... 216
Figure 5.11.1 Program effect heterogeneity on binary outcomes by respondents' gender in November ... 217
Figure 5.11.2 Program effect heterogeneity on incomes and wages by respondents' gender in November ... 217
Figure 5.11.3 Program effect heterogeneity on binary outcomes by respondents' age in November ... 218
Figure 5.11.4 Program effect heterogeneity on incomes and wages by respondents' age in November ... 218
Figure 5.11.5 Program effect heterogeneity on binary outcomes by respondents' vocational training experience in November ... 219
Figure 5.11.6 Program effect heterogeneity on incomes and wages by respondents' TVET experience in November ... 219
Figure 5.11.7 Program effect heterogeneity on binary outcomes by respondents' work experience in November ... 221
Figure 5.11.8 Program effect heterogeneity on incomes and wages by respondents' work experience in November ... 221
Figure 5.11.9 Program effect heterogeneity on binary outcomes by respondents' construction sector experience in November ... 222
Figure 5.11.10 Program effect heterogeneity on incomes and wages by respondents' construction sector experience in November ... 222
Figure 5.12.1 Optimism about employment in the long term by treatment group ... 223
Figure 5.12.2 Optimism about employment in the long term by training level ... 224
Figure 5.12.3 Income aspirations in the long term ... 224
Figure 5.12.4 Estimated treatment effects of the S4C program and placement component on labor market aspirations in the long run ... 225
Figure 5.12.5 Estimated treatment effects of S4C level 2 training on labor market aspirations in the long run ... 225

Figure 5.12.6 Internal migration intentions by treatment group 226
Figure 5.12.7 Internal migration intentions by training level 226
Figure 5.12.8 External migration intentions by treatment group..... 227
Figure 5.12.9 External migration intentions by training level..... 227
Figure 5.12.10 Estimated treatment effects of the S4C program and placement component on migration intentions 228
Figure 5.12.11 Estimated treatment effects of S4C level 2 training on migration intentions 228
Figure 6.2.1 Responses to the questions "To what extent, if at all, would you like to move to…" a) "… another region within Uganda to live?" and b) "… another country to live?" 232
Figure 6.2.2 Response to "What kind of planning or preparation have you done to move to another region within Uganda to live?" (multiple answers were possible) 232
Figure 6.2.3 Response to "What kind of planning or preparation have you done to move to another country to live?" (multiple answers were possible)..... 233
Figure 6.2.4 Response to „As far as you know, does any member of your household have plans to move to another region, district or country?" .. 233
Figure 6.2.5 Current region of residence (subfigure a) and first national destination choice (subfigure b)..... 234
Figure 1.2.6 Current region of residence (subfigure a) and first international destinations choice (subfigure b) 235
Figure 6.2.7 Response to "To what extent, if at all, would you move to another region or district within Uganda for work without having secured a job in that region or district yet?" 236
Figure 6.2.8 Response to "What is the smallest amount of money you would need to earn per month to move to another region or district within Uganda just for work?" 237
Figure 6.2.9 Extent to which respondents would like to move to the Albertine region to work in the oil and gas sector value chain by conditionality on whether they would earn the reservation wage to move to another region that was indicated in Figure 6.2.8..... 238
Figure 6.2.10 Response to question „ Why would you not like to move to the Albertine region to work in the oil and gas sector value chain? " (multiple responses were possible) 239
Figure 6.2.11 Response to question „To what extent, if at all, would you like to move to another district for a 3 to 6 months internship where you would get transport allowance, free meals and equipment?" 239
Figure 6.2.12 Response to question "Why would you not like to move to another district for an internship?" 240
Figure 6.3.1 Distribution of expected median monthly wage in main job by region of residence and sex of respondent..... 244
Figure 6.3.2 Expected and actual median monthly wage in main job by region of residence and sex of respondent 245
Figure 6.3.3 Percentage share of respondents who ranked regions by monthly median wage correctly..... 246
Figure 7.4.1 VfM Conceptual Framework (DFID, 2011) 256
Figure 8.2.1 Employment effects of infrastructure investments 279
Figure 8.3.1 Multistage job-creation effect of human capital development interventions 282
Figure A1.1.1 Number of study participants in baseline, by TVET 306

Figure A1.1.2 Number of observations by wave and study group, all TVETs but no CBET ... 307
Figure A1.1.3 Number of observations by wave and study group, only mixed TVETs and no CBET ... 307
Figure A1.1.4 Placement status by TVET ... 310
Figure A1.2.1 Estimated effects on employment indicators, no CBET participants ... 311
Figure A1.2.2 Estimated effects on employment indicators, no CBET participants and mixed TVETs only ... 311
Figure A1.2.3 Estimated effects on income and wages, no CBET participants ... 312
Figure A1.2.4 Estimated effects on income and wages, no CBET participants and mixed TVETs only ... 312
Figure A1.2.5 Estimated effects on working hours among employed ... 313
Figure A1.2.6 Estimated effects on job search, no CBET participants ... 313
Figure A1.2.7 Estimated effects on job search, no CBET participants and mixed TVETs only ... 314
Figure A1.3.1 Estimated treatment effects on income, by gender ... 315
Figure A1.3.2 Estimated treatment effects on income, by age group ... 315
Figure A1.3.3 Estimated treatment effects on income, by experience ... 316
Figure A1.3.4 Estimated treatment effects on working hours among employed, by gender ... 316
Figure A1.3.5 Estimated treatment effects on working hours among employed, by age group ... 317
Figure A1.3.6 Estimated treatment effects on working hours among employed, by experience ... 317
Figure A1.3.7 Estimated treatment effects on job search, by gender ... 318
Figure A1.3.8 Estimated treatment effects on job search, by age group ... 318
Figure A1.3.9 Estimated treatment effects on job search, by experience ... 319
Figure A2.3.1 Overview of MAXQDA code system used for analysis ... 323
Figure A2.3.2 Detailed MAXQDA code system used for analysis (1/5) ... 324
Figure A2.3.3 Detailed MAXQDA code system used for analysis (2/5) ... 325
Figure A2.3.4 Detailed MAXQDA code system used for analysis (3/5) ... 326
Figure A2.3.5 Detailed MAXQDA code system used for analysis (4/5) ... 327
Figure A2.3.6 Detailed MAXQDA code system used for analysis (5/5) ... 328
Figure A3.1.1 Set-up and timeline of the RtW cohort difference-in-differences design ... 356
Figure A3.1.2 Estimation strategy of the cohort difference-in-differences design ... 356
Figure A3.2.1 The effect of the RtW program on binary employment indicators using Double Machine Learning estimations ... 357
Figure A3.2.2 The effect of the RtW program on income and wages using Double Machine Learning estimations ... 357
Figure A3.2.3 The effect of the RtW program on working hours using Double Machine Learning estimations ... 358
Figure A3.3.1 The effect of the RtW program on income and wages by respondents' age ... 359
Figure A3.3.2 The effect of the RtW program on hours by respondents' age ... 359
Figure A3.3.3 The effect of the RtW program on income and wages by respondents' sustained work experience ... 360
Figure A3.3.4 The effect of the RtW program on hours by respondents' sustained work experience ... 360
Figure A3.3.5 The effect of the RtW program on income and wages by respondents' construction sector experience ... 361

Figure A3.3.6 The effect of the RtW program on hours by respondents' construction sector experience 361
Figure A3.3.7 The effect of the RtW program on income and wages by respondents' experience in electrical or mechanical work 362
Figure A3.3.8 The effect of the RtW program on hours by respondents' experience in electrical or mechanical work 362
Figure A3.3.9 The effect of the RtW program on income and wages by respondents' vocational training level 363
Figure A3.3.10 The effect of the RtW program on hours by respondents' s vocational training level 363
Figure A4.1.1 Estimated treatment effects of S4C program and placement component on binary outcomes (alternative treatment definition) 364
Figure A4.1.2 Estimated treatment effects of S4C level 2 training on binary outcomes (alternative treatment definition) 364
Figure A4.1.3 Estimated treatment effects of S4C program and placement component on incomes and wages (alternative treatment definition) 365
Figure A4.1.4 Estimated treatment effects of S4C program and placement component on working hours (alternative treatment definition) 365
Figure A4.1.5 Estimated treatment effects of S4C level 2 training on incomes and wages (alternative treatment definition) 366
Figure A4.1.6 Estimated treatment effects of S4C level 2 training on working hours (alternative treatment definition) 366
Figure A4.1.7 Estimated treatment effects of S4C program and placement component on labor market aspirations (alternative treatment definition) 367
Figure A4.1.8 Estimated treatment effects of S4C level 2 training on labor market aspirations (alternative treatment definition) 367
Figure A4.1.9 Estimated treatment effects of S4C program and placement component on migration intentions (alternative treatment definition) ... 368
Figure A4.1.10 Estimated treatment effects of S4C level 2 training on migration intentions (alternative treatment definition) 368
Figure A4.2.1 Program effect heterogeneity on binary outcomes by respondents' gender in March 369
Figure A4.2.2 Program effect heterogeneity on incomes and wages by respondents' gender in March 369
Figure A4.2.3 Program effect heterogeneity on binary outcomes by respondents' age in March 370
Figure A4.2.4 Program effect heterogeneity on incomes and wages by respondents' age in March 370
Figure A4.2.5 Program effect heterogeneity on binary outcomes by respondents' vocational training experience in March 371
Figure A4.2.6 Program effect heterogeneity on incomes and wages by respondents' vocational training experience in March 371
Figure A4.2.7 Program effect heterogeneity on binary outcomes by respondents' work experience in March 372
Figure A4.2.8 Program effect heterogeneity on incomes and wages by respondents' work experience in March 372
Figure A4.2.9 Program effect heterogeneity on binary outcomes by respondents' construction sector experience in March 373
Figure A4.2.10 Program effect heterogeneity on incomes and wages by respondents' construction sector experience in March 373

Executive Summary

There has been an increasing interest in rigorous impact evaluation of development activities, as evidenced by the decision of the Royal Swedish Academy of Sciences to award the Sverige Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2019 to development economists Abhijit Banerjee, Esther Duflo, and Michael Kremer. Donors and implementing organizations give increasing importance to impact evaluations for reasons of transparency, efficiency, and institutional learning. A main focus in recent years has been on the evaluation of employment promotion programs, because employment promotion and the improvement of employment conditions and labor income have been identified as important channels to alleviate poverty and increase welfare.

The **importance of employment in the global development agenda** is reflected in the World Bank's 2013 World Development Report on "Jobs" (World Bank, 2012) and the United Nations' Sustainable Development Goal (SDG) #8 "Decent work and economic growth". In German development cooperation, "Economic activity, trade and employment" is one of the three pillars of the "Marshall Plan with Africa" (BMZ, 2017). Therefore, employment promotion has also played an important role for many years in the activities of the implementing agency for development cooperation in Germany, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH. The sector project on "employment promotion in development cooperation" has been advancing the implementation of the German development cooperation agenda in partner countries worldwide for more than twenty years and has produced a series of studies addressing appropriate methodologies for impact evaluation (Kluve, 2011; Kluve and Stöterau, 2014; RWI, 2013, 2014) as well as a study putting these approaches into practice (RWI, 2019).

More recently (starting in 2015), **GIZ's Employment and Skills for Development in Africa (E4D) program**, commissioned by the German Federal Ministry for Economic Cooperation and Development and co-funded by the European Union, the Norwegian Agency for Development Cooperation (Norad), the Korean International Cooperation Agency (KOICA), the Department for International Development (DFID) of the United Kingdom and the companies Sasol, Quoniam Asset Management, and Shell, has been pushing employment promotion in six African partner countries. The E4D program closely cooperates with public and private sector partners to sustainably improve employment outcomes. Approaches comprise vocational education and training, non-cognitive or soft skills training, matching of job seekers and employers, entrepreneurial training, and enterprise development measures. In total, 74 projects have been implemented in Ghana, Mozambique, South Africa, Kenya, Uganda, and Tanzania by December 2020.

The main objective of this research project was to **develop and implement rigorous but practical solutions to evaluate the impacts of a selection of E4D programs on the target group's employment outcomes**. For this purpose, the researchers closely collaborated with local program managers, M&E teams, and relevant stakeholders. To this end, the results of this report provide **important learning outcomes for future rigorous impact evaluations within German development cooperation**.

Three projects of the E4D portfolio were selected for a quantitative evaluation of their impacts on employment outcomes: (1) the "Promoting Youth Employment Through Technical Human Capital Development" program in Kenya (henceforth **KAM program**, named after the implementing partner the Kenyan Manufacturers Association), (2) the "ReadyToWork" program in Uganda (henceforth **RtW program**), and (3) the "Skills for Construction" program in Uganda (henceforth **S4C program**). In addition to the three impact evaluations, a change in the **KAM program's internship stipend funding was studied qualitatively** and S4C survey data were used

to quantitatively **analyze participants' labor mobility intentions and behavior**. In the first half of 2019, mission trips to Kenya, South Africa, and Uganda were conducted, evaluation designs were developed, and costs budgeted. In the second half of 2019, data collections were planned and prepared. In 2020 data were collected and analyzed. In 2021 this report was drafted and finalized.

In addition to the data based evaluations, the report provides a discussion on **Value for Money**, including **a step-by-step guide for cost-effectiveness analyses** with an example analysis based on the KAM program, as well as a discussion on **indirect and induced employment effects**. The report concludes with **lessons learned for future program designs** drawn from the generated evidence and **lessons learned for future impact evaluations** drawn from the experiences of this research project.

The following provides a summary of the report:

1. **The quantitative evaluation of the KAM program examines the impact of a two-component youth employment promotion program in Kenya on employment and labor market outcomes among vocational training graduates**. The KAM program consisted of a two-to-three-days work readiness training and internship placement for a subset of trained beneficiaries. To disentangle the impact of the two program components, two research questions were studied:

   (RQ#1) What are the gains in employment and earnings for vocational training graduates from **participating in the work readiness training and internship placement** of the KAM program?

   (RQ#2) What are the gains in employment and earnings for vocational training graduates from an **internship placemen**t organized by KAM **in addition to participation in the work readiness training** organized by KAM?

   The evaluation uses graduates from no-intervention schools and graduates at intervention schools who did not choose to participate in the KAM program to construct a control group. The analyses apply linear multivariable regression models, which control for participants' background characteristics and baseline employment outcomes, to improve the comparability of individuals of treatment and control groups. The analysis is based on recurring survey data collected by— Innovations for Poverty Action Kenya between 2017 and 2019.

   Overall, the results show that:

   - Vocational training graduates who participated in the work readiness training and internship placement of the KAM program experienced a significant improvement in their labor market outcomes. Particularly striking are the large and persistent effects on decent and formal employment as well as on income. Moreover, results show an improved job search performance, a reduced financial dependency, and an increased probability of having a bank account.
   - The positive effects on labor market outcomes seem to be mainly driven by the work readiness training rather than the internship placement.

2. **The qualitative evaluation of the KAM program investigates how a change in the internship stipend funding from E4D to companies affected companies' ownership and sustainability of internship placements of the KAM program.** In 2017 and 2018, E4D

funded internship stipends for graduates placed in internships through the KAM program. In order to prepare the continuation of KAM program activities after E4D withdraws its funding and to test the implications of the funding withdrawal for the sustainability of the program and its benefits, a pilot program called 'sustainability program' was introduced in 2019 in one region of Kenya, under which the responsibility of stipend funding was transferred from E4D to companies who hired the interns. In all other regions than the pilot region, E4D continued the stipend funding. Based on qualitative semi-structured interviews with all stakeholders (companies, GIZ, and KAM) the following research questions were investigated:

(RQ#1) How has the switch in the internship payment scheme affected the **companies' ownership of internship placements** in the pilot region?

(RQ#2) How has the switch in the internship payment scheme affected **the sustainability of internship placements** in the pilot region?

(RQ#3) How can the **companies' ownership and the sustainability of the internship placements be enhanced**?

Overall, the results show that:

- All companies had full ownership of internship placements because they already had internship programs before E4D's KAM program started and E4D did not take control over the design and implementation of internships.
- The ownership of the change in stipend funding was mixed because the new funding scheme was not adequately introduced to all companies. Companies that were appropriately informed about the change in funding seem to have taken ownership, whereas companies that were not at all or not appropriately informed about the funding switch were more reluctant to take ownership.
- Only eight of 23 companies that were recruited to participate under the sustainability program actually participated and were willing to fund stipends. This resulted in fewer internship placements such that less graduates had the possibility to get industry exposure, training experience, and a chance to be retained in formal employment. The job retention rate of interns in participating companies remained on the same level in relative terms. Hence, overall, the funding switch had a negative influence on employment benefits.
- Willingness to pay KAM interns and satisfaction with KAM interns were the two central aspects determining whether a company sustained KAM internship placements without GIZ stipend funding.
- Potential strategies to enhance companies' ownership of such program changes are to (i) introduce changes in a more consistent and transparent manner, (ii) hold more awareness forums and facilitate more personal follow-ups, and (iii) actively involve companies in decision-making processes.
- Potential strategies to increase companies' willingness to pay internship stipends are (i) financial incentivisation and awareness raising about the economic benefits of paying interns and (ii) increasing companies' satisfaction with KAM interns by improving skills trainings and involving companies in skills trainings.

3. **The quantitative evaluation of the RtW program investigates the short-term impacts of a youth employment promotion program on employment and labor market outcomes among recent TVET graduates in Uganda.** RtW is a two-component program consisting of a two-day work readiness training and subsequent internship or job placements for a subset of trainees. To disentangle the impact of the two program components, two research questions were studied:

   (RQ#1) What are the gains in employment and earnings for vocational training graduates **from participating in the RtW training and internship placement** with UMA companies?

   (RQ#2) What are the gains in employment and earnings for vocational training graduates **from an internship placement** with UMA companies **in addition to participation in the RtW training**?

   The evaluation uses graduates from no-intervention schools and graduates at intervention schools who did not choose to participate in the RtW program to construct a control group. The analyses apply linear multivariable regression models, which control for participants' background characteristics and pre-intervention employment outcomes, to improve the comparability of individuals of treatment and control groups. The analysis is based on primary survey data collected in early 2020, two to 12 months after participants graduated from their vocational training institute. A follow-up survey planned for late 2020 was not conducted due to the coronavirus pandemic. The short term of follow-up data collection raises concern about whether any potential program effects could be detected at all and constitutes a main limitation of this study.

   Overall, the results show that:

   - The RtW training and placement program did not affect TVET graduates' employment probability or earnings in the short time considered.
   - Secondary outcomes, such as employment aspirations or migration intentions, were also not affected.

4. **The quantitative evaluation of the S4C program investigates the short to medium term impacts of a skills promoting program on employment and labor market outcomes among young people in Uganda**. The S4C program comprised a two-stage training (level 1 and level 2) and an internship or job placement for some trainees. The S4C level 1 training consisted of a 6-weeks training including general work readiness training, basic construction skills, and an international certification program in health, safety, and environmental standards. The level 2 training consisted of a 4-week training and certification in either rigging or pipe fitting. To disentangle the impact of the different program components, three research questions were studied:

   (RQ#1) What are the **overall effects on employment and earnings** of the S4C program (includes any combination of program components)?

   (RQ#2) What are the **gains in employment and earnings from a placement facilitated by S4C** in addition to participation in the S4C level 1 or level 1 and 2 training?

   (RQ#3) What are the **gains in employment and earnings from participating in the level 2 training** in addition to participation in the level 1 training?

The evaluation uses individuals who registered their interest in participating but who did not end up participating in the S4C program to construct a control group. The analyses of research questions RQ#1 and RQ#2 apply linear multivariable regression models, which control for participants' background characteristics and pre-intervention employment outcomes, to improve the comparability of individuals of treatment and control groups. For the evaluation of RQ#3, a fuzzy regression discontinuity design is used.

In summary, the results show that:

- The S4C program did not affect participants' employment probabilities, but it significantly improved the incomes of employed participants. The income increase appears to be driven by a mixture of higher wages and longer working hours.
- The S4C internship placements improved the probability of decent employment, formal employment, and employment with a contract of participants who were at least 25 years old, TVET educated, and previously had worked in a job for at least six months.
- The S4C program increased participants' employment aspirations and internal migration intentions.
- The S4C placement seems to be the most effective program component which drives the program's overall impact.

5. **S4C survey data were used to study the labor mobility intentions of the E4D target group and the impact of information about regional wage differentials on migration intentions in Uganda.** Many E4D projects offer training or work opportunities in locations where potential beneficiaries are not residing and some E4D projects use program applicants' commitment to travel or move to the training or work site as a eligibility criterion for participation. Hence, labor mobility constitutes a key assumption of E4D's theory of change. To gain a deeper understanding of participants' labor mobility the report provides **empirical evidence on participants' intentions to migrate internally and internationally** as well as the **fators that prevent them from migrating**. The evidence is drawn from the data collected as part of the S4C program in Uganda. The results show that almost 90% of S4C study participants are interested in moving to another region of Uganda or outside of Uganda. These numbers likely reflect participants' general interest towards migration as only one-third of participants have made preparations to put their interest into action. The most commonly preferred destinations within Uganda are Kampala and Western Uganda and the most commonly preferred international choices are countries in Europe and Northern Africa. Further, the report **investigates how information about regional wage differentials in Uganda affects S4C study participants' intentions to migrate** as misperceptions about wage differentials may distort migration intentions and behavior. In this vein, a random subset of S4C study participants was informed about the median monthly wages of each region of Uganda. Results show that many participants respond to the information by updating their wage expectations, migration intentions, and preferred destinations towards higher earnings destinations. While participants are generally willing to migrate, they will not move under all circumstances, but the decision heavily depends on wage levels and having a secured job at the destination. Providing information to potential beneficiaries about mobility requirements to participate in E4D programs, about locations or regions of the country where the trained skills are likely in demand, and about sector specific earnings in locations where the trained skills are in demand could enable

potential beneficiaries to better gauge the benefits from participating in the E4D program and, thus, may result in a more efficient targeting of E4D interventions and an increased intervention effectiveness.

6. **The report discusses the concept of Value for Money, provides a step-by-step guide for cost-effectiveness analyses using intervention net effects, and applies the step-by-step guide to the example of the KAM program.** Value for Money is about the balancing of resources and impact of the invested resources. Value for Money is not an ad-hoc evaluation of costs and impact but is implemented in daily activities of project-level agency staff and agency management. This is important because Value for Money is often confused with the assessment of Value for Money through cost-effectiveness analyses. Value for Money can be used to guide, steer, and justify individual development projects, specific country programs or whole agency portfolios. The most common conceptualizations of Value for Money follow **DFID's four E's framework – Economy, Efficiency, Effectiveness, and Equity –**, which consists of considerations that regard the economic spending on resources, the efficient conversion of inputs to outputs, the effectiveness of outputs in reaching outcomes (i.e., the impact of an intervention), and the equitable targeting of interventions in order to reach the poorest or most marginalized population groups (DFID, 2011). There are no one-fits-all guidelines on how to approach and implement the four E's framework as Value for Money depends on many complex program-, context-, and agency-specific conditions and procedures. A common tool to evaluate (not to implement) Value for Money is the cost-effectiveness analysis. It summarizes a complex intervention in a ratio of total impact to total costs and, hence, allows comparisons of interventions easily. The **steps to be followed to conduct a cost-effectiveness analysis are**:

   I. Identify competing interventions
   II. Identify the appropriate outcome
   III. Estimate the impact of the interventions, i.e. intervention net effects
   IV. The total impact of the intervention
   V. Gather cost data
   VI. Convert the costs into common units
   VII. Calculate the cost-effectiveness ratio
   VIII. Conduct sensitivity analysis that critically reviews and varies previously made assumptions

   To illustrate the computation of the cost-effectiveness ratio, the results of the quantitative KAM program evaluation were used. Depending on the assumptions made, the cost-effectiveness of the KAM program is 0.00021 to 0.00036 jobs per EUR invested or 2,778 to 4,762 EUR per job. Overall, the implementation of Value for Money requires considerable efforts at various levels of the agency and project stages. Cost-effectiveness analyses require more detailed cost reporting than currently conducted and rigorous impact evaluations to learn about the impact of an intervention.

7. **The report discusses a conceptualization of indirect and induced employment effects for human capital-focused interventions and calculates upper and lower employment effect bounds for the KAM and S4C programs based on substitution, displacement, and multiplier effects from the literature**. Employment effects refer to economy-wide changes in employment due to an intervention or investment. In contrast, the impact

evaluations of the KAM, RtW, and S4C programs only measure direct employment effects. Although the benefits of measuring economy-wide effects are salient in theory, i.e., estimates of total net effects, in practice its measurement poses challenges that are difficult to overcome. Top-down approaches based on input-output tables (IOTs) or social accounting matrices (SAMs) commonly used in the employment effect literature, are often not suitable for applications in the context of human capital interventions, such as the skills trainings within the E4D intervention portfolio. Following Kluve and Stöterau (2014), the report discusses a conceptualization of indirect and induced employment effects for human capital-focused interventions. Indirect employment effects accrue to individuals in the target group who are non-beneficiaries and are triggered through direct employment effects, for example, via spillover or substitution effects. Induced employment effects accrue to individuals outside of the target group and are triggered through direct and indirect effects, for example, via displacement or consumption effects. An exemplary calculation of employment effects based on multipliers of human capital interventions from the literature is applied to the net direct employment effects estimated in the KAM and S4C evaluations. The calculation exercise serves the purpose of illustrating how multipliers could be applied, if available, and highlighting how much multipliers can vary and therefore, provide very uncertain and, to some extent, even arbitrary estimates of employment impacts.

8. **The report draws lessons learned for future program designs from the presented evidence.** The results of the three quantitative evaluations engender mixed implications for the effectiveness of specific program components. The results of the KAM program suggest that the skills training rather than the internship placement was effective in improving employment outcomes, whereas the S4C results imply the reverse. The results of the RtW program evaluation do not support either of the KAM and S4C findings. The results of this report and evidence from the literature suggest that the effectiveness of employment promotion programs or specific components thereof highly depend on the local context, program design aspects, and the target group (Betcherman et al., 2004; Blattman and Ralston, 2015; Kluve et al., 2017, 2019). A number of aspects that potentially influenced the respective programs effectiveness are: (a) the technical construction skills transferred in the S4C training may not (yet) be in demand by employers in places where trainees were seeking jobs; (b) KAM skills included organized opportunities for trainees and employers to meet and mingle, such as at job bazaars and networking forums; these networking opportunities potentially served as a platform to individually seek an internship or job placement; (c) the KAM training was delivered by a service provider whereas the RtW training was delivered through a training of teachers at schools, which could have influenced the training quality; (d) internship matching services were in place for the target group irrespective of E4D's KAM program such that the KAM program may not have relevantly added to that; (e) S4C placements were facilitated by a service provider whereas KAM and RtW placements were facilitated by manufacturing associations, which could have influenced the matching and internship quality; (f) the target groups differed across programs – the RtW program targeted final year TVET students, the KAM program TVET graduates from the past five years, and the S4C program targeted a broader group of young people. Based on these aspects and lending from the knowledge base of the literature, the following recommendations were derived:

I. A careful assessment of the context and the needs of the target groups should be conducted prior to the planning and implementation of employment promoting programs.
II. Project partners must have sufficient resources for program implementation and monitoring.
III. Quality assurance can be incorporated in the program design as project outputs. (This could, for example, be refresher trainings for trainers when interventions are delivered through a training of trainers.)
IV. Involve private sector in the design of the program to ensure context suitability and intervention quality.

9. **The report draws lessons learned for future impact evaluations of German development cooperation activities from the experiences made in the preparation, planning, and implementation of the impact evaluations presented in this report.** Rigorous evaluations of development projects require close collaboration between researchers and practitioners. Despite the overarching aim of employment promotion, and poverty reduction, each party typically follows different objectives, and accommodating all objectives is challenging. **Empathy for the respective other perspective and a strong collective effort have made the collaboration for the quantitative evaluations presented in this report a successful example, demonstrating that quantitative evaluations of policy-relevant projects can be brought into practice**. The following lessons learned are drawn from the experiences made in this project and are likely representative for impact evaluations of German development cooperation activities more generally:

I. Integration of project management and impact evaluation:
It is key for the planning and implementation of an intervention to go hand-in-hand with the impact evaluation from the start. The beginning of an impact evaluation is already in the conception phase of the intervention. To integrate project management and impact evaluation well, the practitioners' and researchers' perspectives should be explained and motivated to and discussed with the respective other party. Throughout the project, close collaboration and continual communication is required. The most important level of communication is between the local project manager and the researcher. This creates a workload for staff at the development cooperation agency that is on top of their daily activities. When impact evaluations are initiated, it is important to discuss how the responsible agency staff can manage the additional tasks and how other workloads can be reduced.

II. Specification of research questions:
The exact research interests need to be discussed ideally before an intervention is planned and implemented. If specific aspects of an intervention should be evaluated, then the impact evaluation needs to be set up in a very different way than when the overall impact was to be evaluated. The early commitment to research questions is also important to understand ex-ante what can be learned from rigorous impact evaluations. One single impact evaluation may evaluate whether a certain intervention or intervention aspect *was* effective or not, but it does not necessarily answer why it was effective. Further, the results of one study never provide an all-embracing answer to whether a specific type of intervention works in all contexts. Instead, one study can only add to the policy-relevant scientific literature, which is continually evolving as

more rigorous studies are being conducted and which as a whole can provide answers to certain questions.

III. **Timing of intervention implementation and data collection:**
In order to understand the extent to which treatment and control groups are comparable and where potential treatment effects are stemming from, it is important to hold as many aspects of the intervention implementation constant as possible. In an ideal scenario, the intervention delivery would take place simultaneously for all participants. Another important aspect poses the timing of data collections. Some research designs require the measurement of participants' characteristics prior to the implementation of the intervention. Further, sufficient time is required between the intervention completion and the follow-up data collection for treatment effects to unfold. It is important to allow the impact evaluation to continue beyond the project phase if needed, in order to adequately measure intervention impacts.

IV. **Sample size:**
Reliable results hinge on large sample sizes. The relevant sample size is the one of the final estimation sample. If the impact of specific population groups or program components should be evaluated, then the estimation sample size in the relevant group of beneficiaries is the relevant sample size.

Many of these lessons learned are no news but have been stated in previous RWI evaluation reports and elsewhere (e.g., Bachmann et al., 2019). The repeated efforts to bring rigorous impact evaluation into practice are commendable. They highlight that mainstreaming rigorous impact evaluation in German development work is a long and complex path. But they also highlight that the **process of mainstreaming rigorous impact evaluation is underway and is enhancing through continual learning of researchers and practitioners under improving institutional parameters**.

1 Introduction

As evidenced by the Royal Swedish Academy of Sciences' decision to award the Sverige Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2019 to development economists Abhijit Banerjee, Esther Duflo, and Michael Kremer, there has recently been an increasing interest in rigorous impact evaluations of development activities. Donors and implementing organizations give increasing importance to impact evaluations for reasons of transparency, efficiency, and institutional learning. A main focus in recent years has been on the evaluation of employment promoting programs, because employment promotion and the improvement of employment conditions and labor income have been identified as important channels to alleviate poverty and increase welfare.

The importance of employment in the global development agenda is reflected in the World Bank's 2013 World Development Report on "Jobs" (World Bank, 2012) and the United Nations' Sustainable Development Goal (SDG) #8 "Decent work and economic growth". In German development cooperation, "Economic activity, trade and employment" is one of the three pillars of the "Marshall Plan with Africa" (BMZ, 2017). Therefore, employment promotion has also played an important role for many years in the activities of the implementing agency for development cooperation in Germany, GIZ. The sector project on "employment promotion in development cooperation" has been advancing the implementation of the German development cooperation agenda in partner countries worldwide for more than twenty years and has produced a series of studies addressing appropriate methodologies for impact evaluation (Kluve, 2011; Kluve and Stöterau, 2014; RWI, 2013, 2014) as well as a study putting these approaches into practice (RWI, 2019).

More recently (starting in 2015), GIZ's Employment and Skills for Development in Africa (E4D) program, commissioned by the German Federal Ministry for Economic Cooperation and Development and co-funded by the European Union, the Norwegian Agency for Development Cooperation (Norad), the Korean International Cooperation Agency (KOICA), the Department for International Development (DFID) of the United Kingdom and the companies Sasol, Quoniam Asset Management, and Shell, has been pushing employment promotion in six African partner countries. The E4D program closely cooperates with public and private sector partners to sustainably improve employment outcomes. Approaches comprise vocational education and training, non-cognitive or soft skills training, matching of job seekers and employers, entrepreneurial training, and enterprise development measures. In total, 74 projects have been implemented in Ghana, Mozambique, South Africa, Kenya, Uganda, and Tanzania by December 2020. The objective of this research project is to rigorously evaluate the impacts of a selection of E4D projects on the target group's employment outcomes.

The selection of projects followed two broad criteria categories: (1) learning and (2) methodological requirements. For the category of learning the criteria included whether the project consisted of an E4D-typical intervention that is implemented in similar ways across E4D countries and is likely to be implemented in future project phases. The methodological criteria included the possibility to form comparison groups that allow to model counterfactuals, the number of beneficiaries for statistical power, the timing of intervention implementation, and, following from this, the possible timelines of data collections and data analysis.

Three projects of the E4D portfolio were selected for a quantitative evaluation of their impacts on employment outcomes: (1) the "Promoting Youth Employment Through Technical Human Capital Development" program in Kenya (henceforth KAM program, named after the implementing partner the Kenyan Manufacturers Association), (2) the "ReadyToWork" program in Uganda

(henceforth RtW program), and (3) the "Skills for Construction" program in Uganda (henceforth S4C program). All three programs follow a similar structure, comprising two components – a skills training and, for a subset of participants, an internship placement. These interventions are typical for skills development interventions within the E4D portfolio, were prolonged in the subsequent program phase which started in 2020, and, because they were implemented in different countries, the evaluation results can speak to the external validity of the research findings.

For the selection of projects, mission trips to Kenya, Uganda, and South Africa were conducted in early 2019 where country portfolios were reviewed and specific program details gathered in alignment with the selection criteria. Subsequently, research questions were specified, evaluation designs developed, and costs budgeted. Following this, the KAM, RtW, and S4C projects were selected. Because the quantitative evaluation of one research question for the KAM program turned out not to be feasible, a qualitative research design was developed and implemented instead. The KAM program evaluation included the implementation years of 2017 to 2019, the RtW program of 2019, and the S4C program the years of 2019 and 2020. For the KAM program, data collected by Innovations for Poverty Action Kenya were used for the quantitative evaluation. For the RtW and S4C programs primary data were collected in early and late 2020, respectively. In the second half of 2020 and the first half of 2021, the analyses and reporting were conducted.

In addition to the evaluations of selected E4D programs, the S4C survey was used to study participants' labor mobility intentions and behavior. For many E4D programs labor mobility constitutes a key assumption of it's theory of change. To gain a deeper understanding of participants' labor mobility this report provides descriptive evidence on participants' intentions to migrate internally and internationally. Further, the report investigates how information on wage differentials across regions affects S4C study participants' intentions to migrate as misperceptions about wage differentials may distort migration intentions and behavior.

The report further discusses the concepts of Value for Money and indirect and induced employment effects. Based on employment multipliers from the literature, upper and lower employment effect bounds that take into account substitution, displacement, and multiplier effects are calculated at the example of the KAM program. Finally, a cost-effectiveness analysis of the KAM program is conducted.

The report is structured as follows. Sections 2 and 3 present the qualitative and quantitative evaluation results of the KAM program in Kenya. Sections 4 and 5 present the quantitative evaluation results of the RtW and S4C programs in Uganda. Each of these sections explains the content of the respective program, the research methodology, and the process of data collection. For the quantitative evaluations, the main focus is on the presentation of descriptive results on employment outcomes and estimates of the programs' impact on employment outcomes. Section 6 presents the results of the labor mobility analysis and the impact of wage differentials on migration intentions. Section 7 discusses Value for Money and section 8 indirect and induced employment effects. Section 9 draws lessons learned from the evidence generated in this report for the design of future E4D programs. Section 10 draws lessons learned for future impact evaluations from the experiences made in the preparation planning and implementation of the impact evaluations presented in this report.

2 Quantitative Evaluation of the KAM program

2.1 The KAM program

The *Promoting Youth Employment Through Technical Human Capital Development* program is an employment promoting program within the E4D initiative implemented by the Kenyan Association of Manufactures (KAM) in collaboration with national training providers and member companies. KAM is an association that represents the manufacturing and value adding industries. Hereafter the program is referred to as *KAM program*. It aims at improving access to jobs and economic opportunities for youth as well as at increasing jobs in the manufacturing sector. According to KAM, the program addresses the future of the manufacturing sector by ensuring that there are skilled human resources in the trades, and it tackles the high unemployment rate in Kenya which is driven by a lack of experience and practical skills of job starters. The program consists of two major program components. First, in a two-to-three-days training youths undergo a work readiness and mentorship workshop and obtain the opportunity to create linkages with trades through job bazaars, HR practitioners networking forums, industry leader forums, as well as Technical Vocational and Education Training (TVET) consultative forums. Subsequently, participants are being placed in internships in manufacturing companies which are members of KAM for a period of 3 to 6 months in which they are attached to a mentor. Similar work-readiness programs have also been implemented in Mozambique, Tanzania, and Uganda.

The KAM program was launched in October 2017 and ran until October 2019. It targeted TVET certificate or diploma graduates of selected trades who graduated in the past 5 years.[1] KAM facilitated the implementation of the project by reaching out to TVET institutions all over Kenya to identify graduates of the selected trades who are (i) not older than 24 years, (ii) willing to move within Kenya, and (iii) match the companies' needs. Interested and eligible graduates were then trained in a two-to-three-days work-readiness training. These work-readiness trainings were conducted by third party training providers and included soft- and life-skills training and job application training. At the last day of the training, participants could attend a job bazar or similar events where company representatives already had the opportunity to interview participants. After the work readiness trainings, KAM matched the graduates with companies for internships, either based on the companies' preferences indicated after the job bazars or by directly sending CVs of potential candidates to companies. Upon completion of an internship period of about three months, the companies could decide to terminate the internship or to extend it for another three months. After the end of the internship, companies could offer a regular employment position to the intern.

The needs-based approach of the matching process had multiple implications for the selection of beneficiaries as well as the timing of the implementation of program components. First, not all program participants that completed the KAM work-readiness training also received an internship placement through KAM because not everyone was offered a position and some who have been offered a placement rejected it. Second and strongly related to the first point, the allocation of internships among KAM participants did not occur randomly but was based on participants' merit and skills. Therefore, participants with better labor market prospects were more likely to be placed or at least to receive their placement earlier than those with lower prospects.

[1] *The program covered the following trades: heavy and light machinery operating, welding, electrical (installation), electronics (instrumentation), mechanical technology and maintenance, construction (masonry and concrete works), carpentry, pipefitting and plumbing, as well as industrial painting and brushing.*

Third, the length of internships and, therefore, the time between completing the internship and being interviewed for this study, during which potential program impacts could unfold, differed across participants. Points two and three indicate that there exists considerable heterogeneity in the intensity of the program as well as the time of program completion. A more detailed description of the sample and timing of the internships as well as how these aspects were addressed in the analysis follow in sections 4 and 5.

Throughout this report section the two-to-three-days' work readiness trainings are referred to as the *KAM training* and the internship placements facilitated through KAM as *internship placements*. Internship placements sought individually or by other organizations are referred to as *self-sought internships.* The overall program, including both components is referred to as the *KAM program*, although not everyone who participated in the KAM program necessarily has completed the internship. Therefore, we will use the following terms to differentiate between the different groups of study participants: (i) KAM beneficiaries who participated in the training and received a KAM placement are henceforth called *training and placement group*, (ii) participants who only participated in the KAM training but did not receive a KAM placement are called *training only group*, (iii) and control group participants who did not receive any KAM benefits are referred to as *control group*.

2.2 Research questions

The following evaluation aims at identifying and understanding the effects of the KAM program within the E4D project on employment and welfare of KAM beneficiaries over time. Since the KAM program comprised two components – the first consisting of the KAM work readiness training and the second being the KAM organized placement in companies – the analysis of this research requires to consider the potentially differing effect mechanisms of these two components.

As described in section 2.1, the KAM work readiness training lasted only two days whereas the internship placement was planned for 3 months. Effectively the internships lasted between 1 week up to 15 months and had an average length of about 3.5 months. The internship placement linked beneficiaries to potential employers and enabled to gain practical skills and experience in the respective trade. Thereby, the internships are tackling the so-called *skills gap,* which is considered a major challenge for youths in Eastern Africa, by bridging the skills that are demanded by employers and offered by young graduates. Therefore, the main impact on employment and labor market outcomes was, a priori, expected to primarily arise from the internship placement component. However, during the work-readiness training KAM beneficiaries could also get in touch with human resource officers from different companies. This could have had a lasting effect even for KAM beneficiaries who did not receive an internship placement through KAM, for example, through an increased professional network or by obtaining an internship on their own initiative during the networking sessions.

In order to disentangle the impact of the two program components, the following research questions were studied:

(RQ#1) What are the gains in employment and earnings for vocational training graduates from **participating in the work readiness training and internship placement** of the KAM program?

(RQ#2) What are the gains in employment and earnings for vocational training graduates from an **internship placemen**t organized by KAM **in addition to participation in the work readiness training** organized by KAM?

RQ#1 addresses the effect of receiving both program components in comparison to not participating in the KAM program at all. It does so by comparing survey participants who participated in the KAM training and were afterwards placed by KAM in a company for an internship to survey participants who did not receive any benefits from KAM. This design disregards the effect of only participating in the KAM training as well as the effect of participating in the KAM training and afterwards completing a self-sought company placement. Therefore, RQ#1 should not be understood as the overall KAM program impact. In order to obtain the overall KAM program impact, one would need to compare KAM beneficiaries who received any of the two program components to survey participants who did not receive any benefits from KAM.

RQ#2 aims at isolating the placement effect from the KAM training effect and to measure the additional impact of participating in an internship placement beyond the potential benefits from the KAM training alone. To study this research question, survey participants who participated in the KAM training and were afterwards placed by KAM in an internship are compared with survey participants who only benefited from the KAM training. For this research question it is important to acknowledge two important aspects.

First, the additional gain from internship placements, i.e., first participating in the KAM training and subsequently completing an internship, might be different to the pure internship placement effect, i.e., only completing an internship without previously being trained. The reason for this is that interaction effects are likely to take place between the training and the placement which is also the reason why the program was designed in this two-component structure in the first place. One could consider the following example: Individuals who have previously been trained in work ethics are more likely to also have followed these ethics and have behaved in a respective manner during their internships in comparison to individuals who have not been trained in work ethics. Consequently, the probability of being retained because the employers appreciated the interns' behaviour may be higher for those individuals who received the training prior to the internship.

Second, not having received a KAM internship placement does not exclude the possibility that individuals obtained self-sought internships after having participated in the KAM training. Hence, the second research question does not estimate the additional effect of receiving any internship placement compared to only participating in the KAM training. Instead, the research design evaluates the effect of receiving a KAM organized internship placement compared to participating only in the KAM training or participating in the KAM training and undergoing a non-KAM organized internship.

2.3 Research design

2.3.1 Estimation methods

For the purpose of evaluating the impacts of the KAM program on labor market outcomes, a reoccurring survey comprising a total of 2,114 participants was conducted over the years 2017 to 2019. The survey included individuals who after their graduation from TVET enrolled in the KAM program as well as individuals who are still enrolled in or already graduated from TVET but who never participated in the KAM program. The evaluation sample consists of a total of three main groups:

1. **beneficiaries who graduated from a TVET in the last 5 years and participated in the KAM training and received a company placement organized by KAM[2],**
2. **beneficiaries who graduated from a TVET in the last 5 years and participated in the KAM training but who did not receive a company placement organized by KAM[3],**
3. **non-beneficiaries, who never participated in the KAM program and thus did not receive any of the KAM program's benefits[4].**

For the estimation of program impacts, the first group is considered the treatment group and the latter two groups constitute the comparison groups. Specifically, group two constitutes the comparison (or control) group for RQ#1 and group three constitutes the comparison group for RQ#2.

As described in section 1, participation in the KAM program did not follow a random allocation but eligible graduates could decide individually whether they wanted to enrol in the KAM program. Further, internship placements were driven by companies' needs. In consequence, KAM beneficiaries who were placed in internships, KAM beneficiaries who only participated in the training, and non-KAM beneficiaries may differ in ways that can also affect their job search and employment outcomes. Having as detailed knowledge as possible about such potential differences is of great importance when constructing an adequate counterfactual.

For example, and with regards to RQ#1, it seems plausible that graduates who already have a secured job after their graduation are less likely to enrol in programs like KAM, whereas graduates who are less certain about their employment prospects might be more likely to enrol to increase their employability. If the employment outcomes of these two groups were compared to each other, the measured program impact may be negative or at least downward biased just because beneficiaries and non-beneficiaries are not directly comparable to each other but differ with respect to relevant baseline characteristics. Considering RQ#2, differences in KAM beneficiaries' labor market prospects might influence who at the end received a KAM organized company placement and who did not, i.e., who only participated in the training. Hence, measured positive program effects may rather stem from the placed beneficiaries' labor market advantages than from KAM's positive program impact, at least partly. Failures to account for these inherent differences in treatment and control group participants can lead to biased impact estimation and misleading conclusions.

To make treatment and control groups comparable, linear multivariate regression models are used. Such models control for background characteristics of participants at the time of the baseline, which are expected to drive the decision to participate in KAM as well as subsequent labor market outcomes. These include demographic and socioeconomic status variables like age, gender, household size, marital status, respondents' dependency on the household head, whether respondents' have any children or dependents, as well as the highest school grade completed by the female household head. Moreover, the analyses also control for the respective outcome variable measured at the time of the first interview. Controlling for baseline outcomes (i.e., the

[2] *This group also includes KAM beneficiaries who received an internship placement through KAM but did not complete it.*

[3] *This group also includes KAM beneficiaries who participated in the KAM training but afterwards individually sought a company placement.*

[4] *Nevertheless, individuals of this group might still have participated in any other career services offered by other institutes and might have obtained a company placement either individually sought or sought by another institute.*

respective outcome at the time of the first interview) is similar to employing a difference-in-differences (DiD) approach and aims at making treatment and control group observations as comparable as possible. However, in the underlying setting, controlling for baseline outcomes within multivariate regression models can be expected to be more efficient than the traditional DiD approach. The traditional DiD approach relies on the assumption that baseline outcomes reflect participants labor market potential. In this application the assumption would, however, likely be violated because the study population at least partially consists of students whose baseline employment outcomes are not representative for their full labor market potential as they did not participate (full-time) in the labor market prior to the program. The estimation strategy that is applied in this evaluation does not rely solely on pre-intervention employment outcomes, but also controls for the sociodemographic background characteristics listed above.

Additionally, the analysis takes into account the different TVET courses in which participants were enrolled in. To control for any course specific characteristics which might influence the observed outcome variables TVET course fixed effects will be included in the multivariate regression models. The idea behind this approach is that the quality of TVET education and, hence, participants' future labor market prospects might vary across TVET courses. If at the same time treatment group participants were to be more frequently enrolled in low-quality TVET courses than control group participants to improve their, a priori, lower employability, the estimated program effects would underestimate the impact of the KAM program. Including fixed effects for the different TVET courses taken by the study participants (e.g., electrical engineering, building and construction, or carpentry and joinery) any time-consistent effect that these TVET courses might have on the observed outcome variables are controlled for.[5]

Moreover, to further reduce the probability that the programs visited by treatment and control group participants might not be comparable with each other the analysis is applied to two different sample compositions as robustness checks. First, control group participants who were enrolled in the CBET program will be excluded. Second, the sample will be restricted to study participants of mixed TVETs and again exclude those control group participants who were enrolled in the CBET program. Mixed TVETs are defined as TVETs for which the study sample contains both KAM beneficiaries as well as control group participants. TVETs for which the sample only covers KAM beneficiaries might be inherently different from TVETs for which the sample exclusively contains control group participants. Mixed TVETs and consequently also graduates from mixed TVETs are expected to be more comparable to each other. The estimation approach based on these alternative samples is further explained in section 2.9.4 and results are shown in Appendix A1.2.

2.3.2 Outcome indicators

Primary outcomes

The objective of the KAM program is to improve labor market prospects and outcomes of young Kenyans. Hence, this impact evaluation assesses direct employment benefits as primary outcome variables. These include employment status, type of employment (e.g., self-employment, formal employment, etc.) and earnings.

The different employment measures are dichotomous indicators (1 if employed, 0 if unemployed) and include the following outcomes:

[5] *Figure 2.5.7 provides an overview of the different TVET courses study participants were enrolled in and differentiates between treatment and control group participants.*

1. **Employment status,** which is defined as any paid work. It includes any kind of job indicated by the respondents.[6]
2. **Decent employment,** which is defined as paid work of at least 20 hours per week that earns at least 6,209.93 KES per month. It includes any kind of job indicated by the respondents.
3. **Self-employment**, which relies on the participants' report of whether they are self-employed.
4. **Formal employment,** which relies on the participants' report of whether their employment position is formal. The survey company Innovations for Poverty Action (IPA) described formal employment to be "more structured with a contract and clearly stipulated terms of service while informal employment has no set structures**".**
5. **Full time employment,** which relies on the participants' report on whether their employment position is full time.

The income measures considered are:

1. **Total monthly income** in KES, which is the sum of earnings from all paid work for participants who have paid work and zero for participant who are unemployed, i.e., do not have paid work.[7]
2. **Total monthly income among employed** in KES, which is the sum of earnings from all paid work for participants who have paid work (estimation sample excludes participants without paid work).
3. **Hourly wage among employed** in KES, which is the average hourly wage rate or earnings from paid work for those who have paid work (estimation sample excludes participants without paid work).

Secondary outcomes

In addition to employment and earnings, the following secondary outcome indicators were assessed:

Aspirations for further education:

– Aspirations to obtain further education is based on the enumerators' perceptions of whether respondents showed any interest of going back to school.

---

[6] *Respondents were asked about the number of jobs they held during the 12 months before joining the KAM program, since joining the KAM program, or since the last interview in the baseline, first follow-up, or any subsequent follow-up interview, respectively. For each indicated job additional questions about the job-specific income, working hours, tasks, etc. were asked. We do not know whether the indicated jobs were internships or traineeships. But we do know that each study participant who was identified to be employed indicated to have at least one job and have an income larger than zero.*

[7] *Study participants were only asked about their average earnings, weekly working hours, as well as the number of weeks worked in an average month. For their average earnings they could indicate a daily, weekly, monthly, or quarterly payment frequency. However, they were not asked about the number of days worked per week. Consequently, for participants who indicated to be paid on a daily basis we had to make assumptions regarding participants' average number of working days per week. Our main analysis is based on the assumption that an average workday has 12 working hours. The estimation results are almost identical when instead applying an assumption with lower daily working hours that also takes into account whether respondents work full or part-time.*

Employment aspirations:

- Employment aspirations in the short term indicates whether respondents would prefer working in an employed or salaried position rather than in self-employment immediately after the KAM program and is coded as 0 = self-employment and 1 = salaried position.
- Income aspirations in the short term, which is the response to the question "What is your expected monthly salary/income [in the short term]?", where short term refers to the time immediately after the KAM program.
- Employment aspirations in the long term indicates whether respondents would prefer working in an employed or salaried position rather than in self-employment one year after the KAM program and is coded as 0 = self-employment and 1 = salaried position.
- Income aspirations in the long term, which is the response to the question "What is your expected monthly salary/income [in the long term]?", where long term refers to the time one year after the KAM program.

Family structure:

- Respondents' marital status, which is a binary variable and is coded as 0 = never married/single and 1 = ever married.
- Binary variable indicating whether respondents have children, which is coded as 0 = no children and 1 = any children.
- Number of children respondents indicated to have at the time of the survey.
- Binary variable indicating whether apart from any own children respondents have dependants to take care of, which is coded as 0 = no dependants and 1 = any dependants.
- Number of dependants respondents indicated to have at the time of the survey.
- Respondents' dependency on the household head, which is a binary variable and is coded as 0 = respondent indicated to be independent from household head and 1 = respondent indicated to depend on household head.

Banking and savings behaviour:

- Binary variable indicating whether respondents have a bank account (0 = no, 1 = yes).
- Binary variable indicating whether respondents have any savings (0 = no, 1 = yes).
- Binary variable indicating whether respondents add money to their savings on a regular basis (0 = no, 1 = yes).
- Total amount of all personal savings in KES respondents indicated to have at the time of the survey.
- Binary variable indicating whether respondents received any loans since the baseline or the last follow-up interview (0 = no, 1 = yes).
- Total amount of loans received since the baseline or the last follow-up interview.
- Share of outstanding loans calculated as the number of loans for which the payment is currently ongoing or has not started yet over the total number of loans received in the respective time period.
- Binary variable indicating whether respondents gave any loans since the baseline or the last follow-up interview (0 = no, 1 = yes).

2.4 Data collection and sample

2.4.1 Data collection through IPA

E4D contracted with IPA-Kenya to develop and carry out a concept for the impact evaluation and monitoring of E4D programs in Kenya. In addition to survey data from E4D program beneficiaries, IPA-Kenya collected data from a comparison group consisting of students and graduates from TVET institutions who did not participate in E4D programs. The TVET institutions, of which comparison group members were recruited from, were selected based on offering courses in fields comparable to those covered by the KAM program. For the evaluation of the KAM program, this comparison group was used.

IPA Kenya collected data from program and comparison group members throughout 2017, 2018, and 2019 via standardized telephone interviews. The data collection includes a baseline survey, which was completed during the first one to three months of the program, as well as follow up surveys conducted at approximately six-month intervals after the baseline survey. The baseline survey focused on a 12 months recall period and data prior to starting the E4D program (IPA, 2020).

Table 2.4.1 presents the structure of the dataset as received by IPA Kenya. The first column refers to the total number of observations, resulting from the number of participants that are observed throughout the different survey waves as indicated in columns 2 to 5 of Table 2.4.1.

Table 2.4.1
Total survey participation

	N° of observations					N° of individuals		
		Survey waves covered						
	Total	Baseline	Follow-up 1	Follow-up 2	Follow-up 3	Treament	Control	Total
	1,440	x	x	x	x	360	-	360
	3,366	x	x	x		225	897	1,122
	54	x	x		x	18	-	18
	854	x	x			215	212	427
	32	x			x	16	-	16
	171	x				84	87	171
Treatment	2,715	918	818	585	394	918	-	-
Control	3,202	1,196	1,109	897	-	-	1,196	-
Total	**5,917**	2,114	1,927	1,482	394	-	-	**2,114**

Notes: For seven participants of the treatment group information is only available on follow-up 1, 8 treatment participants could not be reached, and 16 reached individuals of the treatment group never joined the program. Those observations are excluded. The control group was interviewed only three times. Therefore, a follow-up 3 is not available for the control. - Source: Own calculations based on KAM survey.

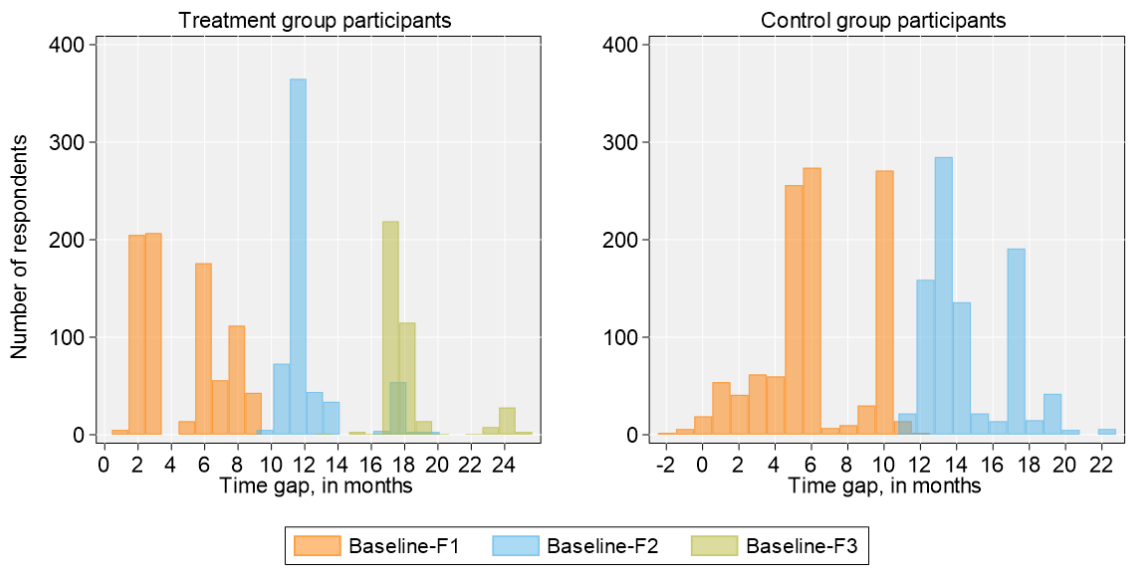
For example, 1,440 data points were collected in total from participants who were observed at each of the four surveys, whereas 3,366 data points were collected in total for participants who were observed only at baseline and follow-up 1 and 2. The last three columns refer to the number of participants included in the dataset and their treatment group affiliation. For example, the 3,366 data points belong to 225 participants of the treatment group and 897 participants of the control group who were observed at baseline, follow-up 1, and follow-up 2. The survey data

includes a total of 5,917 observations referring to 2,114 study participants who were interviewed at different points in time. There are more participants in the comparison (57%) than in the treatment group, i.e., KAM beneficiaries (43%). The columns *Baseline*, *Follow-up 1*, *Follow-up 2*, and *Follow-up 3* of Table 2.4.1 indicate in which survey the respective number of observations were collected or the respective participants were interviewed. Among KAM beneficiaries four survey waves were conducted although not all participants were interviewed in each wave. Participants of the comparison group participated in only three surveys. The sample attrition from baseline to the second follow-up (F2) is noticeably smaller among the comparison group than among the treatment group. Among KAM participants only 64% of the people interviewed at baseline were followed-up in F2, for the comparison group it was 75% of the baseline participants.

2.4.2 Timing of the data collection

The baseline survey was planned to take place during the first one to three months of the KAM program and the follow-up surveys were scheduled in intervals of approximately six-months after the baseline survey. However, this was only partly achieved. Figure 2.4.1 displays the time gap between the follow-up surveys and the baseline for treatment (left) and control (right) group and reveals that many respondents of the treatment group were called for the second interview (F1) less than 3 months after baseline. Further, whereas for the treatment group the time gap between the baseline and the third interview (F2) varies between 10 and 14 months, it varies between 11 and 22 months among control group members.

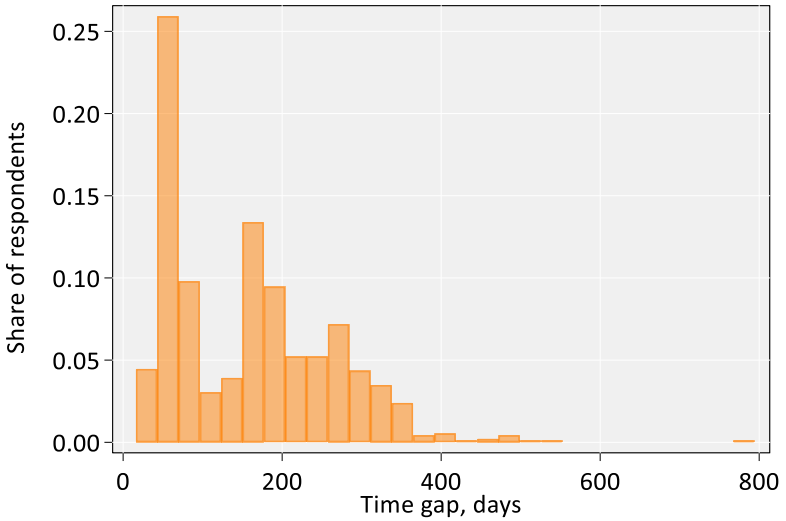
Figure 2.4.1
Time gap between follow-ups and baseline by treatment and control group



Notes: The graphs indicate the number of months between the baseline and follow-up interviews. The left graph refers to interviews conducted with the treatment group and the right to the ones conducted with the control group. - Source: Own calculations based on KAM survey.

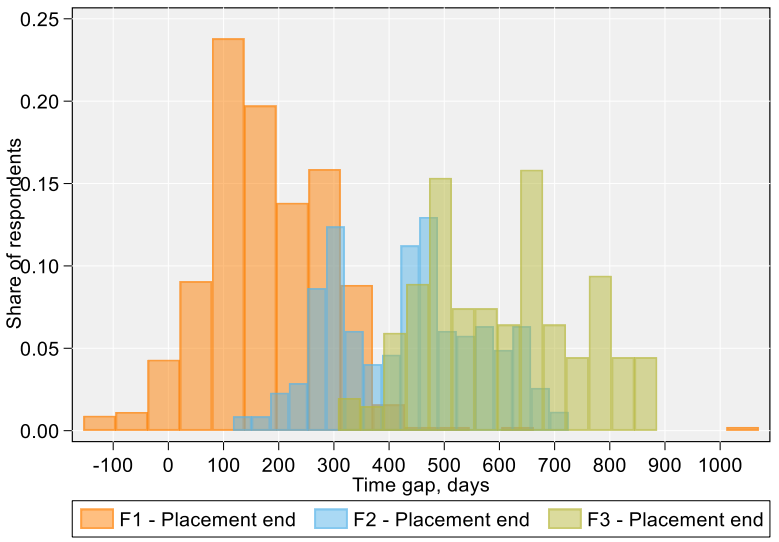
Figure 2.4.2 shows the time in days between joining the KAM supported program and the baseline survey for treatment group members. The baseline survey took place between one and 794 days after joining the KAM supported program. The distribution in Figure 2.4.2 reveals that many study participants were interviewed much later than the originally planned three months after the baseline.

Figure 2.4.2
Time gap between KAM start and baseline for treatment group



Notes: The graph shows the time gap (in days) between KAM start (work readiness training) and the interview date of the baseline survey for treatment participants. - Source: Own calculations based on KAM survey.

Figure 2.4.3
Time gap between placement end and survey wave for treatment group

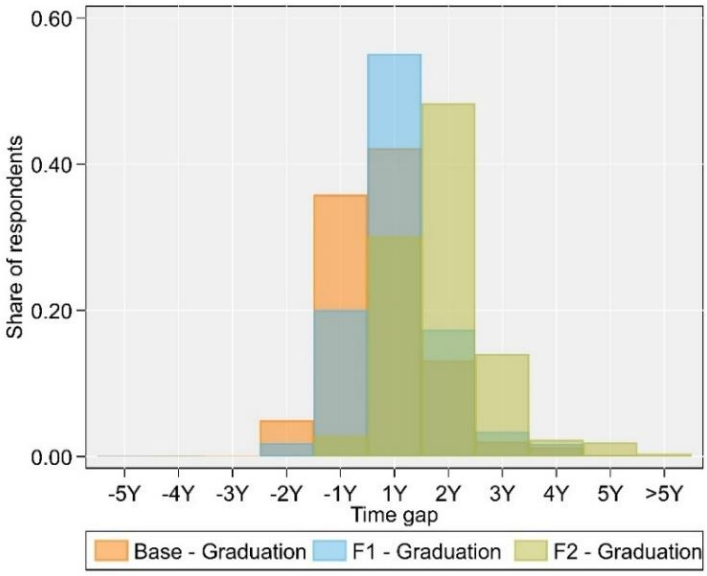


Notes: The graph shows the time gap (in days) between placement end and the interview date of the follow-up surveys for treatment participants. The information about the placement end was only collected in follow-up 1. - Source: Own calculations based on KAM survey.

Figure 2.4.3 presents the time gap in days between program completion, i.e., having finished the placement, and each follow-up survey for treatment group members and, similarly to Figure 2.4.2, shows that this time gap varies considerably across respondents and does not correspond to the planned six months intervals. For some KAM participants the first follow-up took place one year after having completed the KAM internship while at this time other KAM participants already completed their third follow-up.

The surveys took place at different stages of treatment and control group participants' education. Based on the program's eligibility criteria, we know that KAM beneficiaries, i.e., the treatment group, graduated from their TVET program in the past 5 years. Unfortunately, the exact graduation dates of KAM beneficiaries are unknown because the questionnaires only included this information for the control group. For the control group, the Figure 2.4.4 shows that a large number of participants were still enrolled in their TVET program at the time of the baseline survey and some of them were still enrolled even during the follow-up surveys, which can be detected by the negative values in the graph.

Figure 2.4.4
Time gap between graduation and survey wave for control group



Notes: The graph shows the time gap (in days) between the control participants' day of graduation and their respective interview dates of the baseline and the follow-us. The displayed bins cover observations by 365 days. Consequently, 1Y refers to having had the graduation between 0 and up to 365 days prior to the interview, whereas -1Y refers to having had the graduation between 1 and up to 365 days after the interview. - Source: Own calculations based on KAM survey.

2.4.3 Data specificities and modifications

There are some specificities about the structure and information content of the data that require attention with respect to the design and implementation of the impact evaluation:

1. The treatment group was surveyed four times and the control group was surveyed only three times.
2. The timing of the survey waves strongly varies across and within treatment and control group observations.

3. Treatment and control group participants were surveyed at different stages of their education. As stated in section 2.4.1, IPA Kenya collected data from a comparison group consisting of students and graduates from TVET institutions who did not participate in E4D programs and were offering courses in fields comparable to those of the KAM program. Some of the participants of the comparison group were part of alternative programs, for example the National Competence Based Education and Training program (CBET). Whereas KAM addresses individuals who graduated during the past 5 years, some participants of the comparison group were still about to graduate (see also Figure 2.4.4).
4. The questionnaires of the treatment and control groups and across survey waves differed in questions that are relevant for the impact evaluation. For example, treatment group participants' educational attainment and year of graduation is unknown and, therefore, we cannot make any conclusive statement about whether treatment and control individuals are comparable regarding this fundamental dimension.
5. Some important information was missing overall. For example, the number of hours worked per day is missing. Since we only know the days and hours worked per week this required strong assumptions in order to calculate an approximation of total earnings for subjects who indicated to be paid per day.[8]
6. The baseline survey was conducted after the KAM program started and parts of the information collected were not enquired using recall methods; this includes, for example respondents' household characteristics (e.g., number of children, size, assets), aspirations, or savings behaviour.
7. The data were partially inconsistent and required specific detail and sometimes even potentially arbitrary decisions in the data cleaning process. For example, in more than 1,000 cases respondents indicated to be unemployed but also indicated a positive number of jobs, wage income, or hours worked. For another 141 observations respondents indicated to be employed but the number of jobs hold is 0. Further inconsistencies were present with respect to the date at which individuals (i) started the KAM supported program, (ii) completed the work readiness program, (iii) started their internship placement, and (iv) ended their placement. For more than 300 KAM beneficiaries the recorded KAM start date is after the KAM end date and 81 KAM beneficiaries indicated a placement start date that was before their previously indicated KAM program start date. Lacking reliable information on the time between start and end dates of the different program components and the survey waves limits the understanding and evaluation of the potential treatment effects.

These aspects of the data make it challenging to study and evaluate the similarity, and therefore comparability, of the treatment and control groups, which is key for causally attributing changes in labor market outcomes to the KAM program.

To address some of the before mentioned issues, a number of important data manipulations were conducted. First, the original grouping of follow-up surveys (F1, F2, and F3) was rearranged

[8] *We use two different approaches. The first approach applies a fixed definition of 12 working hours per day. The second approach is more flexible and takes into account the distribution of the observations as well as the additional information on whether the job is fulltime. The exact thresholds of the second approach are displayed in Table A1.1.1.*

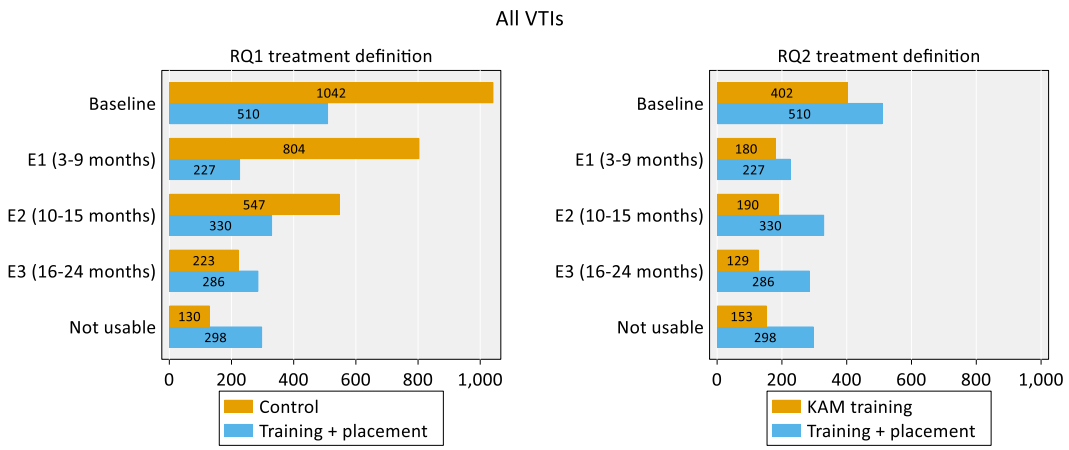
to three synthetic follow-up survey waves based on the time span between the baseline interview date and the respective follow-up interview date and irrespective of the original survey wave. If an individual was observed twice within one of these new follow-up periods the interview that was conducted closer to the originally planned timing (6, 12, or 18 months after baseline) was used.

1. Endline 1 (E1) captures follow-up observations that took place close to 6 months after the baseline interview (caliper between 3 and 9 months after baseline).
2. Endline 2 (E2) captures follow-up observations that took place close to 12 months after the baseline interview (caliper between 10 and 15 months after baseline).
3. Endline 3 (E3) captures follow-up observations that took place close to 18 months after the baseline interview (caliper between 16 and 24 months after baseline).

Second, control group participants whose TVET graduation took place more than 6 months after the baseline survey were dropped from the analysis to make control and treatment group observations more comparable with respect to their graduation date. This excludes 154 control group participants and a total of 455 control observations.

Figure 2.4.5 displays the number of treatment and control observations for the two research questions after (i) modifying the survey waves, (ii) dropping control group observations with very late TVET graduation dates, and (iii) applying our definitions of treatment and control group participants as explained in section 2.3.1 for the total sample.[9]

Figure 2.4.5
Number of observations by wave and study group, all TVETs



Notes: The graphics show the number of observations by treatment and control group for the baseline survey and each modified follow-up wave. The left graphic displays the number of observations by treatment and control group according to the RQ#1 definition. The right graphic displays the number of observations by treatment and control group according to the RQ#2 definition. "Not usable" refers to interviews conducted less than 3 months after the baseline or individuals who were interviewed twice within the same caliper. - Source: Own calculations based on KAM survey.

[9] *An overview on the number of observations for the two alternative analyses, i.e., (i) the total sample excluding CBET control participants and (ii) the subsample of mixed TVETs excluding CBET control participants, is displayed in Figures A1.1.2 and A1.1.3.*

2.5 Background characteristics of study participants

This section describes individual and household level characteristics of study participants across treatment and control groups with respect to participants' sociodemographic background and pre-intervention employment characteristics.

The pre-intervention employment indicators were assessed using baseline survey data and were enquired via a set of recall questions for which respondents described their employment situation 12 months prior to joining the KAM program for KAM beneficiaries or 12 months prior to the baseline interview for control group participants.[10] In contrast, recall methods were not used for questions regarding participants' sociodemographic background characteristics. Since the baseline survey was conducted after KAM beneficiaries already received the KAM training, the participants' economic situation prior to the KAM program is unknown. Therefore, information referring to the time of the baseline survey could, in principle, already been affected by the KAM training, however, changes due to program participation can be expected to be negligible at this point.

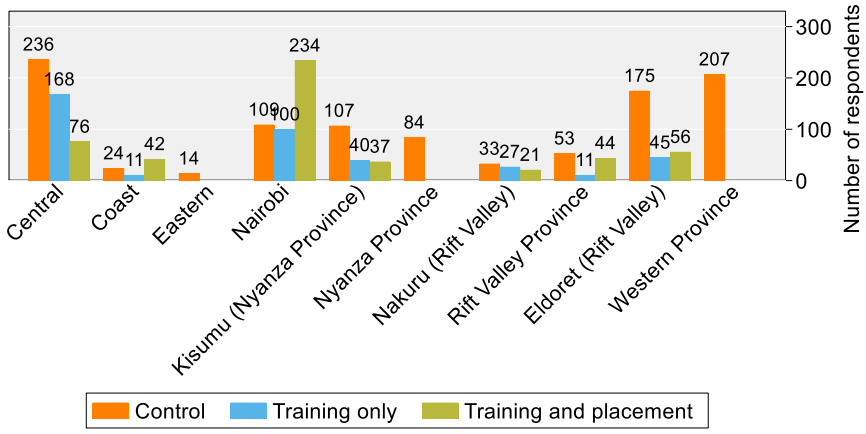
The examination of individual background characteristics serves two main purposes. First, it describes the study population for whom treatment effects are measured. Second, it allows a detailed investigation of the similarity and, thus, comparability of treatment and control group members. In case significant differences in background characteristics are detected across groups the differences can be controlled for in the impact estimations.

2.5.1 Regions and trades of study participants

KAM reached out to TVET institutions all over Kenya. Figure 2.5.1 shows that KAM beneficiaries mostly attended TVET institutes located in the provinces Central and Nairobi. Smaller number of beneficiaries attended institutes located in the provinces Coast, Nyanza, and Rift Valley. The distribution of KAM beneficiaries is not always mirrored by the control group. For example, whereas none of the KAM beneficiaries attended institutes in Eastern Province or Western Province, some of the control group participants did. The different TVET institutes in which study participants were enrolled in are listed in figure A1.1.1 of the Appendix.

[10] *With regards to respondents' employment situation such recall questions were only applied for the question of the job roster, i.e., how many jobs they hold 12 months prior to joining the KAM program or prior to the interview as well as the characteristics of each of these jobs. Instead, the question about respondents' employment status referred to the current situation. Thus, if in the following we talk about the employment status, this does not refer to the employment status as indicated by the respondent but to a calculate employment status based on respondents' answers regarding number of jobs hold, working hours, and total income.*

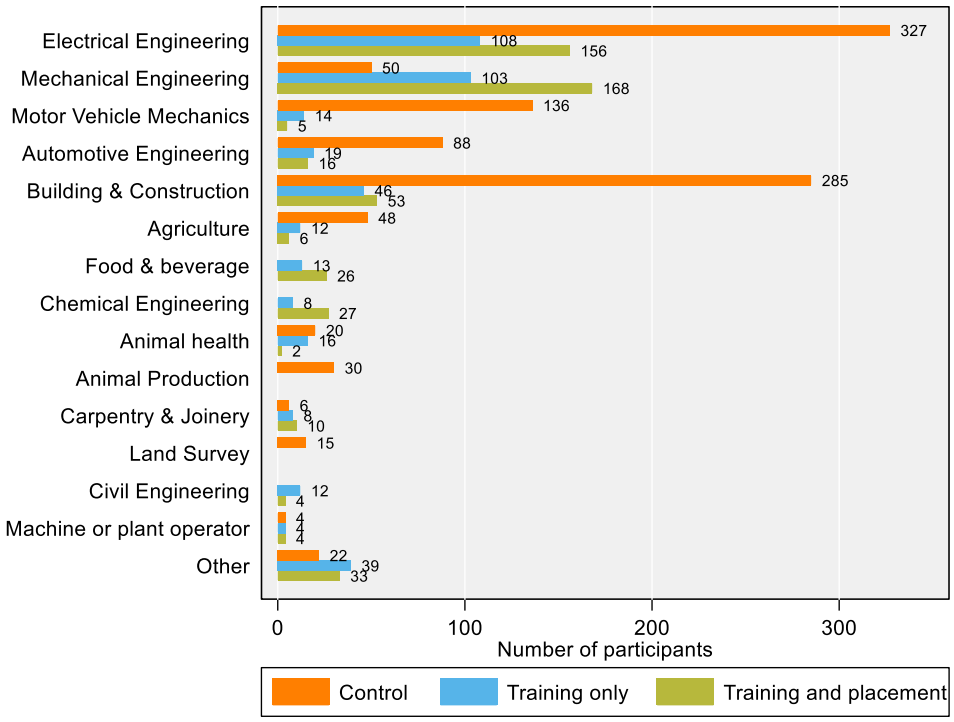
Figure 2.5.1
Number of study participants by region and by study group



Notes: The graphic displays the number of baseline respondents by region and by (i) study participants who did not participate in the KAM program (Control), (ii) KAM beneficiaries who only participated in the training (Training only), and (iii) KAM beneficiaries who participated in the training and received an internship placement through KAM (Training and placement). - Source: Own calculations based on KAM survey.

Figure 2.5.2 presents the different TVET courses study participants were enrolled in and differentiates between treatment and control group participants.

Figure 2.5.2
Study participants by TVET course visited and study group



Notes: The graphic displays the number of baseline study participants by TVET course and by treatment and control groups. - Source: Own calculations based on KAM survey.

The KAM program covered a variety of different trades. The majority of KAM beneficiaries were enrolled in mechanical or electrical engineering and another larger part were enrolled in building and construction. The distribution of trades was slightly different among control group participants. While some of the trades in which KAM beneficiaries were enrolled in are not covered by the control group (e.g., food and beverage, chemical engineering, civil engineering) other trades were only taken by control group participants (e.g., animal production, land survey).

2.5.2 Sociodemographic characteristics

Table 2.5.1 shows the study participants' sociodemographic background characteristics by treatment and control groups in columns (1) to (3) and comparisons of these background characteristics across treatment and control groups in columns (4) and (6). In order to provide insights into the background characteristics for the whole sample, Table 2.5.1 is based on the baseline data from all study participants irrespective of the type of TVET visited. However, as some of the robustness checks will focus on the subsample of mixed TVETs, Table A1.1.2 displays the background characteristics of study participants of mixed TVETs only across treatment and control groups. Overall, the characteristics are very similar.

Sample characteristics

Column (1) presents mean background characteristics (or the percentage share for binary indicators) of respondents who did not participate in the KAM program at all. Column (2) presents background characteristics of respondents who only received the KAM training, and column (3) presents the characteristics of respondents who participated in the KAM training and received an internship placement organized by KAM.

A small share of participants is female (11-18%) and participants' average age is 25 years. The large majority of participants is single (78-89%) and in line with this only few indicated to have children (14-19%) or dependants to take care of (6-8%). For participants who have children in school age, almost all of their children are also enrolled in school (99%).

About half of all KAM beneficiaries indicated to currently depend on their household head, whereas it was almost three-quarter among control group participants. Accordingly, most control group participants indicated to live with their family (62%). This share is noticeably lower among KAM beneficiaries (47%) and an almost equally large share indicated to live alone (46%).

Female household heads of KAM beneficiaries mostly obtained high levels of education, i.e., secondary school or higher (41-43%). This is also the case for female household heads of control group participants, although for a lower share of 34%. Further, one quarter of control group participants and even one-third of KAM beneficiaries indicated that they do not have a female household head or refused to answer. Only a small share of female household heads in all three participant groups did not obtain any formal education or completed only primary standard.

Table 2.5.1
Sociodemographic characteristics of study participants, all TVETs

	Control (no treatment)	Training only	Training and placement	Std. Diff.	Std. Diff.
	(1)	(2)	(3)	(4)	(5)
	% or Mean	% or Mean	% or Mean	(1) - (3)	(2) - (3)
Respondent characteristics					
Gender	11.3	16.5	18.3	-0.069	-0.018
Age at interview	23.7	25.5	25.4	-1.703	0.151
Family situation					
Single/never married	89.3	78.4	83.1	0.061	-0.048
Has children, yes/no	14.2	19.4	18.8	-0.046	0.006
Children, number	0.2	0.3	0.2	-0.054	0.039
Children aged 6-14 are all enrolled in school	98.8	99.5	99.2	-0.004	0.003
Children aged 6-18 are all enrolled in school	97.9	99.0	98.4	-0.005	0.006
Other dependants (excl. children), yes/no	7.3	6.0	8.2	-0.009	-0.023
Other dependants, number	0.2	0.1	0.2	-0.013	-0.043
Dependency on household head, yes/no	72.5	48.3	51.0	0.215	-0.027
Education of female household head					
None or pre school	5.3	4.0	3.1	0.021	0.008
Primary standards 1 to 6	6.1	2.7	3.5	0.026	-0.008
Primary standard 7	7.0	5.0	5.1	0.019	-0.001
Primary standard 8 or secondary forms 1 to 3	22.6	12.9	13.1	0.095	-0.002
Secondary form 4 or higher	34.0	42.8	40.6	-0.066	0.022
No female household head or refusal	25.0	32.6	34.5	-0.096	-0.019
Living situation					
Live alone	27.9	41.5	46.1	-0.182	-0.045
Live with family	61.8	51.7	46.9	0.149	0.049
Live with room mates	10.0	6.5	6.9	0.031	-0.004
Live with non-relatives	0.3	0.2	0.2	0.001	0.001
Number of observations	1,042	402	510		
Test of joint orthogonality					
F-test statistics				12.942***	0.734
p-value				0.000***	0.778

*Notes: Information displayed as indicated by study participants during the baseline survey. Observations include participants from all TVETs. The value displayed for t-tests are the differences in the means across the groups (***, **, and * indicate significance at the 1, 5, and 10 percent critical level). - Source: Own calculations based on KAM survey.*

Differences in background characteristics across groups

Columns (4) and (5) display statistics that allow to evaluate the similarity and, thus, comparability of the three participant groups. Column (4) compares KAM beneficiaries who participated in the full KAM program with control group participants, and column (5) compares KAM beneficiaries who participated in the full KAM program with KAM beneficiaries who only received the KAM training. Both columns show standardized differences (Std. Diff.) as measures of similarity for each background characteristic. The standardized difference takes the difference in means of the respective two groups and weights it by the variance of the indicators.[11] Standardized differences with an absolute value of 0.2 or less are considered small. Thus, for the evaluation of the similarity of treatment and control groups, every absolute difference that is larger than 0.2 is considered as not similar.

The two bottom rows of Table 2.5.1 present measures that evaluate the similarity across the respective treatment and control group considering all presented characteristics jointly. A p-value below 0.1 would imply that the listed background characteristics of the two considered groups are significantly different from each other and, therefore, the two groups would not be comparable without applying additional econometric techniques that restore comparability.

The p-value of below 0.1 in column (4) indicates that KAM beneficiaries who participated in the KAM training and received a placement and respondents who did not participate in the KAM program at all significantly differ from each other when considering all observed characteristics jointly. Examining the individual standardized differences, reveals that the groups differ in almost all of the observed background characteristics. Only the enrolment status of children, having dependants, as well as the average number of dependants do not pass the threshold of 0.2. Overall, control group participants have a lower share of females and are younger. Presumably as a consequence of the younger age level, control group participants are also more frequently single, do less often have own children, more frequently depend on their household head, and tend to live with their family rather than living alone.

Column (5) presents the comparability of KAM beneficiaries who participated in the training and received a placement with KAM beneficiaries who only received the training. The p-value of above 0.1 at the bottom of the table suggests that the two groups are similar when taking into account all observed characteristics jointly. With respect to differences in individual variables, the only detectable difference regards beneficiaries' marital status. KAM beneficiaries who participated in the training and received a placement through KAM are more often single than KAM beneficiaries who only received the training.

2.5.3 Employment status and earnings prior to the KAM program

Tables 2.5.2 and 2.5.3 present mean pre-intervention employment and income indicators of study participants across the three groups for the full sample. Table 2.5.2 refers to all participants of all TVETs, whereas Table 2.5.3 only includes observations from participants who indicated to be employed at the time of the first survey. Again columns (1) to (3) display the background characteristics for each group and columns (4) and (5) compare the characteristics across groups. For control group participants the pre-intervention period refers to the one-year-period prior to the baseline, whereas for KAM beneficiaries it refers to the one-year-period prior to joining the KAM program. Tables A1.1.3 and A1.1.4 provide the balance tables for the subsample of mixed TVETs.

[11] *The variance is defined as a measure of how much the indicator spreads around its mean.*

Sample characteristics

Table 2.5.2 reveals that already before starting the KAM program more than 80% of the KAM beneficiaries have had any form of paid work. In the control group 71% indicated to had have any form of paid work. Most of this paid work refers to fulltime employment which is reflected in the total hours worked per week, which range between 34 and 47 hours. Only a small share of study participants reported to have a decent or a formal employment and the share is noticeably lower in the control than in the treatment groups. The share of study participants who indicated to be self-employed ranges between 18 and 25%. The average total income per month ranged between roughly 9,500 KES among control group participants and 13,000 KES among KAM beneficiaries and the average hourly earnings varied between 63 KES in the control group and 91 KES in the treatment group that only received the KAM training.

Table 2.5.2
Pre-intervention employment characteristics of study participants, all TVETs

	Control (no treatment)	Training only	Training and placement	Std. Diff.	Std. Diff.
	(1)	(2)	(3)	(4)	(5)
	% or Mean	% or Mean	% or Mean	(1)-(3)	(2)-(3)
Employment					
Employment status	70.9	82.3	81.6	-0.106	0.008
Decent employment	11.3	24.6	17.5	-0.062	0.071
Self-employment	19.9	25.2	17.7	0.023	0.075
Formal employment	7.6	21.4	22.1	-0.144	-0.007
Fulltime employment	53.3	74.2	76.2	-0.229	-0.020
Total hours worked per week	34.2	42.4	47.2	-12.994	-4.838
Income					
Total monthly income	9,529.1	13,175.9	11,231.0	-1,701.938	1,944.851
Average hourly earnings	63.1	91.3	89.7	-26.625	1.599
N	1,042	402	510		
Test of joint orthogonality					
F-test statistics				18.184***	3.701***
p-value				0.000***	0.000***

Notes: Information displayed as indicated by study participants during the baseline survey. Observations include participants from all TVETs. The values displayed in columns (4) and (5) are the differences in the means across the groups. The stars indicate the significance at the 1 (***), 5 (**), and 10 (*) percent critical level. - Source: Own calculations based on KAM survey.

Table 2.5.3 considers participants who were employed during the one-year-period prior to the intervention in the treatment group and participants who were employed during the one-year-period prior to the baseline in the control group. It thereby provides more detailed information on the average number of jobs held, hours worked, and hourly wages. On average, study participants held 1.2 jobs and worked between 48 and 58 hours per week. The average total income among employed varied between approx. 13,400 and 16,000 KES per month and the average wage per hour varied between 89 and 111 KES.

Table 2.5.3
Pre-intervention employment characteristics of employed study participants, all TVETs

	Control (no treatment) (1) Mean	Training only (2) Mean	Training and placement (3) Mean	Std. Diff. (4) (1)-(3)	Std. Diff. (5) (2)-(3)
Employment					
No. of jobs	1.2	1.3	1.3	-0.107	-0.074
Total hours worked per week	48.2	51.5	57.9	-9.633	-6.417
Income					
Total monthly income	13,436.1	16,002.118	13,768.8	-332.664	2,233.317
Average hourly earnings	89.1	110.9	110.1	-20.934	0.865
N	739	331	416		
Test of joint orthogonality					
F-test statistics				7.745***	3.951***
p-value				0.000***	0.004***

Notes: Information displayed as indicated by study participants during the baseline survey. Observations include participants from all TVETs who had a job at the time of the baseline. The stars indicate the significance at the 1 (***), 5 (**), and 10 (*) percent critical level. - Source: Own calculations based on KAM survey.

Differences in background characteristics across groups

Again, column (4) displays the standardized differences between KAM beneficiaries who participated in the full KAM program and control group participants, while column (5) displays the differences between KAM beneficiaries who participated in the full KAM program and KAM beneficiaries who only received the KAM training.

The test statistics at the bottom of each table indicate that treatment and control group participants are significantly different from each other in both comparisons when considering all presented characteristics jointly. Column (4) of Tables 2.5.2 and 2.5.3 indicates that study participants who did not receive any KAM benefits are noticeably less often employed, have a lower number of jobs, do less often have a decent, formal, or fulltime employment, and have a lower number of weekly working hours than KAM beneficiaries who participated in the training and received a placement. Column (5) shows that KAM beneficiaries who only participated in the training are on average significantly more often employed in a decent job or are self-employed but have a lower number of weekly working hours than full program beneficiaries.

2.5.4 Restoring the comparability of treatment and control groups

The similarity and, thus, comparability of treatment and control groups is important to attribute measured employment effects to the actual impact of the KAM program. If treatment and control groups are not comparable the measured treatment effects could be caused by participants' different background characteristics and labor market potential rather than the KAM program itself. For example, if treatment group participants are on average older than control group participants and if being older goes along with having had more time to obtain practical experience, then estimated program effects may be due to participants' age rather than due to the KAM program if the different age structure across groups were to be ignored in the estimation approach. To restore the comparability of treatment and control groups, the estimation strategies

outlined in section 2.3.1 are applied and control for all background characteristics that differed across treatment and control group participants, i.e., participants' background characteristics that correlate with treatment group membership and labor market outcomes simultaneously.

2.6 Descriptive statistics on the KAM program

This section provides descriptive statistics to of (i) the outreach of the KAM program, (ii) the type of internships completed by all study participants, as well as (iii) KAM beneficiaries' satisfaction with the program in general and the KAM internship in particular.

2.6.1 Outreach of the program

Table 2.6.1 presents descriptive statistics of the outreach of the KAM program. In total, 918 graduates started the KAM program, i.e., participated in the E4D work readiness training, out of which only 160 were female. Despite the short duration of the training, approximately 12% of the participants did not complete the work readiness training. Two thirds of all KAM participants received an internship placement. The majority of these placements were organized by KAM, however, 83 KAM participants indicated to have sought their placement individually. Only about half of all participants who started the KAM program reported that they completed a placement. Less than one quarter stated to have completed the full KAM program, which requires them to also receive a program certificate once they completed the internship. Overall, Table 2.6.1 illustrates that drop-outs occurred throughout the different program steps and that the number of drop-outs was largest at the internship placement stage. Drop-out rates do not noticeably differ for female participants.

Table 2.6.1
Outreach of the KAM program

Number of participants who …	Total	% of baseline participants	Females	% of female baseline participants	Reference period
started the work readiness training	918	100.0	160	100.0	Baseline
completed the work readiness training	803	87.5	137	85.6	Follow-up 1
who received a placement	588	64.1	105	65.6	Ever
Individual sought	83	9.0	12	7.5	Follow-up 1
Sought by KAM	483	52.6	86	53.8	Follow-up 1
completed the placement	431	46.9	80	50.0	Ever
completed full KAM program	194	21.1	34	21.3	Follow-up 1

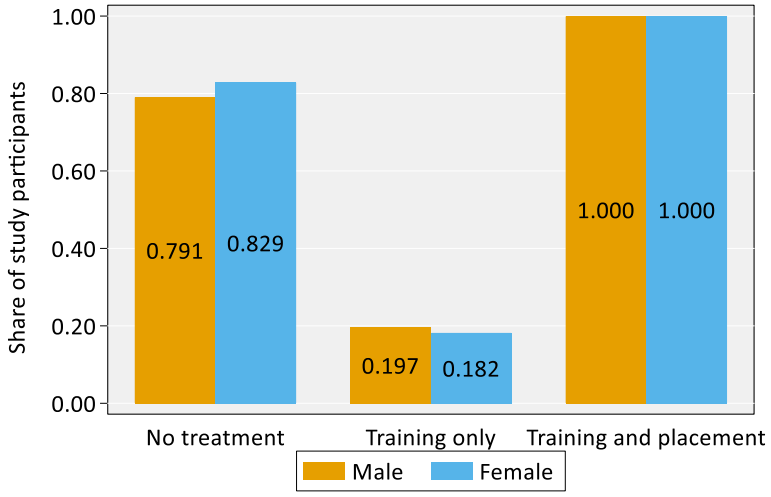
Notes: Some information was only asked in Follow-up 1 whereas other information was asked in all four waves (see last column). The number of participants who started the E4D work readiness training is defined as everyone who participated in the baseline survey and indicated to be part of the KAM program. Having completed the full KAM program is defined as having (i) participated in the work readiness training, (ii) completed the placement, and (iii) received the KAM certificate. - Source: Own calculations based on KAM survey.

2.6.2 Internship placements

Figure 2.6.1 displays the share of placed individuals in each treatment and control group. As per definition all study participants of the treatment group ("Training and placement") received a placement. Among KAM beneficiaries who only participated in the training, 20% of male and 18%

of female participants obtained an internship, either because they sought it individually or they received it through another program. A considerably large share of the control group, i.e., study participants who did not participate in the KAM program, were placed in other forms of internships (79% among male and 83% among female participants).

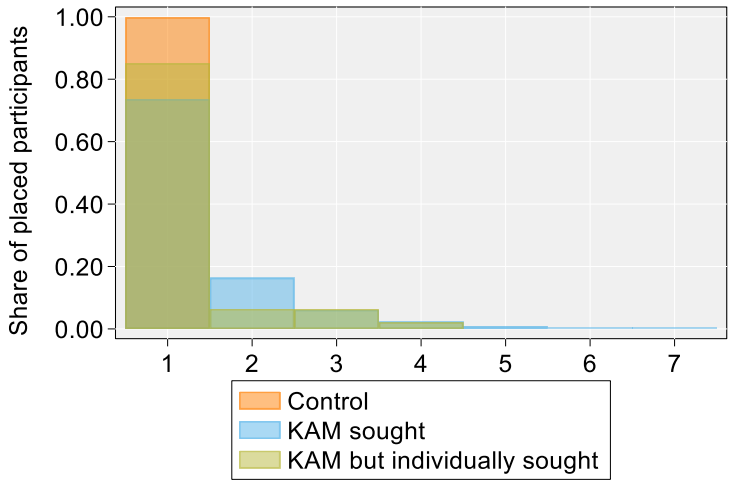
Figure 2.6.1
Placement share by gender and study group



Source: Own calculations based on KAM survey.

Figure 2.6.2 shows that while all control group participants and most KAM beneficiaries indicated that their internship consisted of only one section, some of the beneficiaries who received an internship through KAM as well as some of those beneficiaries who individually sought an internship went through two to seven different sections within the company. [12]

Figure 2.6.2
Number of placement sections within internships by study group

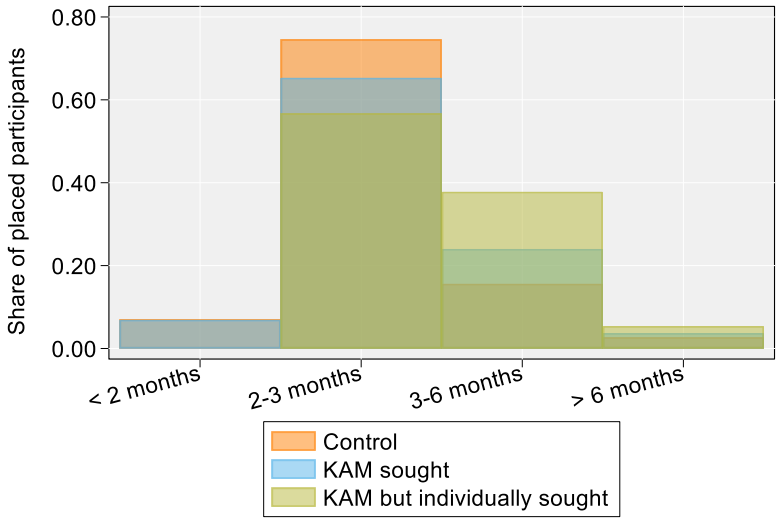


Source: Own calculations based on KAM survey.

[12] *Among KAM participants the information was collected during F2 whereas among the control group the information was collected in F3.*

Generally, and independently of who facilitated the placement, most internships lasted between two and three months as shown in Figure 2.6.3. The placement length of KAM-sought internships was two to three months in 65%, more than three and up to six months in 24%, and longer than six months in 4% of the cases. Although KAM intended internships to last three months, 7% of the KAM internships were even shorter than two months. Further, Figure 2.6.3 shows that placements of KAM beneficiaries who individually sought their internship more often lasted between three and six months (38%).

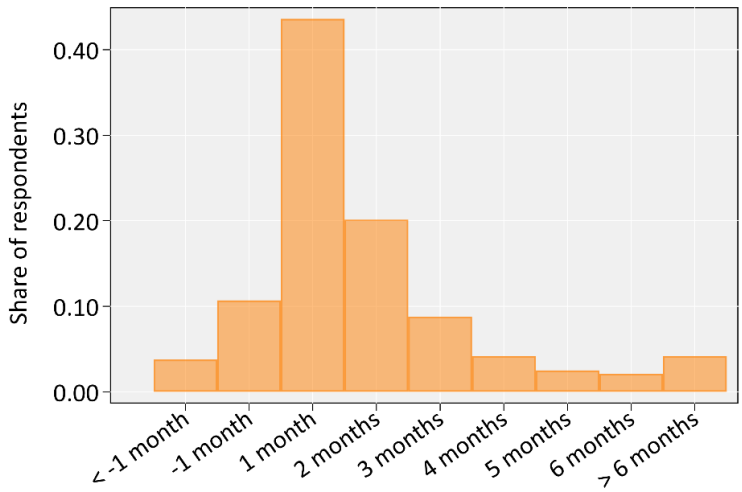
Figure 2.6.3
Placement length by study group



Source: Own calculations based on KAM survey.

Figure 2.6.4 displays the time that passed between KAM beneficiaries' participation in the work readiness training and starting their internship. The graph only refers to KAM beneficiaries who received a placement sought by KAM, i.e., observations from KAM participant who individually sought their placement are not considered .

Figure 2.6.4
Time gap between KAM start and placement start

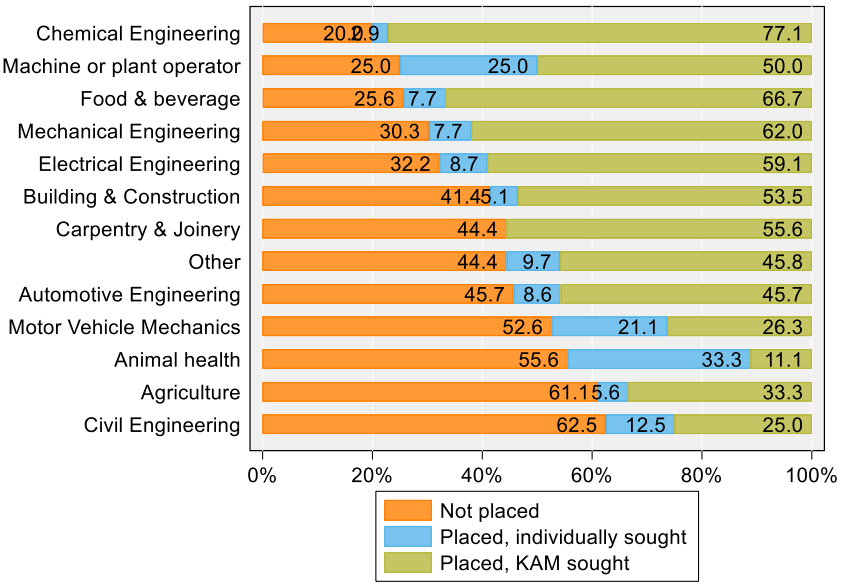


Source: Own calculations based on KAM survey.

However, including KAM beneficiaries who individually sought their placement does not change the overall picture. Most KAM sought internships (44%) started one month after beneficiaries' participation in the work readiness training and an additional 20% of the internships started two months after the training. The displayed data on the timing of the training and the internship start relies on the information provided by the study participants. Fourteen percent of KAM beneficiaries started their internship before having started the KAM program, which might be due to reporting errors of study participants and a lack of consistency checks in the questionnaire.

Whether KAM beneficiaries received a placement or not strongly varied across the TVET courses in which they were enrolled in prior to participating in the KAM program. Figure 2.6.5 presents the share of KAM beneficiaries who ever received an internship placement through KAM. The placement share is calculated for each TVET course separately. Whereas three quarter of KAM beneficiaries in the sector of chemical engineering received a KAM-sought placement, this was only the case for one quarter of KAM beneficiaries in the sector of civil engineering. Apart from the beneficiaries' course of study also the visited TVET facility seems to play an important role in receiving an internship placement (see Appendix figure A1.1.4).

Figure 2.6.5
Placement status by TVET course taken

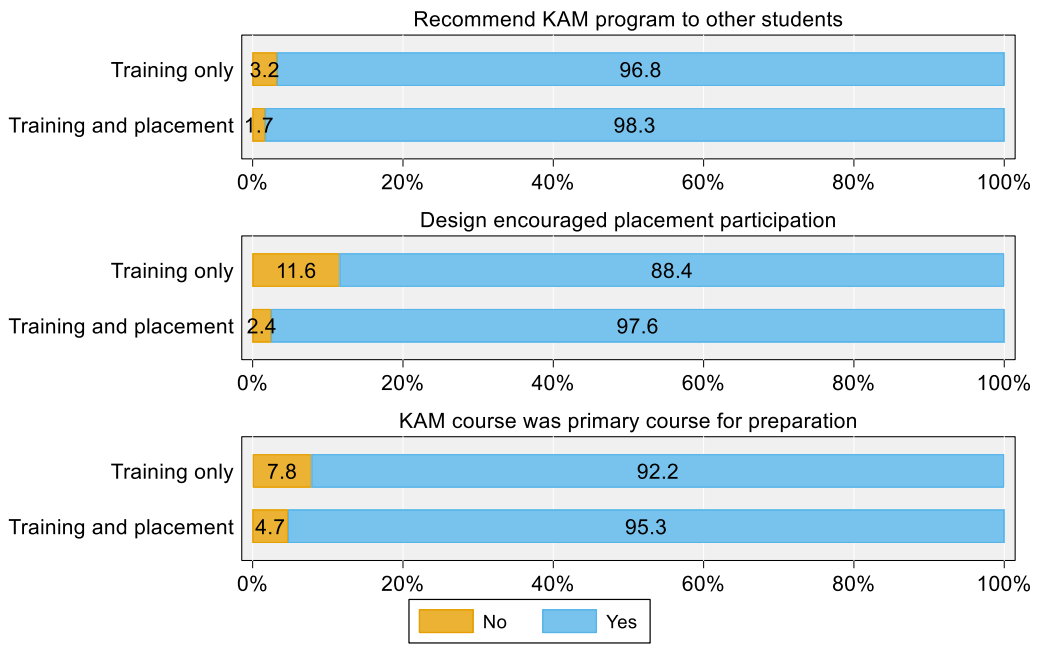


Source: Own calculations based on KAM survey.

2.6.3 Participants' satisfaction with the KAM program

Overall, KAM beneficiaries were very satisfied with the program. As displayed in the upper graph of Figure 2.6.6, almost everyone would recommend the KAM program to other students. The middle graph shows that the large majority of beneficiaries (97%) indicated that the KAM program design encouraged their participation in an internship. Not surprisingly this share is somewhat lower among KAM beneficiaries who only participated in the training but did not receive a placement through KAM (88%). The bottom graph shows that for more than 90% of all KAM beneficiaries the KAM work readiness training was the primary course for preparation for the internship placement.

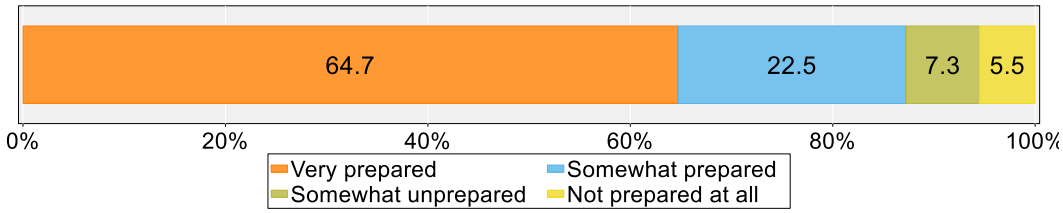
Figure 2.6.6
Feedback of KAM participants on KAM program



Notes: The first indicator on whether the respondent would recommend KAM to other students was asked in all three follow-up waves. The remaining two indicators solely refer to answers given during Follow-up 1 because they were only included in the questionnaire of Follow-up 1.
Source: Own calculations based on KAM survey.

Figure 2.6.7 shows in how far KAM beneficiaries felt that the KAM training prepared them for the subsequent company placement. Almost 90% of the program participants felt prepared or very prepared for their internship placement due to the participation in the work readiness training. Yet, 7% indicated to have felt somewhat unprepared and 6% felt not prepared at all. Answers from KAM participants who individually sought their placement are not considered in Figure 2.6.7. Including their answers does not change the overall picture.

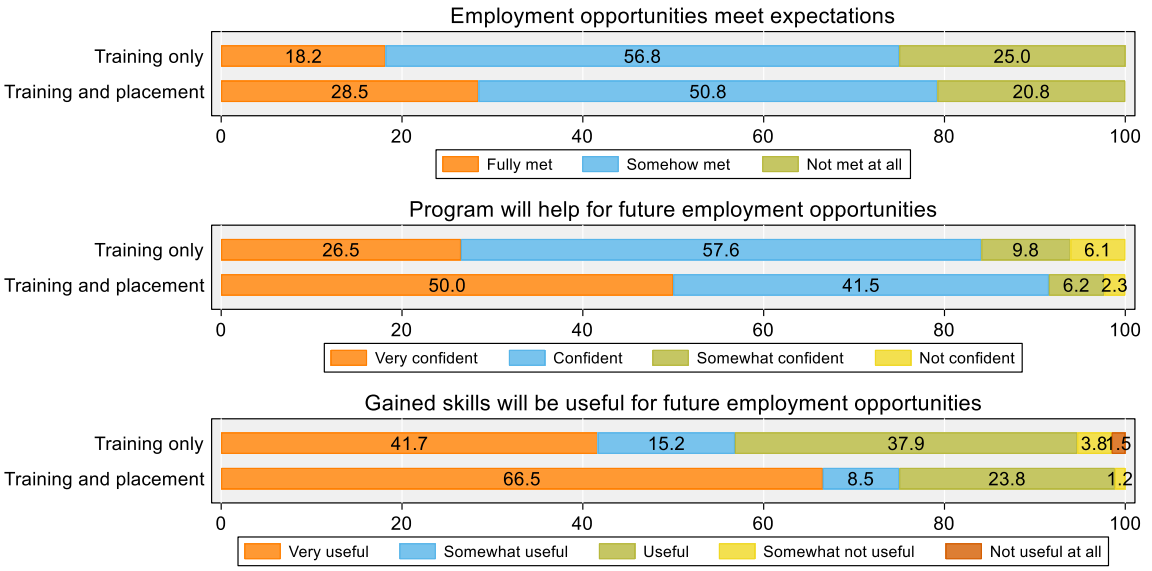
Figure 2.6.7
Level of preparation for placement through the KAM training



Source: Own calculations based on KAM survey.

Figure 2.6.8 presents how KAM participants who were interviewed in the third follow-up survey expect the KAM program to impact their future employment opportunities and differentiates between participants of the training only group and participants who received both program components. The large majority indicated that the employment opportunities that they had since participating in the KAM program full or somehow meet the employment expectations they had when leaving the program. Still, about one-fifth of participants of the full program and one-fourth of the training only participants reported that their expectations were not met at all. Moreover, participants of the full KAM program were very confident that the program (50.0%) and the gained skills (66.5%) will be useful for future employment opportunities. Perceptions are less optimistic among participants who only participated in the KAM training.

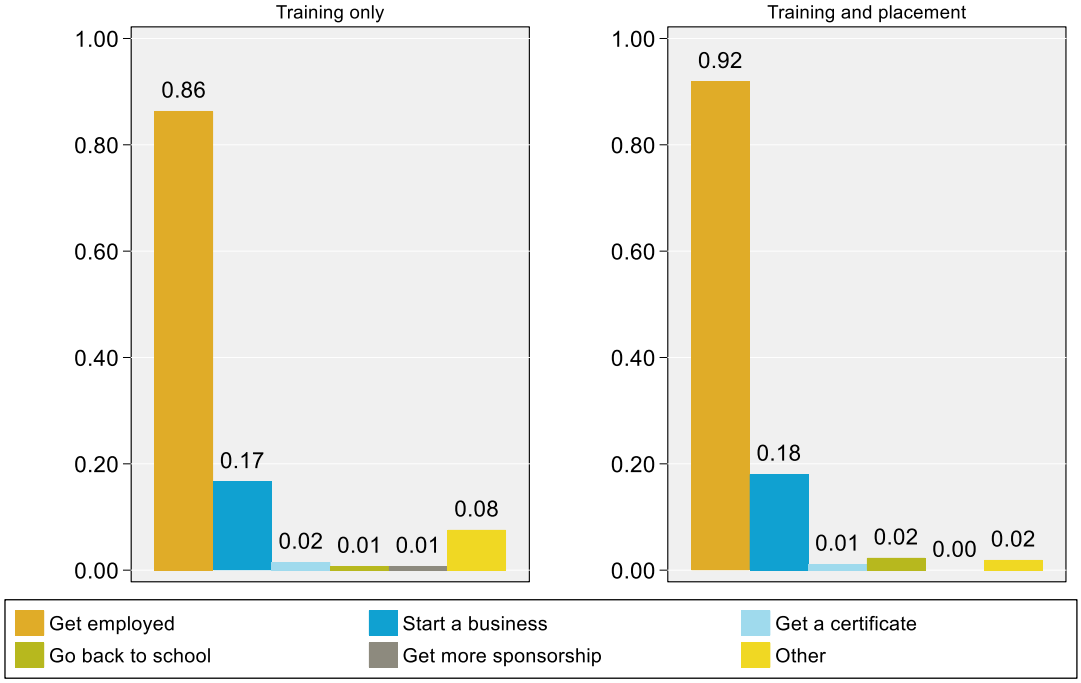
Figure 2.6.8
Perceived employment opportunities after KAM program



Notes: These questions were only included in interviews of the third follow-up survey. - Source: Own calculations based on KAM survey.

In the third follow-up wave, participants were also asked about the key goals they had at the time leaving the KAM program. The answers are displayed in Figure 2.6.9. The most prominent goal was to get employed which was mentioned by 86% of participants who only participated in the training and by 92% of participants of the full KAM program. Between 17 and 18% of the participants indicated that they want to start a business and very few participants (≤ 2%) indicated that they want to get a certificate, go back to school, or get more sponsorship.

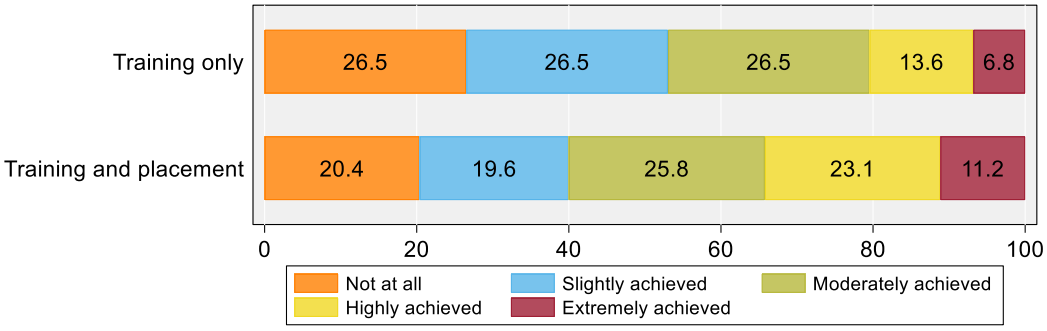
Figure 2.6.9
Key goals at the time leaving the KAM program



Notes: This question was only included in interviews of the third follow-up survey. - Source: Own calculations based on KAM survey.

Figure 2.6.10 additionally shows that the majority of KAM participants perceived that they were able to achieve these goals at least to some extent. As already observed in the previous graphs, the overall picture is more optimistic among KAM beneficiaries who received both program components than among beneficiaries who only received the training. One third of the participants who were trained and placed indicated that they achieved their goals to a high or extreme extent, while it is only one-fifth among participants who only received the training.

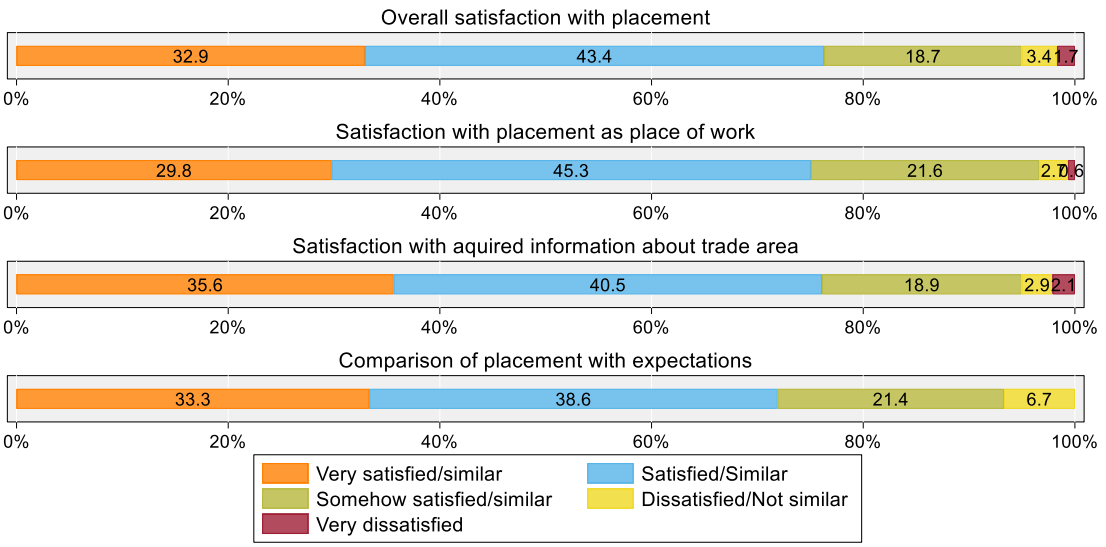
Figure 2.6.10
Achievement of key goals



Source: Own calculations based on KAM survey.

Figure 2.6.11 shows that about three quarters of KAM beneficiaries indicated to be satisfied or very satisfied with their placement. This holds for participants' overall satisfaction with the placement as well as for their satisfaction with the internship as place of work and the acquired information about the respective area of trade. Two thirds of KAM participants indicated that the placement was similar or very similar compared to their prior expectations. About 7% of the participants reported that the placement was not similar to what they had expected. However, this information was only enquired in the first follow-up wave and only covers KAM beneficiaries who at that time had already received a placement. As before, only answers from KAM participants whose placement was sought by KAM were used.

Figure 2.6.11
Indicators of satisfaction with placement (I)

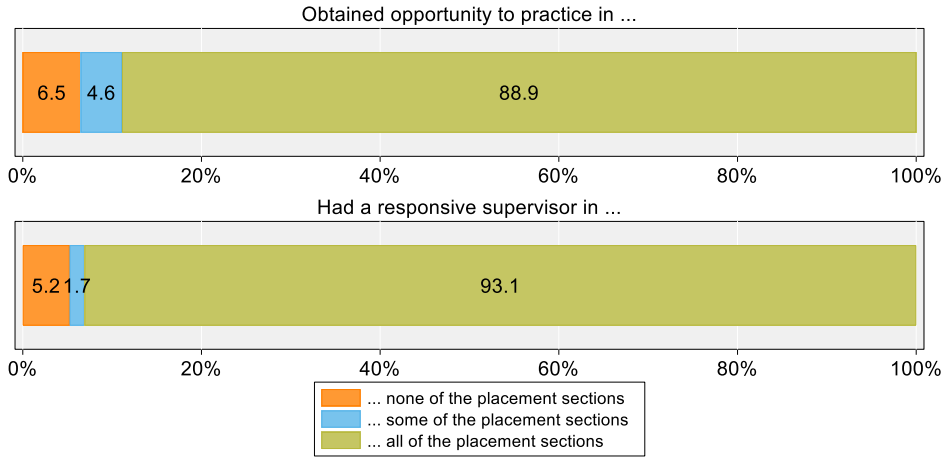


Source: Own calculations based on KAM survey.

As depicted in Figure 2.6.2, some participants passed through different internship sections. Figure 2.6.12 reveals that, overall, participants' satisfaction was very positive and very homogenous across placement sections. About 89% of the participants who received a placement through KAM reported to have obtained opportunity to practice in all placement sections, whereas 5% indicated to have obtained opportunity to practice only in some and 7% in none of the placement sections. A similar picture emerges when participants were asked about their supervisor. The overwhelming majority indicated to have had a responsive supervisor in all of their placement sections. Only 5% indicated that they have not had a responsive supervisor in any of the sections and 2% only in some.[13]

[13] *Again, information was only conducted in F1 and only covers KAM participants who already received a placement at this point in time. Only answers from KAM participants whose placement was sought by KAM were used for this graphic, i.e., answers from KAM participants who individually sought their placement are not taken into account. However, including their answers does not change the overall picture.*

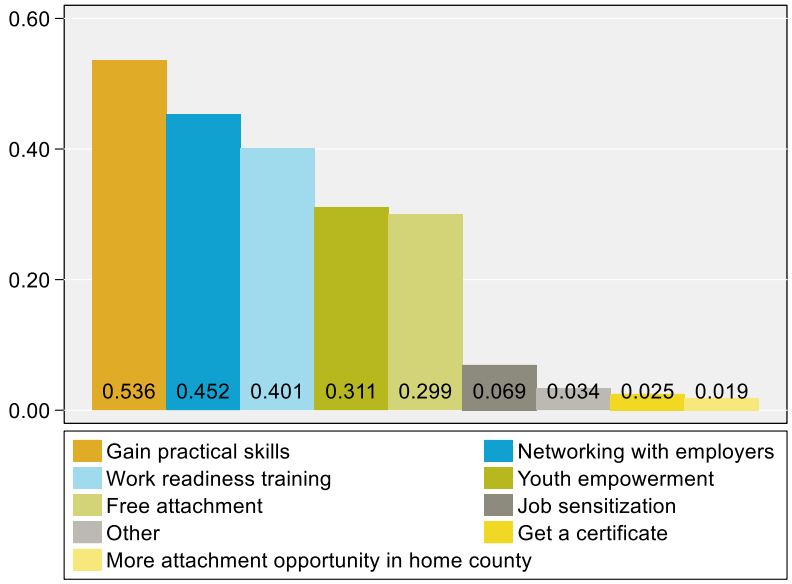
Figure 2.6.12
Indicators of satisfaction with placement (II)



Source: Own calculations based on KAM survey.

Figure 2.6.13 depicts KAM participants' reasons to recommend the KAM program to others. The most prominent answers were gaining practical skills (54%), networking with employers of the respective trade (45%), and the content provided in the work readiness training (40%). Other frequent responses included youth empowerment (31%) and free attachment (30%), an answer option provided by IPA for which we could not identify what it is referring to. Aspects mentioned by only very few of the respondents (<7%) were job sensitization, getting a certificate, or obtaining more attachment opportunities in the home county.

Figure 2.6.13
Reasons to recommend the KAM program

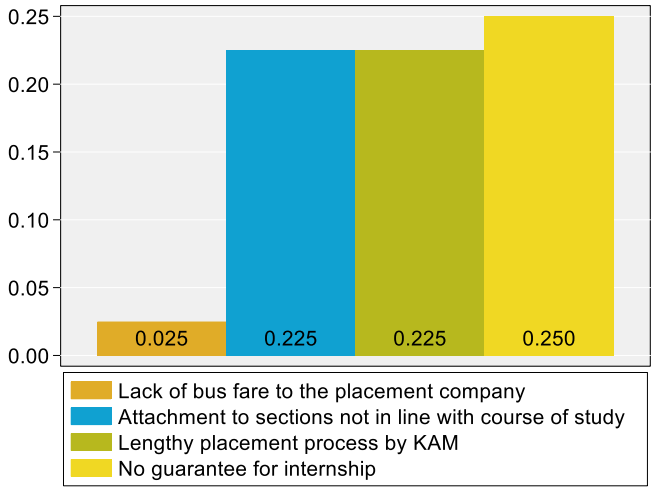


Notes: The graphic indicates all aspects mentioned by KAM participants across survey waves F1, F2, and F. - Source: Own calculations based on KAM survey.

Reasons why KAM participants would not recommend the program are presented in Figure 2.6.14 and mostly refer to the organization of the program. One quarter claimed that

participating in the program did not guarantee them an internship placement, 23% criticized that KAM's placement process was too lengthy, and 23% stated that the attachment to the placement sections was not in line with their course of study. About 3% mentioned that they would not recommend the program because of lacking bus fares to the placement company.

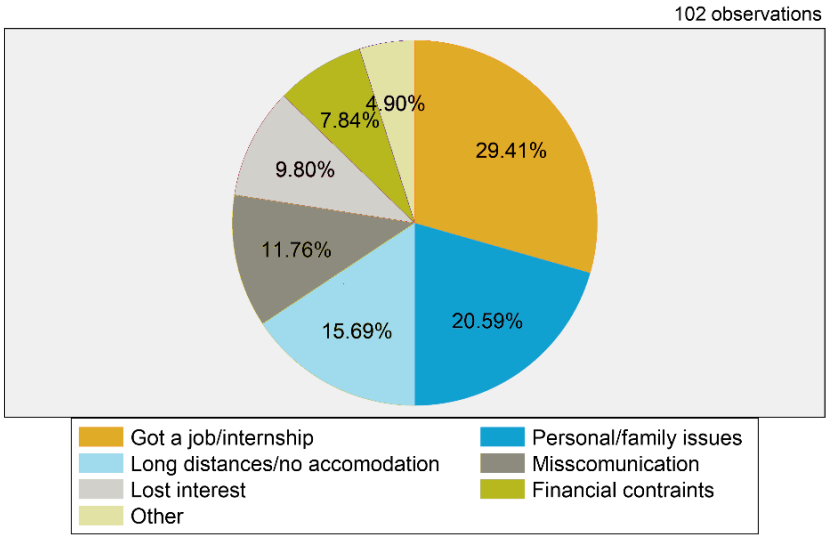
Figure 2.6.14
Reasons not to recommend the KAM program



Notes: The graphic indicates all aspects mentioned by KAM participants across survey waves F1, F2, and F3. - Source: Own calculations based on KAM survey.

Figure 2.6.15 displays the reasons for dropping out of prior KAM participants. Most KAM participants indicated that they dropped out because they got a job or a more attractive internship (30%). The second most common reason for dropping out were personal or family issues (21%). Long distances or the lack of accommodation (16%) as well as financial constraints (8%) were other reasons indicated by former participants. Moreover, 12% of the participants who dropped out mentioned miscommunication either between them and KAM, them and the company, or KAM and the company as a reason for their drop out.

Figure 2.6.15
Reasons for dropping out of the KAM program



Source: Own calculations based on KAM survey.

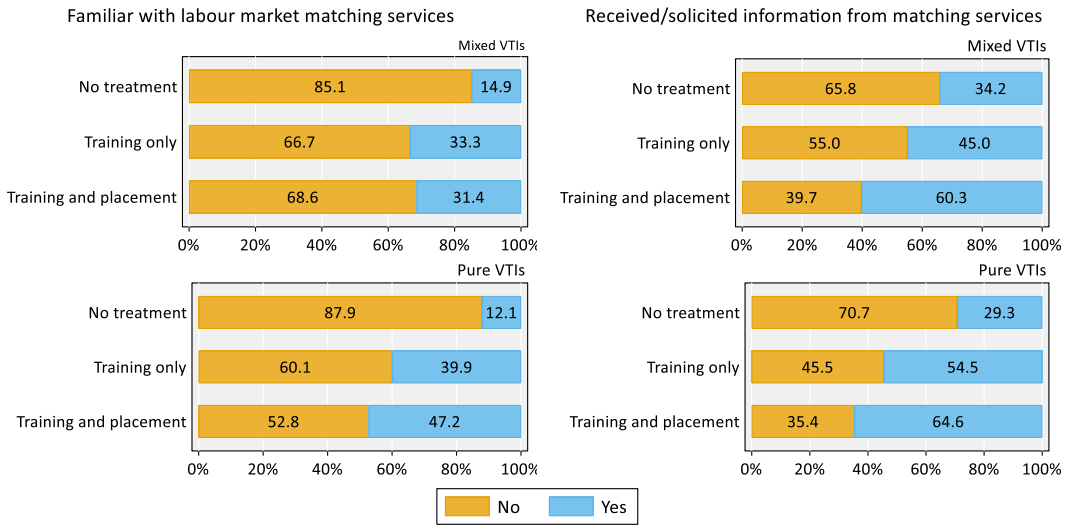
2.7 Descriptive statistics on participants' job search

This section describes participants' job search experience by treatment and control groups, including their participation in and knowledge of alternative employment promoting programs and participation in interviews. In addition, this section reports KAM beneficiaries' self-reported evaluation of their KAM program participation on their job search.

2.7.1 Alternative employment promoting programs

Figure 2.7.1 presents study participants' knowledge and usage of alternative employment promoting programs. The main take-away is that among KAM beneficiaries a larger share is familiar with labor market matching services and also received or solicited information from such matching services than among control group members. KAM beneficiaries were additionally asked about their participation in online job linkages, job bazaars, partner events, open days, other internship arrangements, as well as coaching and mentorship services. Unfortunately, this information was not collected among control group participants. Consequently, a comparison between treatment and control group was not possible.

Figure 2.7.1
Knowledge and usage of labor market matching services by study group

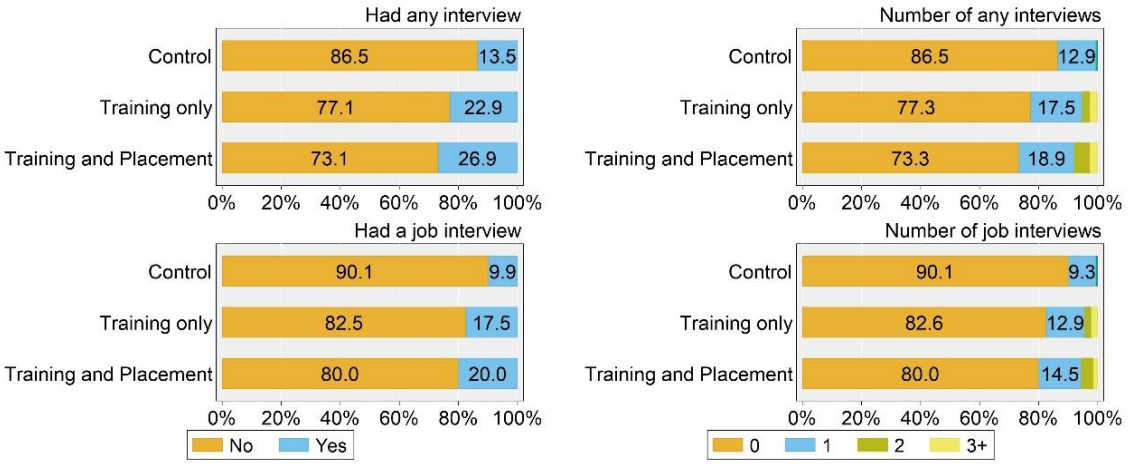


Source: Own calculations based on KAM survey.

2.7.2 Participation in interviews

Figure 2.7.2 presents descriptive statistics on whether study participants had any interviews 12 months prior to the intervention. For control group participants the pre-intervention period refers to the one-year-period prior to the baseline, whereas for KAM beneficiaries it refers to the one-year-period prior to joining the KAM program. The upper left graph shows that 27% of the KAM placement group, 23% of the KAM training only group, and 14% of the control group participated in at least one interview, respectively. Most of these interviews were job interviews, as can be seen in the lower left graph. The two right-hand side graphs show the *number* of (job) interviews that participants had in the pre-intervention period. Most individuals who had an interview only had one. The small discrepancies in percentages between the left-hand side and right-hand side graphs are due to rounding.

Figure 2.7.2
Participation in interviews at baseline by study group



Source: Own calculations based on KAM survey.

2.7.3 Participant-reported evaluation of KAM on job search

During the follow-up waves KAM beneficiaries were asked about how they felt that their KAM program participation affected their chances of getting job interviews and their performance in the interviews. Table 2.7.1 shows the results of this self-reported evaluation of the KAM program. About half of all KAM beneficiaries indicated that the "skills developed in the E4D program" or the "work readiness training" helped them to get an interview. About one third indicated that the "E4D network" or the "practical job training" helped them to get an interview. Between 22 to 28% reported that their interview performance improved thanks to the "E4D network" or the "skills developed in the E4D program".

Table 2.7.1
Impact of KAM program on job interviews

Number of participants who indicated that …	Total	% of baseline participants	Females	% of female baseline participants
… E4D network improved interview performance	198	21.6	44	27.5
… E4D network helped to get an interview	301	32.8	51	31.9
… skills developed in E4D improved interview performance	259	28.2	55	34.4
… skills developed in E4D program helped to get an interview	455	49.6	80	50.0
… work readiness training helped to get an interview	429	46.7	77	48.1
… practical job training helped to get an interview	306	33.3	48	30.0
… they had an interview with the placement company	39	4.2	11	6.9

Notes: Refers to any job indicated by KAM beneficiaries in any of the three follow-up waves. Statements are shown as phrased by IPA in their questionnaire. - Source: Own calculations based on KAM survey.

2.8 Descriptive statistics on labor market outcomes

This section reflects on participants' change in labor market performance throughout the KAM program, including employment status, income, hourly wages, and working hours. Further, this section reports KAM beneficiaries' self-reported evaluation of their KAM program participation on their current employment situation.

2.8.1 Changes in labor market characteristics between baseline and follow-up

In the baseline survey, information on jobs held and incomes earned in these jobs for the 12 months prior to the intervention, i.e., prior to joining the KAM program among KAM beneficiaries and prior to the baseline interview among control group participants, were collected.[14] This pre-intervention period is used as a reference to compare labor market outcomes after the intervention, i.e., after the delivery of KAM program benefits, with those prior to the intervention.

Changes in employment status

Figures 2.8.1 and 2.8.2 present job transitions over time by comparing paid work post-intervention with paid work pre-intervention across treatment and control groups. Figure 2.8.1 refers to participants' employment status in general and Figure 2.8.2 especially looks at decent employment. In this vein, study participants are allocated into four distinct groups. Respondents who were employed (decently) during the 12 months prior to the intervention but were not employed (decently) when they were interviewed in subsequent follow-up interviews are considered to have *lost employment*. Participants who did not have a (decent) job at either time are referred to as *stayed unemployed*. Those who did not have a (decent) job at the pre-intervention time but did have one post-intervention are considered to have *gained employment*. Participants who had a (decent) job at both times are referred to as *stayed employed*.[15]

About two-thirds of participants stayed employed (63-74%) and about one-fifth gained employment when comparing respondents' employment status at baseline with their status 3 to 9 months after the first interview. Considering the follow-up 16 to 24 months after the baseline survey, the share of respondents who stayed employed was higher (69-82%) whereas the share who gained employment reduced (10-16%), especially, among KAM beneficiaries. Only very few participants stayed unemployed or even lost their employment and this share is reducing over time in all three study groups.

Across groups, participants of the two KAM beneficiary groups had a higher rate of staying employed than control group participants within all three observed time periods. Further, KAM beneficiaries had a higher rate of gaining employment 3 to 9 months after the baseline than control group participants. However, when long-term changes are considered, the no intervention control group had a higher rate of gaining employment than both KAM beneficiary groups for the times 10 to 15 months and 16 to 24 months after the baseline. When comparing the share of participants who lost employment across groups and survey waves Figure 2.8.1 illustrates that these shares were highest among control group participants at each follow-up. Interestingly, the share of individuals who lost employment strongly reduced over time for KAM beneficiaries who

[14] *The survey question in the baseline was "How many jobs have you held in the last 12 months before joining KAM?" for the treatment group and "How many jobs have you held in the last 12 months before this interview?" for the control group.*

[15] *The employment does not need to be in the same job.*

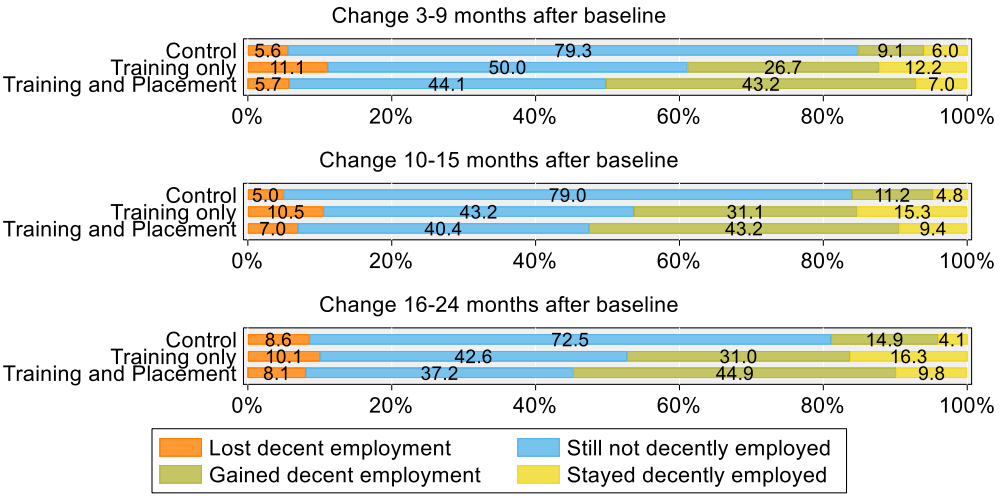
were trained and received a placement, whereas it slightly increased for KAM beneficiaries who were only trained.

Figure 2.8.1
Change in employment status by study group

Source: Own calculations based on KAM survey.

Focusing on decent employment in Figure 2.8.2 shows a positive picture, too. Among KAM beneficiaries about 30% of those who only participated in the KAM training and slightly more than 40% who participated in the full KAM program found a paid work of at least 20 hours per week that earns at least 6,209.93 KES per month after the baseline. Among control group participants, the share who gained decent employment is much lower and varies between 9 and 15%.

Figure 2.8.2
Change in decent employment status by study group



Source: Own calculations based on KAM survey.

The share of study participants who stayed decently employed varies between 7 and 10% among participants of the full program, between 12 and 16% among the training only group, and 4 and 6% among the control group. Consequently, the share of participants who remained without decent employment is noticeably smaller among participants of the full KAM program (37-

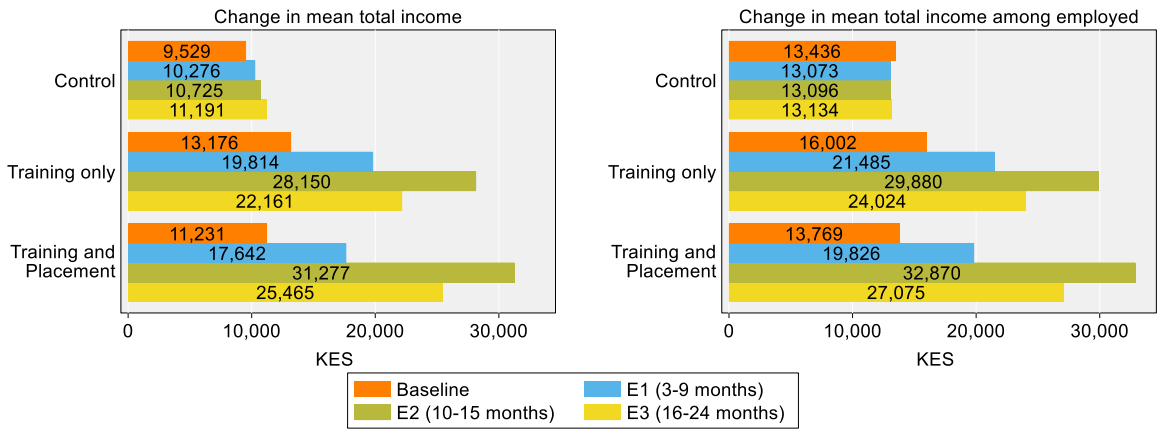
44%) and the training only group (43-50%) and much larger among control group participants (73-79%). A small share of participants lost their decent employment, and the share is highest among KAM beneficiaries who received both program components (10-11%).

Changes in income

The change in participants' mean monthly income in KES from all paid jobs is displayed in Figure 2.8.3. The graph compares participants' mean monthly income at pre-intervention (collected at baseline) with the income indicated in subsequent interviews, which took place 3 to 9 months, 10 to 15 months, and 16 to 24 months after the first interview. The left-hand side graph refers to mean monthly income among all respondents (i.e., irrespective of being employed or unemployed) and shows that while control group participants experienced only a slight but steady increase in their incomes, KAM beneficiaries reported a strong increase which was especially pronounced 10 to 15 months after the baseline. After 15 months, the income of KAM beneficiaries somewhat reduced in comparison to the follow-up at 10 and up to 15 months but was still noticeably higher than prior to the intervention. Comparing incomes reported at baseline with incomes reported 16 to 24 months later, control group participants experienced an increase in mean monthly income of 1,662 KES (=11,191 - 9,529), KAM beneficiaries who only participated in the training experienced an increase of 8,985 KES (=22,161 - 13,176), and KAM beneficiaries who participated in the training and received a placement through KAM experienced an increase of 14,234 KES (=25,465 - 11,231).

The right-hand side of Figure 2.8.3 presents a similar picture for mean monthly income among respondents who were employed at the time the respective interview. However, for control group members, the improvement in incomes vanishes which suggests that the improvement in incomes among all control group members stems from an increase in the employment rate rather than incomes among employed control participants.

Figure 2.8.3
Change in mean total income by study group

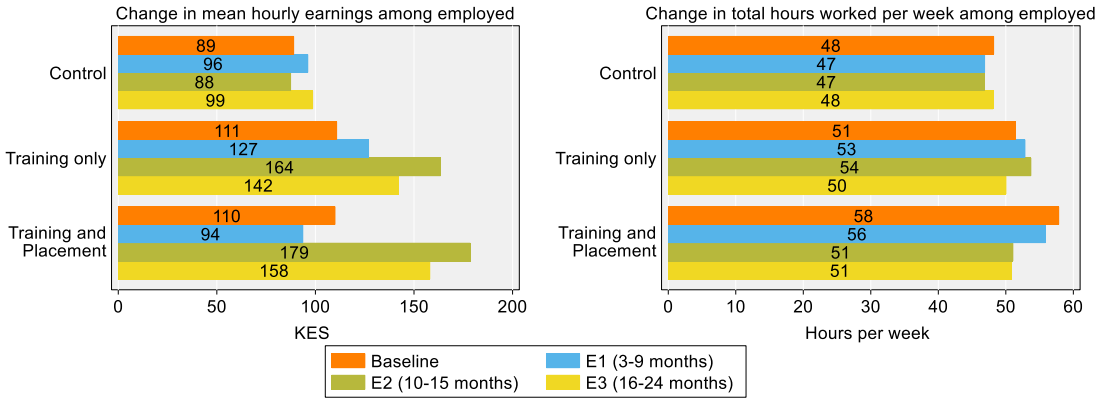


Source: Own calculations based on KAM survey.

Changes in hourly wages and working hours

In order to understand how and why total income, particularly among employed participants, changed from before to after the KAM program, it is important to understand how the average hourly wage and the number of hours worked changed among employed participants. Figure 2.8.4 illustrates that hourly wages changed considerably, whereas the mean working hours among employed changed only moderately and even reduced in some cases. Among employed control group participants, mean hourly earnings changed from 89 to 99 KES on average 16 to 24 months after the baseline. For the same period KAM beneficiaries reported an increase of 31 KES if they participated only in the training and an increase of 48 KES if they additionally received a placement by KAM. Regarding participants' working time no meaningful change among control group participants can be observed. Among KAM beneficiaries who only received the training, working hours increased 10 to 15 months after the baseline (+ 3 hours) but reduced thereafter (-1 hour). In contrast, KAM beneficiaries who were trained and received a KAM internship reported a considerable reduction in working hours by seven hours 10 to 24 months after the baseline. Consequently, the changes in mean incomes among employed pictured in Figure 2.8.3 are to a large extent due to changes in hourly wages.

Figure 2.8.4
Change in mean hourly earnings among employed by study group



Source: Own calculations based on KAM survey.

2.8.2 Participant-reported evaluation of KAM on current jobs

During the follow-up waves KAM beneficiaries were asked about their perception on how the KAM program impacted their job search and about how the program affected their current employment situation. The results are summarized in Table 2.8.1. Whereas perceived impacts were rather limited for participants' job search, participants attribute a large effect of the KAM program on their current employment situation. Especially, the skills developed during the program were perceived as very helpful with respect to getting a job and improving performance in the job. In addition, the E4D network in general was stated as useful by more than half of all KAM beneficiaries. Further, 20% of all KAM beneficiaries indicated to be currently employed in the company where they completed the internship. The perceived effectiveness of the KAM program was lower among female participants as summarized in the two right-hand side columns of Table 2.8.1.

Table 2.8.1
Impact of KAM program on current jobs

Number of participants who indicated that …	Total	% of baseline participants	Females	% of female baseline participants
… E4D network improved job performance	628	68.4	101	63.1
… E4D network helped to get a job	513	55.9	82	51.3
… skills developed in E4D program improved job performance	738	80.4	117	73.1
… skills developed in E4D program helped to get a job	772	84.1	123	76.9
… work readiness training helped to get a job	690	75.2	104	65.0
… practical job training helped to get a job	561	61.1	87	54.4
… they are employed in the placement company	189	20.6	41	25.6

Notes: Refers to any job indicated in any of the three follow-up waves. - Source: Own calculations based on KAM survey.

2.9 Estimated program effects on labor market outcomes

This section shows the results of the estimated program effects on labor market outcomes based on the estimation methods discussed in section 2.3.1. Specifically, it presents the estimated impact of having participated in both the KAM training and a KAM organized company placement in comparison to (i) control group participants who did not receive any KAM benefits and (ii) participants who only participated in the KAM training but did not receive a KAM placement. The results are based on observations from study participants of all observed TVET facilities.

The figures in section 2.9 show the estimated treatment effects of the KAM training and placement program impact, also referred to as *coefficients* or *point estimates*, as well as their respective 95% confidence interval. Confidence intervals indicate a range of values in which the true parameter, i.e., the true program effect, lies with a probability of 95% and can be understood as a measure of precision of the estimated treatment effects. The smaller the confidence interval the better we can pin down the size of the treatment effect. If a confidence interval does not include zero, the estimated treatment effect is said to be significantly different from zero. In such a case the estimated effect is considered to have a significant positive or negative effect (depending on the direction of the effect) on the respective outcome.

All presented employment measures are categorical indicators which only have two categories or levels. Therefore, treatment effect estimations on employment outcomes can be interpreted as percentage point changes in the respective outcome. For example, a coefficient of 0.1 for the indicator of employment status would be interpreted as a 10-percentage point increase in the employment rate due to the KAM program.

Income and wage indicators, which were measured in KES, were logarithmized using the natural logarithm (referred to as Ln(indicator) in the following figures). Using logarithms of income and wage variables has the advantage that the estimated effects can be expressed as percent changes and hence can be compared across contexts and different currencies. An estimated effect of 0.1 would then indicate a 10% increase, for example in monthly total income, due to the KAM program.[16]

Working hours are measured on a continues scale in hours per week. Treatment effect estimations on working hours can be interpreted in hours. A coefficient of 0.1 would, for example, be interpreted as an increase in working hours of 0.1 hours per week, i.e., 6 minutes, due to the KAM program.

2.9.1 Program impact on employment status

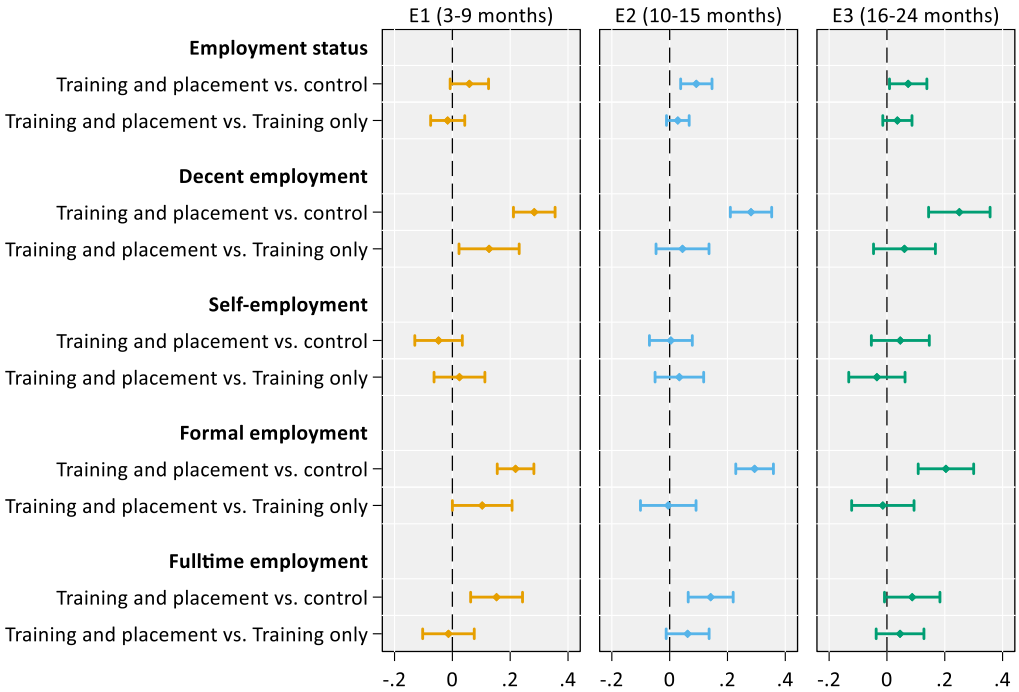
Figure 2.9.1 presents the estimated treatment effects of the KAM training and placement on (i) having paid work (employment status), (ii) having a decent employment, (iii) being self-employed, (iv) having a formal employment, and (v) having a fulltime employment for the sample of all TVETs over the different follow-up waves. For each employment indicator the graph shows the estimated impact according to RQ#1, i.e., *training and placement vs. control*, and RQ#2, i.e., *training and placement vs. training only*.

In comparison to the no intervention control group, participation in the KAM training and internship placement significantly increased participants' probability to be employed 10 to 24 months after the baseline survey (Endline 2 and Endline 3). For the time shortly after the baseline survey, i.e., 3 to 9 months after (Endline 1), no significant effect can be observed yet. The average program impact on participants' probability of having paid work is 7.3 percentage points compared to the control group 16 to 24 months after the baseline (E3). Further, the KAM program had immediate and lasting positive impacts on participants' probability of having a decent as well as formal employment compared to the control group. KAM beneficiaries who received the training and the placement showed a 25.0 and 20.4 percentage points higher probability of having a decent and formal employment 16 to 24 months after baseline, respectively. A significant positive impact on fulltime employment can only be observed for the time of three up to 15 months after the baseline but fades out thereafter.

In comparison to the KAM training only group, the additional effect of the KAM internship placement is small and often close to zero. Only during the first follow-up, 3 to 9 months after baseline, there are significant impacts on decent and formal employment of on average 12.7 and 10.3 percentage points, respectively. These effects cease to exist at later follow-ups.

[16] *For models with a logged outcome variable and a dummy explanatory variable, estimates cannot be directly interpreted as percentage changes (although they are still a close approximation for values smaller than 0.3). The correct interpretation in percentage changes is given by the following conversion:* $(e^{coef} - 1) * 100 = program\ effect.$

Figure 2.9.1
Estimated effects on employment indicator



Notes: "Training and placement" refers to treatment group which received the skills training plus placement. "Training only" refers to the control group which received only the skills training. "Control" refers to the control group which did not receive any KAM benefits. - Source: Own calculations based on KAM survey.

Overall, the results of Figure 2.9.1 suggest that the KAM program had a positive and lasting impact on employment outcomes of participants who participated in both program components and that the additional impact of a KAM-sought internship is rather limited. A discussion of the implication of these results and their robustness to alternative estimation methods will be discussed in sections 2.9.3 and 2.12.

2.9.2 Program impact on income

The effect estimates on (i) total income earned from all jobs among all respondents, (ii) total income earned from all jobs among employed respondents, (iii) and the average hourly wage from all jobs among employed respondents are displayed in Figure 2.9.2. As before, for each employment indicator the graph shows two effect estimations in alignment with RQ#1 and RQ#2.

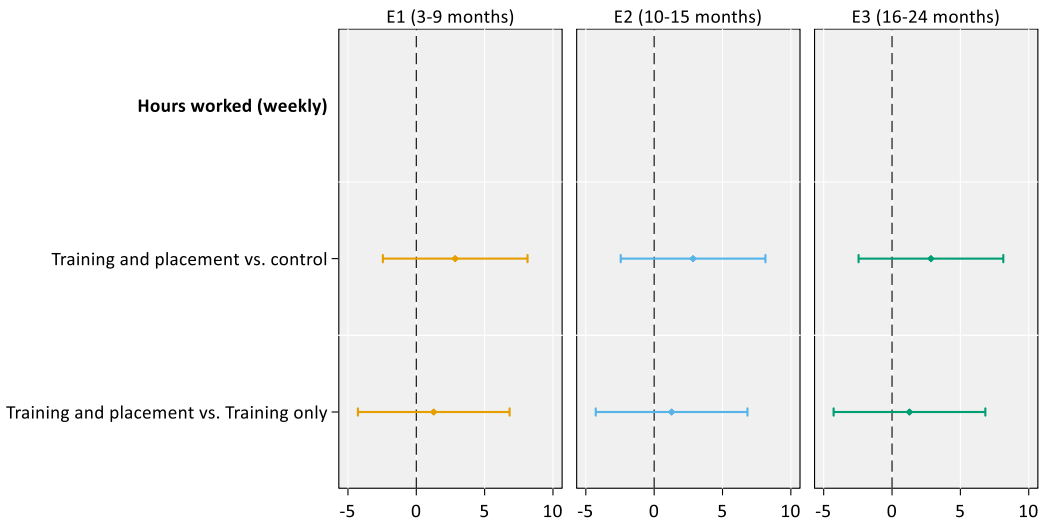
Figure 2.9.2 shows a very similar picture as for participants' employment status. The KAM program had a positive and lasting impact on income and wages for KAM beneficiaries who participated in both components compared to the no intervention control group. Participation in the full KAM program increased income among employed participants by 52.8% on average 16 to 24 months after baseline, compared to study participants who did not receive any KAM benefits. In contrast, we do not detect any additional impact of the KAM-sought internship compared to only participating in the KAM training.

Figure 2.9.2
Estimated effects on income and wages

*Notes: The coefficient for ln(monthly total income) among employed study participants on "Training and placement vs. control" for Endline 3 translates to an income increase of (exp(0.528)-1)*100 = 69.55%. - Source: Own calculations based on KAM survey.*

Figure 2.9.3 presents the estimated KAM program effects on working hours among employed respondents. The treatment effects are small for both comparisons and the confidence intervals include zero. Although the KAM program did not significantly affect the weekly working hours of those that were employed at the time of the survey, the positive coefficients indicate that the positive effect on monthly incomes might be the results of both, increased working hours and increased hourly wages.

Figure 2.9.3
Estimated effects on weekly working hours among employed



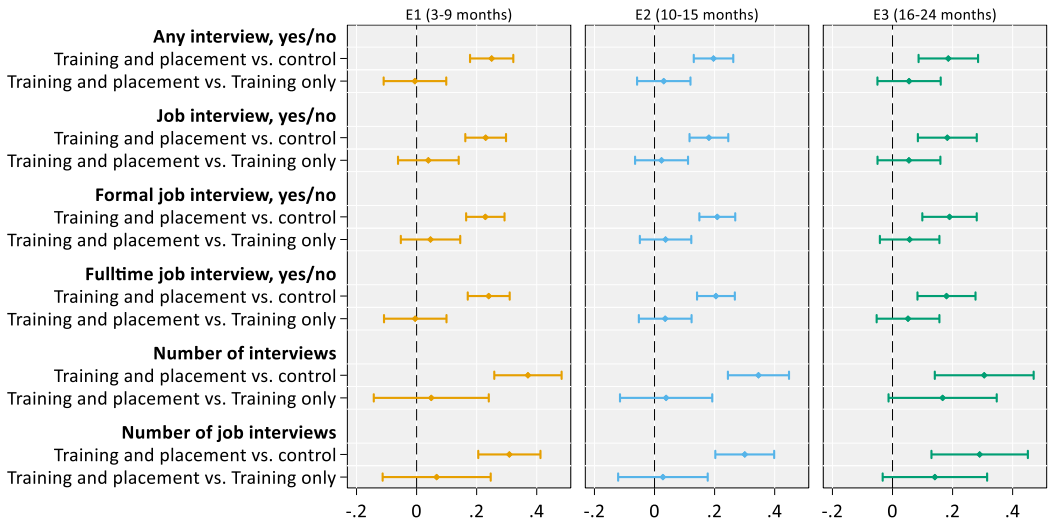
Source: Own calculations based on KAM survey.

2.9.3 Program impact on job search

Figure 2.9.4 shows the estimated impacts on job search indicators, such as whether participants had (i) any interview[17], (ii) a job interview, (iii) a job interview for a formal employment, or (iv) a job interview for a fulltime employment. Moreover, the number of interviews in general and job interviews in particular is assessed. All questions referred to interviews conducted since participants started the KAM program for the treatment group and since the baseline survey for the control group. Hence, interviews conducted in the course of the KAM internship itself are not covered.

The positive coefficients of the analysis comparing the training and placement group with the control group indicate that full KAM program participation significantly increased participants' probability in having had any form of an interview, a job interview, a formal job interview, and/or a fulltime job interview compared to not having received any KAM benefits. In comparison to the control group, KAM beneficiaries of the full program showed an 18 percentage points higher probability of having had an interview for a fulltime job 16 to 24 months after the baseline. The KAM program also significantly increased the number of (job) interviews for those who participated in the training and received an internship by 0.31 (0.29) in the third follow-up. We do not observe any additional impact of having received a KAM placement compared to the training-only group. The coefficients are small in magnitude and their confidence intervals always include zero.

Figure 2.9.4
Estimated effects on job search



Source: Own calculations based on KAM survey.

The strongly positive results of Figure 2.9.1 suggest that the previously found positive impact of participation in the full KAM program might have occurred thanks to beneficiaries' improved performance when searching for jobs after having participated in the KAM training. The job bazaar, which took place at the last training day and at which trainees could connect with

[17] *The survey question was "Have you had any job/internship interviews since you started the SOGA program on [date}?" for the treatment and "Have you had any job/internship interviews since we last interviewed you on [date]?" for the control group.*

companies, might have substantially improved chances for job interviews even for those trainees' who only participated in the KAM training and were not subsequently placed in an internship by KAM.

2.9.4 Robustness checks

The results presented in section 2.9.1 to 2.9.3 suggest that the KAM training and placement program had a significant and positive impact on employment and labor market outcomes for those who received the full KAM program and compared to study participants who did not participate in the KAM program at all.

In order to alleviate potential concerns about the comparability of treatment and control group participants with respect to obtained vocational education and, thus, labor market prospects, the estimations in sections 2.9.1, 2.9.2, and 2.9.3 controlled for the different TVET courses in which participants were enrolled in. To further test the robustness of the results, the analyses in sections 2.9.1 to 2.9.3 were repeated in Appendix A1.2 using alternative treatment and control group compositions. In one robustness check, control group participants who were enrolled in a CBET program rather than the traditional TVET program were excluded from the analysis.[18] In a second robustness check, the sample was restricted to study participants of mixed vocational training institutes and control group participants who were enrolled in CBET were also excluded. Consequently, in the second robustness check, participants from TVET facilities that were exclusively visited by KAM beneficiaries or exclusively visited by control group participants were excluded.

Figures A1.2.1, A1.2.3, and A1.2.6 in Appendix A1.2 present the results for the first robustness check, which excludes CBET students and graduates from the sample of control schools. Note, that there are no CBET students at treatment schools. As this change in the sample only regards the control group composition, the analysis focuses on the comparison of the KAM training and placement group and the no intervention control group.[19] Dropping CBET control group participants from the sample drastically reduces the number of observations in the control group (Figures A1.1.2 and A1.1.3). For the third follow-up wave conducted 16 to 24 months after the baseline only two observations remain. Hence, the robustness checks for RQ#1 (training and placement vs. control) is based on follow-up waves one and two only. The estimated treatment effects on employment, income and job search outcomes using this smaller control group sample are similar to those presented in sections 2.9.1 to 2.9.3.[20] The effects on total income (among all participants and among employed participants only) and on hourly earnings among employed start to materialize only ten months after the first interview, i.e., the effects are not significant

[18] *The Competency Based Education and Training program CBET is an alternative mode of training, which was implemented in some Kenyan TVET institutions in recent years. It shall address concerns that the previous Kenyan TVET system was largely based on theoretical training which gave less attention to the assessment of competences as required in the workplace. The new CBET program aims to improve the quality of training by adapting an industry and business demand-led model and establishing approved and industry-validated occupational standards for all vocational jobs and trades.*

[19] *Estimation results of RQ#2 are not based on control group participant and therefore remain unaffected when changing the sample of study participants who did not receive any KAM benefits. Hence, the graphs only show the estimation coefficients for RQ#1.*

[20] *For the time 10 up to 15 months after the baseline we still observe a significant increase in the probability of having a decent, formal, as well as fulltime employment among KAM beneficiaries who received both program components compared to study participants who did not receive any KAM benefits (22.0, 26.2, and 11.2 percentage points, respectively).*

in the short term 3 to 9 months after the baseline. For job search outcomes, the positive impact remains almost unchanged 3 to 15 months after the baseline.

Figures A1.2.2, A1.2.4, and A1.2.7 in Appendix A1.2 present the results for the second robustness check. When the estimation sample is restricted to mixed TVETs and excludes CBET control group participants, the estimated effect sizes remain similar for most outcomes to those in section 2.9.1 to 2.9.3 for RQ#1 but are mostly insignificant. The loss of precision and, thus, significance, is likely due to the reduction in sample size. For participants' employment status, only the effect for formal employment remains significant for the time of 10 to 15 months after baseline. KAM beneficiaries who participated in the full program are more likely than the control group to have a fulltime employment 3 to 9 months after the baseline, but this effect vanishes for the second follow-up survey. For income, there are no significant short-term impacts 3 to 9 months after baseline. However, 10 to 15 months after the baseline, total income among employed participants was significantly higher in the KAM training and placement group compared to the control group.[21] The hourly earnings among employed as well as the total income among all participants were only slightly higher among KAM program participants compared to the control group and the difference is found to be insignificant. The effects on job search vanish 10 months after the baseline.

Concerning the estimation results of RQ#2, the restriction to mixed TVETs does not change the overall null-effect of having received an internship sought by KAM in addition to the KAM training. Only for the time of 16 to 24 months after the baseline survey a significant increase in the probability of decent employment (16.9 percentage points) can be observed.

Figure A1.2.5 displays the estimated treatment effects on working hours among employed participants for the two alternative sample compositions. The results continue to not show any significant impact on working hours for both research questions.

2.10 Program impact for sub-groups of study participants

The results presented in section 2.9 refer to the average treatment effects for the whole study population (or those in mixed TVET facilities only). However, the treatment effects for specific sub-populations may differ from those of the whole study population. Therefore, this section estimates the impact of the KAM training and placement program for sub-samples of participants with specific background characteristics. The considered background characteristics are:

1. Respondents' sex: **female** versus **male**,
2. Respondents' age: **less than 25 years** versus **25 years and older at baseline**,
3. Respondents' prior work experience: **had sustained work experience**, defined as having worked six months in one job, prior to the intervention versus **no previous sustained experience**.

The sub-samples are constructed from the total sample of all TVETs, including control group participants from CBET programs. Otherwise, the sub-sample sizes would be too small. For each of the main outcome variables, two separate estimations were conducted: one for the sub-population fulfilling the respective characteristic and one for the sub-population not fulfilling it.

[21] *For the time 10 to 15 months after the baseline, having participated in the full KAM program increased the income among employed participants by 58.5% on average and compared to study participants who did not receive any KAM benefits.*

A potential heterogeneity in the treatment effects can be assessed by comparing the estimated treatment effects of the respective sub-populations.

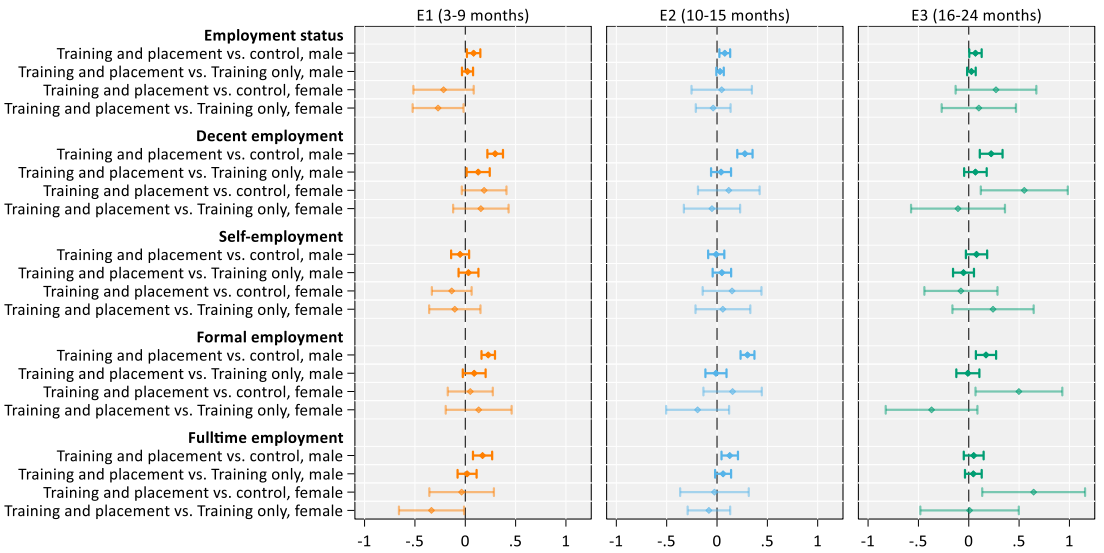
The figures presented in this section only include the estimated treatment effects on employment indicators. The results for incomes, wages, and working hours are presented in Appendix A1.3.

2.10.1 Program effect heterogeneity by respondents' gender

The results of the sub-sample analysis with respect to gender are displayed in Figure 2.10.1. Results displayed in darker colour refer to the sub-sample of respondents who are male, whereas the results in lighter colour refer to the sub-sample of female respondents. Since the study sample covers much fewer women (N=291) than men (N=1,821), the results for the female sub-sample are estimated with much lower precision than those of the male sub-sample resulting in much broader confidence intervals for estimates in the female sub-sample.

Figure 2.10.1 shows that in comparison to male control group members, male KAM beneficiaries experienced an immediate and persisting positive impact on employment status and the probability of having a decent, formal, and/or fulltime job (RQ#1). In contrast, the program impact on women tends to be negative in the short-term and only materializes after 16 months, if at all. Compared to the control group, a positive and significant impact of the KAM program on females' probability of decent, formal, and/or fulltime employment can be detected 16 to 24 months after baseline. Although these long-term effects are of much larger size than for men, ranging between an increased probability of 50 and 65 percentage points, they are less precisely estimated. With regards to RQ#2, the effect on employment status and the probability of fulltime employment is negative for female KAM beneficiaries who received a KAM placement in comparison to women of the training-only group.

Figure 2.10.1
Estimated treatment effects on employment indicators, by gender



Source: Own calculations based on KAM survey.

Appendix A1.3 presents the program effects by respondent's gender on income, wages, working hours, and job search and reveals a very similar picture: Most of the positive impacts are only found for the male sub-sample. Among females a positive effect of the full KAM program can

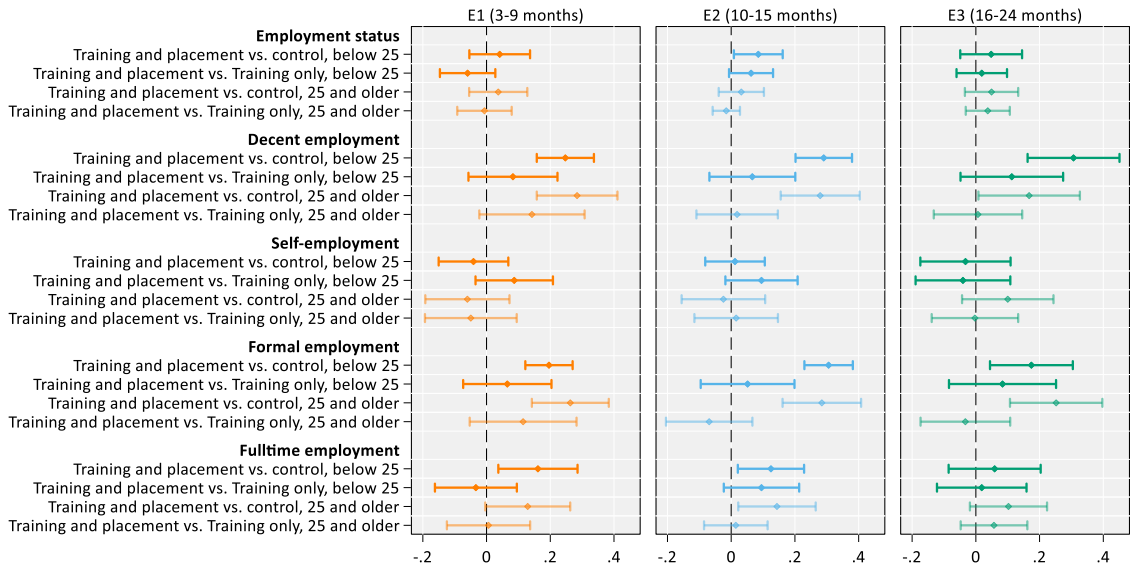
only be observed for monthly income among employed and only for the time 16 to 24 months after the baseline (RQ#1). Comparing female KAM beneficiaries who received the full program with female beneficiaries who only received the training, the treatment effect on total income is significantly negative in the first follow-up (RQ#2). Also, the persisting positive impacts on job search indicators in the setting of RQ#1 can exclusively be observed for male KAM beneficiaries. However, female participants of the full KAM program significantly increased their working hours 10 months after the baseline in comparison to female control group participants, whereas there is no effect on working hours for males.

2.10.2 Program effect heterogeneity by respondents' age

Figure 2.10.2 presents KAM program effects on employment indicators in (i) a sub-sample of respondents who are younger than 25 years old in darker color and (ii) a sub-sample of respondents who are 25 years or older in lighter color. Although, the coefficients vary in size and, sometimes, even in their direction, the overall pattern of effect sizes is similar across the two sub-samples.

In both sub-samples, the effect of the full KAM program is significantly positive for decent, formal, and fulltime employment in comparison to the no intervention control group. The effect of the full KAM program in comparison to the KAM training only is insignificant for all observed indicators.

Figure 2.10.2
Estimated treatment effects on employment indicators, by age group



Source: Own calculations based on KAM survey.

Appendix A1.3 presents the program effects by respondents' age group on income, wages, working hours, and job search. For income and wages the coefficients are very similar across younger and older participants who are employed at the time of the interview. However, for total monthly income irrespective of the employment status, the results in Appendix A1.3 suggest that the KAM program impact starts to materialize later for older participants in comparison to the no intervention control group. For participants' weekly working hours, we do not detect any significant impact neither for the younger nor for the older subsample and effect sizes are almost identical across the samples. In comparisons of KAM training and placement beneficiaries

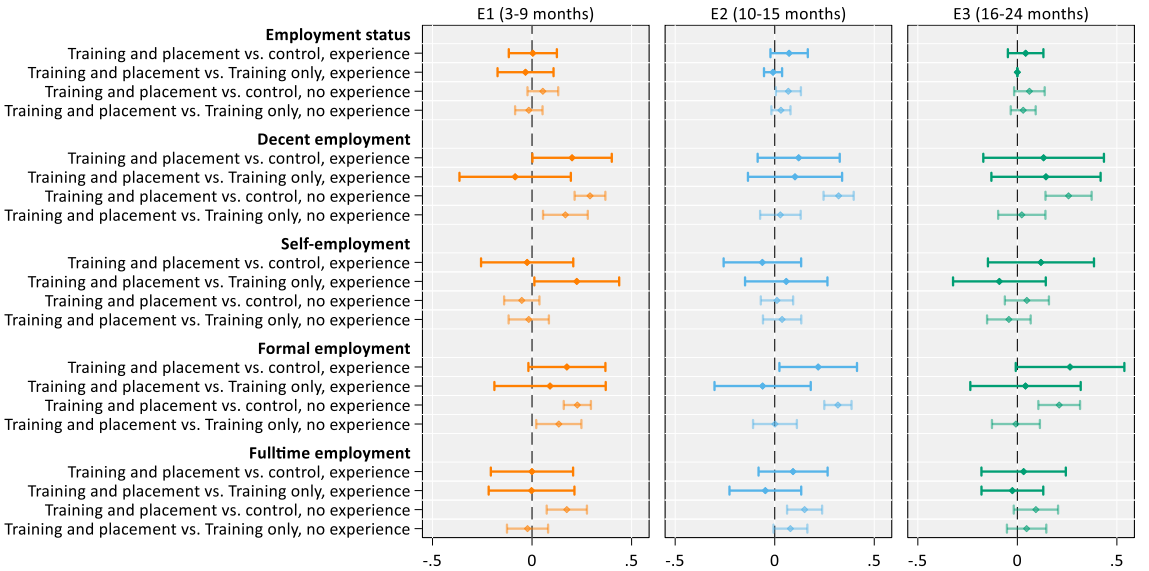
with KAM training beneficiaries only, the effects on job search are larger for older participants, especially 16 to 24 months after the baseline. The comparison of KAM beneficiaries who participated in the full KAM program with beneficiaries who only received the training suggests a negative but insignificant impact of the KAM internship for older beneficiaries in the short-term which vanishes 10 months after the baseline survey.

2.10.3 Program effect heterogeneity by respondents' work experience

Figure 2.10.3 presents the KAM program impact on employment in (i) a sub-sample of respondents who had sustained work experience at baseline in darker colors and (ii) a subsample of respondents who did not have sustained work experience in lighter colors. Prior sustained work experience is defined as having worked six months in one job prior to the intervention. Since most KAM beneficiaries (84%) did not have sustained work experience prior to the program start, the results for the experienced sub-sample are estimated with much lower precision than those of the no-experience sub-sample resulting in much broader confidence intervals for estimates in the sub-sample of experienced participants.

Most of the confidence intervals displayed in Figure 2.10.3 include zero for participants with experience, suggesting no impact for both research questions. In contrast, the effect of the full KAM program in comparison to the control group is large and persistent for KAM beneficiaries without prior work experience. The estimated effect for decent and formal employment even suggests a significant positive impact of additionally receiving an internship placement through KAM 3 to 9 months after the baseline for KAM participants without experience.

Figure 2.10.3
Estimated treatment effects on employment indicators, by experience



Source: Own calculations based on KAM survey.

Appendix A1.3 displays the program effects by respondents' prior work experience on income, wages, working hours, and job search and reveals a very similar conclusion as Figure 2.10.3. The positive program impact on participants income can almost solely be observed among participants without prior experience. For job search indicators, the positive impact of participation in the full KAM program is persistent over all follow-up waves among participants without experiences but vanishes after 10 months for those who already had sustained work experience. With

respect to working hours, there is no significant impact in any of the sub-samples but again the results suggests that the effects might differ across samples, especially in the short run: 3 to 9 months after the baseline survey the estimated coefficient is positive for participants with experience, indicating an increase in working hours, whereas it is negative for those without prior experience, indicating a decrease in weekly working hours.

2.11 Program impact on secondary outcomes

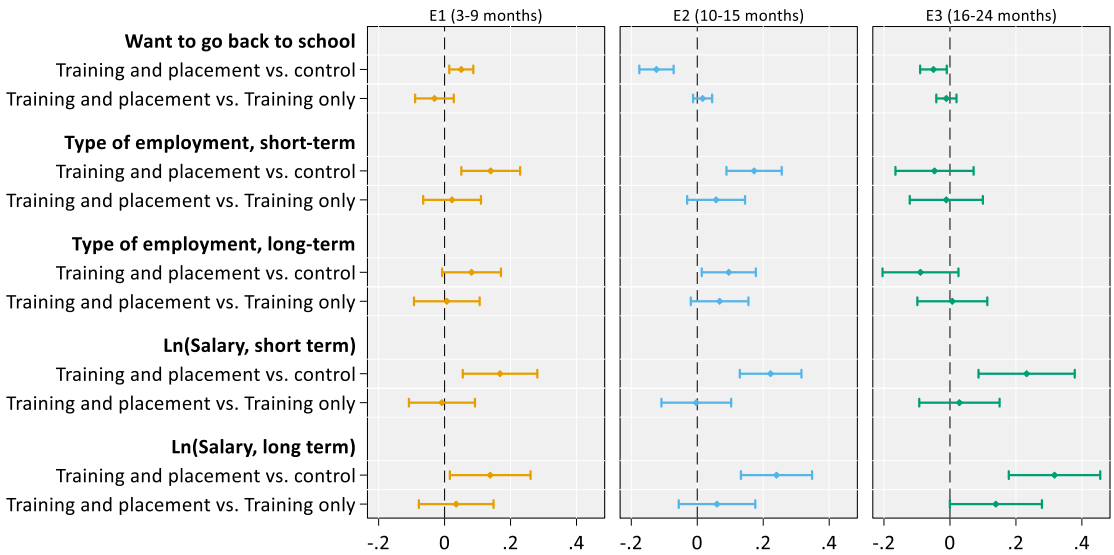
Besides employment, earnings, and job search, the KAM program may have affected participants' lives in several other ways. This section examines whether participation in the KAM program influenced participants' (i) aspirations with respect to future education and labor market outcomes, (ii) household and living conditions, as well as (iii) banking and savings behavior. All subsequent results are based on estimations including observations from study participants of all TVETs.

2.11.1 Employment and earnings aspirations

Figure 2.11.1 shows the estimated KAM program impact on respondents' aspirations regarding (i) the wish or plan to go back to school, (ii) whether they aspire salaried or self-employed employment in the short- and long-term, as well as (iii) their salary expectations for the short- and long-term.

While KAM beneficiaries had a significantly higher desire to return to school immediately after the baseline this reversed in subsequent follow-up waves for which we observe that KAM beneficiaries of the full program were significantly less likely to desire to return to school than the control group. With respect to the aspired type of employment, we observed a similar reversion in the effect over time. For the time 3 to 15 months after the baseline, KAM beneficiaries were significantly more likely to aspire a salaried position than control group participants.

Figure 2.11.1
Estimated treatment effects on participants' aspirations



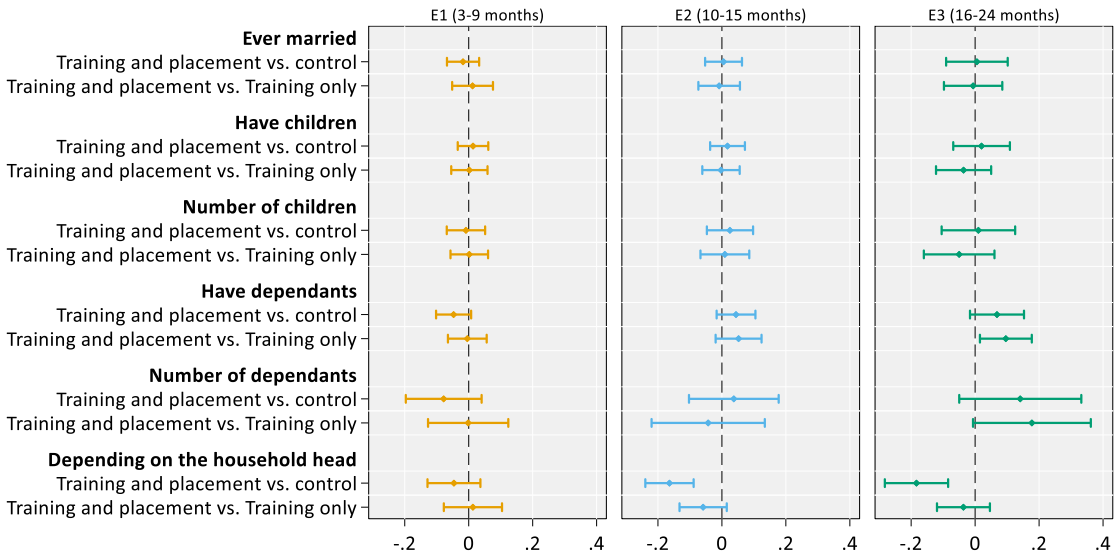
Source: Own calculations based on KAM survey.

However, the effect reversed 16 to 24 months after the baseline, when KAM beneficiaries were more likely to aspire a self-employed position, though, this effect is not significant. With respect to salary aspirations in the short- and long-term, participants of the full KAM program had significantly higher aspirations for their future salary than control group participants. For comparisons of KAM training and placement beneficiaries with KAM training only beneficiaries (RQ#2) there are no significant differences in participants' aspirations for their future education and labor market outcomes.

2.11.2 Family structure

Figure 2.11.2 displays the effect estimates on participants' family structure. We do not detect any impact of the KAM program on participants' marital status nor on having children or dependents. However, Figure 2.11.2 results suggest that participation in the KAM training and placement program significantly reduced participants' probability to depend on the household head 10 to 24 months after the baseline in comparison to control group participants. For KAM training and placement beneficiaries in comparison to KAM training only beneficiaries, the effect on dependency on the household head is also negative but insignificant.

Figure 2.11.2
Estimated treatment effects on family structure

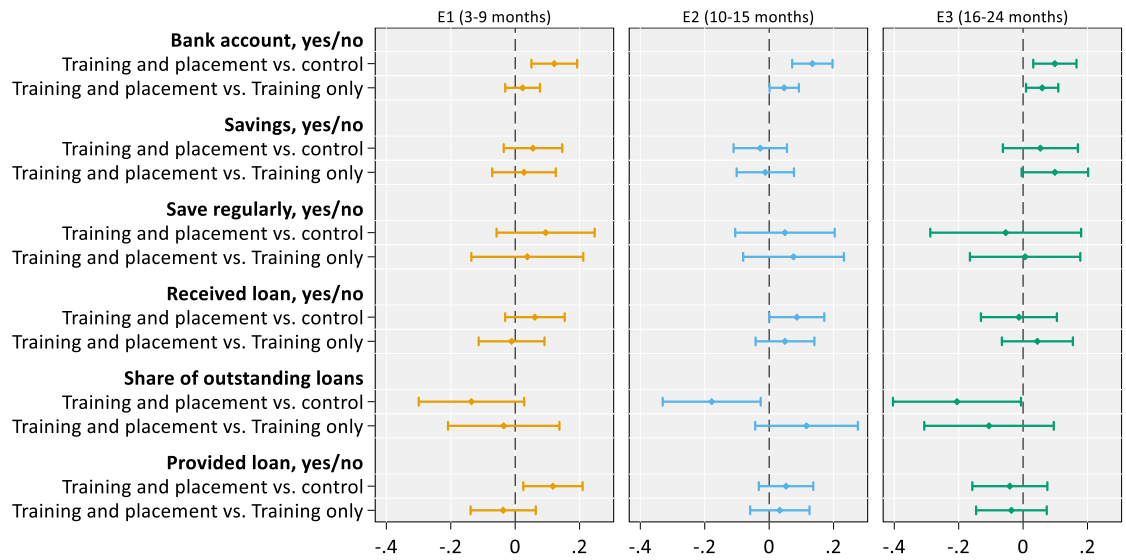


Source: Own calculations based on KAM survey.

2.11.3 Banking and savings

The estimated program impact on participants' banking and savings behaviour is presented in Figure 2.11.3 The KAM program had a positive and lasting impact on beneficiaries' probability of having a bank account. Remarkably this effect is not only significant when comparing beneficiaries of the full KAM program to control participants but also when comparing them to the training only group, suggesting an important additional impact of participating in a KAM internship placement. While there seems to be no program effect on beneficiaries' savings behaviour 10 to 24 months after the baseline, having participated in the KAM training and placement program significantly reduced beneficiaries' share of outstanding loans compared to the control group.

Figure 2.11.3
Estimated treatment effects on banking and savings behaviour



Source: Own calculations based on KAM survey.

2.12 Summary and discussion

This study aimed at assessing the impact of the KAM program on employment and labor market outcomes among vocational training graduates in Kenia. The KAM program consisted of a two-to-three-days training in work readiness and a subset of trained beneficiaries were matched into internships at KAM member companies. This impact evaluation focuses on the effectiveness of receiving both program components as a measure to overcome the so-called *skills gap,* which is considered a major challenge for youths in Eastern Africa and refers to the lack of practical experience of vocational training graduates and related difficulties in the school-to-work transition.

KAM beneficiaries had graduated from their respective vocational training institute during the past 5 years and control group participants also already completed or were about to graduate from their TVET facility at the time of the baseline survey. Study participants were interviewed four (KAM beneficiaries) or three (control group) times over a period of up to two years. The surveys collected information about the participants' current and retrospective employment status, earnings, and sociodemographic background characteristics.

Two research questions were studied. The first research question examined the gains for vocational training graduates from participating in the KAM training and internship placement organized by KAM (RQ#1). The second research question addressed whether vocational training graduates benefitted from any additional gains if they received an internship placement organized by KAM in addition to participation in the KAM training alone (RQ#2). The impact of the KAM program was measured for job search, employment and income indicators and their heterogeneity with respect to gender, age, and prior sustained work experience. In addition, participants' employment aspirations, family structure, as well as banking and savings behavior have been assessed.

Overall, the main results suggest that the KAM program improved the employment situation of its beneficiaries in several dimensions. With regards to **RQ#1** the analysis shows that vocational training graduates who participated in the KAM training and internship placement experienced a noticeable improvement in their labor market outcomes. Particularly striking are the large

effects of 25.0 and 20.4 percentage points in decent and formal employment which were sustained until 24 months after the baseline survey. Also, the effect sizes on fulltime employment were positive and large, however, not significantly so anymore 24 months after the baseline survey. Additionally, treated participants significantly improved their incomes by 52.8% and these improvements were also sustained until 24 months after the baseline survey. While working hours increased only by a few hours per months and not significantly so, the improvement in incomes was driven by a higher pay per hour worked, which significantly and sustainably increased by 37.2% until 24 months after the baseline survey. These effect sizes are remarkable. In a recent meta-study, Kluve et al. (2019) systematically reviewed and compared 113 impact evaluations of youth employment interventions worldwide. The estimated program effects of the KAM program are comparable to the standardized average effect sizes found in Kluve et al. (2019) of 0.04 standard deviations in income and 0.03 standard deviations in hourly wages. The standardized effect sizes for participants' employment probability are smaller in magnitude (0.02 standard deviations) than those found in Kluve et al. (2019) (0.06 standard deviations).[22] The results of the KAM program evaluation provide additional evidence for the positive impact of youth employment interventions in low-income countries and contrasts to rather sobering results of programs in developed countries (e.g., Card et al. 2017). Further, Kluve et al. (2019) and Card et al. (2017) suggest that the impact of youth employment interventions increases in the long-term. However, the KAM program effects are present already 3 to 9 months after the baseline survey and do not increase throughout the considered time periods.

In addition to employment and income, the KAM training and placement program also significantly improved the probability for an interview for a formal job by 19.0 percentage points and for a full-time job by 18.0 percentage points sustainably until 24 months after the baseline survey. Further, the program significantly and sustainably reduced participants' financial dependency on the household head by 18.3 percentage points and increased the probability of having a bank account by 9.9 percentage points. The effect on bank accounts is remarkable given the importance of the banking sector on growth and behaviour and the attention that the concept of "banking the unbanked" received in the recent policy and academic debate with respect to financial inclusion, reduction of transaction costs, and more efficient payment of government transfers (e.g., Barajas et al. 2020, Dupas et al. 2018, Grohmann, Klühs, and Menkhoff 2018, and Demirguc-Kunt et al. 2017).

Generally, the effect sizes on employment and income tend to be larger for participants without prior work experience, although the coefficients of the two subsamples with and without experience do not differ from each other significantly. By gender, the effects for decent, formal, and full-time employment are particularly large for women 16 to 24 months after the baseline survey, although the effect sizes for women are smaller than those for men at earlier follow-up periods. However, given the small sample sizes, we cannot statistically differentiate the treatment effects of men and women (i.e., the effects by gender do not statistically differ from each other). One disadvantage of the heterogeneity analysis is that the sample is split into smaller sub-samples which, due to the smaller sample sizes, have less statistical power to detect effect sizes. The

[22] *To summarize individual study results and to allow for a comparison of treatment effects across outcomes and studies, Kluve et al. (2019) used standardized mean difference (SMD). The SMD is defined as the ratio of the estimated treatment effect for a specific outcome relative to the outcome's standard deviation within the study sample. In order to compare the effects of the KAM program to those included in the meta-study, the same approach was applied to the treatment effects estimated within this evaluation.*

subsamples of women and experienced participants are particularly small and, therefore, the effect sizes of these groups are imprecisely measured and insignificant.

While these results are promising, not all effect estimations were robust to excluding CBET graduated from the control group observations. Overall, the effects remain significant and of similar magnitude for the time 10 to 15 months after the baseline survey but previously found effects in the longer term until 24 months after the baseline survey now turn insignificant. When additionally restricting the sample to treatment and control group observations that attended the same TVET institutions most effects cease to exist. The only results that remain significant are short term effects (3-9 months) on participants' job search outcomes and their probability of having a fulltime employment as well as having a formal employment for the time 10 to 15 after the baseline survey.

Concerning **RQ#2**, the estimation results suggest that the impact of receiving an internship placement in addition to participating in the KAM training alone is small and insignificant. Although, there were significantly positive effects of 12.7 and 10.3 percentage points on decent and formal employment in the short term, i.e., 3 to 9 months after the baseline survey, the effects did not sustain in the longer term until 24 months after the baseline survey. The heterogeneity analysis suggests that the short-term effects on decent and formal employment were driven by participants without prior work experience. Except for these short-term impacts on decent and formal employment, there were no significant effects of the KAM internship placement in addition to the training on other employment and income outcomes. Similarly, there were no effects on job search outcomes, such as the probability or number of job interviews. However, 16 to 24 months after the baseline survey, KAM beneficiaries were significantly more likely to have dependents and a bank account by 9.6 and 5.9 percentage points, respectively. These findings are in line with Card et al. (2017) whose meta-analysis focused on the effect of active labor market programs on participants' employment probability and took into account different types of programs. Their meta-analysis shows that the effects of job search assistance programs are smaller in magnitude in comparison of training programs of unspecified length.

The heterogeneity analysis of RQ#2 further shows that KAM's placement component had significant negative short-term impacts on women's general employment status (i.e., not decent or formal employment), fulltime employment, and total monthly income 3 to 9 months after the baseline survey. However, these negative effects cease to exist at later follow-up periods. One potential explanation for the adverse short-term effects could be that the KAM sought internship placements were less suited for female than for male participants. However, a battery of alternative explanations, specifically with respect to gender roles and norms, could explain the adverse short-term effects for women.

Jointly, the results of RQ#1 and RQ#2 imply that the training component of the KAM program was more effective in improving labor market outcomes than the labor market attachment component. One potential mechanism could be the change in job search outcomes. While the overall KAM training and placement program increased the probability and number of job interviews, the placement alone did not affect job search outcomes in addition to the effects of the KAM training. An important design aspect of the KAM training might have been the job bazaar, which took place at the last training day and at which trainees could meet with company representatives and human resources staff. This feature of the KAM training might have improved the trainees' chances for job interviews. Individuals who did not receive a company placement directly arranged by KAM were not considered as trained and placed but as trained only. However, of the trained only group 20% of the participants completed an internship placement and in addition some directly acquired employment. Unfortunately, these internship and job placements are

coded as self-sought and it cannot be identified whether some of the placements of trained only participants were actually facilitated through the KAM training and the included job bazaars.

Another interesting result is that the treatment effects were more pronounced among participants without prior work experience. For unexperienced participants the KAM program might have had more leverage in bridging the skills gap and linking participants to employers, in particular through the job bazaar, than for participants with work experience who have better networks and practical skills already.

The central limitation of the study is that it cannot be ensured that the applied estimation methods address all differences between the treatment and control group that confound and, thus, bias the impact estimation. For example, there is limited information available about the educational attainment of participants and the timing of the surveys of the treatment and control group. However, the educational attainment and the timing of the surveys constitute important dimensions along which treatment and control groups should be comparable for our effect estimations to be valid. To alleviate this limitation to the extent possible, the analysis controls for relevant background characteristics of participants that can be observed, including the TVET course in which participants were enrolled in. However, information about the time at which treatment group participants completed their TVET program is unavailable and, thus, KAM beneficiaries who graduated five years ago may be compared with control group members who are just about to graduate. The labor market prospects of these two types of participants are likely very different and could drive the measured treatment effects to some extent. The presented estimates could therefore be up-ward biased and the true effects could be smaller or even zero. To affirm the results presented in this report and to obtain additional evidence on the impact of training and placement interventions such as the KAM program it is recommendable to combine future interventions that are similar to KAM with rigorous impact evaluations.

3 Qualitative evaluation of KAM's `sustainability program'

3.1 Introduction: Sustainability problematic of development cooperation projects

Globally, many donor-funded development projects are not sustained once donors withdraw from a project. Thus, in contrast to the donors' intentions, most projects fail to bring long lasting sustainable benefits to their beneficiaries and target groups (cf. Ochunga and Awiti 2017; Cekan 2020; Hofisi and Chizimba 2013; Morán and Ferguson 2013; Karanja 2014). A major reason is often the lack of ownership among local stakeholders. If no 'local owners' exist, no one takes responsibility to continue the project activities after external development assistance ends (Ballantyne 2003; Rey-Moreno et al. 2014; Oino et al. 2015; Okun 2009). Therefore, a central interest of international development is "to learn how to design projects so that the activities and benefits that are achieved during the life of the project will continue after the donor's funding has ceased" (Bossert 1990, 1015). This study applies this quest to the context of GIZ's "*Promoting Youth Employment Through Technical Human Capital Development"* program in Kenya (GIZ 2017; KAM 2017).

3.2 Overview of case background: KAM TVET project

The project under consideration was officially launched in January 2017 and has since been implemented by the Kenyan Association of Manufacturers (KAM). KAM is an association that represents over 40 percent of the manufacturing and value adding industries in Kenya (GIZ 2017; KAM 2020a; KAM 2020b; KAM 2017; Table A 2.1.1 [3]). The project is shortly referred to as the 'KAM project'. It aims to improve access to jobs and economic opportunities for youth in Kenya by placing technical vocational education and training (TVET) graduates for practical hands-on training in internships[23] in manufacturing companies[24] which are members of KAM (cf. GIZ 2018; GIZ 2017; Table A 2.1.1 [4]). The objective of the practical training during internships is to increase TVET graduates' employability and ideally enable them to find formal employment in the internship-offering or other manufacturing companies (GIZ 2017; Table A 2.1.1 [1]). In doing so, the project promotes youth employment in Kenya and addresses the skills mismatch that exists between the labor market requirements in and around natural resource-based industries and what is taught in training institutions (ibid.).

KAM facilitates the implementation of the project by reaching out to TVET institutions all over Kenya to identify graduates with technical profiles that match the companies' needs. The graduates are then trained in two-to-three-days' 'work-readiness trainings'[25]. After the work readiness trainings, KAM matches the graduates with companies, either by organizing job bazars[26] or by directly sending CVs of potential candidates to companies. After an internship period of three to six months, the companies themselves can decide whether they offer a formal employment position to the intern (cf. ibid.; Table A 2.1.1 [2]; [3]; [1]).

---

[23] *Internships are offered in areas such as electrical or mechanical engineering, installation, welding and fabrication, plumbing, or carpentry.*

[24] *The internship offering companies have sectoral affiliations (among others) to the food, beverage, textile, leather, pharmaceutical, plastics, as well as automotive sector.*

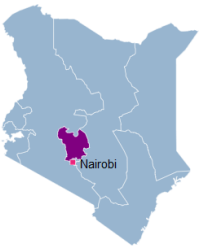
[25] *The work-readiness trainings are conducted by third party training providers and are focused on soft-skills training and job applications (e.g., CV writing, conduct in job interviews) (Table A 2.1.1, [2]; [1]).*

[26] *Job bazars are held a few times a year and usually follow a work-readiness training. At the job bazars company representatives have the possibility to interview the candidates they are interested in. Subsequently, companies can short list the interns and KAM then allocates them to the respective companies for internships (Table A 2.1.1, cf. [12]; [6]; [18]; [13]; [5]; [16]).*

The project follows a partnership approach in which KAM facilitates the implementation of the project and GIZ provides support offering technical expertise, capacity building as well as funding for all project activities including the KAM staff who is implementing the project (Table A 2.1.1, cf. [1]; [4]; [2]).

During the first phase of the project (2017/2018), the internships were funded by GIZ with stipends for the interns of 10.000 KES (approx. 80€) per month. In the beginning of 2019, a pilot program called 'sustainability program' was introduced in one pilot region in Kenya. Under the 'sustainability program', the GIZ funding of the internship stipends was stopped, and the companies were asked to take over the funding of the internship stipends on a voluntary basis according to their own employee policy. In case a company in the pilot region decided not to pay the KAM interns, no further interns were supposed to be matched to this company (Table A 2.1.1, cf. [2]; [1]).

The region in which the sustainability program was introduced is known as 'Central Chapter'[27]. Within this region, particularly companies from one county, called Kiambu county[28], participated. The Central Chapter was chosen as a pilot region because the Central Chapter Committee (CCC), a board of eight elected companies representing the KAM member companies in the region, was particularly willing to pilot the 'sustainability program' (cf. section 3.7.1.3 for more details).



Outside of the pilot region, the GIZ stipend scheme continued over the course of 2019 until the end of January 2020 when the first phase of the project ended. In the second program phase, which started in May/June 2020, the GIZ stipends were discontinued for all regions and the sustainability program, in which the companies are supposed to take over the internship funding themselves, was up-scaled to all other project regions (Table A 2.1.1, cf. [4]; [2]; [3]; [1]).

The reason for introducing the sustainability program (i.e., switching the funding responsibility for the stipends to the company level) was to pilot a 'sustainability strategy' that had the purpose of investigating whether companies in the pilot region would sustain the internship placements without the help of GIZ funding. The objective was to convince many companies to take over the responsibility of funding the KAM interns' stipends, so that the internships would be sustained once GIZ's participation in the project comes to an end (Table A 2.1.1, cf. [4]; [2]; [1]).

3.3 Research questions and design

This study investigates the following three research questions:

1. How has the switch in the internship payment scheme affected the companies' ownership of the internship placements in the pilot region?
2. How has the switch in the internship payment scheme affected the sustainability of the internship placements in the pilot region?
3. How can the companies' ownership and the sustainability of the internship placements be enhanced?

---

[27] *KAM divides its member companies into seven Chapters, which represent different regions in Kenya. One of them is the Central Chapter, which consists of nine counties (Kiambu, Kirinyaga, Murang'a, Nyeri, Embu, Meru, Isiolo, Tharakanewi, Laikipia) (Table A 2.1.1, cf. [1]; [19]; [20]).*

[28] *Kiambu county comprises over 80% of all manufacturing companies in the Central Chapter and holds all companies that were recruited to participate in the sustainability program (Table A 2.1.1, cf. [1]; [19]; [20]).*

The first research question evaluates the companies' ownership of the internship placements because local ownership of a development activity is considered a major determinant of sustainability. Some studies even argue that ownership is a necessary prerequisite for sustainability (cf. Brolin 2017; Edgren 2003; Ballantyne 2003; Oino et al. 2015). Additionally, the first research question also includes an evaluation of the companies' ownership of the funding switch *itself*. However, ownership alone is not sufficient for development activities to be sustainable (Ostrom et al. 2002). Therefore, the second research question evaluates how the sustainability of internship placements is affected by the funding switch. Finally, to improve the up-scaling of the switch to other regions, the third research question provides recommendations on how both ownership and sustainability can be enhanced in the future. The findings of the third research question are directly discussed following the analyses of research question one and two, respectively.

During a five-week field research period in February and March 2020, qualitative semi-structured interviews with 22 company representatives in 12 companies in the pilot region were conducted.[29] Semi-structured interviews provide a certain structure that allows to cover the most important topics in each interview, while offering the flexibility to follow the direction the interviewees take (cf. Aurini et al. 2016, Bryman 2012). Six companies were interviewed that participated under both funding schemes (GIZ stipend and sustainability program) and six companies that only participated under the GIZ stipend but not under the new funding scheme.

The balanced sample of participating and non-participating companies was selected through a *stratified purposeful sampling* approach (Aurini et al. 2016; Bryman 2016), which allowed to investigate the underlying reasons of why some companies participated under the sustainability program and, thus, sustained the KAM TVET internship placements and why others did not. By directly interviewing company representatives, the companies' requirements to start paying their interns in the future and different perspectives on the funding switch could be examined.

In addition to the interviews on company level, four in-depth background interviews with GIZ and KAM staff were conducted to gather background information about the project and the introduction of the sustainability program. Since it became apparent during the field research that the Central Chapter Committee (CCC) plays a key role in the promotion of the KAM TVET program in general and for the sustainability program in particular, the former and current chair of the CCC were also interviewed and a CCC meeting was observed.

All interviews were recorded[30]. The interviews at company level were then transcribed verbatim, while the background interviews were protocolled. The raw data of the interviews was analyzed with a *hybrid approach of deductive (theory-driven) and inductive (data- driven) thematic analysis* that helped to identify overarching themes and patterns in the data (cf. Fereday and Muir-Cochrane 2006). To identify themes and patterns in the data, the interview data was coded with the help of MAXQDA, a 'Computer Assisted Qualitative Data Analysis Software'[31] (cf. Kuckartz and Rädiker 2019). Finally, the coded interview data was clustered into

[29] *22 people in 12 companies were interviewed because often more than one person per company was interviewed. An overview of the interviews (e.g., interview type, the interviewees position, etc.) is included in the appendix (see Appendix A2.1).*

[30] *ince the interviews were recorded written consent was obtained for the interviews on company level and verbal recorded consent was obtained for the background interviews with GIZ/KAM staff. Blueprints of the provided consent forms to the interview partners at company level are included in the appendix (see Appendix A2.2).*

[31] *The code system that was developed in MAXQDA and is included in the appendix (see Appendix A2.3).*

categories, cross-analyzed over the different transcripts, and then interpreted to answer the research questions.

3.4 Theoretical background information on ownership and sustainability

A factor model on ownership and sustainability was developed to provide a theoretical framework for the research processes and objectives of this study. The factor model identifies and categorizes factors and prerequisites that are mentioned frequently in the ownership as well as sustainability literature as being helpful or sometimes even necessary to ensure ownership and/or sustainability of development projects or project activities.

Table 3.4.1
Overview of factor model

Ownership success factors		Sustainability success factors	
Local stakeholder involvement & participation		**Local stakeholder involvement & participation**	
O.1	Projects/activities must be in the interest of and useful for local stakeholders	S.1	Projects/activities must be in the interest of and useful for local stakeholders
O.2	Local stakeholders should already be involved in the design phase of projects/activities	S.2	Local stakeholders should already be involved in the design phase of projects/activities
O.3	Possible evolvement of local ownership of a project/activity over time		
O.4	Local stakeholders should be actively involved in decision making processes in all project stages/activities	S.3	Local stakeholders should be actively involved in decision making processes in all project stages/activities
O.5	Financial involvement of local stakeholders - a strong indicator of local ownership	S.4	Financial involvement of local stakeholders - one way to ensure project sustainability
What donors must do to ensure ownership		**What donors must do to ensure sustainability**	
O.6	Donors should limit their direct involvement in projects/activities	S.5	Donors should show flexibility to changing priorities of local stakeholders
O.7	Donors should be accountable to and transparent with local stakeholders		
At project level		**At project level**	
O.8	Incorporating projects/activities into local structures, values, and knowledge	S.6	Incorporating projects/activities into local structures, values, and knowledge
O.9	Local leadership and management of projects/activities should be encouraged	S.7	Local leadership and management of projects/activities should be encouraged
O.10	Capacity building for local stakeholders to promote ownership	S.8	Capacity building for local stakeholders to ensure project sustainability
		S.9	Thorough monitoring and evaluation can be helpful to achieve sustainability
		S.10	Regular sustainability assessments are helpful to plan for and enhance future sustainability
		S.11	Implementation of sustainability strategies and exit plan for project phase-out
		S.12	Ensuring a continuation of funding is essential for project sustainability
Relationships between different stakeholders		**Relationships between different stakeholders**	
O.11	Open and transparent communication between all stakeholders involved		
O.12	Clear allocation of roles and responsibilities among all stakeholders involved	S.13	Close coordination and shared understanding between all stakeholders involved
		S.14	Wider networks and partnerships should be built to ensure project sustainability

Source: RWI.

Table 3.4.1 provides an overview of these factors, which are clustered under broader thematic areas. Regarding the structure of the factor model, it has to be mentioned that while ownership and sustainability are two different concepts, they are also linked in the sense that (local) ownership of a development initiative is seen by many authors as a necessary *prerequisite* for sustainability (cf. Brolin 2017, Edgren 2003; Kaplan 2013; Ostrom et al. 2002). Thus, studies that focus purely on ownership or sustainability partly describe the same or similar factors that promote ownership *and* sustainability.

In the following, the factor model is very briefly described and put into context. More detailed descriptions of each factor are available under the links below (cf. Appendix A2.4).

A factor that is considered important in the ownership as well as sustainability literature is that projects or project activities should be in the interest of and useful for local stakeholders (Factor O.1 ; Factor S.1). Local stakeholders should, therefore, be *actively* involved throughout all project stages (Factor O.4; Factor S.3), but especially in project design phases where the direction and objective of a project is decided (Factor O.2; Factor S.2). Financial involvement of local stakeholders is also regarded as vital in both concepts (Factor O.5; Factor S.4) and it is suggested that projects should be incorporated into pre-existing local structures, and forms of organization (Factor O.8; Factor S.6). Additionally, scholars unanimously find that local leadership and management of project (activities) should be encouraged (Factor O.9; Factor S.7) and that capacity building for local stakeholders is key for success (Factor O.10; Factor S.8).

Furthermore, the literature gives recommendations on what donors must do to ensure ownership and/or sustainability. For example, studies researching ownership recommend that donors should limit their direct roles and responsibilities (Factor O.6) and should be accountable to and transparent with local stakeholders (Factor O.7). For sustainability it is important that donors do not just push through 'donor-led top-down projects', but that they show flexibility and responsiveness to adapt the project activities to the changing priorities of local stakeholders (Factor S.5). In regard to relationships between different stakeholders of a project, the literature on ownership recommends an open and transparent communication (Factor O.11) as well as a clear allocation of roles and responsibilities among all stakeholders involved (Factor O.12). Studies researching sustainability advocate similarly for a close coordination and shared understanding between stakeholders (Factor S.13) as well as for an establishment of wider networks and partnerships to ensure project sustainability (Factor S.14).

Since ownership *alone* is not sufficient for a development project to be sustainable, the model also includes factors that are just relevant for either ownership or sustainability. Various studies focused on project sustainability recommend implementing sustainability strategies and exit plans for a project phase-out (Factor S.11), which includes ensuring a continuation of project funding (Factor S.12) and regular monitoring and evaluation (Factor S.9), particularly in the area of sustainability assessments (Factor S.10). Studies focused on ownership find that local ownership can also evolve over time, especially when projects are designed by donors (Factor O.3).

Where suitable, these factors are applied in section 3.7.3 and section 3.9.3 to assess whether the project and the pilot scheme of the funding switch were designed and implemented in accordance with what is recommended in the literature to achieve local stakeholder ownership and project sustainability. Further, the factor model serves as a basis for the recommendations on how both ownership and sustainability in the KAM project context can be improved in the future (cf. section 3.7.4; section 3.8.5 and section 3.9.4).

3.5 Evaluation of overall KAM TVET program by companies

Before turning to the evaluation of the three research objectives, the companies' general impressions of the KAM TVET program are presented.

The analysis showed that three quarters of the interviewed companies are very satisfied with the KAM TVET internship program (Table A 2.1.1, cf. [17]; [6]; [7]; [18]; [20]; [8]; [5]; [10]; [9]; [16]). For example, many companies appreciate that the KAM program has a pool of *pre-selected* interns *readily* available for them:

"An advantage [refers to the KAM program] is you'll have them [refers to interns] readily available because they have […] vetted [i.e. pre-selected] the applications. For these other routes [refers to the companies searching for interns themselves] you have to […] maybe go to the institutions, ask them to give you the people that they had […] [But here] you have the people readily available, that's the advantage to it." (Table A 2.1.1, [12], pos. 251; cf. also [19]; [17]; [18]; [8]; [5]; [10]; [9]; [16]).

A few companies highlighted that they value particularly that interns from specific *skill areas* are available on request:

"We like working with them [refers to KAM] […] they have quite a big database [for interns]. So, you request, you give your specifications, you get what you want. Yeah, instead of waiting for people, who will drop their CV's. […] for example, I told [person at KAM], I want people with this skill background, and he will definitely give me the people with that background." (Table A 2.1.1, [10], pos. 182; cf. also [18]; [9]).

Furthermore, some companies emphasized that they appreciate the structured and easy way of the KAM TVET program to connect companies and interns through job bazars. The job bazars simplify the companies' search and selection processes of interns (Table A 2.1.1, cf. [6]; [18]; [14]; [5]).

More generally, some of the interviewed companies also evaluated the KAM TVET program positively because it gives TVET graduates the opportunity to get practical training experience and industry exposure and, therefore, helps to bridge the (technical) skills gap in Kenya:

"In my opinion, the program is a good one because it helps the students to get hands-on experience and the skills that they don't get from the training institutions. The training institutions of the country are quite [uhm] under equipped, so they do more theory than practicals. And therefore, when these people come to the job market, they're not ready for the job market. [And that is] the point at which KAM is intervening." (Table A 2.1.1, [8], pos. 129; cf. also [17]; [9]; [6]; [15]).

"It's doing a very good job [refers to the TVET program] because […] there is a big [skills] gap. You can't get a single person as a plumber, you can't get a single person as a machinist, a good one. You can't get a good mechanic, who understands things […] so we need that program." (Table A 2.1.1, [20], pos. 45; cf. also [6]; [7]; [9]; [16]).

3.6 Suggestions for improvement of overall KAM TVET program by companies

While some companies view the KAM TVET program as a way to bridge the (technical) skills gap, some other companies and particularly the CCC chairs argued that it would be more effective to intervene on other levels. For example, a few companies suggested that it would be better to implement a program that works directly with training institutions[32] or on the national level[33]:

[32] *e.g., in the sense that training curricula are reviewed, the teaching quality is improved, or skills needs are better aligned with the private sector (Table A 2.1.1, cf. [19]; [20]; [8]; [12]; [13])*

[33] *e.g., in the sense that national TVET forums are organized, the government is engaged to change the national curricula or by advertising Technical Vocational Education and Training (TVET) to make it more attractive for youth (Table A 2.1.1, cf. [20]; [13])*

"So what […] we were trying to run and really drive through w[ith] the funding body [was] […]: Okay, let's start off with first establishing, which institute is 'bona fide' [meaning: in good faith] and has the correct mandate. […] Now, once we know those institutes, which want to work with us, we give them the support in the sense that we will now tell them […] if you give us a plumber, then they must know this much, please. […] So, it's a win-win in the sense [that] they [refers to training institutes] can market better to say, look, we are accredited by KAM and KAM members [refers to the companies] will take your students on, if you are [up] to the standard." (Table A 2.1.1, [19], pos 40; cf. also [20]; [9])

"What we're trying to achieve is to get output from these [training] institutions, which is of a standard, which is acceptable to the manufacturers. That is the core line and that is the bottom line." (Table A 2.1.1, [19], pos. 28)

"I think, […] we need a national engagement forum […] specifically dealing with TVET. […] And then, issues that come up are discussed every two months or something like that in a forum. […] Yeah […] you need that engagement, so that you keep ideas alive and you also help […] the government to formulate policies on issues like the syllabi. That kind of things." (Table A 2.1.1, [20], pos. 119-127; cf. also [13])

A few other companies would prefer to implement an industrial attachment program[34] that places students *during* their education. Matching graduates (≙ interns) was seen by those companies as intervening "too late":

"The point at which KAM is intervening, to me, I think is a bit late […] because they're intervening, when the students have left the training institutions. To me, these internships should come in at the industrial attachment level, when the students are still within the training institutions. So that by the time they're graduating, they already have the skills and the know-how on what happens in a natural workplace. […] But now, we're intervening at a point, where this person […] has graduated […] Yet, when they go to the industries, the industries have again to re-train them […] for them to be able to start delivering." (Table A 2.1.1, [8], pos. 129-131; cf. also [16])

Furthermore, while some companies emphasized that KAM is generally very supportive and helpful when problems arise (Table A 2.1.1, [5]; [10]; [18]; [6]), many suggested that the program could be improved if KAM followed up personally and more frequently with KAM interns (i.e., through meetings or company visits, etc.) (Table A 2.1.1, cf. [17]; [6]; [18]; [11]; [13]; [14]; [8]; [5]; [16]). The companies argued that it would help to:

a) see what challenges the interns are facing (Table A 2.1.1, cf. [6]; [11]; [14]; [8]; [5]; [16]).

b) assess the performance of the interns (Table A 2.1.1, cf. [17]; [8]; [5])

c) make the interns more responsible (e.g., in regard to their behavior during an internship) if they are also accountable to KAM (Table A 2.1.1, cf. [6]; [18]; [5]).

Some further suggestions by the interviewed companies to improve the KAM TVET program were, first to hold the job bazars on a more regular basis instead of sending CVs to the companies to select interns (Table A 2.1.1, [19]; [5]).

"With the [job] bazars, I think that was much easier because it's a day you spend, and you finish and close with that thing. As opposed to: 'Please look for me at this person I've sent'. Then I said: 'No, I don't think this one is qualified, please send me another CV'. Back and forth, back and forth" (Table A 2.1.1, [5], pos. 167; [19])

[34] *Industrial attachments are practical training periods where students work at a company (usually 3 months) before going back to school. On the contrary, internships are for graduates who have finished their education (cf. [2]).*

Second, to allocate the interns to companies closer to their homes. It would be less costly for the interns if they did not have to move and it would also make them more inclined to stay for a job in the respective company after an internship:

"Let's look for [interns] just around us as opposed to a wider area. If they were from within Kiambu, they […] don't have to worry about accommodation, they just go home. There are others, […] this is their home, so they would be probably more inclined to getting a job here." (Table A 2.1.1, [19], pos. 52; [17]; [7])

Third, to consider broadening the skills range offered by the KAM program to also include non-technical skills (e.g., HR, finance, etc.) (Table A 2.1.1, cf. [19]; [16]).

3.7 Research question 1: Ownership of internship placements and funding switch

3.7.1 Companies' ownership of internship placements

The first research objective was to evaluate how the switch in the internship funding scheme has affected the companies' ownership of the internship placements in the pilot region.

3.7.1.1 Applied ownership definition for internship placements

When local ownership at project level is discussed, it is often defined as "*processes where local stakeholders take control and responsibility for the design, implementation, and monitoring of a [development] activity*" (Ballantyne 2003, 3; cf. also Rey-Moreno et al. 2014). Following this definition, local stakeholders can be said to 'own' a development project or activity when they have a commitment towards it that leads them to take over control and assume responsibility for achieving a successful outcome (cf. Edgren 2003; Rey-Moreno et al. 2014).

Applying this definition to this research context means that the companies in the pilot region can be considered to have ownership of the internship placements at their company if they take control and responsibility for:

a) **the design of the internship placements**
(i.e., the companies set their own objectives and agendas how the internships are structured and organized (e.g., in which departments the interns work or about the duration of the internships))

b) **the implementation of the internship placements**
(i.e., in regard to how the interns are trained and supervised or in regard to who is hired as an intern or regular worker after an internship)

c) **the monitoring and evaluation of the internship placements**
(i.e., the companies take responsibility for assessing and monitoring the interns)

Generally, if the companies in the pilot region have ownership of the internship placements, a certain level of commitment towards, responsibility for, and control over those placements should be visible.

3.7.1.2 Evaluation of companies' ownership of internship placements

The analysis of the interviews showed that all companies have, without exception, full ownership of the internship placements at their premises. Each company has their own internal procedures of designing, structuring and organizing the internships. Differences exist in regard to the departments the interns are deployed to, the internship durations (usually between 3-6 months) and whether the interns receive an extra induction training.

Additionally, each company has different ways to implement the internships, e.g., in regard to how the interns are trained and supervised. A few companies expect the interns to take over the same responsibilities as any other worker, while in some others they are very closely supervised. Again, in other companies the interns' responsibilities grow over time:

"For me, [an] intern needs to not look like a trainee but [needs to] look like an employee. That's from my end. So, I expect them to have deliverables that I can review, that I can kinda [uh] appraise." (Table A 2.1.1, [9], pos. 80; cf. [17])

"You see, it's like a classroom set-up, it's only that it's practical. So, they are guided, they're not left alone. The whole area, you know, […] they're not responsible anywhere, basically they're not responsible." (Table A 2.1.1, [18], pos. 195; cf. [11]; [10]; [13]; [14]; [8])

"On the initial stages yes, they will be with somebody but later they will be left by themselves." (Table A 2.1.1, [12], pos. 151; cf. [6]; [5]; [16])

The companies also use different recruitment procedures for internships or subsequent permanent job placements. The variance in these procedures also implies that each company has ownership of the internship placements offered at their company.

For example, there are those companies that have rather sophisticated procedures of hiring interns (i.e., including standardized interview processes) (Table A 2.1.1, cf. [12]; [6]; [18]; [13]; [10]; [16]). Then, there are other companies that have less sophisticated hiring procedures for internships (i.e., interns are taken on a 'first come first serve' basis, often without a screening process) (Table A 2.1.1, cf. [17]; [8]; [14]; [9]; [5]).

Regarding the companies' recruitment procedures for a formal position after the end of an internship, there are those companies that prefer an 'internal hire' (i.e. if possible, they hire former interns). Whereas other companies have more formal and competitive recruitment processes which everybody has to undergo, no matter whether someone previously completed an internship at the company[35]:

"I start with my people, I mentor them, they come around the company, they are able to understand my system. They're able to understand how we work; they understand the culture and all that. So basically, we want people from the system." (Table A 2.1.1, [13], pos. 186; cf. [17]; [7]; [14]; [5]; [9])

"We usually don't promise hiring them [refers to interns], because our responsibility is to give them the exposure. If there is an opening that requires us to hire them on a permanent basis, then we would give them an equal chance with the others [refers to external applicants] to be interviewed through the process. And therefore, they competitively compete for the position. So, coming for internship is not an automatic [uh] pre-qualification that one is going to be hired." (Table A 2.1.1, [8], pos. 45; cf. [12]; [18]; [10]; [16]).

Finally, according to the definition applied above, ownership of a development activity also entails taking responsibility for monitoring and evaluation procedures of the activity.

The research showed that the interviewed companies all take responsibility for monitoring and evaluating their interns, mostly because they need such records to decide which interns they want to retain after an internship. However, the stringency and proficiency of the companies' monitoring and evaluation (M&E) systems also vary considerably between companies.

[35] *However, it is worth mentioning that most of the companies using the more competitive approach argued that former interns still have some advantage over external applicants because they are familiar with the company (Table A 2.1.1, cf. [12]; [18]; [10]; [16]).*

Thus, while all companies have ownership of the internship placements at their premises, a potential pitfall is that it results in very different internship design, implementation and monitoring procedures. These differences caused inconsistencies across the KAM TVET program, e.g., in regard to:

a) the interns' training and treatment at the companies,

b) the interns' opportunities to start an internship placement in the first place or

c) the interns' chances of being retained in a company upon completion of an internship.

Thus, it should be noted that local ownership is not a panacea for development projects but that it comes with its own challenges and difficulties (cf. Hasselkong and Schierenbeck 2017).

3.7.1.3 Underlying reasons for companies' ownership of internship placements

One of the reasons why the interviewed companies all have ownership of the internship placements at their specific premises, is that most of them already had an internship (or more often an industrial attachment) program at their companies long before GIZ and KAM implemented the TVET program in 2017:

"We have done attachments for so long. I think all through, since the company started, we've been offering attachments." (Table A 2.1.1, [18], pos. 201; cf. [14]; [5]; [16]; [6]; [20]; [8]; [9])

"KAM just came in, in a part of what [...] was already pre-existing, internships have been long ongoing. So, KAM just came in, I think, [uh] to try to assist in that, in that space." (Table A 2.1.1, [16], pos. 309)

This also explains why the design, structure and implementation of the internship placements vary so much between the different companies. Since almost all companies already had attachment and/or internship programs in place, they are also taking in other industrial attachés or interns apart from those of the KAM program (Table A 2.1.1, cf. [12]; [17]; [14]; [18]; [20]; [16]; [19]; [6]; [10]; [13]; [8]; [5]). About half of the companies take in significantly more interns/attachés *outside* of the KAM program each year, than they had KAM interns in total since 2017 (Table A 2.1.1, cf. [16]; [5]; [8]; [14]; [18]; [6]). Thus, many companies integrate the relatively small number of KAM interns into the *existing* internship structures at their companies.

Additionally, after the implementation of the KAM TVET program KAM and GIZ never interfered to a level where they took control over the design and implementation of the internships at company level. While GIZ and KAM developed guidelines for the internship program (Table A 2.1.1, [4]; [2]; [3]), they never strictly implemented them but encouraged the companies to use their own policies (Table A 2.1.1, [2]; [3])[36]. The companies confirmed that practice:

"It's totally up to the company to decide that this is the requirement or specification that I need, and this is how I am going to train them. Yeah, there was no guideline that this is how you're supposed to do it" (Table A 2.1.1, [5], pos. 177; cf. [12]; [11]; [14]; [8]; [10])

KAM and GIZ offered only advice and guidance when the companies required it:

"*We would always communicate [with KAM] if we had challenges or anything that we felt, we needed to do something about."* (Table A 2.1.1, [12], pos. 235; cf. [13]; [14]; [10])

Thus, the companies kept their ownership of the internship placements even after the implementation of the KAM TVET program.

[36] *The only areas where KAM is regularly involved is matching of interns and companies and monitoring of the internship and job placement numbers at the companies (cf. section 3.2)*

3.7.1.4 Relevance of funding switch for companies' ownership of internships

However, that GIZ and KAM neither started the internship program at the companies nor dictated how the internships should be designed and implemented, also means that the funding switch from the GIZ stipend to the sustainability program had no effect on the companies' ownership of the internship placements. One company put it aptly by stating:

"I think it's only the funding part that has changed but everything else in terms of processes […] still remains the same." (Table A 2.1.1, [6], pos. 187)

This is contrary to the original research assumptions which assumed that the funding switch would affect the companies' ownership of the internship placements.

3.7.2 Companies' ownership of funding switch

While all companies kept their ownership of the internship placements after the implementation of the funding switch, a different picture emerges when assessing the companies' ownership for the funding switch *itself*.

3.7.2.1 Applied ownership definition for funding switch

Local stakeholder ownership is often also connected to control and responsibility over funding (cf. Olsson et al. 2008; cf. also Ostrom et al. 2002).

As per the definition presented in section 3.7.1.1, the companies can be considered to have ownership of the funding switch if they take control over and responsibility for its implementation. That is, they support the premise of the sustainability program and are willing to take over the responsibility of funding the KAM interns themselves.

3.7.2.2 Evaluation of companies' ownership of funding switch

The evaluation showed that ownership of the funding switch is rather mixed among the interviewed companies. About half of the companies were not willing to take ownership of the switch, meaning they were not willing to invest their company's internal resources to fund the stipends of KAM interns themselves (Table A 2.1.1, cf. [12]; [18]; [13]; [14]; [5]).

3.7.2.3 Reasons for mixed ownership of funding switch among companies

There are several possible reasons for the companies' mixed ownership of the funding switch. One explanation why about half of the interviewed companies have not taken ownership of the funding switch could be that with the implementation of the sustainability program (i.e., the funding switch), GIZ and KAM intervened for the first time *directly* in the design and structure of the internships at company level. That is, they created regulations that companies either had to take over the funding of the stipends for interns or they would no longer receive interns via the KAM TVET program (Table A 2.1.1, cf. [4]; [3]; [2]):

"If […] the industry […] is not willing to give anything, let's leave them aside, because what we are testing is, are they willing to give something? It doesn't matter if it is 3.000 KES or 20.000 Kenya KES." (Table A 2.1.1, [4], min. 00.10.58f.)

A further, and more probable reason for the companies' mixed ownership is that the switch to the sustainability program was inconsistently introduced to them.

The six companies that *have* taken ownership of the funding switch (i.e. they are already paying the KAM interns' stipends or in one case were about to start) (Table A 2.1.1, cf. [6]; [9]; [16]; [10]; [8]; [17]) tended also to be the ones that were properly informed about the introduction of the switch. For example, they went to a forum where the switch was explained or they were personally contacted by KAM (e.g., through a company visit or phone call) (Table A 2.1.1, cf. [6]; [7]; [9]; [16]; [10]; [8]).

On the contrary, the other half of the companies that have not taken ownership of the switch (Table A 2.1.1, cf. [12]; [18]; [13]; [14]; [5]) were either not at all (i.e., they heard about it for the first time during or just before the interview for this study) or not appropriately informed (i.e., they only received an email without a proper explanation) about the implementation of the funding switch:

"We have not heard about it [refers to funding switch]. We are not informed about that one. We are not aware." (Table A 2.1.1, [14], pos. 192; cf. also [12]; [13])

"They just told us effectively through an email, that effectively they're [refers to KAM] not going to be paying, that [uh] we are the ones, who are supposed to meet that."
(Table A 2.1.1, [18], pos. 132; cf. also [5])

This led to a situation where these companies either felt not appropriately engaged in the decision to implement the sustainability program or misunderstood the underlying reasons for it:

"What KAM is supposed to do, they're supposed to engage stakeholders in depth [emphasis] about it [refers to implementation of funding switch]. Because well, we would like to [pause] trust KAM is good. But […] how deeply have they really engaged with the manufacturers, with their members to buy the idea that I can pay the stipend to these people?" (Table A 2.1.1, [13], pos. 167; cf. also [5])

"KAM [are] the ones, who […] set a standard of 10.000 [KES], they are unable to pay and now they are telling the employer, pay! You see [laughs], [they]'re pushing [their] budget requirements to us. […] they told people, we'll be sending you on internship, we'll be paying you. So, suddenly they decided we don't have this money in our pockets, let someone else foot the bill." (Table A 2.1.1, [18], pos. 162; cf. also [12]).

In both cases, the consequence was resistance against a takeover of the funding. Thus, ownership of the funding switch *itself* seems to be influenced by an understanding of the reasons for its implementation and the feeling to have been involved in the decision-making process.

An explanation why some companies did not feel appropriately involved in the decision to implement the funding switch, and thus did not take ownership, could be that the design and the implementation of the sustainability program was decided between GIZ, KAM and representatives of the CCC[37] without involving the broad base of the companies in the pilot region:

"So that was, I know, a discussion between both KAM and GIZ in [regard to], how can this program be sustained? What are the different ways that we can sustain the program (Table A 2.1.1, [2], min 00.06.21f.; cf. also [4])

"The Central Chapter [Committee] was like: Yeah, I think we can try this here [refers to implementing the funding switch] and we are willing because we have seen the benefit of having interns, even if GIZ is not there to support." (Table A 2.1.1, [4], min. 00.08.54f.)[38]

[37] *The Central Chapter Committee (CCC) is a board of eight elected companies representing the KAM member companies in the region in front of KAM, regional government entities and other stakeholders (Table A 2.1.1, cf. [19]; [8]; [20]).*

[38] *Addendum: While the CCC representatives confirmed that they agreed to stop the GIZ stipends, they did not want to take over the funding themselves but hoped that the funding, which was formerly used for the stipends, would be reinvested to accredit the training institutions to help them to produce better trained graduates (Table A 2.1.1, cf. [19]; [20]).*

While all companies in the pilot region were invited to forums where the decision to implement the switch was explained, only one quarter to one third of the companies actually attended those forums. Moreover, those that did attend, did often not keep their expressed commitment at the forums to take interns under the new funding scheme. (Table A 2.1.1, cf. [1]; [3])[39]

Thus, although KAM and GIZ seem to have failed to involve the broad base of the companies in the pilot region in the decision to implement the funding switch, they aimed to be transparent about its implementation. The remaining question is why they followed up with some companies that did not attend the forums, whereas with others they did not.

3.7.3 Comparison of ownership success factors to research findings on ownership

In Table 3.7.1, the success factors on ownership which were presented in section 3.4 are compared to the research findings on the companies' ownership of internship placements and their ownership of the funding switch itself (cf. also Appendix A2.3 for a more detailed description of the success factors on ownership). Table 3.7.1 provides a clear overview where the project acted in accordance with recommendations from the literature to enhance local ownership and where there is potential for future improvement.

Table 3.7.1
Comparison of ownership success factors to research findings on ownership

Ownership success factors			Application of ownership success factors to research context
Local stakeholder involvement and participation			
O.1	Projects/activities must be in the interest of and useful for local stakeholders	P	By placing TVET graduates in internships to enhance their practical skills and to improve their employability, the project is tackling the skills gap/mismatch that local stakeholders (here the companies) identified.
		O	The project failed to confirm with the majority of companies whether the switch to the sustainability program was in their interest and useful for them.
O.2	Local stakeholders should be involved in the design of projects/ activities	O	By deciding the parameters of the funding switch between KAM, GIZ and the representatives of the CCC, the broad base of companies was not involved in the design phase of the sustainability pilot scheme.
O.4	Local stakeholders should be *actively* involved in decision making processes in all project stages/activities	P	By leaving the design, implementation and monitoring of the internships to the companies, they are actively involved in all stages of the internship activities.
		O	The project failed to actively involve the broad base of companies in the decision to implement the funding switch. This led to a misunderstanding of the reasons for the implementation and the feeling to not have been involved, which in turn hampers ownership.
O.5	Financial involvement of local stakeholders – a strong indicator of local ownership	P	By implementing the funding switch, the project followed recommendations from the literature to financially involve local stakeholders in a development activity to promote their ownership.

[39] *One explanation why some companies did not keep their commitments to take interns under the new funding scheme is: "You realize a HR manager does not have the ultimate decision for finances. [...] So, after they fill this commitment form, they have to go and talk to the boss. And the CEO might be like, we are supposed to lower costs, not incur more costs. In that case, although they had committed, they might find it tough." (Table A 2.1.1, [3], min. 00.48.03ff.).*

Table 3.7.1 continued

What donors must do to ensure ownership			
O.6	Donors should limit their direct involvement in projects/activities	P	GIZ/KAM limited their direct involvement in the internship program at company level, because they never took control over the design and implementation but only offered advice and guidance when required and encouraged the companies to otherwise use their own policies.
		O	By implementing the funding switch, GIZ/KAM created the regulation that the companies have to take over the internship funding themselves to receive further interns over the program. Thus, they intervened in the internship design that was formerly left to the companies.
O.7	Donors should be accountable to and transparent with local stakeholders	o	By facilitating forums to introduce the funding switch to the companies, GIZ/KAM tried to be open and transparent about its implementation. However, since not many companies attended those forums, the project failed to (personally) follow up with companies that did not attend the forums.
Ownership success factors			**Application of success factors to research context**
At project level			
O.8	Incorporating projects/activities into local structures, values, and knowledge	✓	By implementing the KAM TVET internship program in companies that already had internship and/or industrial attachment programs at their premises, the project built on existing structures that were already owned by the companies.
Ownership success factors			**Application of success factors to research context**
At project level			
O.9	Local leadership and management of projects/activities should be encouraged	✓	By encouraging the companies to manage and lead the design, implementation and monitoring of the internships at company level according to their own policies, the project follows the literature's recommendations on promoting local stakeholder ownership.
O.10	Capacity building for local stakeholders to promote ownership	✓	By offering support and capacity building to the companies when it is needed and asked for, the project helps to empower the companies to become drivers and owners of the project activity, instead of mere passengers.
Relationships between different stakeholders			
O.11	Open and transparent communication between all stakeholders involved	o	By communicating (the objective of) the funding switch well to some companies but not to others, the project only partly succeeded to ensure an open and transparent communication regarding the funding switch.

Legend: P≙ project fulfilled ownership recommendations; o ≙ project only partially fulfilled ownership recommendations; O≙ project failed to fulfill ownership recommendations. - Source: RWI.

3.7.4 Research question 3 (part 1): Lessons learnt and potential solutions to enhance companies' ownership of the funding switch

The aim of the third research objective was to provide recommendations on how the companies' ownership and the sustainability of the project activities can be enhanced in the future. This section will present lessons learnt as well as potential solutions from the analysis of the first part of research question three on ownership. The recommendations can be used to enhance ownership of the funding switch among companies in the pilot region as well as to inform the upscaling process of the sustainability program to other KAM project regions in Kenya.

The prior analysis showed that all interviewed companies have ownership of the internship placements at their companies. Thus, the recommendations focus on potential improvements of

the companies' ownership of the funding switch itself (i.e., on improving the companies' willingness to take responsibility for funding KAM interns).

Potential strategies to enhance companies' ownership of the funding switch in the future are:

a) **parent manner to companies** (e.g., through more awareness forums and more per**to introduce the switch to the sustainability program in a more consistent and trans**sonal follow-up from KAM).

In the future, GIZ and KAM should ensure that *all* companies are appropriately informed about the switch in the internship funding scheme and, more importantly, also *why* such a switch is taking place. According to the literature on ownership, local stakeholder ownership can be fostered when donors communicate new project objectives and activities in an understandable and transparent manner (cf. Factor O.7).

To facilitate that, an increased number of **awareness forums** should be held in the future to inform more companies appropriately about the funding switch and the reasons for it (cf. Factor O.11). This is particularly important because rather few forums were held to introduce the switch in the pilot region (≈ 2 to 3 between February and June 2019, Table A 2.1.1, [2]; [3]). According to GIZ and KAM, such efforts are already planned for the up-scaling of the sustainability program to other project regions:

"And of course, having more of those awareness forums with the industr[ies], just trying to show them the importance of taking these people and having them commit." (Table A 2.1.1, [4], min 00.24.54f.; cf. also [2])

Additionally, since those forums had a rather low attendance rate in the pilot region [3], it might be particularly effective to facilitate **more personal follow-up by KAM**. Especially with companies that do not attend the forums because most companies mentioned that they value KAM's advice and input. For example, personal visits by and interactions with KAM encouraged some companies which had no policy of paying interns before to take ownership of the funding switch and to start paying KAM interns themselves (Table A 2.1.1, cf. [6]; [9]; [8]).

b) **to actively involve companies in decision-making processes about (implementing) the funding switch.**

However, forums and personal follow-up by KAM should not just be facilitated to inform the companies about decisions already made, but ideally the companies should be actively involved in decision-making processes regarding the implementation of the sustainability program (i.e. the funding switch) in the future. This is exactly what some companies expect:

"My view would be, that before you change [refers to implementing the funding switch] that you check in, you engage and say, guys well, we think we can't sustain for whatever reasons. [Uh] can we partner with you? Can we reduce it to 50/50 or is there something? That's what I'm talking about, engagement. You don't just stop [refers to stopping the stipend funding]. […] Tell us your challenges, ask us how we can be of help because we are the beneficiaries of this at the end of the day." (Table A 2.1.1, [16], pos. 317-319; cf. also [15]; [13]; [5])

Consulting with and actively involving local stakeholders in decision-making processes regarding new project designs is identified as key in the academic literature (cf. Factor O.1; Factor O.2; Factor O.4) to foster local stakeholder ownership. Only if local stakeholders are involved in the design and decisions to implement new project activities, will they be able to shape these activities according to their needs, which is crucial to facilitate their ownership. Thus, to assist in that process, spaces should be created where all stakeholders, including the internship offering companies, can meet and discuss their demands and expectations regarding the switch to the sustainability program (cf. Factor O.11).

If these suggestions are considered in the up-scaling process of the sustainability program to other KAM project regions, more companies might be willing to take ownership of paying stipends in the future. These strategies can also still be implemented in the pilot region, because ownership can evolve over time, especially in projects in which donors are responsible for developing new project ideas (cf. Factor O.3). Thus, it is not too late to (appropriately) inform certain companies and to convince them to take ownership of the switch.

Additionally, since ownership is a prerequisite for sustainability (cf. section 3.4), increasing the companies' ownership of the funding switch also helps to ensure the sustainability of the internship placements. Under the current premise of the sustainability program, only those companies that take responsibility for funding the internships themselves receive further interns via the KAM program (cf. section 3.7.2.4). Thus, if companies take ownership of the sustainability program, more internship placements can be sustained.

3.8 Research question 2: Sustainability of internship placements after the switch to the sustainability program

The second research objective was to evaluate how the switch in the internship funding scheme has affected the sustainability of the internship placements in the pilot region.

3.8.1 Applied sustainability definition for internship placements

The here applied definition of sustainability focuses on the continuation and long-term effectiveness of development projects (cf. Maier et al. 2016). An often used definition developed by the OECD/DAC Committee defines a development project or an activity to be sustainable when there is a "*continuation of benefits from a development intervention after major development assistance has been completed*" (OECD 2002, 36; cf. also Okun 2009; Brinkmann et al. 2014; Maier et al. 2002; Oino et al. 2015). Thus, project sustainability is achieved when not only the activities of a project continue but, more importantly, when the flow of benefits to the beneficiaries persists after the donor withdraws funding (cf. Cekan 2016c; Karanja 2014; Okun 2009).

Applying this definition to the context of this research means that the internship placements in the pilot region can be considered sustainable if:

a) **the companies continue to offer internship placements after the funding switch without external GIZ funding** and

b) **the internships still produce benefits for the beneficiaries**
(i.e., benefits are in this context positive employment outcomes of interns at internship and job level)

Sustainability of employment benefits can be approached from multiple dimensions. An alternative approach to the one that is subject of this research can be the sustainability of employment benefits that occur at the individual level over a lifetime. Such individual-level employment benefits may be sustained over a lifetime irrespective of the sustainability of internship placements and are, although important, therefore not further discussed in this evaluation.

3.8.2 Sustainability of internship placements after funding switch

The analysis in the pilot region shows that only eight of 23 companies that were recruited to participate in the sustainability program actually took over the funding of KAM interns'

stipends.[40] Thus, about one year after switching the funding responsibility to the company-level, only about one third of the companies sustained the KAM program internship placement activities (cf. Appendix A2.5, Table A2.5.2 for self-verification).

3.8.3 Sustained employment benefits for interns after funding switch

Furthermore, the majority of the eight companies that participated in the pilot scheme took a smaller number of interns than under the GIZ stipend scheme (cf. Appendix A2.5, Table A2.5.2, C9; C17; C3; C4; C2; C1)[41] because taking interns is now associated with costs for the companies:

"So, if they [refers to KAM] are paying and the company is not the one that is paying, you feel [...] more comfortable to give them [refers to internship placements] because it's not a cost to you. But now, [...] you're saying the industry should also come in and chip in, then the industry will only go based on their requirement. [...] So, that is the approach that most of the industries actually are taking." (Table A 2.1.1, [7], pos. 18; cf. also [8]; [10]; [9])

Some of the interviewed companies which participated in the new funding scheme further stated that while GIZ was funding the interns they were willing to take in KAM interns beyond their actual labor requirements:

"We would bring the numbers from KAM [refers to GIZ stipend phase]. […] If they did have ten, they wanted to bring, we would have no issues because there was no cost implication to us." (Table A 2.1.1, [8], pos. 101; cf. also [7]; [9])

The job retention rate of KAM interns among those eight companies remained in relative terms after the funding switch on the same level (i.e., three companies retained a higher percentage of interns for a job after the funding switch (cf. Appendix A2.5, Table A2.5.2, C3; C2; C1) three companies retained a lower percentage (cf. Appendix A2.5, Table A2.5.2; C4; C9; C17) and two the same percentage of interns (cf. Appendix A2.5, Table A2.5.2; C8; C14). Most companies stated that whether they hire someone permanently depends on the performance and behavior during the internship rather than on the fact that they invested in a stipend. Thus, since the performance or behavior of the interns under the new funding scheme was not considerably better or worse than under the GIZ stipend scheme, the employment outcomes remained on the same level (Table A 2.1.1, cf. [10]; [18]; [13]; [9]; [16]).

This shows that switching the funding responsibility of the stipends to the companies seems not to have affected the interns' employment benefits at job level relative to the total number of internship placements. However, the implementation of the funding switch had a rather negative influence on the sustainability of the program's employment benefits in absolute terms. As considerably fewer internships were offered, fewer graduates had the possibility to get industry exposure, training experience and a chance to be retained for formal employment.

---

[40] *The data base for this assumption is only partially reliable because it could not be verified with exact certainty how many companies were recruited to participate in the pilot of the sustainability program and out of those, how many companies actually participated, i.e., took on interns and paid them a stipend for more details).*

[41] *The internship and job placement numbers presented in Table A2.5.2 (cf. Appendix A2.5) are only partially reliable because different sources (KAM M&E records, verbal accounts of KAM/GIZ, company records and IPA data) reported partly contradicting. Thus, it was difficult to verify which numbers are correct. To simplify the analysis always the highest recorded internship and job placement numbers per company were compared (cf. Appendix A2.5, Table A2.5.2).*

3.8.4 Reasons for companies' (non-)participation in the sustainability program

Whether a company participates in the pilot of the sustainability program is influenced by a multitude of different and often interdependent reasons and pre-conditions at the company-level. For the context of this report, only the companies' main reasons for participation or non-participation, are discussed, respectively. Additionally, in Appendix A2.6, some more detailed observations regarding the reasons for (non-)participation of the interviewed companies as well as some further general patterns applicable across the interviewed companies are elaborated.

The main reasons why some companies participated in the pilot and others did not, can be clustered into two, often interdependent, categories:

a) Companies' overall attitude towards paying stipends to (KAM) interns.

There are companies that are generally in favor of paying interns (Table A 2.1.1, cf. [6]; [7]; [17]; [9]; [16]) and then there are other companies that are, to varying degrees, against it (Table A 2.1.1, cf. [12]; [18]; [13]; [5]; [10]; [14]; [19]; [8]). The different reasons that were brought forward by the interviewed companies for or against paying interns a stipend are discussed in section 3.8.4.1. However, the attitude towards paying interns is not the only indicator that determines why some companies sustained KAM internship placements while others did not.

b) Companies' perception of and satisfaction with former (KAM) interns.

The second category, which is often intertwined with the first one, is based on previous experiences the companies made with KAM interns.[42] The individual reasons that make companies (dis-)satisfied with the KAM interns are discussed in section 3.8.4.2. Similar to the first category, just looking at the satisfaction with former KAM interns cannot explain participation patterns.

3.8.4.1 Companies' motives (not) to pay stipends to interns

In the following, the general reasons that were brought forward by the interviewed companies for or against paying interns are presented. The general reasons in favor of paying interns and attachés can be broadly sorted into three, however not mutually exclusive, categories. The categories altruistic, self-serving and policy-related reasons are presented in Table 3.8.1.

Table 3.8.2 presents the justifications provided by the interviewed companies for not wanting or not being able to pay stipends to interns and/or attachés. Here, the reasons can be clustered into two broad categories, which are company-related and intern-related reasons.

The most important finding that can be deducted from Tables 3.8.1 and 3.8.2 is that companies seem to calculate whether an investment in a stipend is beneficial for them.

[42] *The interviewed companies were often strongly influenced by individual positive or particularly negative experiences with KAM interns, independent of the actual ratio of good or bad experiences they made with all KAM interns.*

Table 3.8.1

Companies' reasons to pay stipends to interns[43]

'Altruistic reasons' (focused on the well-being of interns/attachés)	
Arguments	**Quotes**
Some companies argued that they give a stipend to interns and/or attachés because they *need* a stipend to pay for accommodation, food, and transport (Table A 2.1.1, cf. [6]; [7]; [12]; [19]; [9]):	*"They stay far from their parents […] during the [internship] period. […] So, whatever little that you pay them monthly, they can use that to be able to pay rent and get something to eat."* (Table A 2.1.1, [6], pos. 265)
Some even emphasized that without stipends, the interns would not be able to attend work or buy something to eat (Table A 2.1.1, cf. [17]; [16]; [6]):	*"We chip in [..] because sometimes […] you become so much [uh] interested to know really, why did you [refers to intern] come late? Are you sick or what has happened? Only to realize that they did not have bus fare. Or others sometimes, maybe during lunch hour, somebody does not have money to go […] to the cafeteria and have something to eat."* (Table A 2.1.1, [17], pos. 251)
A few companies argued that particularly *interns*, which have already graduated, need a stipend urgently because they have more financial obligations than attachés that still go to school (Table A 2.1.1, cf. [16]; [6]):	*"Now, for this person [refers to a graduated KAM intern], […] he has cleared school, maybe he's the head of a family, he needs that amount to get even with his bills."* (Table A 2.1.1, [16], pos. 276)
A few companies even saw paying a stipend as a 'duty' of the company, a way of giving back to interns who invest their time and resources to work for the company (Table A 2.1.1, cf. [9]; [16]; [17]):	"*These individuals have invested their time and resources to study and, I mean, build capacity. These individuals have shown interest to work and [uh] make themselves available."* (Table A 2.1.1, [16], pos. 341)
'Self-serving reasons' (focused on companies' own interests)	
Arguments	**Quotes**
Some companies argued that they pay a stipend because it motivates the interns to perform better, which in turn is beneficial for the company (Table A 2.1.1, cf. [17]; [7]; [16]; [20]).	*"We […] like to make them feel very comfortable [refers to paying a stipend]. So that you will not have a reason why you're not able to […] perform, like others [uh] are performing."* (Table A 2.1.1, [17], pos. 251).
Many companies also revealed that they pay a stipend because it is an investment that is profitable for them in the long run (Table A 2.1.1, cf. [20]; [12]; [17]; [6]; [7]; [9]; [16]):	*"When you pay, you're sought to get something out of the young man. And you get something from the young man […], when he gets exposed, could be he enjoys your work environment and he […] might want to work with you in the future. And that is value."* (Table A 2.1.1, [20], pos. 59).
'Policy-related reasons' (focused on pre-existing rules about internship payment in companies)	
Arguments	**Quotes**
Many companies stated that stipends are paid to interns because it is the company's policy to do so, without questioning the underlying reasons (Table A 2.1.1, cf. [14]; [12]; [19]; [17]; [10]; [16]; [6]):	*"You know, our practice is that 1.000 [KES] are paid. Everybody, even if it's TVET […]. This is our regiment; we have to do that."* (Table A 2.1.1, [14], pos. 312)
A common justification was that a policy of paying stipends was implemented by the management (Table A 2.1.1, cf. [14]; [12]; [16]):	*"The management from some years back, they decided to give […], so we are just continuing that."* (Table A 2.1.1, [14], pos. 188)

Source: RWI.

[43] *Noteworthy is that the companies cited here are not necessarily in favor of paying KAM interns. In some cases, the interviewed companies also argued in favor of paying their 'own' interns/attachés but against paying KAM interns (cf. Appendix A2.6 for more details on individual companies' reasons for or against participating in the new funding scheme).*

Some companies regard stipends as a worthwhile investment, e.g. because it directly increases the interns' motivation and productivity or their willingness to work for the company in the future (cf. Table 3.8.1; 'self-serving reasons'), while other companies do not seem to see any value in interns, or at least not enough that it would justify a stipend provision. This group of companies typically argues that interns still need to be trained, which already requires an investment from the company's side, or they fear that the interns will not stay at their company after the internship (cf. Table 3.8.2; 'intern related reasons').

This is consistent with the literature on project sustainability, in which it is argued that local stakeholders (especially in a private sector setting) only take over the funding of an activity and, thus, contribute to its sustainability, when they see that it is economically or financially beneficial for them (cf. Factor S.4), or when they perceive it as otherwise useful (cf. Factor S.1).

A further important observation is that companies that are *not* willing to pay stipends to their interns/attachés or that only pay a very minimal stipend can generally afford this attitude because there are many other channels through which they can recruit interns:

"We don't need to get [interns] actually from KAM, we'll get them elsewhere in the marketplace." (Table A 2.1.1, [13], pos. 186; cf. also [14]; [18])

Most companies receive plenty of direct applications from well-educated interns/attachés that are willing to work for free or very little pay:

"I have guys with degrees here and they're working for us with that minimum." (Table A 2.1.1, [13], pos. 256, cf. also [14])

Thus, even companies that depend on interns for their workforce are not forced to pay a stipend because there are enough people willing to do an internship for no or very little pay.

However, the analysis also shows that all interviewed companies that do *not* pay any or only a very small stipend[44] struggle with high dropout rates, especially among KAM TVET interns:

"TVET [refers to KAM interns] because they have completed their school, now they are expecting to be earning [...] and that is the reason, they don't stay." (Table A 2.1.1, [14], pos. 320; cf. also [15]; [5]; [16]; [13])

Consequently, these companies face higher costs in the long run because they frequently have to re-train new interns:

"Again, [...] when that now happens [refers to the fact that interns leave after a short time], it also is a cost. [...] It will affect production, […] we will employ some more, we train them, before they get assimilated [...] [they] are going." (Table A 2.1.1, [13], pos. 343-349).

[44] *Among the 12 interviewed companies, five either do not pay a stipend to anyone (Table A 2.1.1, cf. [13]; [18]) or only pay small stipends to the (KAM) interns (lunch to 3.000 KES per month) (Table A 2.1.1, cf. [5]; [16]; [14]). The other seven companies pay substantial amounts of stipends (between 8.000 KES to 15.000 KES per month) (Table A 2.1.1, cf. [12]; [19]; [17]; [9]; [10]; [8]; [6]). However, four of them do not pay these stipends to the KAM interns (Table A 2.1.1, cf. [17]; [10]; [19]; [12]).*

Table 3.8.2

Companies' reasons to not pay stipends to interns[45]

'Company-related reasons' (focused on companies' (financial) abilities or attitude)	
Arguments	**Quotes**
Analogous to the policy-related argument in Table 3.8.1, other companies also justified not paying stipends to interns/attachés with a policy related argument (cf. [18]; [13]; [8]; [5]):	*"Normally, in our policy […] we don't pay stipends."* ([18], pos. 132)
When asked if it would be possible to change that policy, the interviewees (mostly HR managers) from both camps often argued that this would be a decision that is up to the management (cf. [18]; [5]; [12]; [14])[46]:	*"You see now, this a tall order from the management, to say that we can afford to pay other people […]. And the approvals are given from the management, top management."* (Table A 2.1.1, [18], pos. 134)
One company argued that it would be unfair to pay a stipend to the KAM interns, when other interns/attachés at the company are not paid (cf. [5])[47]:	*"I already have 20 attachés […] that I am not paying a stipend. […] And then I have interns here that I'm told that I must pay a stipend. You see now, it's going to be an issue."* ([5], pos. 81-83)
A further argument against paying interns was that paying stipends is not incorporated in the company's budget (cf. [18]; [13]; [8]):	*"You see, […] paying stipends […] [t]his was not in our budget."* ([18], pos. 132)
It was also claimed that the funding switch was introduced on such short notice that the budgets had already been preplanned, which left no room to include stipends ([8]; [18]; [10]):	*"When you are paying, [….] then you have to pre-factor that in your training budget at the start of the year or it eats into your budget that you had not planned for."* ([8] pos. 69)
Interestingly, only one company stated that they cannot at all afford to pay stipends:	*"This factory was close to being closed down […]. Due to the financial situation […] we were not in a position to pay for it"* ([5], pos. 326-328)
'Intern-related reasons' (focused on interns' (ascribed) abilities or behavior)	
Arguments	**Quotes**
Some companies argued against paying stipends because they felt that an internship is rather about skills development and learning and not about getting paid (cf. [5]; [18]; [14]; [19]):	*"As far as I understand, as far as an internship is concerned, it's you are there to try and gain skills."* ([5], pos. 117)
Another point raised by some companies was that they already invest their time and resources by training the interns, which would make paying a stipend an additional burden (cf. [13]; [19]; [12]):	*"Don't forget one thing, […] when we take the intern, […] the person who is training, you know, who is doing the normal work, will not work when he is training […]. So, there's a cost, that will be associated with it."* ([13], pos. 167)
A few companies remarked that investing in a stipend is not beneficial for them because the work of an intern is not very productive (cf. [10]; [13]; [12]):	*"It is hard […] for a company to get […] interns and [uh] you're paying them. They are not very much productive at that time, eh?"* ([10], pos. 108)
Some companies stated that instead of paying an intern, they would rather hire and pay someone with experience whom they do not have to train (cf. [5]; [12]; [11]):	*"It's not a must that I pick [refers to KAM interns], so what do I get from it? You know, if I want to hire, why can't I just go and hire people with experience? And I don't have to train them. So, what do I get, what are my benefits [laughs] in this?"* ([5], pos. 225)
A few companies do not want to invest extra resources in interns that, in their opinion, will opt for other opportunities in the job market rather than staying in their company after the internship (cf. [8]; [13]):	*"Some of them [refers to the own and other companies] don't want to incur an expense because […] it means you're training people for the open job market."* ([8], pos. 75)

Note: For numbers in brackets refer to Table A 2.1.1 – Source: RWI.

[45] *The companies cited here are either against paying interns/attachés in general or specifically against paying KAM interns (cf. Appendix A2.6 for more details on individual companies' reasons for (non-)participation in the pilot scheme).*

[46] *It is noteworthy that some of the HR managers that referred to management for the ultimate decision power, said that they would personally be in favor of an introduction/ increase of stipends: "On a personal note, I think maybe the 150 [KES per day] is not very nice, being that this guy has finished school." (Table A 2.1.1, [16], pos. 263; cf. also [15]; [5]).*

[47] *However, companies that do pay interns but that do not pay attachés or that pay them on different levels said that this has not caused problems at their companies (Table A 2.1.1, cf. [8]; [6]; [17]).*

3.8.4.2 Companies' satisfaction with (former) KAM interns

The analysis showed that five out of the 12 interviewed companies are generally very satisfied with KAM interns (Table A 2.1.1, cf. [5]; [6]; [9]; [10]; [18]), four companies have a mixed perception of them (Table A 2.1.1, cf. [16]; [17]; [14]; [8]) and three are rather dissatisfied (Table A 2.1.1, cf. [19]; [12]; [13]).

The most relevant finding of the analysis of companies' satisfaction with KAM interns is that negative perceptions of KAM interns seem to outweigh positive or mixed ones. While a company's decision to participate in the sustainability program is based on versatile motives, it is particularly evident that especially those companies with the strongest negative perceptions are reluctant to participate in the sustainability program (Table A 2.1.1, cf. [19]; [12]; [13]).

It is noteworthy that those companies frequently used anecdotal evidence of individual negative experiences that were then generalized to all KAM interns which influenced their perception as well as behavior (Table A 2.1.1, cf. [19]; [12]; [13]) (cf. Appendix A2.6 for a visualization and more detailed information).

Table 3.8.3 outlines the underlying reasons which led to companies' dissatisfaction with KAM interns. Negative perceptions of KAM interns are first and foremost based on soft-skills and attitude issues and not on the interns' (lack of) technical skills.

This is not to say that a lack of technical or practical skills among KAM interns was not mentioned in the interviews. However, almost all companies reported that a good working attitude and soft-skills compensate the lack of such.

Table 3.8.3
Companies' reasons for dissatisfaction with KAM interns

Reasons for dissatisfaction with KAM interns	
Arguments	**Quotes**
Quite a few companies complained that KAM interns (compared to attachés) are too focused on being paid and getting a job after the internship, instead of on learning new skills and acquiring working experience (cf. [12]; [14]; [8]; [16]):	*"They [refers to KAM interns] were having a vision of getting money. [...] Money, money, just money, not to train. […] they should come here and get experience. But they were not in that line, they were in the line of employment and getting money."* ([14], pos. 290-292) *"I tend to think they're [refers to the KAM interns] [more] money-driven, than skill-driven because they feel we have the skill."* ([12], pos. 53)
It was also claimed that the money focus is particularly strong among KAM interns (compared to other graduates) because KAM interns are aware that they are supposed to receive a stipend (cf. [5]; [12]; [19]; [16]):	*"The biggest issue now, which comes with now the interns, eh? [...] Now [that] there is no more support as far as stipend is concerned. Now the students, they already know, they already have a mentality that, if I'm going through KAM, there's something that I'm supposed to get."* ([5], pos. 71)
Some companies also argued that expectations of KAM interns are too high when they start an internship, i.e., they expect to start at the top or expect to do 'white-collar work' (cf. [19]; [12]; [17]; [16]; [13]):	*"They feel they need to be white collared; they don't feel they need to dirty their hands. [Groans] they don't want to work low levels, they don't understand manufacturing. They really want office work, which is nine to five, bang."* ([19], pos. 30)
A further frequent complaint was that since the KAM interns are graduates from 'tertiary' training institutions, they would sometimes undermine or not respect their supervisors or other superior workers, especially those with a lower education (cf. [12]; [17]; [13]; [16]):	*"You [refers to the KAM interns] feel [uh] you're more educated than I am, you've been placed under me, I'm the one who is training you. […] So, they tend to undermine the people that they find below them."* ([12], pos. 179)

Note: For numbers in brackets refer to Table A 2.1.1 – Source: RWI.

On the contrary, the companies that are (rather) satisfied with former KAM interns appreciate them exactly for qualities other companies complained about (e.g., their focus on getting employed or their soft-skills).

Table 3.8.4 describes the individual reasons brought forward by various companies that contributed to a high satisfaction with KAM interns.

Table 3.8.4
Companies' reasons for satisfaction with KAM interns

Reasons for satisfaction with KAM interns	
Arguments	**Quotes**
Contrary to some dissatisfied companies, the satisfied ones spoke positively about the KAM interns' focus on getting a job and impressing the employer (cf. [5]; [18]; [6]; [15]; [10]):	*"Someone who has completed [refers to graduated KAM interns], this is someone who knows what they want in life. And also, you know, they want to do their best because they know probably […] I can get employed."* ([5], pos. 51)
Additionally, some companies declared also that in comparison to attachés (≙ students), KAM interns would show a higher readiness for the job market and would have better skills because they are already graduated (cf. [18]; [6]; [15]; [10]):	*"There is a big advantage, like I mentioned, these are people, who are ready for the job market. […] So basically, you can't compare them with people, who are going on with their studies outright. They are the best […]. So [pause, uh] I would say they are good [refers to skills], they're better than the [uh] attachment ones."* ([18], pos. 167)
Furthermore, the rather satisfied companies have praised particularly the KAM interns for their mature behavior and soft-skills (e.g., for their discipline, punctuality, and work-ethic). Some ascribe this to the association with KAM and the work-readiness training the interns receive before they are matched to the companies (cf. [6]; [10]; [15]; [18]; [5]; [11]):	*"The main difference [refers to company-internal interns] […] for KAM, they are, you know, they've gone through [uhm] work preparedness training. [...] The soft-skills training, which gives the students an edge. Plus, number two, when a student is referred by an organizational body like KAM, they know very well that […] we can also report any misbehaviors to KAM because they are your main reference. So, they are more disciplined and structured"* ([6], pos. 97-99)

Note: For numbers in brackets refer to Table A 2.1.1 – Source: RWI.

Tables 3.8.3 and 3.8.4 illustrate that the satisfaction with KAM interns depends very much on the companies' expectations towards interns (i.e., job focus vs. learning focus) and on their individual experiences with former KAM interns. Thus, the overall satisfaction with and impression of KAM interns depends in many cases on individual interns the companies had the pleasure or misfortune of meeting.

3.8.5 Research question 3 (part 2): Lessons learnt and potential solutions to enhance the sustainability of the internship placements in the future

In section 3.7.4. lessons learnt as well as potential solutions to improve companies' ownership of the funding switch were discussed. Thus, this section (incl. the following two sub-sections) will focus on the second part of the third objective and provide recommendations on how the sustainability of the internship placements can be enhanced.

This is particularly important because the implementation of the funding switch had a rather negative influence on the sustainability of the internship placements (i.e., only one third of the companies sustained the internship placements without GIZ funding (cf. section 3.8.2)) as well as on the program's employment benefits (i.e., internship numbers and employment outcomes of interns decreased among the participating companies (cf. section 3.8.3)).

Since the sustainability program's premise dictates that only those companies that take over funding for the KAM interns' stipends receive further interns through the program, it is paramount to increase companies' willingness to participate in the new company-funded internship scheme in the future.

The analysis revealed that two central categories (i.e., willingness to pay KAM interns a stipend and their satisfaction with them) determine whether a company participated in the new funding scheme and, thus, sustained the KAM internship placements without GIZ funding.

Strategies to increase companies' willingness to pay stipends and their satisfaction with KAM interns are presented in the following two sub-sections. With these recommendations at hand the sustainability of the internship placements as well as the thereof resulting employment outcomes may be enhanced.

3.8.5.1 Strategies to increase the companies' willingness to pay stipends in the future

Five potential strategies to increase companies' willingness to pay stipends were identified:

a) to set financial incentives that increase companies' willingness to invest their own resources.

One possibility to implement such financial incentives could be to offer *commitment rewards* for taking a certain number of interns a year (e.g., in the form of a percentage off the KAM member subscription):

"So, you see, maybe […] they can tell the employers that, if you take on maybe more than 20 interns in a year, then maybe you get five percent off the subscription [refers to the fee for the KAM membership subscription]. That at least would make the employers more willing […] to do it. Or any other [laughs] form of incentive." (Table A 2.1.1, [5], pos. 221, 215).

Another way to incentivize companies to fund stipends could be to implement a *shared payment plan* for a transition period. Various companies suggested different versions of such shared payments (Table A 2.1.1, cf. [9]; [16], [19]; [12]):

1. Splitting the stipend between the companies and KAM/GIZ through a reimbursement program[48]:

   "The factories [....] pay a 100% [of the stipend] at first but then KAM does 50% reimbursement to the company [...] once the student has finished. […] I think, it could be more successful and more sustainable than leaving the whole of the stipend to the factories right away." (Table A 2.1.1, [9], pos. 208-210).

2. Implementing a 'take two interns get one free' approach to ensure that more interns receive practical training and industry exposure:

   "Now to give these guys [refers to TVET graduates] an opportunity to train further, we can then try and say, look, if you pick one person from here [refers to KAM program] then they [refers to KAM] are going to give you a second one, they [refers to interns] take on the same job. […] One they pay [refers to company] […] and the other one gets a stipend from KAM. […] What has happened is, at least, even that guy has got [uh] training." (Table A 2.1.1, [19], pos. 106-114)

[48] *An advantage of a reimbursement program would be that certain conditions could be attached to the reimbursement (e.g., companies must deliver intern numbers for monitoring purposes in a timely manner).*

3. Shortening the GIZ stipend period to a 'trial period' after which the companies can decide if they want to keep the intern and invest in a stipend:

> *"Sometimes as an employer, you will not invest in somebody you're not sure what they're bringing back to you. That's why I would really appreciate, if their stipend thing [refers to GIZ stipend] could continue for a short time span […] don't prolong it, but just make it shorter. And then […] you would already know if you want to invest in them [refers to KAM interns]."* (Table A 2.1.1, [12], pos 129)

Implementing the new funding scheme through some form of shared payment plan in the future likely incentivizes more companies to participate as it diminishes the feeling among the companies that the financial burden has been, from one day to the other, passed on to them:

"It would have made a lot more sense for the sustainability program [i.e., the funding switch] […] to be gradual. In the sense that if [the] industry was told, we are giving you interns but for this program to be sustainable, let's have a matching fund, we pay 5.000 [KES], you pay 5.000 [KES]. It is a softer landing. There is more buy-in than saying, ok 10.000 [KES], do!" (Table A 2.1.1, [2], min. 01.02.16ff.)

In the sustainability literature the benefits of cost sharing schemes between donors and recipients are also emphasized – at least for a transition period. Scholars argue that in order to ensure a project's sustainability, local stakeholders should contribute to the best of their ability to the project funding, while donors can provide the supplementary funding required to ensure the most sustainable outcomes (cf. Factor S.4).

Additionally, setting such financial incentives (e.g., commitment rewards for taking a certain number of interns a year) could also help to ensure that more companies take a larger number of interns under the new funding scheme, which would in turn ensure higher employment outcomes for interns:

"The company gives up something and KAM gives up something, if the numbers are to grow [refers to intern numbers]. Because if that stipend is left to the factories alone, and I have mentioned, what's gonna happen is, it's going to only limit the numbers." (Table A 2.1.1, [9], pos. 208)

b) to show companies that paying interns a stipend can be (economically) beneficial for them.

The analysis showed that some companies that pay stipends to their interns/attachés do so because they see that such an investment is profitable for them (i.e., they conduct a cost-benefit analysis). These companies understand that stipends can - in the short term - increase interns' motivation and productivity and - in the long term - make them more willing to work for the company in the future (cf. section 3.8.4.1; Table 3.8.1, 'self-serving reasons`). Thus, other more reluctant companies should be made aware of these benefits.

c) to help companies understand that not paying stipends is associated with higher costs for them in the long run.

Additionally, the analysis revealed that companies which do not pay a (substantial) stipend often struggled with high dropout rates, particularly of KAM interns. Consequently, these companies face higher costs in the long run because they frequently have to re-train new interns (cf. section 3.8.4.1).

Not all interviewed companies which complained about high dropout rates seemed to be aware of the connection between payment of stipends and intern dropout rates. Thus, raising awareness about this tradeoff, might help to convince companies that neglecting to pay stipends is associated with higher costs in the long run.

d) to appeal to companies' humanitarian responsibility.

A further, however, potentially less effective, strategy could be to appeal to companies' 'humanitarian responsibility', in the sense that interns *need* a stipend to be able to sustain themselves during the internship (e.g., to afford rent or to pay for food and transport). This could be successful in cases where companies already have a policy of paying their 'own' interns/attachés but are reluctant to take over the funding of stipends for interns coming in through the KAM program.

e) to advertise and lobby for the funding switch with higher level management personnel in the companies.

Last but not least, the analysis revealed that most HR managers (which are currently KAM's main contact persons in the companies) often do not have the decision-making power to start or to increase interns' stipends. Such decisions are in most cases left to higher management personnel (e.g., managing directors, company owners or CEOs):

"You realize a HR manager does not have the ultimate decision for finances. [...] So, after they [...] commit[...] [refers to taking interns under the new funding scheme], they have to go and talk to the boss. And the CEO might be like, we are supposed to lower costs, not incur more costs. In that case, although they had committed, they might find it tough." (Table A 2.1.1, [3], min. 00.48.03ff.).

Finally, a few companies argued that (the advantages of) the KAM TVET program require(s) more visibility and better lobbying, particularly among company managers. Company managers would only be willing to invest in a program which they have heard 'good things' about (cf. [5]; [16]) (see section 3.5 for an elaboration on the programs advantages from the company point of view).

Thus, in the future it might be more fruitful to introduce novelties, such as the funding switch, directly to higher management personnel with financial decision-making power, instead of addressing HR managers that might or might not lobby for a takeover of the funding with their superiors.

An opportunity to address managing directors and CEOs directly and to further lobby for the advantages of the KAM TVET program could be at KAM's annual general meetings (AGMs) (Table A 2.1.1, cf. [3]). Additionally, the members of the Central Chapter Committee (CCC) could support the lobbying for the program among higher management personnel because they are all CEOs and/or managing directors themselves:

"*If you talk from owner to owner of the business, it's [a] completely different affair, so that helps [refers to convincing companies to participate in the new company-based funding scheme].*" (Table A 2.1.1, [20], pos. 71; cf. also [5])

While the strategies provided above might be a good start to increase companies' willingness to pay stipends to KAM interns in the future, it should also be noted that there is only so much KAM or GIZ can do if some companies are not willing to pay their interns.

Some of the interviewed companies that were strongly against paying KAM interns a stipend would have continued to take them in, either if the GIZ stipend returned or if they were allowed to take interns without paying them (Table A 2.1.1, cf. [18]; [12]). Therefore, if the sole objective of the program was to continue internship placement activities under the KAM TVET program, then it could be considered to allow companies to offer internships without paying stipends. If this was allowed, probably more companies would sustain the internship placements at their premises and more interns would benefit from training experiences and industry exposure.

However, this would mean that only interns who can afford to finance themselves or that have parents/guardians who can support them can participate in the program. This was the case in one company that received KAM interns under the new company-based payment scheme but that did (against the actual rules) not provide them with any monetary stipend:

"I normally tell them from day one: 'Just know that we do not offer any stipend'. And some of them, they will be like: 'Ok, then I cannot [uh] I'm sorry, but it would not be easy for me'. And you just understand. But for some they're like: 'That's fine, I live around so that's fine with me'. [….] Or they say: 'It's ok I've already communicated to my guardian and they're ok with just giving me the allowance as if I was in school'" (Table A 2.1.1, [5], pos. 99-105)

Thus, it should be very *carefully* considered whether the payment premise of the sustainability program should be lifted in the future.

3.8.5.2 Strategies to increase companies' satisfaction with KAM interns

To increase the sustainability of the internship placements, it is also essential to increase companies' satisfaction with KAM interns. The analysis showed that those companies which were dissatisfied with former KAM interns were also not willing to take in KAM interns under the new funding scheme. As the companies' satisfaction with and their perception of KAM interns was mostly concerned with the interns' soft-skills, improvements should start there (cf. section 3.8.4.2).

Four potential strategies to improve companies' satisfaction with KAM interns were identified:

a) to increase the frequency and improve the content of the already existing 'work-readiness trainings' in regard to soft-skills training and expectation management.

The project already offers short (two to three days) 'work-readiness trainings', which include soft-skills and career training components, before KAM interns are placed. (cf. section 3.2). However, despite the previous training, some companies were still dissatisfied with the soft-skills of KAM interns (cf. section 3.8.4.2) and requested to prolong the work-readiness trainings and to and to improve their content (Table A 2.1.1, cf. [12]; [13]).

Some of the interviewed companies suggested that the work-readiness trainings should focus more on topics such as discipline, punctuality, communication, pro-activeness, and general behavior in the workplace:

"What KAM should take up is [uh] they need to train this people [refers to KAM interns] […] in terms of […] [work] ethics, […] personal grooming, hygiene, discipline, punctuality […] because the employer will mainly give them skills in terms of work. […] if you have a disciplined person, you will be able to invest more time, not in disciplining them, but in improving their skills on a day-to-day base." (Table A 2.1.1, [12], pos. 243; cf. also [5]; [13]).

Other companies argued that **sensitizing interns** in regard to their general attitude or mindset would be even more important than soft-skills training (Table A 2.1.1, cf. [17]; [6]; [13]; [14]). According to the interviewed companies, the interns' sensitization should be particularly focused on the following areas:

1. Interns should understand that internships are more about learning and gaining experience but not about getting paid:

   "At the back of their minds, they need to know, we're not just here for the stipend, it is more than that. It is learning." (Table A 2.1.1, [12], pos. 207; cf. also [14]; [16]; [19])

2. Interns should be advised that they cannot expect to start working in a high-level position and that unpopular tasks will have to be completed:

"Before dispatching a team [refers to KAM interns being sent on internships], prep them. […] introduce them to […] workplace issues. […] show them people lifting things at [the] workplace […]. Show them people cleaning [the] workplace […]. You don't just come from the university and you become the head of the engineering division. […] So, that aspect is what we need to inculcate in our graduates." (Table A 2.1.1, [16], pos. 321-325; [12])

3. Interns should be made aware that working in manufacturing companies is not a white-collar job:

"Where maybe I thought KAM and GIZ could chip in is […] they have not prepared them what to expect on the job market or at the industry. […] These people [refers to KAM interns] are actually prepared for white-collar jobs. […] But coming to be employed in an industry like here, the environment is completely, totally very different." (Table A 2.1.1, [13], pos. 138-140; cf. also [19])

4. Interns should be taught that they must respect their supervisors and co-workers, particularly those that are less educated than the TVET graduates:

"Advise them [refers to KAM interns] when you go there [refers to an internship in a company] […] teach other people to […] behave well, respect old age, even if you find somebody is a cleaner and he never went to school, respect him!" (Table A 2.1.1, [17], pos. 284; cf. also [12]; [13]; [20])

5. Interns should be prepared that formal job opportunities in manufacturing companies are rare and that most of them will have to rely on the skills acquired during the internship to start their own businesses:

"They [refers to interns] also should be made to understand that […] 70% of the economy is supported by SME's, you expect […] 70% of them will be assimilated in th[at] kind of environments. This is [uh] the link that KAM should come up with and also explain to them." (Table A 2.1.1, [13], pos. 149; cf. also [17]; [15]; [20])

b) to increase the awareness about the work-readiness trainings among companies.

Not only the contents of work-readiness trainings but also the awareness about them among companies should be improved. Companies that were aware of the work-readiness trainings were generally more satisfied with KAM interns' behavior and soft-skills (Table A 2.1.1, cf. [6]; [10]; [5]; [18]; [20]), whereas most companies that complained about the lack of soft-skills among KAM interns were not aware of the trainings (Table A 2.1.1, cf. [13]; [12]; [17]; [16]). Therefore, raising the awareness about the training could positively influence the companies' perception of the interns' behavior and soft-skills.

c) to involve the companies in the work-readiness trainings.

Some companies offered to participate in the soft-skills or expectation management trainings. For example, some companies offered present their expectations at work readiness trainings:

"During the training, during the three days training [refers to work-readiness training]. Probably the industries can be given a session on the last day of the training because they know exactly what is happening on the ground. […] let's say, for example, […] we get our engineer to come and talk to the students during the job readiness workshop, I think the students will connect." (Table A 2.1.1, [6], pos. 279; cf. also [16]; [17]).

d) to establish peer-learning events for new interns.

Finally, a further suggestion was to establish peer-learning events where successful interns that have received a permanent position after an internship are invited to the work readiness trainings to share their experiences with future and current interns:

"Let people learn from the experiences made because […] those who have succeeded [can share] success stories." (Table A 2.1.1, [16], pos. 359; cf. also [6]; [17]).

In conclusion, focusing on increasing the soft-skills and managing the expectations of interns as well as raising the awareness about the already existing work-readiness trainings might increase the companies' willingness to take in KAM interns under the new funding scheme in the future, which in turn improves the sustainability of internship placements.

The analysis has also shown that most interviewed companies value good soft-skills, the willingness to learn, as well as respectful conduct during internships at least as much as technical performance when they decide about retaining interns for a formal position (Table A 2.1.1, cf. [12]; [15]; [18]; [5]; [16]; [17]; [6]; [14]; [10]; [9]). Thus, improving soft-skills and the attitude of future interns potentially enhances retention rates and future employment outcomes of participants as well.

3.9 Evaluation of ownership and sustainability at project level

While the project itself was not the primary unit of analysis of this research, ownership and particularly sustainability at project level (i.e., at the level of KAM and GIZ) have nonetheless meaningful implications for the findings discussed in the previous sections.

3.9.1 Ownership at project level

Ownership of a project is not always concentrated on one individual stakeholder but can also be shared (to varying degrees) by different stakeholders, even between donors and recipients (Edgren 2003; Olsson et al. 2008).

In the case of the analysis of the KAM project, the companies are considered as the primary stakeholders (which have ownership of the *internship placements* at their companies), but they share the ownership of the *project* in general with KAM and GIZ. Both KAM and GIZ seem committed to the project and stand behind its objectives. Further, KAM takes full responsibility for and has control over the implementation of the project, which is in line with the core characteristics of ownership (Table A 2.1.1, cf. [1]; [2]; [4]). Ownership of a national stakeholder like KAM is particularly important, because they can ensure the continuation of project activities in the future.

Equally, GIZ and KAM seem to have ownership of the sustainability program component (i.e., the pilot program of the funding switch). Both developed the pilot scheme together and both stand behind the idea and are committed to its implementation. In general, they evaluate the pilot scheme positively and are convinced of its eventual success, despite concerns about low placement numbers so far and the way the implementation took place (Table A 2.1.1, cf. [1]; [2]; [3]; [4]).

A shared ownership of different program components is possible in this case because roles and responsibilities for the different components are clearly defined (cf. Factor O.12). GIZ is responsible for funding and oversight of the project, KAM is responsible for project implementation (e.g., allocation of interns to companies) and the companies have full rein over the design and implementation of the internships at their premises (cf. section 3.2). This is an excellent example of how shared ownership of a project can be managed.

3.9.2 Sustainability at project level

Evaluating the sustainability of the whole project, i.e., not only the continuation of internship placements at company level but the continuation of higher-level project activities that enable the internships to take place (e.g. matching of interns and companies), has important implications for the overall findings of this report. In the *project* sustainability literature, a development

project is considered sustainable, when there is a "*continuation of benefits from […] [the] intervention after major development assistance has been completed*" (OECD 2002, 36). This definition implies that the benefits (here internship opportunities & employment outcomes) for the chosen group of beneficiaries (here TVET graduates) must continue after donor assistance ends. Following this definition, for benefits to be sustained, the *relevant project activities that enable the existence of benefits* (here selection of TVET graduates at VTIs and matching of potential interns and companies) must continue as well (cf. Okun 2009; Oino et al. 2015; Kuria and Wanyoike 2016; Cekan 2016b).[49]

The evaluation of the sustainability of the whole project, revealed that GIZ and KAM have succeeded in providing a good basis for ensuring project sustainability once GIZ's participation in the project comes to an end.

For example, the project was incorporated into existing institutional and administrative structures of KAM and the companies. Furthermore, management and implementation of the project is largely left to the local implementing agency KAM, whereas GIZ supports the implementation through capacity building and technical expertise to ensure that KAM has the necessary abilities to carry out the implementation (Table A 2.1.1, cf. [1]; [2]; [3]; [4]). Thus, if KAM decides to continue the project activities after GIZ withdraws, they would have the abilities to do so without significantly weakening the project activities. With the implementation of the sustainability program (i.e., the funding switch), sustainability became a key objective of the project (at least on company level) and the commissioning this sustainability assessment reaffirmed the willingness to achieve project sustainability.

However, while these are important pre-conditions for sustainability, there is also a central short coming that impinges on the future sustainability of the overall project.

Currently, the project seems to have no exit strategy at project level when GIZ withdraws its funds. While the pilot of the sustainability program (i.e., the funding switch) is indeed a 'sustainability strategy', it is only focused on ensuring sustainability of the internship placements at *company level*. However, there seems to be no plan how a continuation of the *relevant project activities* can be facilitated once the project ends and GIZ stops funding KAM's work. Other than assumed by GIZ (Table A 2.1.1, cf. [1]), this research has shown that the KAM project implementation team would currently not be able to continue its work (i.e., selection and placement of interns) without GIZ funding (Table A 2.1.1, cf. [2]).

Thus, as long as no (sustainability) strategy is implemented that ensures funding to continue the relevant project activities (i.e., matching of interns and companies; selection of interns at training institutions) after GIZ withdraws, the benefits for the chosen group of beneficiaries (here TVET graduates) will not be sustained at the end of the next project phase. This would make the project unsustainable - no matter whether the companies would be willing to fund stipends for the interns or not. Consequently, not only should the sustainability of the internship placements be enhanced in the future, but also the sustainability of the overall project, because the former is irrelevant without the latter.

[49] *Noteworthy is here that sustainability can have different dimensions. The focus of this section lies on sustainability at project level (i.e., it is evaluated whether the project and its activities are sustained). As discussed in section 3.8.1, sustainability can also occur at the individual level of the beneficiaries that already profited from an intervention. Following an individual definition of sustainability benefits may be sustained over a lifetime (here: improved employability of TVET graduates and long-term formal employment).*

3.9.3 Comparison of sustainability success factors to research findings on sustainability

In the following, the success factors on sustainability which were presented in section 3.4 are compared to the research findings on the sustainability of the project in general.

Table 3.9.1 compares the relevant sustainability success factors to the research findings and provides a clear overview on aspects that are in accordance with recommendations from the literature to promote sustainability as well as on aspects that have potential for improvement.

Table 3.9.1
Comparison of sustainability success factors to research findings on sustainability

Sustainability success factors			Application of sustainability success factors to research context
At project level			
S.6	Incorporating projects/project activities into local structures, values and knowledge	✓	By building the project activities on existing administrative and management structures within KAM (i.e., the Chapter structure) and within companies (i.e., pre-existing internship programs), the project built a good basis for achieving project sustainability.
S.7	Local leadership and management of projects/activities should be encouraged	✓	As suggested in the literature, the management and implementation of the project is largely left to the local implementing agency KAM. Thus, after GIZ withdraws, KAM has the necessary knowledge to continue the project implementation, which means the project activities have a good prospect to become sustainable after donor withdrawal.
S.8	Capacity building for local stakeholders to ensure sustainability	✓	By supporting the implementation of the project through capacity building, GIZ ensures that KAM has the necessary abilities to sustain the project activities after GIZ withdraws.
S.10	Regular sustainability assessments are helpful to plan for and enhance future sustainability	✓	Regular sustainability assessments that go beyond the normal monitoring and evaluation activities are recommended in the literature to identify deficiencies and develop an action plan for sustainability. The project followed that recommendation by commissioning this research.
S.11	Implementation of sustainability strategies and exit plan for project phase-out	✓	In the literature, implementing 'sustainability strategies' is regarded as a key for sustainability. By implementing the funding switch to ensure that the internship activities would still be funded once GIZ withdraws, sustainability became a key objective *on company level.*
		×	However, the project lacks an exit strategy at *project level*, which ensures a phase-out of GIZ's involvement and a takeover by KAM or any other stakeholder that continues to fund the implementation of the project activities.
S.12	Ensuring a continuation of funding is essential for project sustainability	✓	By switching the funding responsibility of the stipends to the companies, the project took a major step to ensure sustainability of the internship placement activities after donor withdrawal.
		×	However, the project has yet to ensure a continued funding of the relevant project activities (specifically of KAM's work, i.e., the whole project implementation) after GIZ withdraws from the project. Ensuring a continuation of funding is regarded as one of the most central components to ensure project sustainability.
Relationships between different stakeholders			
S.13	Close coordination and shared understanding between all stakeholders involved	✓	A close coordination and a shared understanding of project objectives between the donor and implementing agency are regarded as helpful for sustainability. This seems to be the case here since new activities (such as the pilot of the funding switch) are negotiated and discussed jointly between GIZ and KAM.
S.14	Wider networks and partnerships should be built to ensure project sustainability	✓	By introducing the project to county governments and including it in the agendas of the KAM Chapter Committee meetings, the project managed to build wider networks and partnerships with relevant stakeholders that can potentially support the sustainability of the project.

Legend: P ≙ project fulfilled sustainability recommendations; O≙ project failed to fulfill sustainability recommendations - Source: RWI.

3.9.4 Lessons learnt and potential solutions to enhance sustainability at project level

To ensure that the relevant project activities can be sustained after GIZ withdraws, an exit strategy at project level must be developed that includes a plan on how continuous project funding can be ensured. Only then can the benefits for the target population (here the employment benefits for the interns) be sustained.

In the literature, the implementation of exit strategies for local takeover as well as the development of a plan to ensure the continuation of project funding after project phase-out it is regarded as key to ensure project sustainability (cf. Factor S.11 and Factor S.12). Different studies suggest various possibilities to ensure a continuation of project funding. For example, it is suggested to raise funding at the national level of recipient countries (e.g., from governments), from partner organizations (e.g., local implementing organizations) or from local stakeholders or end-users of the project itself (cf. Factor S.12).

However, the analysis indicated that it is unlikely that the required funding (which includes aspects like financing a team that facilitates the selection of interns as well as the matching of companies and interns) can be raised from the internship offering companies because many companies could not even be convinced to take over the funding of internship stipends.

Thus, other possibilities such as increasing KAM member fees or convincing KAM to continue the project funding need to be explored. As a last resort, if no further funding can be ensured, enhancing bilateral connections between training institutions and companies (without involving KAM) could be a possibility to ensure the sustainability of employment benefits for interns (Table A 2.1.1, cf. [3]; [6]).

4 Quantitative Evaluation of the RtW program

4.1 The Ready to Work Program

ReadytoWork (RtW) is an employment promoting program supported by E4D that seeks to reduce the skills gap in the local workforce in natural resource-based industries and related sectors in Eastern Africa by preparing youth for the world of work and easing the education to employment transition. In a two-day training youths learn essential life skills that prepare them to enter the job market. Subsequently, they are linked to companies for industry exposure via internships. E4D partners with private sector firms and local industry organizations to implement the program in the target countries. Work-readiness programs have been implemented in Kenya, Mozambique, Tanzania and Uganda.

The RtW program evaluated in this chapter was launched in Uganda in 2016 in partnership with the Absa Bank Uganda Limited (Absa Bank, formerly called Barclays Bank Uganda Limited) and the Uganda Manufacturers Association (UMA). The program targeted students in their final year at selected vocational training institutes (VTIs) and technical colleges (TCs).

Absa Bank conducted a training of trainers in work readiness with teachers at the selected schools in all four regions of Uganda and supported the implementation of RtW trainings with students. RtW trainings lasted two days and covered sections on work, people, money, and entrepreneurship skills. Among other things this included resumé preparation, interviewing skills, communication skills, appropriate behavioral attitudes in the workplace, and work ethics. Students could participate in the RtW training on a voluntary basis. Upon completion of the RtW training, participants received a certificate by Absa Bank. The RtW trainings took place in participating schools and the schools were flexible to conduct RtW trainings through the trained teachers at times of their convenience and as many as wished or required. At all schools in which teachers were trained RtW trainings were conducted. The participating schools are described in more detail in section 4.4.

Subsequent to the RtW training, UMA facilitated internship placements of the trained students with UMA member companies and in some cases students were even directly offered a job by UMA member companies. The matching of students to companies was achieved either through a job fair organized by UMA in which students and member companies could participate or, alternatively, companies could file individual requests for interns or staff. For the second alternative, companies indicated their preferences for the candidate's skill set and UMA then identified trained students who matched the requested profile and encouraged them to apply for the position. In both matching processes candidates went through a regular selection procedure and could be rejected by the company after all.

The matching process followed a needs-based approach, which had multiple implications for the coverage and selection of beneficiaries and the timing of the program roll-out. First, not all students that completed the RtW training were placed in internships or jobs because not all students were offered a position and some students who have been offered a position rejected it. Reasons to reject positions included workplaces being located too far away from residences or transport and opportunity costs were perceived as too high. Second, and closely related to the first point, internship placements were based on merit or skills, such that, among trained students, potentially those, who had better labor market prospects already, were placed in internships or jobs by UMA. Third, internship placements did not immediately follow the RtW training

or graduation from the respective VTI or TC. Fourth, the length of internships differed across participants. Points three and four suggest considerable heterogeneity in the time of program completion. A more detailed description of the sample and timing of internships follows in sections 4.4 and 4.5.

Throughout this report chapter we refer to the two-day skills training as the *RtW training* and the internship and job placements facilitated through UMA as *internship placements* because direct job placements were less common and the program had been designed as an internship matching, but not a job placement program. The overall program, including both components, i.e., the training and placement (although not everyone completed an internship placement), is referred to as the *RtW program*.

4.2 Research Questions

The main objective of the impact evaluation of the RtW program is to understand and measure how the RtW program has improved labor market outcomes of graduates of VTIs and TCs.

Because the RtW program consisted of two components – the first being the RtW training itself and the second being internship placements with UMA member companies – the response to this research question requires to acknowledge the distinct effect mechanisms of the two components.

A priori, the main impact on employment and labor market outcomes was expected to arise from the internship placement component because the RtW training lasted two days only, in which multiple important soft and life skills topics were raised but not intensively trained. In contrast, the internship placements linked beneficiaries to potential employers and offered industry exposure for several months. Therefore, the internship placement component offered greater potential to gain practical skills and experience, thereby bridging the skills that are demanded by employers and offered by young graduates, which is often referred to as the *skills gap* and considered a major challenge for youth employment in Eastern Africa.

In light of the particular interest in the impact of the internship placement program, the research questions can be more specifically posed as:

(RQ#1) What are the gains in employment and earnings for vocational training graduates **from participating in the RtW training and internship placement** with UMA companies?

→ (RtW training + UMA internship) vs. (no treatment)

(RQ#2) What are the gains in employment and earnings for vocational training graduates **from an internship placement** with UMA companies **in addition to participation in the RtW training**?

→ (RtW training + UMA internship) vs. (RtW training)

RQ#1 tackles the effect of completing both program components in comparison to not participating in the RtW program at all. It disregards the effects of only participating in the RtW training and, therefore, RQ#1 should not be confused with the overall RtW program impact which would be a mix of the impact of only participating in the RtW training and participating in both the RtW training and internship placements.

RQ#2 intends to isolate the placement effect from the RtW training effect and to measure the additional impact of participating in an internship placement beyond potential benefits from the RtW training alone. It is important to note that the additional gain from internship placements

can be different to the pure internship placement effect if participants had only been placed in internships but not trained. This is because interaction effects between the training and the placement are likely to occur which is also the reason why the program had been designed in this two-component structure. To see this, consider the following theory of thought: The probability of being retained after an internship was completed may be higher for individuals who have previously been trained in work ethics and have behaved in accordance with these ethics during their internship in comparison to individuals who have not been trained in work ethics and therefore did not behave as it would have been appreciated by the employer.

4.3 Research design

Rigorous impact evaluation is the measurement of changes in relevant indicators, e.g., employment, that can be causally attributed to a policy intervention. At the heart of rigorous impact evaluation is the *counterfactual problem,* which is that it is unknown what would have happened to the respective beneficiary group in the absence of the intervention or policy program. The counterfactual problem makes it inherently difficult to causally attribute changes in relevant outcome indicators to the considered program. In order to emulate counterfactual scenarios, specific research designs and estimation methods are used to establish causality of program impacts.

The estimation methods used for the RtW program evaluation to emulate such a counterfactual scenario are introduced in this section and make use of recent vocational training graduates who did not participate in the RtW program as a comparison group for RQ#1 and beneficiaries who only participated in the RtW training but not in internship placements as a comparison group for RQ#2.

4.3.1 Estimation methods

For the purpose of measuring RtW program impacts a survey of 1,466 former vocational training students, who graduated from VTIs in November 2019 or from TCs in May 2019, was conducted in February 2020. The survey included graduates from schools that participated in the RtW program and schools that were not part of the RtW program. Thus, the evaluation sample consisted of four groups of participants:

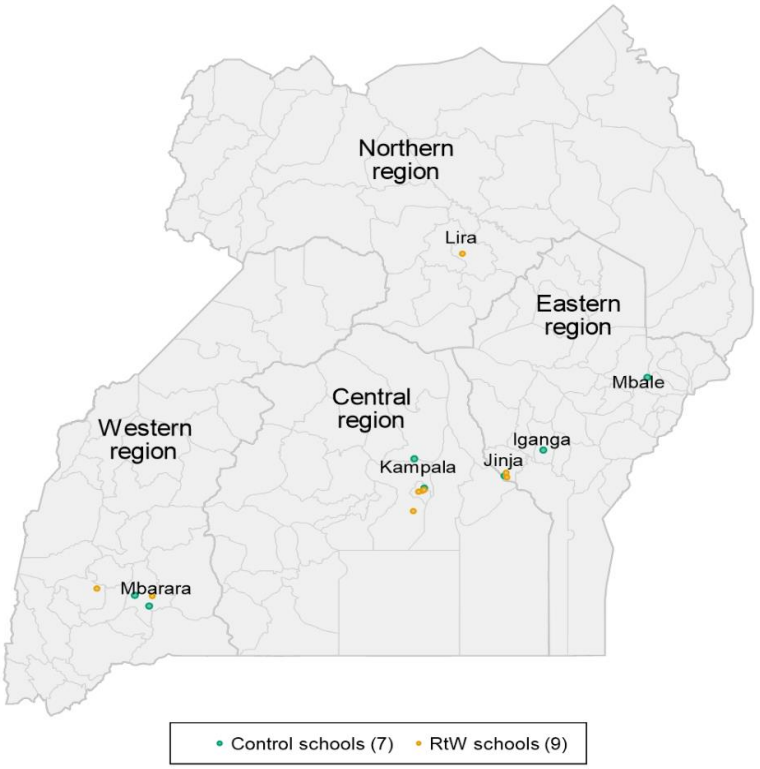
1. beneficiaries at RtW schools who participated in the RtW training and internship placements,
2. beneficiaries at RtW schools who participated in the RtW training but not in an internship placement,
3. non-beneficiaries at RtW schools, and
4. non-beneficiaries at non-RtW schools.

For the estimation of program impacts, the first group is considered the treatment group and the latter three groups constitute comparisons group. Specifically, groups three and four constitute the comparison (or control) group for RQ#1 and group two the comparison group for RQ#2.

The final sample comprised nine RtW schools (= treatment schools) and seven non-RtW schools (= control schools).[50] Treatment and control schools were present in urban and rural areas as well as in all four Ugandan regions, except for the Northern region where only one TC, which participated in the RtW program, was located (Figure 4.3.1).

[50] *Originally, we had planned to recruit nine control schools, but two control schools did not provide us with necessary contact details.*

Figure 4.3.1
Map of Uganda with locations of study schools



Source: Own illustration.

As described in section 4.1, participation in the RtW program did not follow a random allocation but final year students could decide individually whether they wanted to participate in the RtW training. Further, internship placements were driven by companies' needs and followed a standard selection procedure based on skills and merit. In consequence, RtW beneficiaries who were placed in internships, RtW beneficiaries who only participated in the RtW training, and non-RtW beneficiaries may differ in ways that affect job search and employment outcomes, which is important for constructing an adequate counterfactual.

To illustrate this point, consider a scenario in which graduates who already had secured a job for after their graduation decide not to participate in the RtW program whereas graduates who are less certain about their employment prospects do participate. If the employment outcomes of these two groups were compared in early 2020, the measured program impact may be negative or at least downward biased just because beneficiaries and non-beneficiaries are not comparable as easy as that. Further, among RtW beneficiaries, graduates who are being placed in internships potentially have better labor market prospects than RtW beneficiaries who were only trained, which is also the reason they were selected for the internship in the first place. Thus, measured positive program effects may stem from the placed beneficiaries' labor market advantage, but do not or at least not fully stem from RtW's positive program impact. Failures to account for these differences in treatment and control group can lead to biased impact estimations and erroneous conclusions.

In order to make the treatment and control groups comparable, linear multivariable regression models are used which control for background characteristics of participants that are expected to drive the decision to participate in RtW as well as labor market outcomes. These include, for

example, demographic and socioeconomic status variables such as age, gender, graduation from a certificate or diploma level vocational training, trades studied during the vocational training, type and location of school attended (e.g., VTI or TC), household wealth, and education levels of male and female household heads, but also the respective outcome variable measured a year prior to the survey in February 2019. Controlling for pre-intervention outcomes (i.e., the respective outcome in February 2019) is similar to employing a difference-in-differences method but has the statistical advantage of being more efficient than the classical difference-in-differences interaction model approach.[51,52] In addition, the estimation strategy does not rely solely on pre-intervention employment outcomes as some difference-in-differences analyses do, but also controls for the sociodemographic background characteristics listed above.[53]

To check the robustness of the estimated program impact, a second estimation method, called Double Machine Learning, was employed. Compared to the multivariable regression approach, in which the researcher decides to include control variables on theoretical grounds, the Double Machine Learning approach selects relevant control variables from a list of variables that all could potentially be included in the estimation model using a machine learning algorithm. Hence, the estimation model is not specified through the researcher's individual decision but instead a data driven approach is used to adequately model the counterfactual problem. The Double Machine Learning method uses Least Absolute Shrinkage and Selection Operator (LASSO) regressions, a penalized regression method that selects only relevant variables that strongly predict both the participation in RtW and the outcome variable.[54] Once control variables have been selected based on the machine learning algorithm, the RtW program effects are estimated using the selected control variables in a multivariable regression estimation.

Because the results and implications of the Double Machine Learning approach are similar to those based on the linear multivariable regression, they are not presented in the main text but in Appendix A3.2.

---

[51] *Difference-in-differences is an econometric method to measure the impact of an intervention or policy program by comparing the average change over time in the outcome variable, e.g., employment status, for the treatment group with the average change over time in the outcome variable for the control group. The key assumption of this method is that the change over time in the outcome variable for the treatment group would have been the same as the change over time in the outcome variable for the control group in the absence of the intervention or policy program. This method thus relies on the observed outcome variables and, in the case of the RtW impact evaluation, assumes that pre-intervention outcomes reflect participants labor market potential.*

[52] *The efficiency of an estimator refers to the variance of the estimate and, thus, estimates based on more efficient estimators are more precisely measured (while the point estimate based on the two estimators is the same).*

[53] *The difference-in-differences approach assumes that selection into treatment can be captured through pre-intervention employment outcomes and parallel trends in the outcome variable in the absence of the intervention. However, this assumption would be wrong in a sample of students whose baseline employment outcomes are not representative for their full labor market potential because at this time they did not enter the labor market yet. Hence, the required parallel trend assumption likely does not hold.*

[54] *LASSO regressions include a penalty parameter that shrinks coefficient towards zero and set coefficients of irrelevant variables to zero. This procedure allows LASSO to perform variable selection.*

4.3.2 COVID-19 and the timing of data collection

The data collection schedule originally planned in 2019 differed from the one eventually implemented in 2020. The research design was developed prior to the outbreak of the coronavirus 2019 pandemic (COVID-19) and had foreseen to conduct two surveys - one in mid-2019 during participants final year of education and one in late-2020. However, due to delays in ethical clearance and difficulties in the recruitment of participants at non-RtW schools, data were collected just before the outbreak of the COVID-19 pandemic in February 2020 for the first time, and the second survey was cancelled due to the COVID-19 pandemic. The decision to cancel the second survey followed an assessment which weighted the survey's monetary costs against the benefits of collecting employment data that were expected to be severely affected by the pandemic.

As a result of the cancellation of the second survey, the impact evaluation fully relies on data collected in February 2020 and can only measure short-term employment effects as participants had only graduated in 2019 and internship placements typically followed graduation. The estimation sample even includes some beneficiaries who are still in UMA internships at the time of the survey or who have completed internships just shortly prior to the survey (see section 4.4.3 for a detailed analysis). In addition, irrespective of the RtW program, there was overall little time between graduation and data collection for study participants to successfully search for a job. In consequence, potential positive program impacts may have not been materialized at the time of the survey and the measured short-term effects may not reflect the full potential of the RtW program to improve labor market outcomes.

The implications of the cancellation of the second survey potentially differ by participants' date of graduation from their respective VTI or TC. TC students as well as students at the VTIs YMCA, and YWCA graduated in May 2019, whereas students at other VTIs graduated in November 2019. Both VTIs and TCs offer technical and vocational certificate level and diploma level tertiary education. In comparison to VTIs, TCs are national public schools with a focus on engineering trades and a more renowned reputation.

Participants who graduated in May typically participated in the RtW training earlier and were also placed earlier in internships (details on the timing of RtW trainings, internship placements and the survey are presented in section 4.4.3Because in this subsample of early graduates potential program impacts had more time to materialize, the main estimation results of the RtW program impact are therefore presented for the subsample of participants who graduated in May 2019 in addition to the results for the full sample.

An additional important implication of the cancellation of the second survey is that some information that were scheduled to be collected during the follow-up survey in 2020 could not be collected. This includes mostly information on internship placements - such as satisfaction with and learning outcomes of the placements -, details on the job search, and information on key educational outcomes - such as the final grade point average. The final grade point average was oftentimes not available yet at the time of the first survey but may had been relevant for individuals to decide whether they participate in the RtW program and, therefore, could have been used to model the counterfactual.

4.3.3 The pre-COVID-19 research design

The estimation methods originally proposed and planned for in early 2019 differ from the estimation method used and presented in section 4.3.1. The proposed methods comprised a mixed design that combined statistical matching with a difference-in-differences estimation strategy.

In a first step and based on data collected in 2019, RtW beneficiaries and non-beneficiaries at RtW schools would have been matched to students with similar attributes and background characteristics (e.g., trades, school performance, socioeconomic status etc.) at non-RtW schools in order to generate hypothetical beneficiary and non-beneficiary groups of students at non-RtW schools. In a second step and based on data collected in late 2020, the employment and income outcomes of these actual and hypothetical beneficiary (treatment) and non-beneficiary (control) groups at treatment and control schools would have been compared in a difference-in-differences analysis to estimate the RtW program impact and answer RQ#1 and RQ#2. Appendix Figures A3.1.1 and A3.1.2 illustrate the mixed approach of statistical matching and difference-in-differences estimation.

This original research design could not be implemented mainly because of difficulties in the recruitment of participants at non-RtW schools as well as particularly low response rates among participants at non-RtW schools. Out of 1,466 successfully completed interviews only 153 interviews were completed at non-RtW schools. The matching process, upon which the cohort difference-in-differences estimation hinges, became unfeasible as it requires a large sample size of comparison group participants.

4.3.4 Outcome indicators

Primary outcomes

The objective of the RtW program is to improve labor market prospects and outcomes of young people in Uganda. As such, the primary outcomes of this impact assessment constitute direct employment benefits, which include different forms of measuring employment status and earnings.

The different employment measures are dichotomous variables (1 if employed, 0 if unemployed) and include the following indicators:

1. **Employment status,** which is defined as any paid work. It includes self-employment and excludes paid internships.
2. **Decent employment,** which is defined as paid work of at least 20 hours per week that earns at least 148,000 UGX per month. It includes self-employment and excludes paid internships.
3. **Self-employment,** which is defined as any own-account work or own business.
4. **Job with contract,** which relies on the respondents' answer to whether he or she had a contract with the employer (excluding internships).
5. **Formal employment,** which relies on guidelines of the United Nation's International Labor Organization on measuring informality based on the 15th and 17th International Conference of Labour Statisticians (International Labour Organization 2009, 2013). A person who works for someone else with pay and receives at least one of thirteen benefits (e.g., paid sick leave, insurance, social security etc.) or pays taxes is considered to have a formal job (excluding internships). Self-employment is considered formal if the enterprise is registered or taxes on earnings were paid.

The income measures assessed include:

1. **Total income in UGX, which is the sum of earnings from all paid work for participants** who have paid work and zero for participants who do not have paid work (internships are not considered work).
2. Total income among employed in UGX, which is the sum of earnings from all paid work for participants who have paid work (internships are not considered work; estimation sample excludes those without paid work).
3. **Hourly wage among employed** in UGX, which is the average hourly wage rate or earnings from all paid work for those who have paid work (internships are not considered work; estimation sample excludes those without paid work).

In addition to employment status and earnings, also working hours and hourly wages were assessed in order to fully understand the relationship between total income and wages. Specifically, the following two outcome indicators were used:

1. **Total number of hours worked,** which is the sum of hours worked per month in all paid jobs for participants who have paid work and zero for participants who do not have paid work.
2. **Total number of hours worked among employed**, which is the sum of hours worked per month in all paid jobs for participants who have paid work (estimation sample excludes those without paid work).

Secondary outcomes

In addition to employment and earnings, the following secondary outcomes were assessed:

Employment aspirations in the **short term**:

- Employment aspirations in the short term, which is the response to the question "How optimistic are you that you will have work in the short term, i.e., during the next 12 months?" on a scale from 1 (not optimistic at all) to 5 (very optimistic).
- Income aspirations in the short term, which is the response to the question "How much do you aim to earn monthly from that work?"

Employment aspirations in the **long term**:

- Employment aspirations in the long term, which is the response to the question "How optimistic are you in finding work that satisfies your expectations in terms of tasks and salary in the long term, i.e., after 12 months from now?" on a scale from 1 (not optimistic at all) to 5 (very optimistic).
- Income aspirations in the long term, which is the response to the question "What is the monthly salary/income you aim to earn in the long term, i.e., after 12 months from now?"

Migration intentions, which are binary responses (yes/no) to the following two questions**:**

- "Ideally, if you had the opportunity, would you like to move permanently or at least for 12 months to another country, or would you prefer to continue living in this country?"
- "Are you planning to move permanently to another country in the next 12 months, or not?"

4.4 Data collection

4.4.1 Sample size and response rates

The survey of vocational training graduates took place between February 12 and March 5, 2020. The survey targeted to enroll 3,162 final year students in the study, however, only 1,466 interviews were conducted. Of those 1,466 interviews

- **237** respondents are beneficiaries from RtW schools that participated in the **RtW training and were matched to an UMA member company**,
- **541** respondents are beneficiaries from RtW schools that participated in the **RtW training only** and were not macthed in an UMA member company,
- **535** respondents are **non-beneficiaries from RtW schools**, and
- **only 153** respondents are from **control schools**.

This corresponds to a response rate of 47% among respondents at RtW schools (in total contact information for 2,827 students from RtW schools were available) and 14% among respondents at non-RtW schools (in total contact information for 1,075 students from non-RtW schools were available). The main reason for attrition at control schools was that a large share of students for whom we received contact details were, contrary to what had been discussed with school principals, not in their final year but still had a year or two ahead of them until their vocational training was completed and they were eligible for the RtW program. In fewer but still numerous cases phone numbers were erroneous or inactive.

The distribution of treatment groups across school types with different graduation dates is shown in Table 4.4.1. Most participants are from VTIs with graduation dates in November and fewer participants are from TCs and VTIs with graduation dates in May.

Table 4.4.1
Number of respondents by school type and treatment group

	Graduation date	RtW with placement	RtW training only	No intervention control (at RtW and non-RtW schools)
VTI	Nov. 2019	199	426	419
TC	May 2019	33	103	228
YMCA/YWCA	May 2019	5	12	41
Total		237	541	688

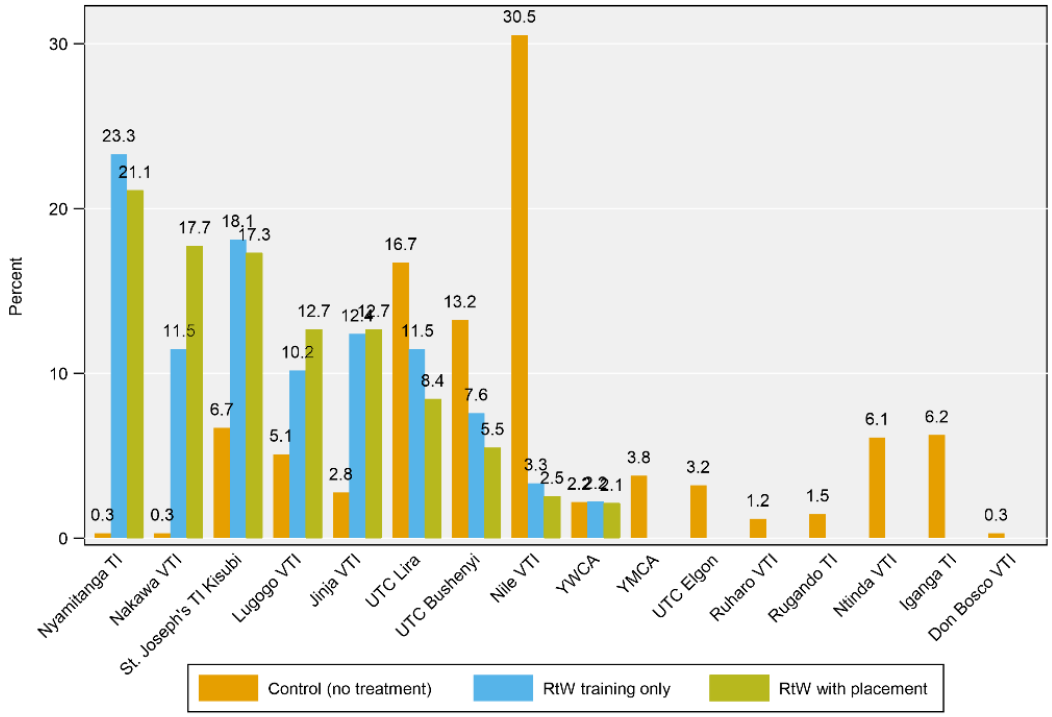
Source: Own calculations based on RtW Survey.

4.4.2 Study participants by treatment group, schools, and trades

In total, the sample covers 16 schools and the distribution of respondents by treatment group varies widely across schools. Figure 4.4.1 illustrates the percentage share of respondents across schools by treatment group and is ordered by the percentage share of beneficiaries who received an internship placement in descending order. In Nyamitanga Technical Institute and Nakawa VTI virtually everyone participated in the RtW program and together with St. Joseph's TI Kisubi these three schools comprise 56% of respondents that were placed in UMA internships and 53% of respondents that received the RtW training only. Note that while the share of trained and placed respondents is similar in Nyamitanga TI and St. Joseph's TI Kisubi, the absolute number of

students placed is much smaller than those trained because in total there were 237 respondents trained and placed and 541 trained only. For example, at Nyamitanga TI 176 respondents were trained of which 50 were placed.

Figure 4.4.1
Distribution of respondents across schools by treatment group



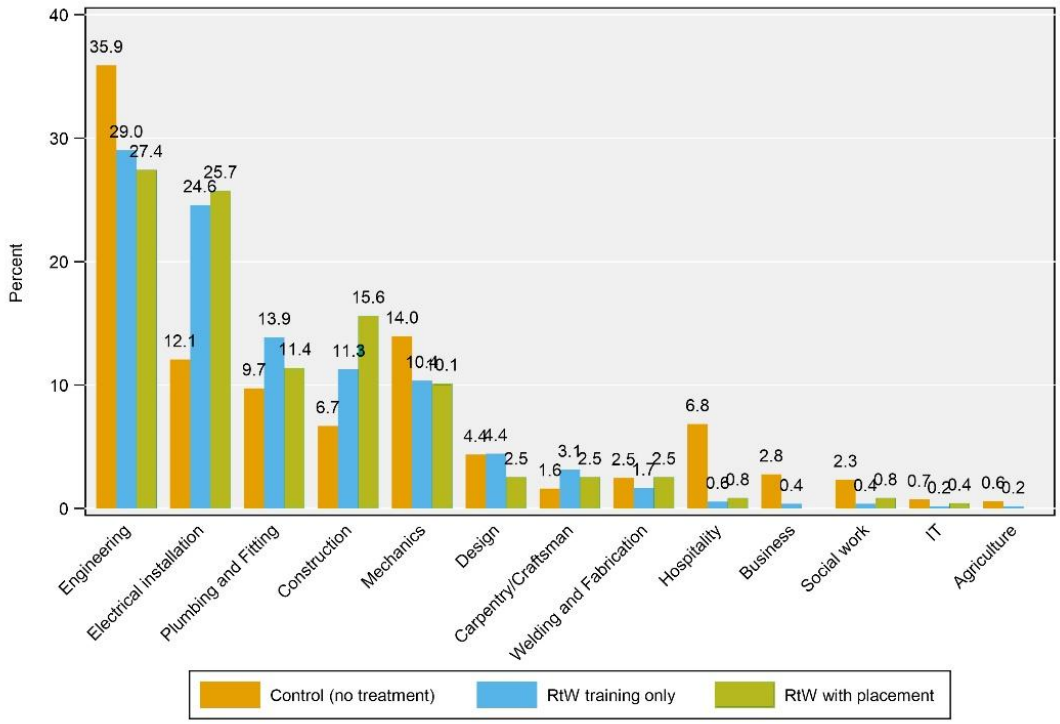
Notes: Nyamitanga TI: N=178; St. Joseph's TI Kisubi: N=185; Nakawa VTI: N=106, Lugogo VTI: N=120; Jinja VTI: N=116, UTC Lira: N=197, UTC Bushenyi: N=145, Nile VTI: N=234; YWCA: N=32; YMCA: N=26; UTC Elgon: N=22; Ruharo VTI: N=8; Rugando VTI: N=10; Ntinda VTI: N=42; Iganga TI: N=43; Don Bosco VTI: N=2. - Source: Own calculations based on RtW Survey.

At St. Joseph's Technical Institute Kisubi, Lugogo VTI, and Jinja VTI, the majority of respondents have participated in the RtW training and considerable shares also received a placement. In contrast, at the Technical Colleges of Lira and Bushenyi, there are more untreated participants than treated ones and also at Nile Vocational Institute only few participants were trained (29 of 234). Overall, Nile Vocational Institute makes up a third of the total no-intervention control group sample.

The remaining 9 schools add relatively few participants to the overall sample and mostly consist of no-intervention control group participants as they are also mostly non-RtW schools.

Figure 4.4.2 illustrates the distribution of respondents across vocational trades by treatment group. Treatment group members have relatively more often studied electrical installation, construction, and plumbing and fitting. Welding and fabrication is balanced across treatment and control groups and so tend to be carpentry or craftsman trades, IT, and agriculture. Control group members have relatively more often studied engineering, mechanics, design, hospitality, business, and social work.

Figure 4.4.2
Distribution of respondents across courses by treatment group



Source: Own calculations based on RtW Survey.

4.4.3 Timing of survey and program roll-out

The time that had passed between program participation and the survey in 2020 was determined by the type of school attended and, thus, the month of graduation, the school-level schedule for RtW trainings, and companies' demand for internships. Figure 4.4.3 illustrates the timing of the survey relative to these relevant events and program activities by type of school, i.e., VTIs, TCs and YMCA/YWCA, and treatment group, considering the following three measures:

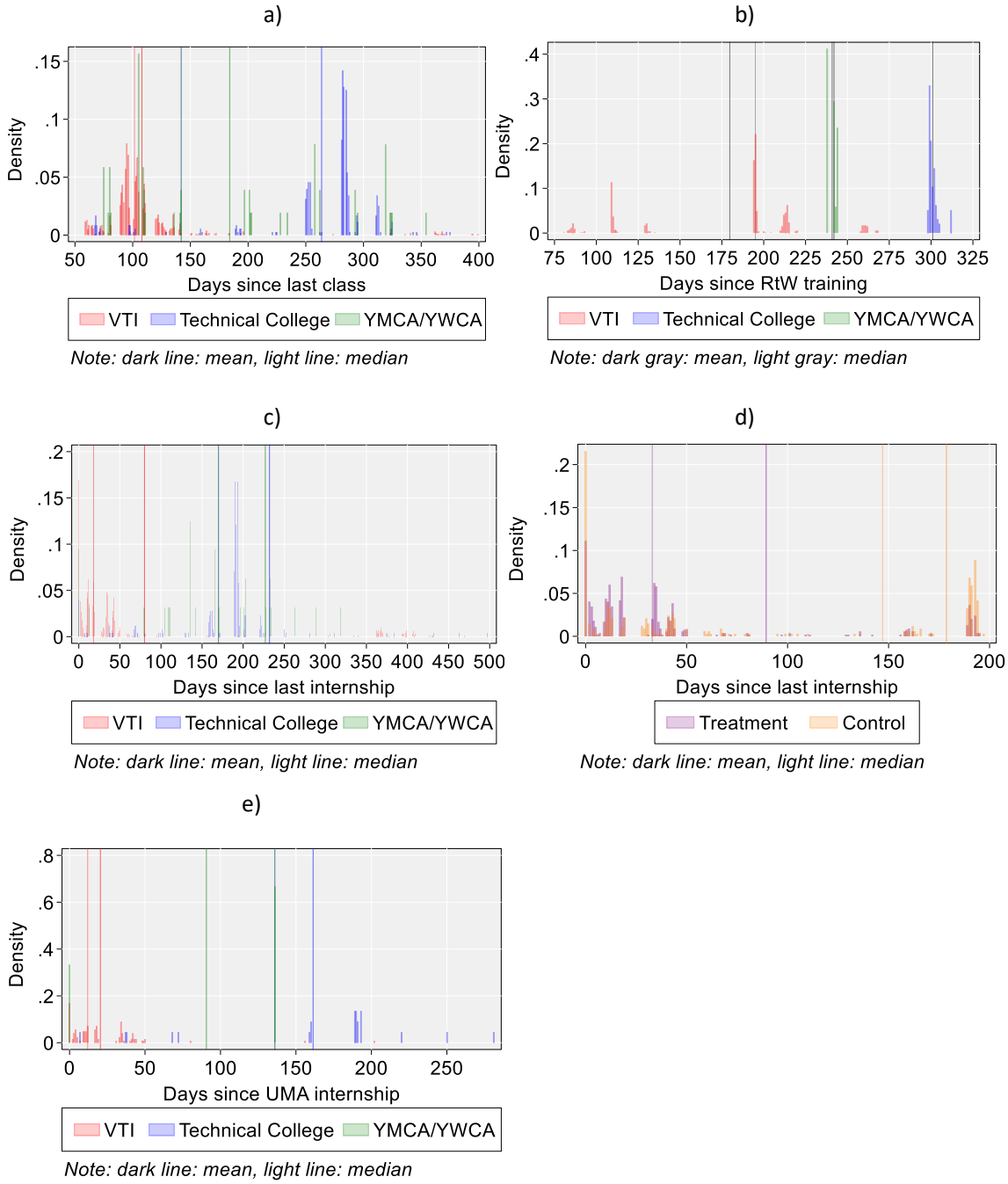
i. the number of days between the last day of class and the survey (Figure 4.4.3.a),

ii. the number of days between the RtW training and the survey (Figure 4.4.3.b), and

iii. the number of days between the last internship was completed and the survey (Figures 4.4.3.c-e).

Overall, the time period between the survey and program activities was very heterogeneous. Because TC and YMCA/YWCA students graduated mostly around May and VTI students around November, the last day of class was much more recent (about 100 days ago) for VTI graduates than for graduates of the other two types of schools (about 200 to 300 days ago) (see Figure 4.4.3.a). Similarly, the RtW training had been 240 to 300 days ago for TC and YMCA/YWCA graduates, whereas the RtW trainings at VTIs were much more spread out – at some VTIs it was also about 260 days before the survey, whereas at other VTIs it had been implemented only three months ago (see Figure 4.4.3.b).

Figure 4.4.3
Time between relevant program activities and date of survey



Source: Own calculations based on RtW Survey.

Further, Figures 4.4.3.c and 4.4.3.e show that VTI graduates had completed any internships as well as UMA internships more recently, i.e., in the last three months, than TC and YMCA/YWCA students did. For comparisons by treatment group, Figure 4.4.3.d shows that both treatment and control group participants completed internships either recently or five to six months ago. However, relatively more treatment group than control group members completed their internships more recently.

The time between completion of the RtW program (i.e., time of RtW training or completion of internship placement) and the survey is important in order to understand the potential treatment effects that can be identified as part of this study. Given the timing of events as described in Figure 4.4.3, only very short-term labor market effects can be recovered from the prospective analysis. Positive program impacts may arise from treatment group members who are being retained in the company upon completion of the UMA placement. However, potential positive program impacts may not have materialized yet at the time of the survey. One important factor in this regard may be that treatment group participants had less time to search for a job as they were focusing on completing UMA placements, whereas in the longer term the benefits from that placement may pay off as the qualifications gained during the internship may have improved beneficiaries' employability. The implications of the differences in job search time for the interpretation of the evaluation results will be discussed in section 4.12.

4.5 Background characteristics of study participants

This section describes individual level characteristics of study participants across treatment groups with respect to participants'

a) sociodemographic background characteristics and

b) pre-intervention employment characteristics.

The pre-intervention employment indicators date back to February 2019, a year before data collection in 2020, and were enquired via a set of recall questions. Similarly, some of the sociodemographic background indicators, such as those on wealth, are based on recall questions in order to represent the economic situation of participants prior to the RtW program and, therefore, are unaffected of the RtW program.

The examination of individual level background characteristics serves two purposes. First, it describes the study population for whom treatment effects are measured. And second, it allows a detailed investigation of the comparability of treatment and control group members.

4.5.1 Sociodemographic characteristics

Table 4.5.1 presents background characteristics of study participants by treatment group in columns (1) to (3) and comparisons of background characteristics across treatment groups in columns (4) and (5).

Sample characteristics – Columns (1) to (3)

Column (1) presents mean background characteristics (or percentage shares for binary indicators) of respondents who received the RtW training and an internship placement. Column (2) refers to respondents who received the RtW training only and column (3) refers to respondents who did not participate in the RtW program at all.

Most participants are males (81-87%) and participants' average age is 23 years. Among RtW beneficiaries (columns 1 and 2), less than 20% have attended a diploma level course rather than a certificate level course in contrast to 36% of the no-intervention control group. On average, participants are single (89-94%) and have 0.8 dependents (e.g., a family member to provide for).

Table 4.5.1
Background characteristics of study participants

	(1) RtW with placement % or Mean	(2) RtW training only % or Mean	(3) No Treatment % or Mean	(4) Std. Diff (2) - (1)	(5) Std. Diff (3) - (1)
Respondent characteristics					
Age	22.7	22.5	23.6	-0.05	0.24
Male	86.1	87.1	80.8	0.03	-0.14
Diploma level (=1 vs. Certificate=0)	15.2	19.8	35.8	0.12	0.49
Single	94.1	94.1	88.7	-0.00	-0.19
No. of dependents	0.95	0.58	0.77	-0.20	-0.10
Household characteristics					
Wealth quintile					
1 (poorest)	16.5	18.9	22.2	0.06	0.15
2	17.7	21.1	19.9	0.08	0.06
3	21.5	21.1	18.6	-0.01	-0.07
4	24.1	18.5	19.8	-0.14	-0.10
5 (wealthiest)	20.3	20.5	19.5	0.01	-0.02
Male head characteristics					
Can read	71.3	71.0	73.8	-0.01	0.06
None	8.4	10.7	10.5	0.08	0.07
Primary	13.9	14.6	12.5	0.02	-0.04
Secondary-O	16.0	14.6	14.8	-0.04	-0.03
Secondary-A	3.8	2.6	4.2	-0.07	0.02
Certificate	10.6	9.8	11.8	-0.02	0.04
Diploma	9.3	10.9	12.9	0.05	0.12
Tertiary education	11.0	11.7	11.1	0.02	0.00
Don't know	6.8	6.1	5.4	-0.03	-0.06
No male head/spouse	20.3	19.0	16.9	-0.03	-0.09
Female head characteristics					
Can read	52.7	62.9	64.2	0.21	0.23
None	19.0	20.9	22.5	0.05	0.09
Primary	17.7	17.4	13.4	-0.01	-0.12
Secondary-O	8.0	12.6	12.1	0.15	0.13
Secondary-A	1.7	2.0	3.9	0.03	0.14
Certificate	2.5	5.2	6.0	0.14	0.17
Diploma	7.2	5.9	8.0	-0.05	0.03
Tertiary education	2.1	6.8	4.7	0.23	0.14
Don't know	8.4	6.7	10.0	-0.07	0.05
No female head/spouse	33.3	22.6	19.5	-0.24	-0.32
Test of joint orthogonality					
F-test statistic				2.20	7.20
p-value				0.02	0.00
Observations	237	541	688		

Source: Own calculations based on RtW Survey.

Wealth quintiles are based on assets owned by participants or their household members and were constructed in a way such that about 20% of participants of the whole sample (i.e., irrespective of the treatment group) make up one quintile category. There seem to be relatively more participants of the control group in the lowest wealth quintile, but overall the differences are small.

71 to 74% of participants live in households in which the male head can read. In much fewer households the female head can read, who can be the spouse of the male head or any other main female decision maker. In the internship placement group, in only 53% of participants' households the female head can read. In contrast, in 63 and 64% of households of beneficiaries in the RtW-training-only group and the no-intervention control group can the female head read, respectively. The most common levels of completed education among male heads are primary and secondary O-level education, closely followed by tertiary education, diploma level, and certificate level. The least common level of education is Secondary A-level (3-4%). Between 20 and 24% of male heads have completed some vocational training (either certificate level or diploma level). Among female heads, the most common levels of education are no or primary education (19-23% with no education, 13-18% with primary education) and the least common one is secondary A-level education. Between 10 and 14% of female heads have completed some vocational training (either certificate level or diploma level).

Differences in background characteristics across groups – Columns (4) and (5)

Columns (4) to (5) present statistics that allow to evaluate the similarity and, thus, comparability of the three groups. Column (4) compares the two treatment groups, RtW with placement and RtW training only, and column (5) compares the RtW beneficiaries that received a placement with participants that did not receive a treatment at all. Columns (4) and (5) show standardized differences (Std. Diff. in Table 4.5.1) as measures of similarity for each characteristic respectively. A standardized difference takes the difference in means of the two groups and weights it by the variance of the indicators (the variance is a measure of how much the indicator spreads around its mean). Standardized differences of 0.2 (or -0.2) or less are considered small. Thus, when the similarity of treatment and control groups is evaluated, everything that is larger than the 0.2 threshold (or smaller than –0.2) would be considered as not similar.

The last two rows of Table 4.5.1 in columns (4) and (5) present measures that evaluate the similarity across treatment and control groups considering all presented characteristics jointly. A p-value below 0.1 would imply that the listed background characteristics of the two considered groups are significantly different from each other and, therefore, the two groups would not be comparable without applying econometric techniques that restore comparability.

The test statistic at the bottom of column (4) indicates that the two treatment groups significantly differ from each other taking into account all characteristics jointly, as indicated by the p-value of below 0.1 at the bottom of Table 4.5.1. Considering the individual standardized differences, there are five characteristics – namely, the average number of dependents, the share of female heads who can read, the share of female heads who have completed a certificate level, the share of female heads who have completed tertiary education, and whether there is a female decision maker in the household – that pass the threshold of 0.2. The standardized differences with respect to female heads indicate lower levels of education of female heads in the treatment group that received an internship placement in comparison to the treatment group that received the RtW training only.

Turning to the comparison of beneficiaries who received the RtW placement with the no-intervention control group in columns (5), Table 4.5.1 indicates marked differences between these groups. The p-value at the bottom of Table 4.5.1 is below 0.01, indicating significant differences when considering all presented characteristics jointly. Also, the individual standardized differences in column (5) are often larger than 0.2 or smaller than -0.2; particularly, with respect to age, diploma level, the female heads' reading abilities, and whether there is a female decision maker in the household.

4.5.2 Employment status and earnings prior to the RtW program

Tables 4.5.2 and 4.5.3 present mean pre-intervention (February 2019) employment and income indicators of participants across the three treatment groups. Table 4.5.2 comprises all participants whereas Table 4.5.3 only includes participants who were employed in February 2019.

Table 4.5.2 illustrates that only very few participants had any paid work in February 2019 (1-3%), which stands to reason given that the sample of participants was still in education in early 2019. However, 59 to 68% of participants were in internships in February 2019. The share of beneficiaries who participated in an internship is about 9% larger in the RtW internship placement group (68%) and RtW-training-only group (68%) than in the no-intervention control group (59%). The differences in internship rates are unrelated to the RtW program given that the statistics presented in Table 4.5.2 refer to the time before the RtW program roll-out. The standardized differences for the internship indicator are smaller than 0.2 (-0.00 and -0.16) and thus do not mark a considerable difference of the RtW-with-placement group and RtW-training-only group in comparison to the no-intervention control group. Taking into account all employment indicators jointly, the p-values indicate that the RtW placement and RtW-training-only groups are not comparable prior to the intervention with respect to their labor market characteristics (p-value<0.1), whereas the RtW placement and no-intervention control group are.

Table 4.5.2
Pre-intervention employment characteristics

	(1) RtW with placement % or Mean	(2) RtW only % or Mean	(3) No treatment % or Mean	(4) Std. Diff (2) - (1)	(5) Std. Diff. (3) - (1)
Employment					
Employment	2.1	0.92	2.8	-0.10	0.04
Decent employment	1.7	0.92	2.5	-0.07	0.05
Self-employment	0.42	0.00	0.00	-0.09	-0.09
Has contract	0.42	0.37	0.87	-0.01	0.06
Formal employment	0.00	0.00	0.00	.	.
Internships	68.35	68.21	59.01	-0.00	-0.16
Income					
Total income	11,781	12,357	17,53	0.01	0.06
Hourly wage	28.2	10.8	98.9	-0.10	0.07
Test of joint orthogonality					
F-test statistic				1.93	0.89
p-value				0.07	0.50
Observations	237	541	688		

Source: Own calculations based on RtW Survey.

Table 4.5.3 takes a closer look at the number of jobs, hours worked, hourly wages, and positions held among those that were employed. However, given that only five, five, and 19 participants of the RtW with placement, the RtW training only and the control group were employed in February 2019, respectively, Table 4.5.3 does not allow to draw conclusions about the employment situation of those that had paid work.

Table 4.5.3
Background characteristics of study participants

	(1) RtW with placement % or Mean	(2) RtW only % or Mean	(3) No treatment % or Mean
Employment			
No. of jobs among employed	1.8	1.8	1.1
Hours worked among employed	186.7	215.4	215.4
Income			
Income among employed	248,000	250,750	518,778
Hourly wage among employed	1,335	1,273	3,660
Positions			
Office worker	0.0	0.0	5.3
Skilled tradesman	60.0	100.0	84.2
Unskilled tradesman	20.0	0.0	5.3
Self-employment	20.0	0.0	5.3
Observations	5	5	19

Source: Own calculations based on RtW Survey.

4.5.3 Restoring the comparability of treatment and control groups

The similarity and thus comparability of treatment and control groups is important in order to attribute measured employment effects to the impact of the RtW program. If treatment and control groups are not comparable the measured treatment effects could be due to participants' different background characteristics and labor market potential. For example, if there were more males in the treatment group than in the control group and if males have better labor market opportunities due to gender discrimination in the job market, then estimated program effects may be due to participants' gender if the different gender compositions across treatment groups were to be ignored in the estimation approach. In order to restore the comparability of treatment and control groups, we apply the estimation strategies outlined in section 4.3.1. That is, we control for relevant background characteristics of participants that correlate with treatment group membership and labor market outcomes. In the appendix, we also present results that are based on a Double Machine Learning algorithm to select relevant control variables rather than selecting them on theoretical grounds.

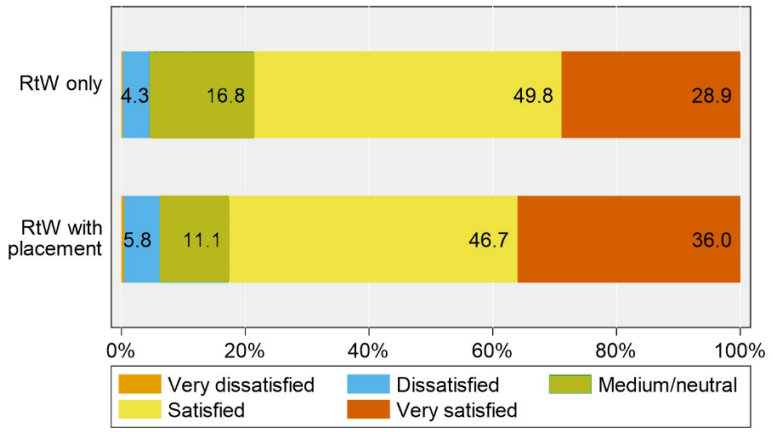
4.6 Descriptive statistics on the ReadyToWork program and UMA internship placements

This section takes a closer look at the study participants who benefitted from the RtW program and provides descriptive analyses about their satisfaction with the RtW training and internship placement rates by beneficiaries' sociodemographic characteristics. Regrettably, there is no information on beneficiaries' satisfaction with internship placements due to the cancelation of the follow-up survey in late-2020.

4.6.1 Participants' satisfaction with the RtW training program

Figure 4.6.1 presents beneficiaries' satisfaction with the RtW training, excluding satisfaction with placement services and experiences. About 80% of beneficiaries indicated to be satisfied or very satisfied with the RtW training and less than 6% were dissatisfied or very dissatisfied. Interestingly, beneficiaries that were placed in internships indicated more often to be very satisfied with the RtW training than beneficiaries who were not placed in internships. Although this question only regarded the RtW training itself, positive experiences gained through the internship placement may have influenced the placed beneficiaries' satisfaction with the RtW training and led to a higher overall ex-post satisfaction rate.

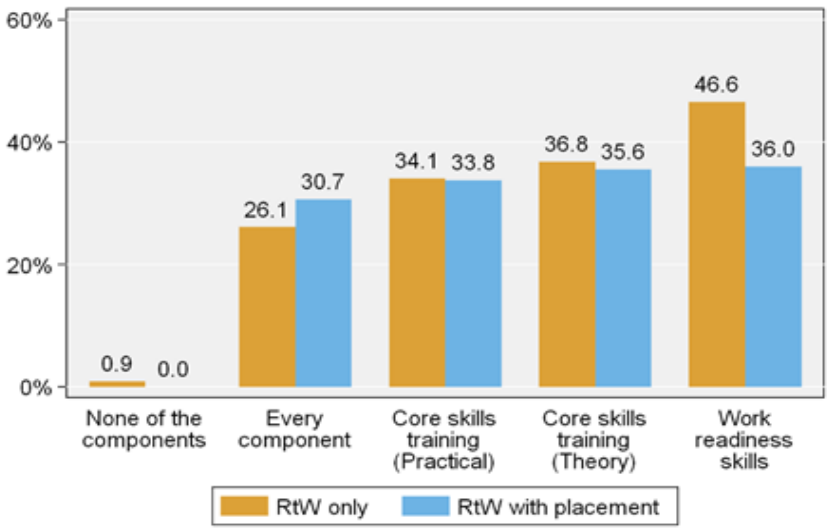
Figure 4.6.1
Beneficiaries' satisfaction with the RtW training, excluding the satisfaction with placement services and experiences



Source: Own calculations based on RtW Survey.

Figure 4.6.2 summarizes which of the RtW training components were considered the most successful ones, including the components of practical core skills training, theoretical core skills training, and work readiness skills.

Figure 4.6.2
Most successful components if the RtW training

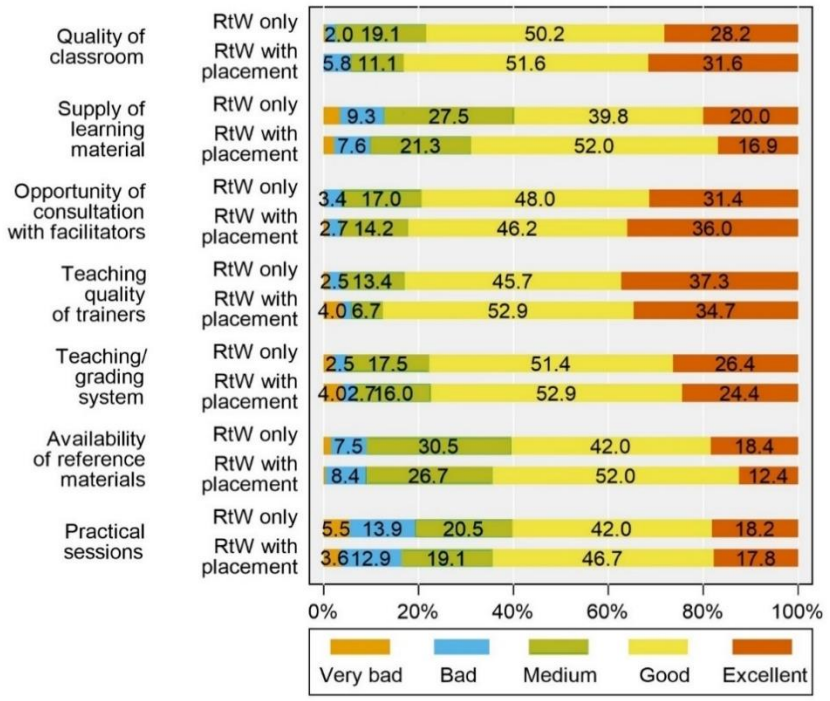


Source: Own calculations based on RtW Survey.

Participants were able to provide multiple responses or indicate none or all components. The work readiness component was considered the most successful part of the training (42%), followed by the theoretic core skills (36%), and finally the practical core skills (33%). A quarter of beneficiaries even considered all three components equally successful. Only few respondents were overall dissatisfied with the RtW training and did not consider any components successful. Satisfaction rates were very similar across the two treated groups, with the exception of the work readiness skills component, which RtW training only participants rated much more often successful (47%) than participants who were placed in UMA internships (36%).

Figure 4.6.3 illustrates beneficiaries' satisfaction with different aspects of the RtW training conditions, including the quality of classrooms, supply of learning materials, opportunity of consultation with facilitators, teaching quality of trainers, teaching and grading system, availability of reference materials, and practical sessions.

Figure 4.6.3
Beneficiaries' satisfaction with different aspects of the RtW training conditions



Source: Own calculations based on RtW Survey.

Across the different aspects of training conditions at least 60% of respondents evaluated each aspect to be good or excellent. Particularly positive evaluations were given for the teaching quality of trainers, the quality of classrooms, and opportunities to consult with facilitators with 80% or more of the respondents judging those conditions as good or excellent. The practical session received the poorest evaluation for which 18% of respondents indicated that the conditions were bad or even very bad and another 20% evaluated the practical session as medium. Further, about 10% of respondents evaluated the supply of learning materials and the availability of reference materials to be bad or very bad and another 25% as medium.

4.6.2 Descriptive statistics of the internship placements

Placement types

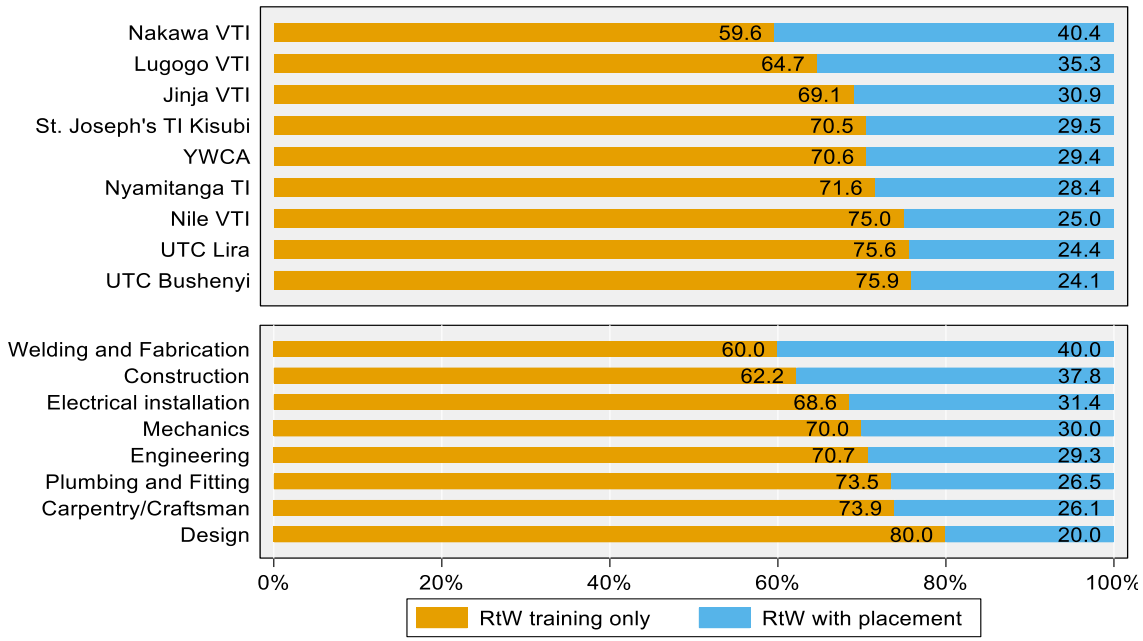
Of the 778 respondents who participated in the RtW training, 237 (30%) were matched to companies by UMA. The majority of placements were internship placements, but a substantial number of beneficiaries was directly placed in jobs:

- **123** beneficiaries were **placed in internships**,
- **87** beneficiaries were **placed in jobs**,
- and in 27 cases it was not possible to identify the type of position.

Who was placed?

Considering all placements, i.e., internships, jobs, and those that could not be defined, this section describes the percentage shares of placed beneficiaries by sociodemographic background characteristics. This means that the sample is restricted to respondents with a certain characteristic, for example, being female, and the share of placed beneficiaries among respondents with that characteristic is presented.

Figure 4.6.4
Percentage share of beneficiaries placed by treatment school and trades

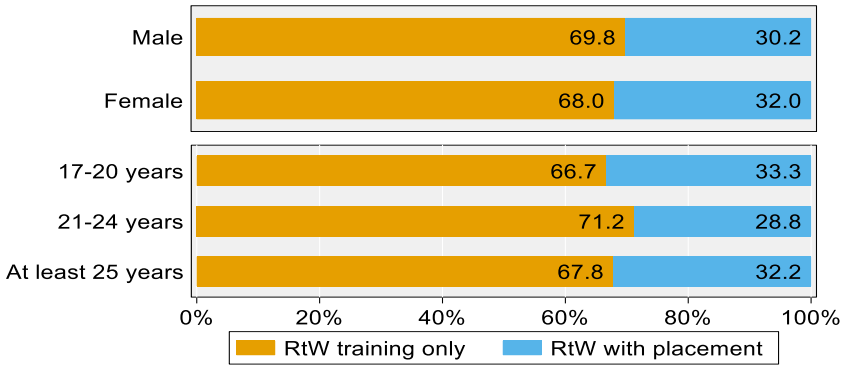


Source: Own calculations based on RtW Survey.

Figure 4.6.4 presents the percentage share of placed beneficiaries in each treatment school and for each vocational training trade. Relative to the total number of beneficiaries at the respective school, most of the placed beneficiaries had attended Nakawa VTI (40% of graduates at Nakawa VTI) and least attended UTC Bushenyi (24.1%). By trade, the largest relative percentage share of placed beneficiaries was among graduates in welding and fabrication (40% of graduates in welding and fabrication), construction (38%), and electrical installation (31%), whereas students in design (20%) and carpentry or craftsman trades were placed the least relative to placement shares in other trades (26%). There were no trades that received none or particularly few placements relative to the total number of students in each trade.

Figure 4.6.5 further disaggregates the percentage shares of placements by beneficiaries' demographic characteristics. The shares of placed beneficiaries among males and females are quite similar, with 30% of males and 32% of females placed, respectively. Across age groups, 33% of 17- to 20-year-olds were placed, 29% of 21- to 24-year-olds and 32% of those 25 years and older.

Figure 4.6.5
Percentage share of beneficiaries placed by gender and age



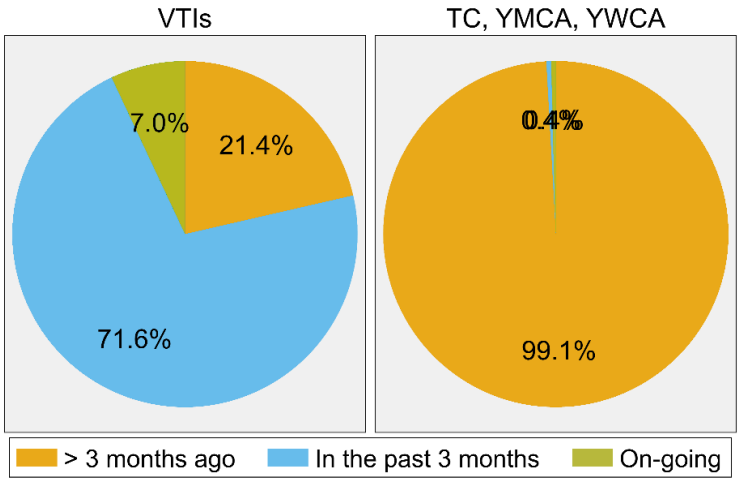
Source: Own calculations based on RtW Survey.

When were beneficiaries placed in internships?

Figure 4.6.6 describes how long ago before the survey UMA internship placements ended and, therefore, speaks to the question of how much time beneficiaries had to search for a job and materialize on the work experience gained during the RtW program. Figure 4.5.6 only includes those beneficiaries who were placed in internships but not those who were placed in jobs or for whom the type of position was not clear.

Of the beneficiaries placed in UMA internships at VTIs with graduation dates in November 2019, 21% completed their internship at least three months ago, 72% completed their internship in the past three months and 7% were still in internships at the time of the survey. In contrast, almost every one of the beneficiaries who graduated in May had completed the internship placement at least three months ago.

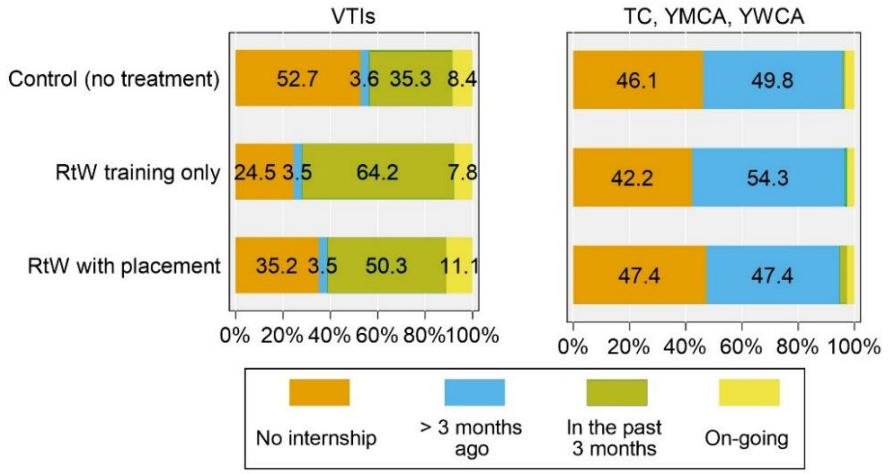
Figure 4.6.6
Timing of UMA placement completion by school type



Source: Own calculations based on RtW Survey.

In order to better evaluate whether beneficiaries that were placed in internships by UMA had a disadvantage in the amount of time they had to search for a job, Figure 4.6.7 describes the percentage share of respondents in each treatment group who completed an internship since their graduation and the time between the survey and internship completion. Figure 4.6.7 does not distinguish between internships that were facilitated by UMA, self-sought, or acquired in any other way (e.g., alternative employment promoting programs).

Figure 4.6.7
Timing of internships (facilitated through UMA and any others) that took place after graduation by treatment group and school type



Source: Own calculations based on RtW Survey.

In Figure 4.6.7, a considerable share of respondents in the RtW-with-placement group appears as not having completed an internship. This is due to the fact that these beneficiaries were placed into jobs directly and, therefore, did not complete an internship after graduation.

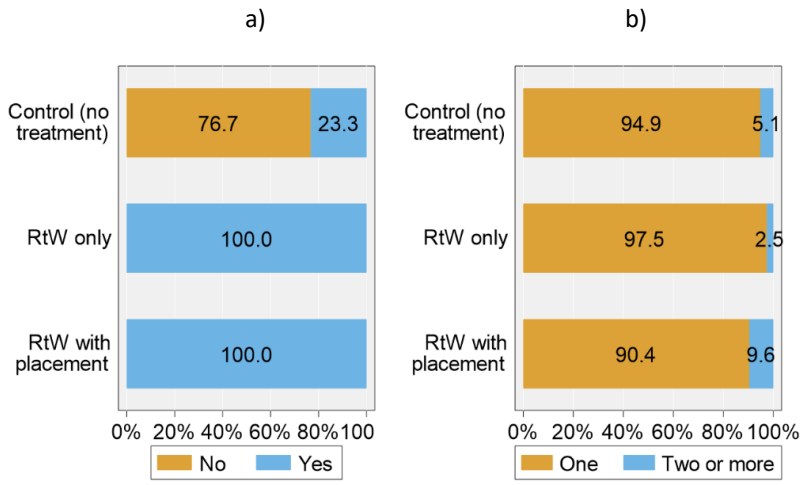
Among VTIs with graduation dates in November, most respondents of the RtW-training-only group had completed an internship in the past three months or were still attending one (72%), followed by those placed into UMA internships (61%) and finally those in the control group (44%). Thus, if time to search for a job can be measured by the time since the last internship was completed, then beneficiaries in both the RtW training only and RtW-with-placement groups were disadvantaged in comparison to the no-intervention control group.

In the sample of respondents with graduation dates in May (TCs, YMCA, and YWCA), only few respondents completed an internship in the past three months or were still participating in one at the time of the survey. The shares of respondents that completed an internship three or more months ago is quite similar across treatment groups, with a slight advantage for the RtW-training-only group in which five percentage points more respondents completed an internship at least three months ago in comparison to the no-intervention control group and the RtW-with-placement group.

4.6.3 Participation in alternative employment promoting programs

In the absence of the RtW program, control group members may have autonomously sought alternative opportunities for training and to gain work experience or their schools may have offered alternative training and employment promoting programs, possibly similarly to the RtW program. Section 4.6.2 revealed that 47% of the control group had participated in an internship after graduation. Figure 4.6.8.a further shows that a little less than a quarter of control group respondents had participated in a career or employment promoting service or activity (outside those routinely offered at schools' career offices). Figure 4.6.8.b shows that those respondents who had participated in any service or activity typically participated in no more than one. This is not only the case for control group members but also for RtW beneficiaries.

Figure 4.6.8
Participation in employment promoting programs or activities (12.a) and the number of activities conditional on participation (12.b)

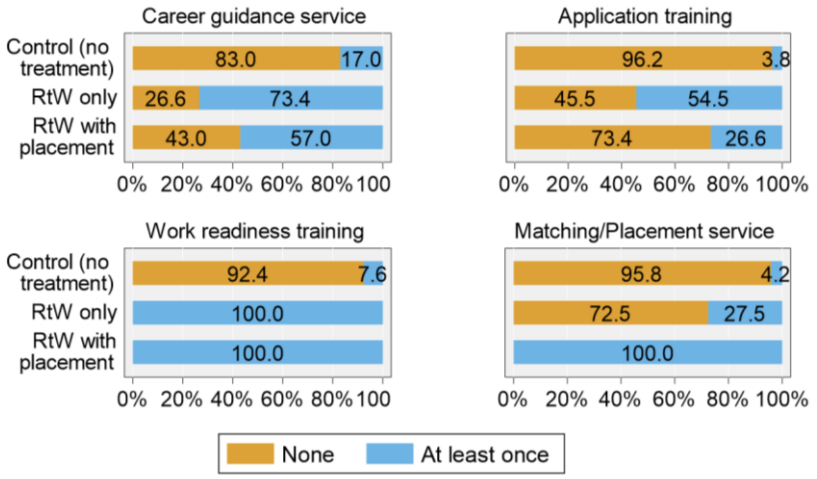


Source: Own calculations based on RtW Survey.

Overall, these numbers illustrate that some control group members have participated in alternative career or employment promoting programs, however, the majority did not. These numbers suggest that the RtW program was not just one of many programs offered to students but that it offered novel and unique opportunities to vocational training graduates.

Figure 4.6.9 provides deeper insights into the types of programs respondents participated in. Of course, all treatment group members attended a work readiness training. But also 8% of the control group did. Most of the control group indicated that they participated in career guidance services, which is a rather broad type of service that may include various activities. Only few control group members indicated participation in an application training (4%) or placement service (4%). At last, quite a few RtW beneficiaries also indicated participation in application trainings, career guidance services and, for those who did not receive a placement through UMA, another placement service. As this was not a single choice question, some RtW beneficiaries may have indicated to have participated in career guidance services and application trainings in addition to the work readiness training as all three may have been components of the RtW training program.

Figure 4.6.9
Type of employment promoting program or activity by treatment group



Source: Own calculations based on RtW Survey.

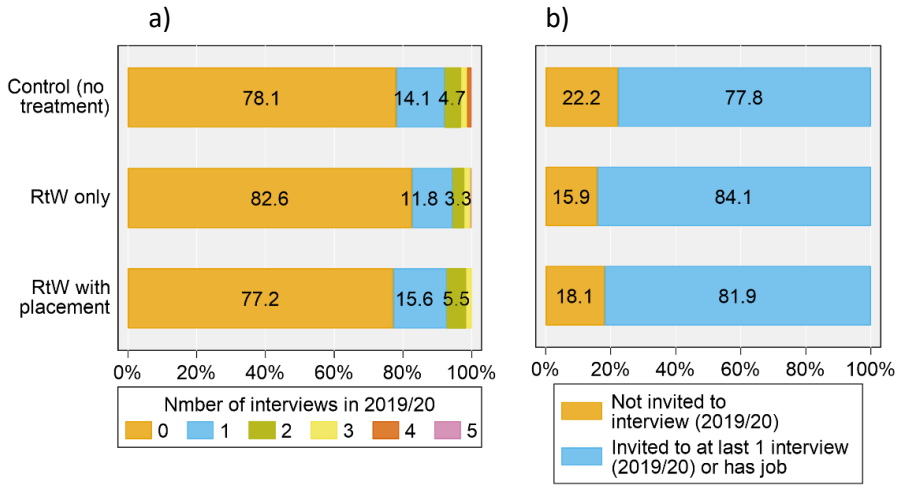
4.7 Descriptive statistics on participants' job search

This section reflects on participants' job search experiences by treatment group, including the number of interview invitations, the number of attended interviews, the success rates of attended interviews, the length of job search, the most successful strategies for job search, and the greatest challenges in searching for a job.

4.7.1 Interviews

Figure 4.7.1 presents descriptive statistics on invitations to job interviews (not including interviews for internships). Figure 4.7.1.a shows that 24%, 17%, and 21% of the RtW placement group, the RtW-training-only group, and the control group had been invited to at least one interview, respectively. Interestingly, relatively more control group respondents were invited to job interviews than respondents in the RtW-training-only group, but not more than in the RtW placement group. Figure 4.7.1.b further shows invitations to job interviews based on an alternative coding of job interviews in which participants who had a job at some point in the year before the survey were coded as "invited to an interview", assuming that people with a job also had an interview. Based on the alternative coding of interview invitations, the pattern from 4.7.1.a reverses and the percentage share of respondents who were invited to at least one interview is largest in the RtW-training-only group (84% versus 78% in the no-intervention control group and 82% in the RtW-with-placement group).

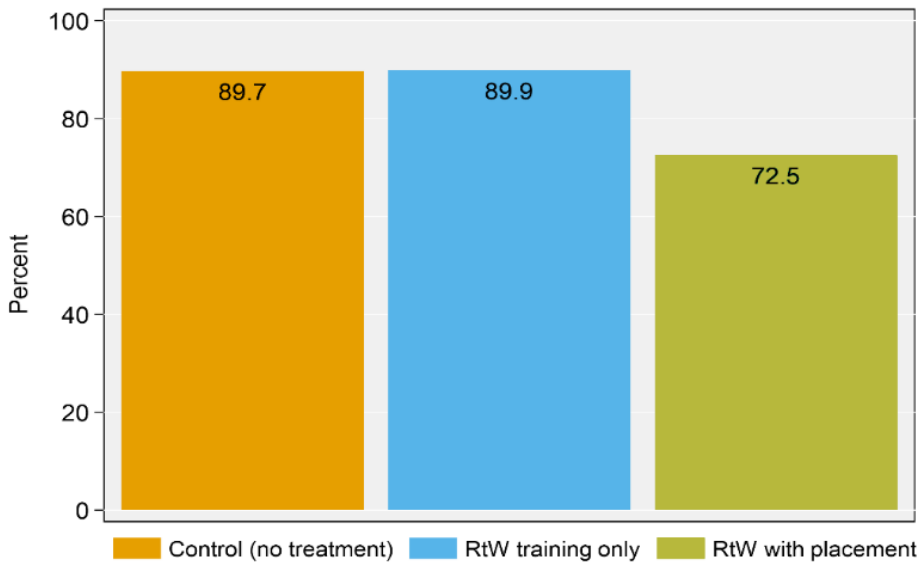
Figure 4.7.1
Number of interview invitations since 2019 by treatment group



Source: Own calculations based on RtW Survey.

Figure 4.7.2 presents the average percentage share of interviews that were attended by the respondents relative to all interviews respondents were invited to. In the control and RtW-training-only group, respondents attended 90% of the interviews on average, whereas respondents that were placed in UMA internships attended only 73% of the interviews they were invited to on average.

Figure 4.7.2
Average percentage share of interview invitations that were attended by treatment group

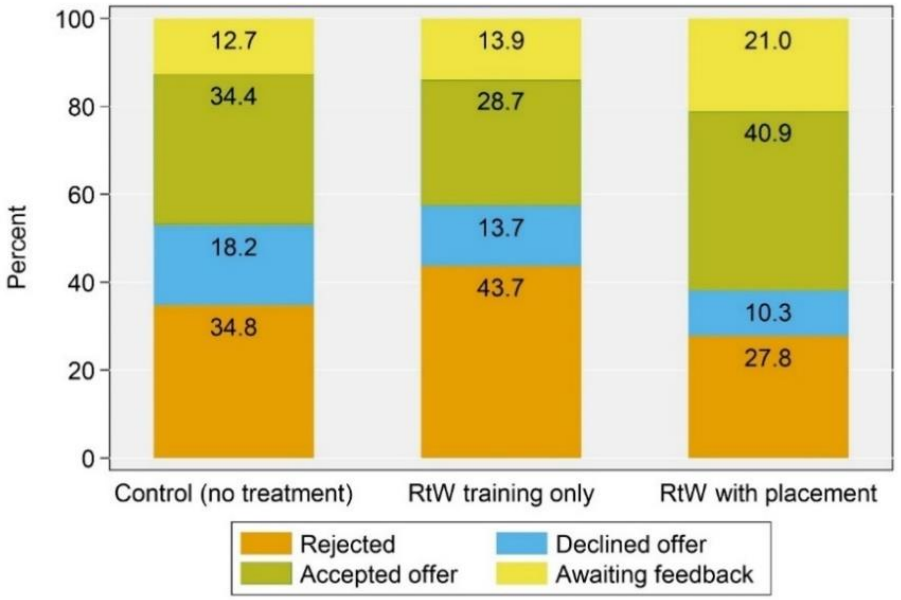


Source: Own calculations based on RtW Survey.

Figure 4.7.3 presents the average percentage shares of attended interviews for which the respondent (i) was rejected, (ii) received an offer and accepted, (iii) received an offer but declined, and (iv) is still awaiting feedback. Relative to the number of attended interviews, the RtW training group received fewest offers (42% versus 53% in the no-intervention control group and 51% in the RtW-with-placement group). As shown in Figure 4.7.1.a, the RtW-training-only group also had the smallest percentage share of respondents who were invited to at least one interview.

Hence, the low rate of offers was not due to the fact that UMA companies had excessively invited candidates for interviews and, therefore, respondents in the RtW-training-only group received a lot of rejections. However, as discussed in section 4.6.2 relatively more respondents in the RtW-training-only group had participated in an internship in the three months prior to the survey in comparison to the no intervention control and RtW-with-placement groups and therefore they may have had less time to prepare for interviews and perform well.

Figure 4.7.3
Average success rates of attended interviews by treatment group



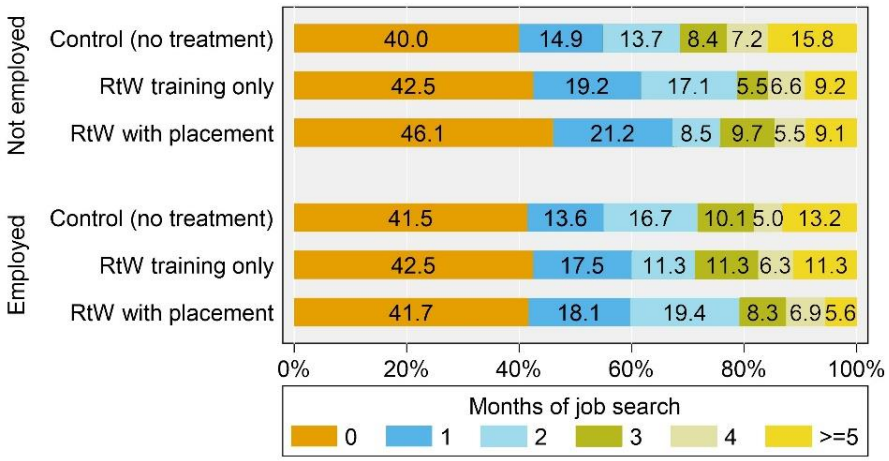
Source: Own calculations based on RtW Survey.

4.7.2 Length of job search

Figure 4.7.4 presents the number of months that participants had searched for a job in the past year by treatment group and employment status. The share of respondents that have searched not at all or less than two weeks (rounded to zero) for a job is similar across treatment groups among employed respondents (around 42%) and largest in the RtW placement group among unemployed respondents (46% versus 40% in the control group and 43% in the RtW-with-placement group).

Among respondents placed by UMA who were employed, 46% did not search for a job, and 48.7% of those unemployed did not search for a job. Among beneficiaries who participated in the RtW training only and among control group members about 40% had not searched for a job in the past year.

Figure 4.7.4
Number of months participants had searched for a job in the past year by treatment group and employment status



Source: Own calculations based on RtW Survey.

The average number of months of job search among respondents who searched at all was 2.9 months. Respondents in the RtW with placement and RtW-training-only groups searched 2.5 and 2.6 months on average, respectively, and respondents in the no-intervention control group searched 3.3 months on average. Table 4.7.1 presents the length of job search in months for each treatment group by employment status.

Table 4.7.1
Average number of months of job search among those that searched at all

	RtW with placement	RtW training only	Control (no treatment)
Employed	2.1 months	2.6 months	3.1 months
Not employed	2.7 months	2.6 months	3.4 months

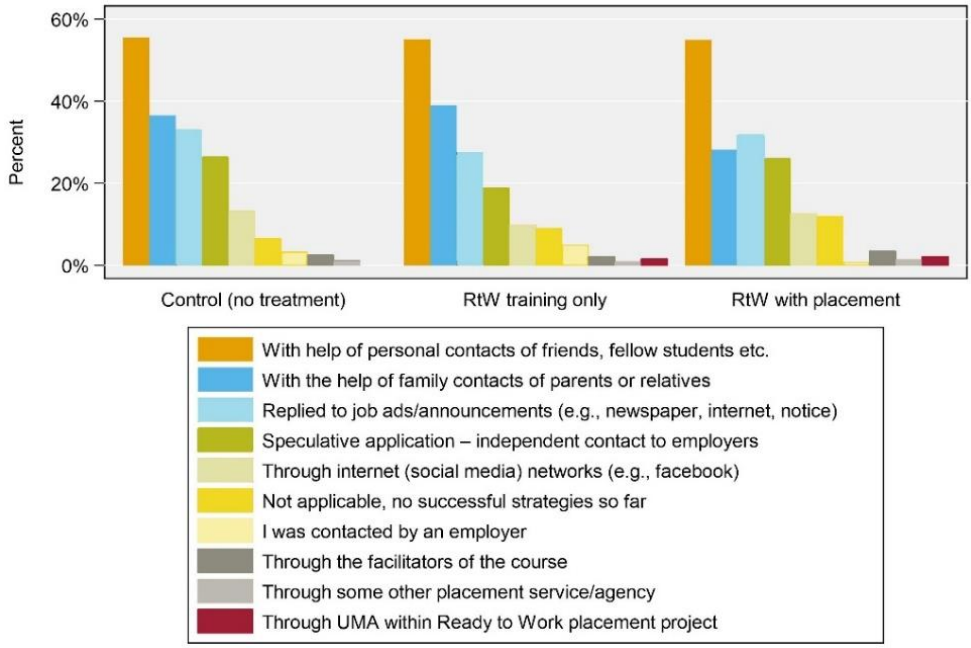
Source: Own calculations based on RtW Survey.

4.7.3 Strategies of job search

Figures 4.7.5 and 4.7.6 describe the most successful strategies and the main challenges of participants' job search, respectively. Across treatment groups the percentage shares of respondents who indicated specific successful strategies and challenges are quite similar.

Figure 4.7.5 shows that for most participants the most successful strategy of job search was with the help of personal contacts of friends or fellow students (55%), followed by family contacts of parents or relatives (36%), replying to a job announcement in the newspaper, internet or notice board (31%), applying to an employer speculatively (24%), and through social media networks (12%) (see Figure 4.7.5). Other strategies were mentioned by less than 10% of participants. Noticeably, only very few RtW beneficiaries (1.5% and 2.1% in the RtW training only and RtW with placement, respectively) felt that job search facilitated through UMA within the RtW project was a successful job search method.

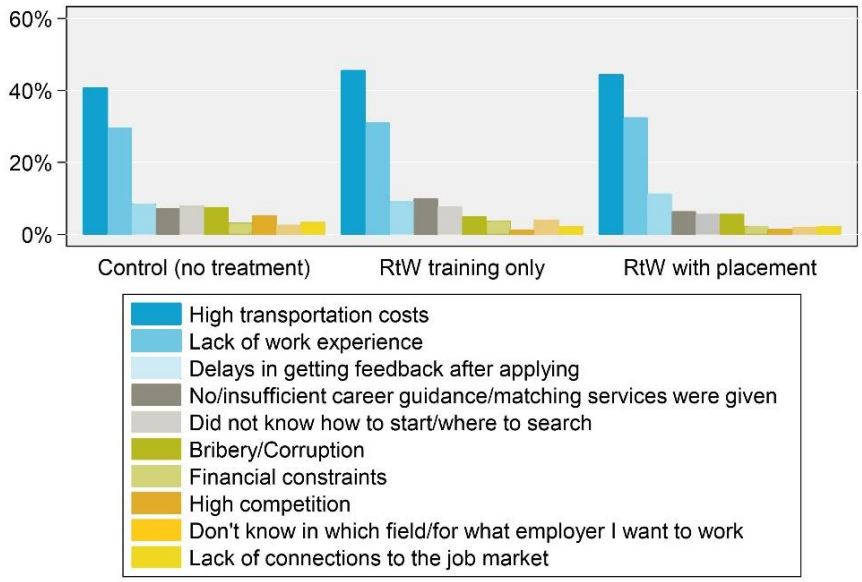
Figure 4.7.5
Most successful strategies of job search by treatment group



Source: Own calculations based on RtW Survey.

Figure 4.7.6 shows that the biggest challenges in job search were high transportation costs (43%, another 3% mentioned financial constraints) and lack of working experience (31%). Other challenges were mentioned by less than 10% of participants. Interestingly, 8.1% of participants mentioned that there was insufficient support through career guidance or matching services, however, there was some heterogeneity across groups. 10% of RtW training only beneficiaries felt that career guidance support was lacking, whereas only 6% and 7% of RtW placement beneficiaries and control group respondents, respectively, thought this was a challenging factor.

Figure 4.7.6
Main challenges of job search by treatment group



Source: Own calculations based on RtW Survey.

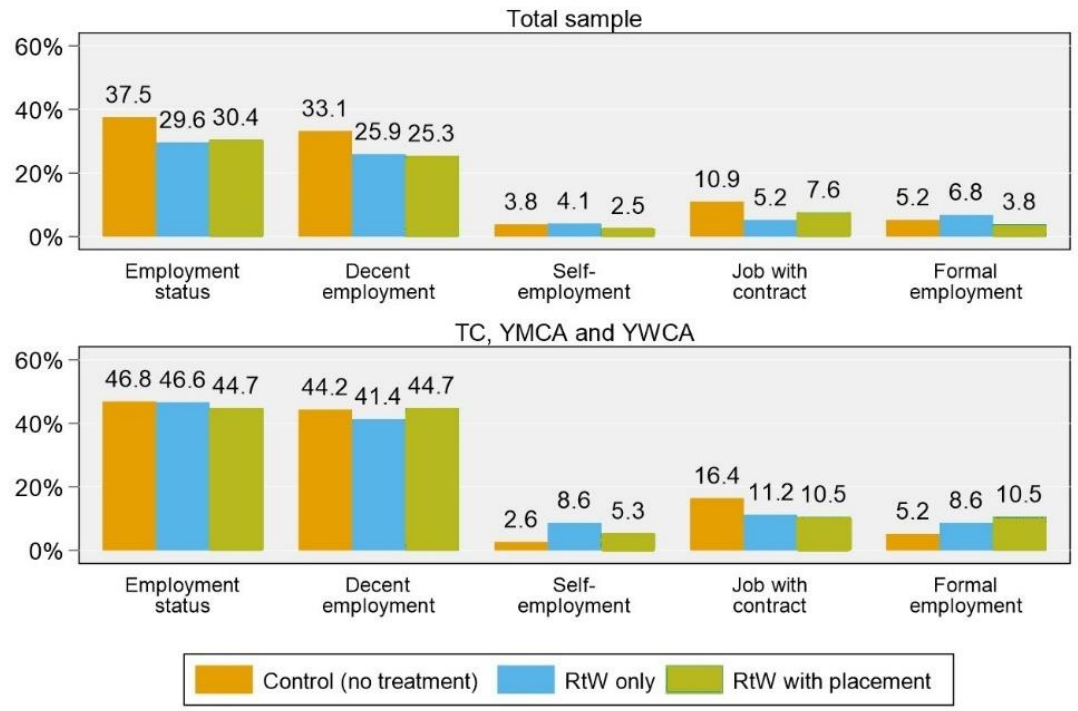
4.8 Descriptive statistics on labor market outcomes

This section describes the labor market outcomes of participants at the time of the survey in February 2020. Because the estimation of treatment effects is based on two samples – one with all participants included and one with only graduates from TC, YMCA, and YWCA included – the descriptive analysis of labor market outcomes also distinguishes between these two samples.

4.8.1 Labor market characteristics in February 2020

In the total sample (i.e., including graduates from all schools), respondents in the no-intervention control group were relatively more often employed than RtW beneficiaries (i.e., who were trained only and who were trained and placed). Figure 4.8.1 shows that 38% of respondents in the control group had any paid work, whereas only 30% of the RtW beneficiaries did. Further, 33% of control group members were in decent employment whereas only 26% of beneficiaries were in decent employment. Only few participants were self-employed (3-4%) or had formal jobs (4-7%). Jobs with a contract were more common among no-intervention control group respondents (11%) than among RtW beneficiaries (5% in the RtW training only and 8% in the RtW-with-placement groups, respectively).

Figure 4.8.1
Employment status at the time of survey in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA



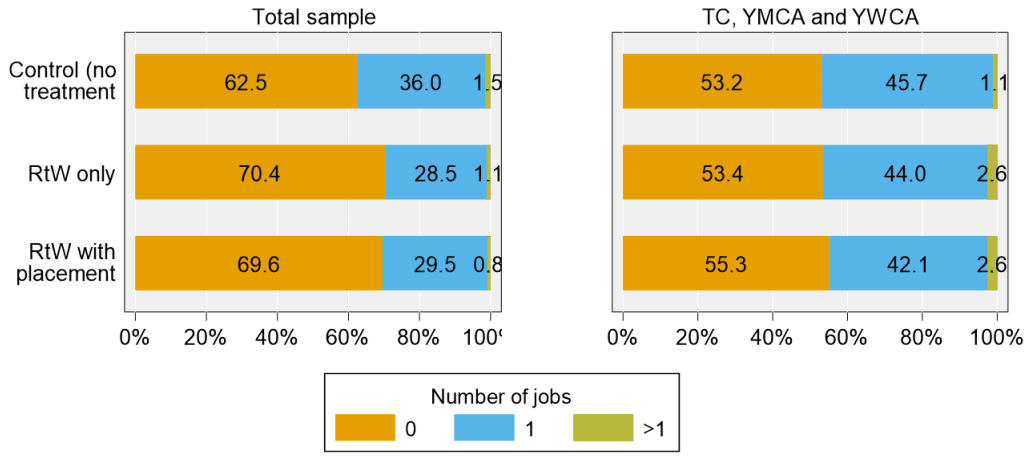
Source: Own calculations based on RtW Survey.

In the sample of respondents who graduated in May 2019 (i.e., only graduates from TCs, YMCA, and YWCA), the employment rates were overall higher than in the total sample and similar across treatment groups. Slightly fewer respondents in the RtW-with-placement group had any paid work (45% versus 47% in the two beneficiary groups), however, respondents in the RtW-with-placement group perform slightly better with respect to decent employment (45%) than those in the no-intervention control group (44%) and those in the RtW-training-only group (41%).

The subsample analysis of May 2019 graduates is based on a small number of observations in the RtW-with-placement group. There are 38 respondents in the RtW-with-placement group who graduated in May 2019 and 17 of those were in decent employment, whereas in the subsample of respondents in the no-intervention control group and RtW-training-only group, 126 and 54 respondents were in decent employment, respectively.

Figure 4.8.2 shows that, when taking into account all paid jobs (referred to as employment status in Figure 4.8.2), very few of the employed respondents had more than one job.

Figure 4.8.2
Number of jobs by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA



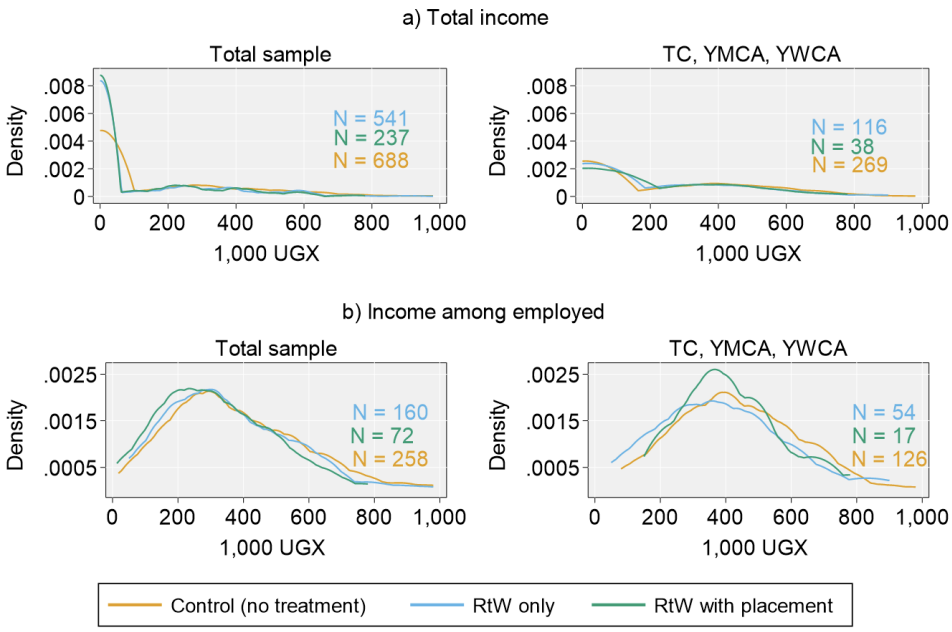
Source: Own calculations based on RtW Survey.

Figure 4.8.3 describes the distribution of monthly income from any paid work at the time of the survey in February 2020 in a sample that includes employed and unemployed respondents (Figure 4.8.3.a) and in a sample that only includes employed respondents (Figure 4.8.3.b). Each Subfigure (4.8.3.a and 4.8.3.b) presents these statistics for the total sample and the sample of respondents with graduation dates in May.

Because many participants did not have a job in February 2020, total income is skewed towards zero (Figure 4.8.3.a). In alignment with the employment statistics in Figure 4.8.1, income in the total sample (4.8.3.a left) is more often zero for the two RtW beneficiary groups than for the no-intervention control group. However, this is not the case for the subsample of participants who graduated from TCs, YMCA, or YWCA, which also confirms the previous observation based on Figure 4.8.1: it is particularly graduates from VTIs in the two RtW beneficiary groups who do not have a job in February 2020.

Considering employed participants, Figure 4.8.3.b shows that incomes were varying quite considerably across participants and in similar ways across treatment groups. Incomes ranged from less than 50,000 UGX to over 900,000 UGX per month. In the total sample, roughly 150,000 to 400,000 UGX were the most common incomes. Among graduates from TCs, YMCA, and YWCA, total incomes were considerably higher, most commonly ranging between 250,000 and 550,000 UGX.

Figure 4.8.3
Monthly income and income among employed (in UGX) at the time of the survey in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA

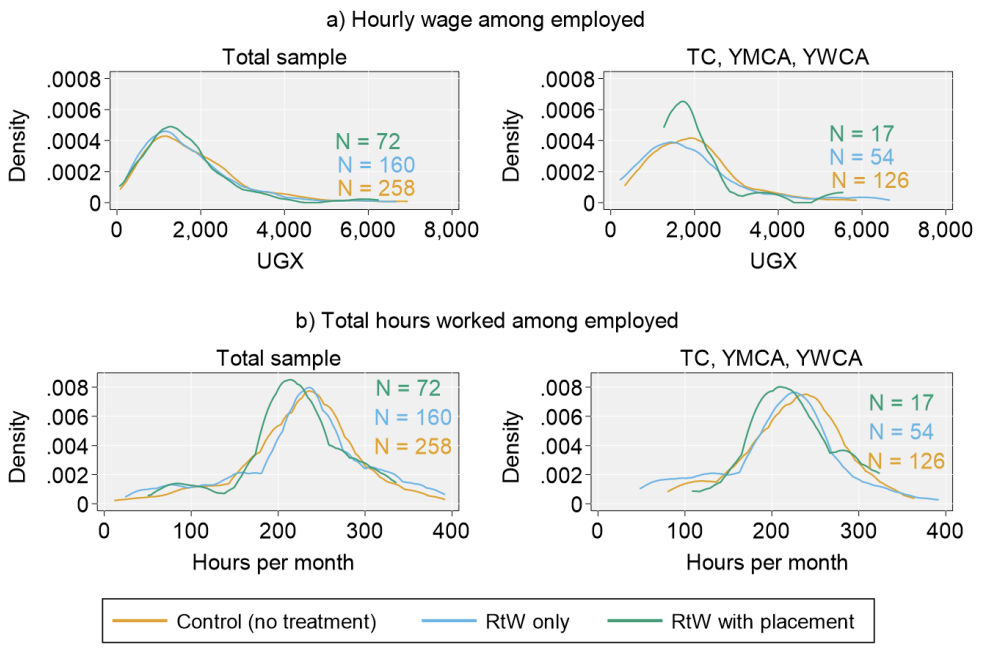


Source: Own calculations based on RtW Survey.

Monthly income was calculated based on the total number of hours worked in all jobs per month and the average hourly wage. Figure 4.8.4.a shows the distribution of hourly wages among employed, which varies between less than 500 UGX to over 6,000 UGX, with most wages ranging between 500 and 2,500 UGX per hour. The number of monthly hours worked (Figure 4.8.4.b) also varies widely between just a few hours up to almost 400 hours. However, most employed participants worked between 180 and 300 hours per month. Defining full-time work as 40 hours per week, these graphs show that a large share of participants (86% of those employed) worked much more than full-time, and a considerable share (40% of those employed) worked even one-and-a-half times the hours typically worked in a full-time job. The monthly working hours indicator considers all paid jobs jointly, but, as evident from Figure 4.8.2, most employed respondents had only one job.

The distribution of hourly wages among employed respondents is relatively similar across treatment groups in the total sample (see Figure 4.8.4.a left). With respect to working hours, respondents in the RtW-with-placement group worked fewer hours overall compared to the other two groups, as indicated by the distribution of working hours being located to the left of the distributions of the no intervention control and RtW-training-only groups. For the subsample of respondents who graduated in May, similar patterns for working hours to those in the total sample can be observed. However, respondents in the RtW-with-placement group who are employed are so few that the distributions of hourly wages and working hours should be interpreted with caution.

Figure 4.8.4
Hourly wage and total hours worked per month among employed at the time of the survey in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA



Source: Own calculations based on RtW Survey.

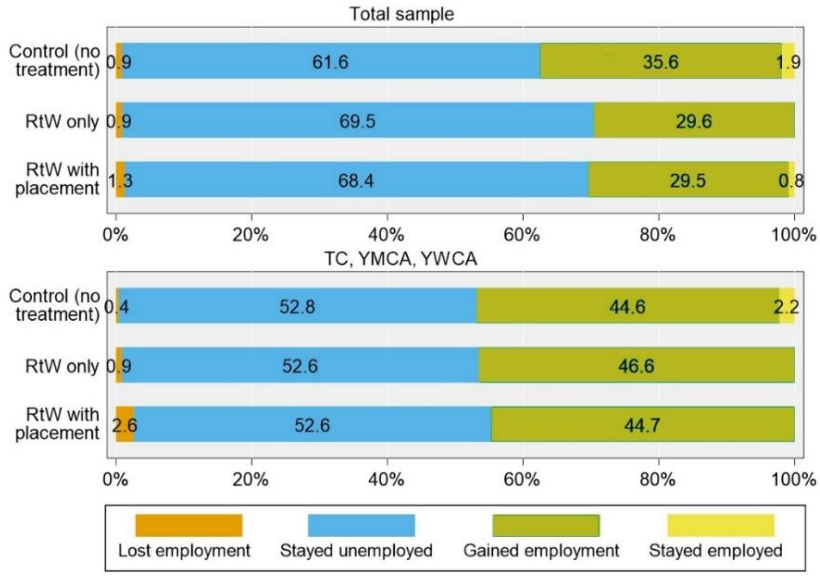
4.8.2 Changes in labor market characteristics between baseline and follow-up

At the time of data collection in February 2020, information on jobs held and incomes earned in the previous years were collected. Based on this recall of labor market outcomes, a pre-intervention point in time was defined as one year before the survey, i.e., in February 2019, when all participants were still attending their vocational training institute or technical college. We use this pre-intervention point in time as a reference to compare labor market outcomes after the intervention with those before the intervention.

4.8.2.1 Changes in employment status

Figure 4.8.5 shows job transitions over time by comparing paid work post-intervention with paid work pre-intervention across treatment groups and samples. For that purpose, respondents are allocated into four distinct groups. Respondents who were employed in February 2019 but were not employed in February 2020 are considered to have “lost employment”. Those who did not have a job at either time are referred to as “stayed unemployed”. Participants who did not have a job in February 2019 but did have one in February 2020 are considered to have “gained employment”. Respondents who had a job at both times are referred to as “stayed employed”.

Figure 4.8.5
Change in employment status between February 2019 and February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA



Source: Own calculations based on RtW Survey.

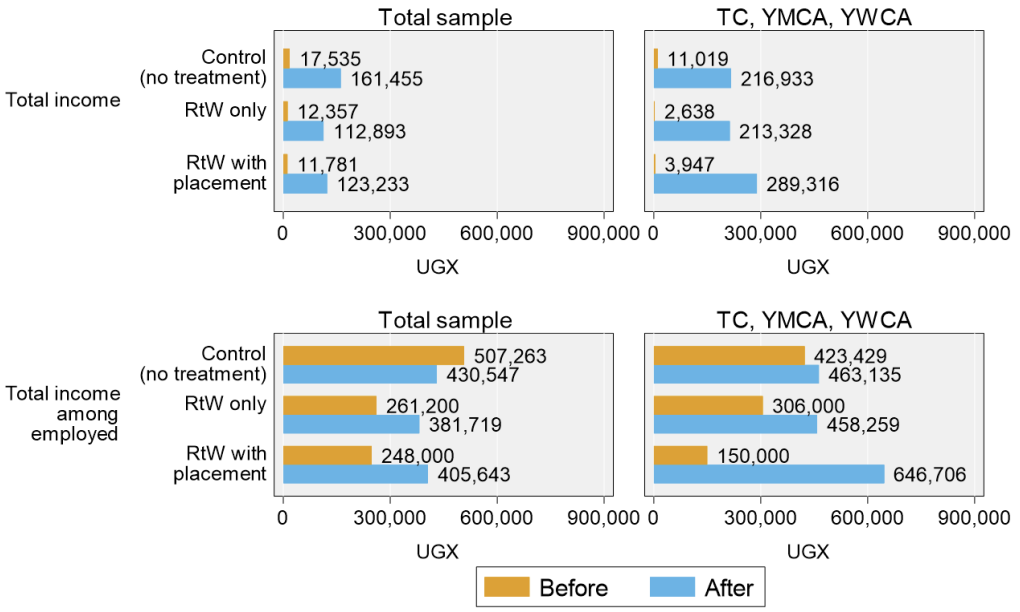
In the total sample, about two-thirds of participants stayed unemployed (62-70%) and about one-third gained employment (30-36%), whereas very few participants stayed employed (0-2%) or lost employment (1%). In the subsample of graduates from TCs, YMCAs, and YWCAs, the percentage share of participants who gained employment is considerably larger (45-47%) than in the total sample. The percentage shares of those who stayed employed (0-2%) and those that lost employment (0-3%) were similarly low.

Across treatment groups in the total sample, the no-intervention control group had a 6-percentage points higher rate of gaining employment and a slightly higher rate of staying employed in comparison to the two RtW beneficiary groups. In the sample of May-graduates, the only difference across treatment groups is that a few more respondents of the no-intervention control group stayed employed and, thus, fewer lost employment. However, as described above, overall, very few participants stayed employed or lost employment in both samples.

4.8.2.2 Changes in income

Figure 4.8.6 describes how participants' mean monthly income from all paid jobs changed from one year before the survey, in February 2019, to the time of the survey in February 2020. The top of Figure 4.8.6 refers to the mean monthly income among all respondents (employed or unemployed) and shows a considerable rise in incomes from February 2019 to February 2020 which mirrors the rise in employment rates that occurred to respondents in all three treatment groups (see Figure 4.8.5).

Figure 4.8.6
Mean monthly income and mean monthly income among employed before the RtW program in February 2019 and after the RtW program in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA



Source: Own calculations based on RtW Survey.

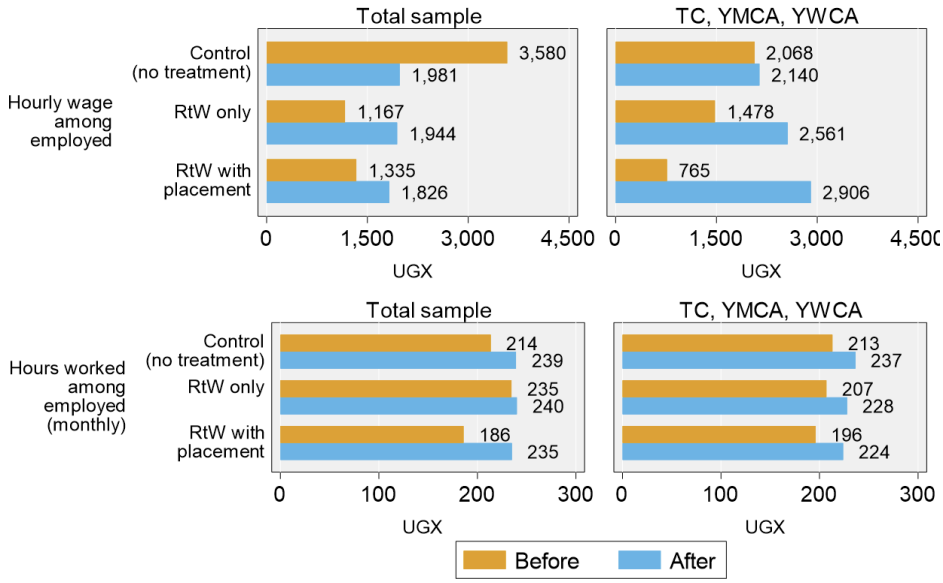
In the total sample, the no-intervention control group experienced the largest increase in mean monthly income (143,920 UGX = 161,455 UGX - 17,535 UGX), followed by the RtW with placement (111,452 UGX) and RtW-training-only groups (100,536 UGX). In contrast, in the subsample of participants who graduated in May, respondents in the RtW-with-placement group experienced by far the largest increase in mean monthly income (285,369 UGX versus 210,690 UGX and 205,914 UGX in the RtW training only and no-intervention control group, respectively).

The bottom of Figure 4.8.6 presents the mean monthly income among respondents that were employed in February 2019 ("before") and in February 2020 ("after"). The mean monthly income among employed in the total sample was by 120,519 to 157,643 UGX larger in 2020 than in 2019, except for the no-intervention control group which earned 103,716 UGX less per month in February 2020 than a year before. In the sub-sample of respondents who graduated in May, the mean monthly income among employed was larger in 2020 than in 2019 in all treatment groups by 39,706 to 496,706 UGX. In both samples, the RtW-with-placement group experienced the highest increase in mean monthly income among employed (157,643-496,706 UGX) and the no-intervention control group experienced the lowest increase in monthly income among employed; in fact, the control group experienced even a negative change in the total sample (between -103,716 and up to -39,706 UGX). For the interpretation of these statistics, it is important to note that the respondents who were employed in 2020 may not be the same who were employed in 2019 and that the sample of respondents that were employed in 2019 is very small (N=21).

4.8.2.3 Changes in hourly wages and working hours

In order to understand how total income and, particularly, income among employed changed from before the RtW program in February 2019 to after the RtW program in February 2020, it is important to understand how the average hourly wage among employed and hours worked per month among employed had changed. Figure 4.8.7 shows that, although hourly wages have changed considerably, the mean working hours among employed increased moderately by 15 to 39 hours across treatment groups. Hence, the changes in mean incomes among employed displayed in Figure 4.8.6 are mostly driven by changes in hourly wages and less by changes in hours worked.

Figure 4.8.7
Mean hourly wage and hours worked per month among employed before the RtW program in February 2019 and after the RtW program in February 2020 by treatment group in the total sample and the subsample of TCs, YMCA, and YWCA



Source: Own calculations based on RtW Survey.

The top of Figure 4.8.7 further shows a reduction in the mean hourly wage for respondents in the no-intervention control group of the total sample. However, the mean hourly wage in 2019 is based on a small number of employed respondents and, therefore, sensitive to outliers.

4.9 Estimated program effects on labor market outcomes

This section presents the impact estimations of the RtW program on labor market outcomes based on the estimation methods discussed in section 4.3.1. Specifically, it presents the estimated impact of having participated in both the RtW training and UMA placement in comparison to a) having participated in no intervention (RQ#1), and b) having participated in the RtW training only (RQ#2).

The figures in this section show estimated treatment effects, also referred to as coefficients or point estimates, as well as 95% confidence intervals of the estimated effects. Confidence intervals are a measure of precision of the effect estimates and provide a range of effect estimates that the true parameter, i.e., the true program effect, might be. Here, we use 95% intervals, which means that the estimated confidence interval will contain the true value of the treatment effect with a probability of 95%. If a confidence interval does not include zero, the estimated

treatment effect is said to be significantly different from zero and, thus, considered to have a significant positive (or negative) effect on the respective outcome.

All presented employment measures are binary indicators which only have two categories or levels. Effect estimations on employment outcomes can be interpreted as *percentage point* changes in the respective employment indicator. For example, a coefficient of 0.1 can be interpreted as a 10-percentage point increase in the employment rate due to the intervention, i.e., RtW training and placement (RQ#1) or RtW placement in addition to the training (RQ#2).

Income and wage outcomes were measured in Ugandan shilling (UGX) and logarithmized using the natural logarithm (referred to as "Ln of [income indicator]" in the figures that follow). Using logarithms of income and wage variables has the advantage that the estimated program impacts can be interpreted as *percent changes* and compared across different contexts and currencies. An estimated effect of 0.1, for example, indicates a 10% increase in income due to the intervention, i.e., RtW training and placement (RQ#1) or RtW placement in addition to the training (RQ#2). For effects greater than 0.3 (or smaller than –0.3), these effects cannot be directly interpreted as percentage changes (although they are still a close approximation) but require conversions that are detailed in the respective figure notes, if applicable.[55]

Working hours are measured as continuous variables and their coefficients can be interpreted as a change in hours worked per month due to the intervention, i.e., RtW training and placement (RQ#1) or RtW placement in addition to the training (RQ#2).

For each outcome, we present *four* effect estimates, which results from the two samples we consider – one with all participants included and one with respondents who graduated in May from TC, YMCA, and YWCA included – and the two comparisons we make. One comparison corresponds to RQ#1, i.e., RtW training and placement vs. no intervention control, and the other corresponds to RQ#2, i.e., RtW training and placement vs. RtW training only.

4.9.1 Program impact on employment status

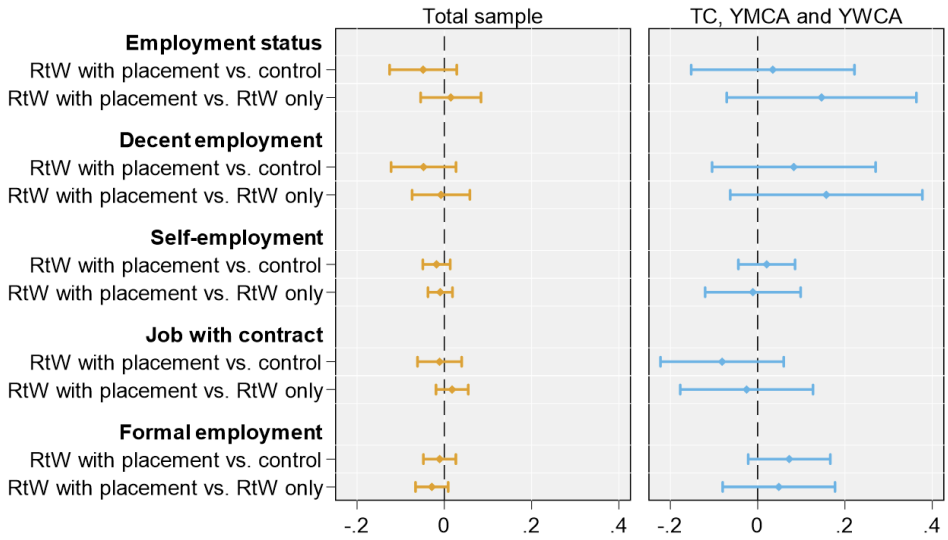
Figure 4.9.1 shows the estimated impacts of the RtW training and placement on (i) having paid work (employment status), (ii) having decent employment, (iii) being self-employed, (iv) having a job with a contract, and (v) being in formal employment for the total sample and the subsample of participants who graduated in May 2019 (TC, YMCA and YWCA).

In the total sample, the estimated treatment effects on employment are close to zero and all 95% confidence intervals include zero. Thus, no significant positive or negative effect of the RtW training and placement program was measured. Considering the sample of respondents who graduated in May, the point estimates tend to be larger than in the whole sample, but the confidence intervals also include zero. The confidence intervals include a larger range of effect estimates in the subsample of May-graduates because the sample is smaller, which reduces the precision of effect estimates and thus increases the confidence intervals.

Overall, the results in Figure 4.9.1 suggest that the RtW program (training plus placement) had no impact on employment outcomes in the short time period that was observed.

[55] *The conversion follows the following formular:* $(e^{coef} - 1) * 100 = program\ effect.$

Figure 4.9.1
Estimated treatment effects on binary employment indicators

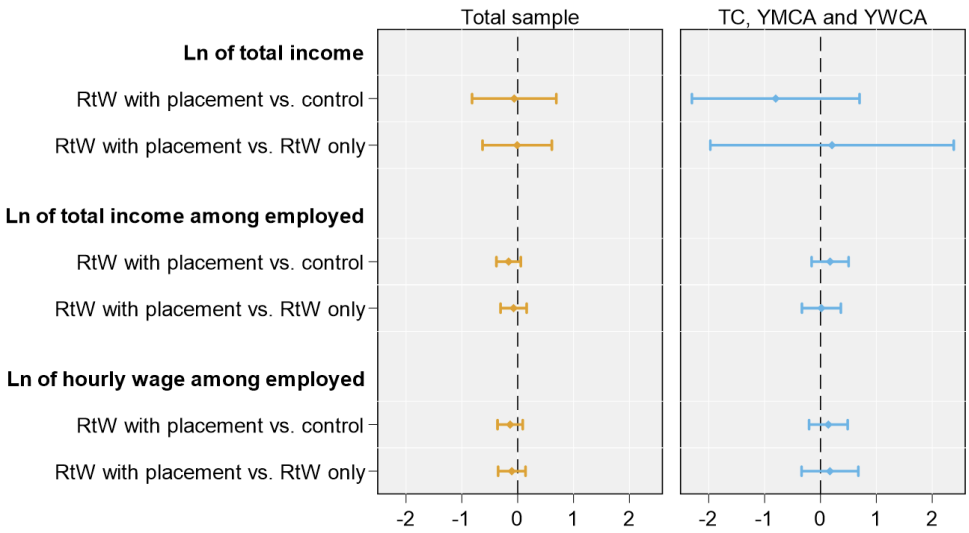


Source: Own calculations based on RtW Survey.

4.9.2 Program impact on incomes and wages

Figure 4.9.2 presents the effect estimates on income earned from all jobs among all respondents and among employed respondents as well as the average hourly wage from all jobs among employed respondents for the total sample and the subsample of participants who graduated in May 2019 (TC, YMCA and YWCA).

Figure 4.9.2
Estimated treatment effects on total income, income among employed and wages among employed. Income and wages were included as the natural logarithm of the respective income and wage measure

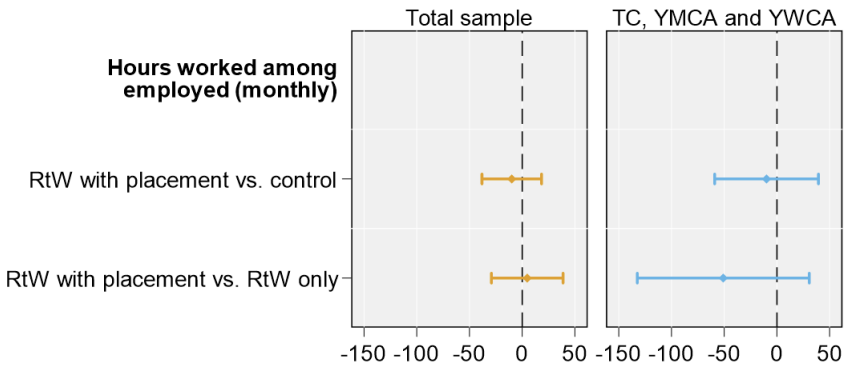


*Notes: The coefficient for ln of total income on "RtW with placement vs. control" for the TC, YMCA, and YWCA sample translates to an income decrease of (exp(-0.801) - 1)*100 = 55%. - Source: Own calculations based on RtW Survey.*

For both, the total sample and the sub-sample of May-graduates, the estimated effect sizes are small and the confidence intervals include zero, suggesting no impact of the RtW program (training plus placement) on incomes and wages.

Figure 4.9.3 presents the estimated program impact on working hours among employed respondents for the total sample and the subsample of participants who graduated in May 2019 (TC, YMCA and YWCA). The treatment effects are overall small, except for the estimate on placements in addition to the RtW training in the subsample of May-graduates, which is negative and large (-50). However, all confidence intervals include zero, suggesting that the RtW program (training plus placement) did not significantly affect working hours of those that were employed.

Figure 4.9.3
Estimated treatment effects on monthly working hours among employed respondents



Source: Own calculations based on RtW Survey.

4.9.3 Robustness checks

Overall, the results presented in this section suggest that the RtW training and placement program had no impact on employment and labor market outcomes on average by February 2020. In order to alleviate potential concerns about the survey measuring outcome indicators too shortly after the program roll-out when potential positive program impacts may not have materialized yet, the estimations in sections 4.9.1 and 4.9.2 already considered the subsample of respondents who had graduated in May 2019, about six months prior to the other respondents. The results of the subsample analysis with May-graduates may suggest two things: first, the additional six months were still not sufficient for potential positive impacts to unfold and even longer-term impact assessments are required, or, second, there were no positive program impacts overall. However, these are speculative considerations, which cannot be answered given the set-up of this study.

To further check the robustness of the results presented in sections 4.9.1 and 4.9.2, an alternative estimation method based on a Double Machine Learning approach (outlined in section 4.3.1) was used to estimate the impact of the RtW training and placement program on the same employment and income indicators. The results based on the Double Machine Learning approach are presented in Appendix A3.2 and they confirm the results presented in Figures 4.9.1 to 4.9.3.

4.10 Program impact for sub-groups of RtW respondents

The estimated treatment effects discussed in section 10 present average treatment effects for the whole study population (or those graduated in May). However, the treatment effects for specific sub-populations may differ from those for the whole study population. In this section, the sample is divided into sub-samples of participants with specific background characteristics and the impact of the RtW training and placement program is estimated for those sub-populations. The background characteristics under consideration are:

1. respondents aged less than **25 years** versus respondents **25 years and older**,
2. respondents who had sustained previous work experience, measured as **having worked for pay in one job for six months**, versus respondents who do not have significant prior work experience,
3. respondents who have any **previous work experience in the construction sector** (could be a paid or unpaid job or internship) versus respondents that do not have construction sector experience,
4. respondents who have any **previous work experience in mechanical or electrical work** (could be a paid or unpaid job or internship) versus respondents that do not have experience in mechanical or electrical work, and
5. respondents who graduated in a **diploma level course versus** respondents who graduated in a **certificate level course**. [56]

These sub-samples are constructed from the total sample, including respondents who graduated in May and who graduated in November, because the sub-sample sizes would be too small if they were constructed from the sample of respondents who graduated in May only.

The figures in this section will only include treatment effects on employment indicators. The results for incomes, wages, and working hours are presented in Appendix A3.3.

4.10.1 Program effect heterogeneity by respondents' age

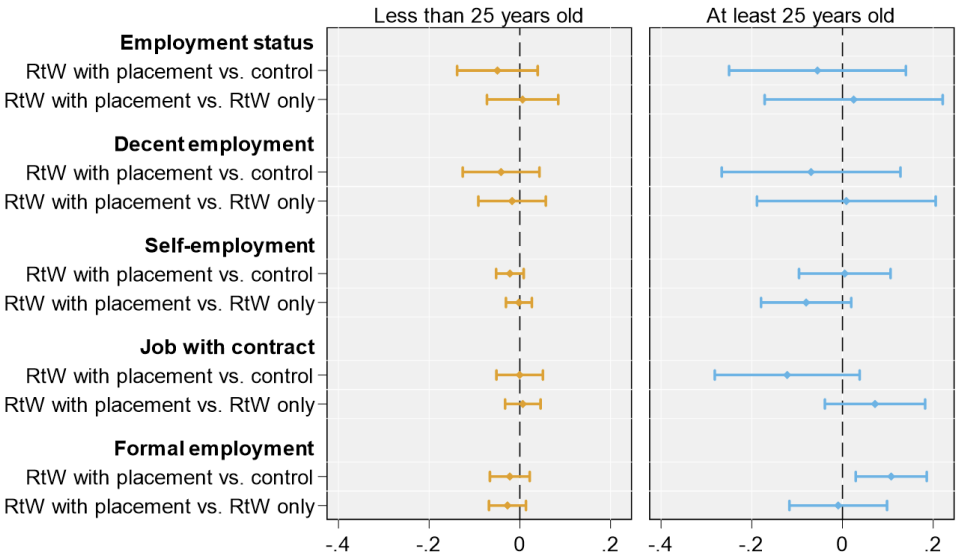
Figure 4.10.1 presents the RtW program impact on employment in a sub-sample of respondents who are younger than 25 years old on the left-hand side and in a sub-sample of respondents who are 25 years or older on the right-hand side. The coefficients tend to be small and all but one confidence interval include zero, suggesting no differential impact of the RtW program by respondents' age.

In the sub-sample of older participants, the effect of the RtW training and placement in comparison to the no-intervention control group on formal employment is significantly positive. However, the effect on the RtW training and placement in comparison to the RtW training only, does not show a similar effect.

Appendix A3.3 presents the program effects by respondents' age on income, wages, and working hours and suggests no significant treatment effect for either age group.

[56] *Further potentially interesting effect heterogeneities exist. For example, gender was not considered because the sample of female participants was too small.*

Figure 4.10.1
Estimated treatment effects on employment by respondents' age



Source: Own calculations based on RtW Survey.

4.10.2 Program effect heterogeneity by respondents' prior sustained work experience

Figure 4.10.2 presents the RtW program impact on employment in a sub-sample of respondents who had worked six months in one job prior to the study on the left-hand side and in a sub-sample of respondents who had not previously worked six months in one job on the right-hand side. The coefficients tend to be small and the confidence intervals include zero, suggesting no impact of the RtW program for respondents with and without sustained work experience.

Appendix A3.3 presents the program effects by respondents' sustained work experience on income, wages and working hours and shows no significant treatment effects for respondents with and without sustained work experience.

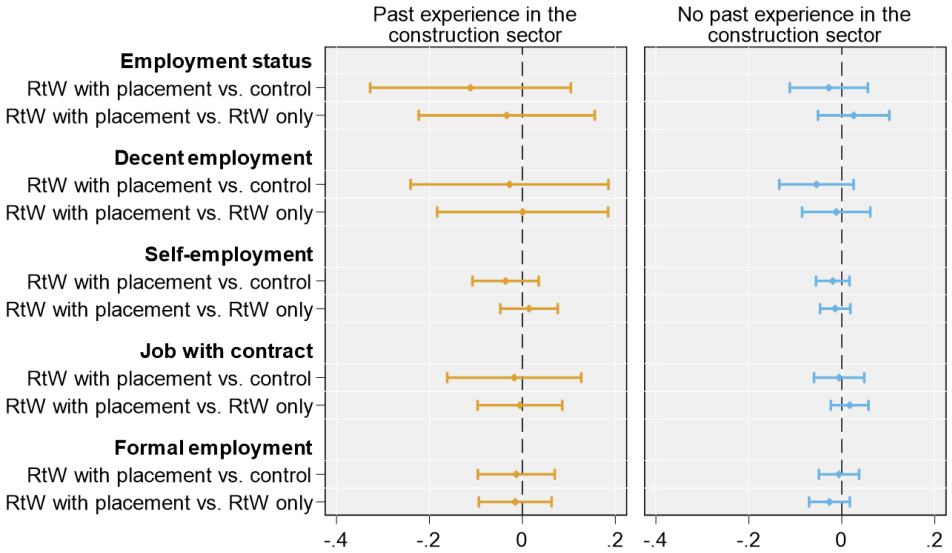
Figure 4.10.2
Estimated treatment effects on employment by respondents' prior work experience

Source: Own calculations based on RtW Survey.

4.10.3 Program effect heterogeneity by respondents' construction sector experience

Figure 4.10.3 presents the RtW program impact on employment in a sub-sample of respondents who had prior work experience in the construction sector on the left-hand side and in a sub-sample of respondents who did not have prior work experience in the construction sector on the right-hand side. The coefficients tend to be small and the confidence intervals include zero, suggesting no impact of the RtW program for respondents with and without experience in the construction sector.

Figure 4.10.3
Estimated treatment effects on employment by respondents' prior construction sector experience



Source: Own calculations based on RtW Survey.

Appendix A3.3 presents the program effects by respondents' construction sector experience on income, wages and working hours and shows no impact of the RtW program for respondents with and without experience in the construction sector.

4.10.4 Program effect heterogeneity by respondents' electrical or mechanical sector experience

Figure 4.10.4 presents the RtW program impact on employment in a sub-sample of respondents who had prior experience in electrical or mechanical works on the left-hand side and in a sub-sample of respondents who did not have prior experience in electrical or mechanical work on the right-hand side. The coefficients tend to be small and the confidence intervals include zero, suggesting no impact of the RtW program for respondents with and without experience in electrical or mechanical work.

Appendix A3.3 presents the program effects by respondents' prior experience in electrical or mechanical work on income, wages and working hours and shows no impact of the RtW program for respondents with and without experience in electrical or mechanical work.

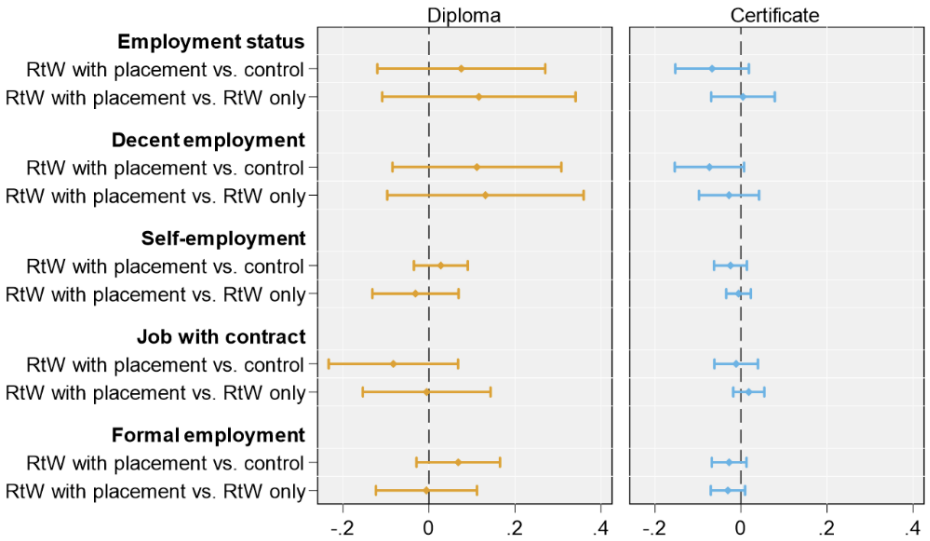
Figure 4.10.4
Estimated treatment effects on employment by respondents' prior experience in electrical and mechanical work

Source: Own calculations based on RtW Survey.

4.10.5 Program effect heterogeneity by respondents' vocational training level

Figure 4.10.5 presents the RtW program impact on employment in a sub-sample of respondents who graduated in a diploma level vocational training course on the left-hand side and in a sub-sample of respondents who graduated in a certificate level vocational training course on the right-hand side. For respondents with a diploma level degree, the coefficients on employment status and decent employment are positive and sizable for both comparisons, the comparison to the no-intervention control group and to the RtW-training-only group.

Figure 4.10.5
Estimated treatment effects on employment by respondents' vocational training level



Source: Own calculations based on RtW Survey.

However, the effects are not statistically significant and therefore, it cannot be concluded that the RtW training and placement program had a positive effect on employment for diploma level students. The remaining coefficients tend to be small and the confidence intervals include zero.

Appendix A3.3 presents the program effects by respondents' vocational training level on income, wages and working hours and shows no impact of the RtW program for respondents with diploma or certificate level vocational training.

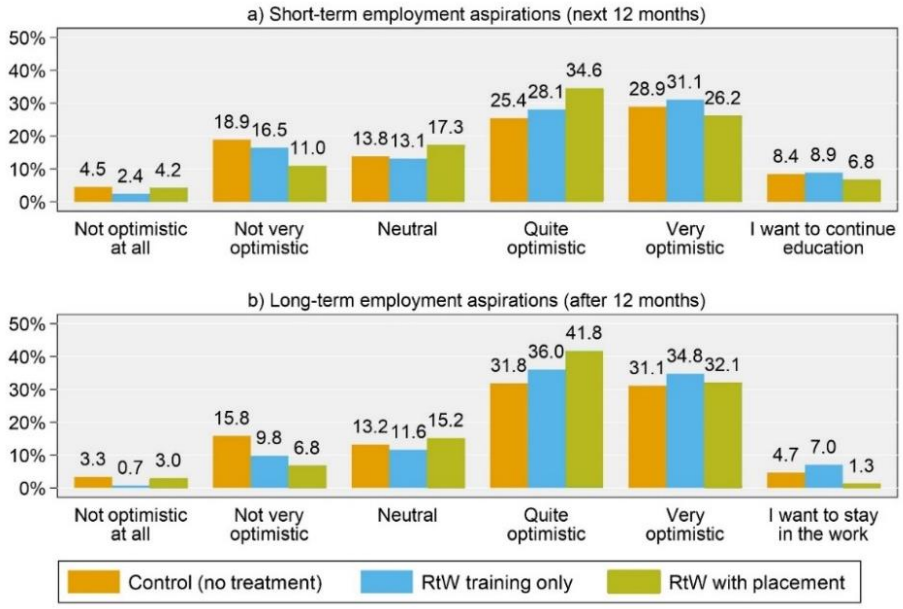
4.11 Program impact on secondary outcomes

Besides employment and earnings, the RtW program may have affected participants' lives in multiple other ways. In this section, we examine whether the RtW training and placement program influenced participants' aspirations with respect to future labor market outcomes as well as their migration intentions.

4.11.1 Employment and earnings aspirations

Figure 4.11.1.a describes respondents' confidence in having work in the next 12 months by treatment group. Overall, 57% of respondents are quite or very optimistic about having a job in the next year and 20% are not very optimistic or not optimistic at all. RtW beneficiaries tend to be more optimistic than control group members. In both RtW beneficiary groups 60% of respondents are optimistic to have a job, whereas in the no-intervention control group 54% are optimistic to have a job. Accordingly, 24% of the no-intervention control group is not optimistic, whereas only 19 and 15% of the RtW training only and the RtW-with-placement group are not optimistic.

Figure 4.11.1
Response to the question in a) "How optimistic are you that you will have work in the short term, i.e., during the next 12 months?" and b) "How optimistic are you in finding work that satisfies your expectations in terms of tasks and salary in the long term, i.e., after 12 months from now?" by treatment group

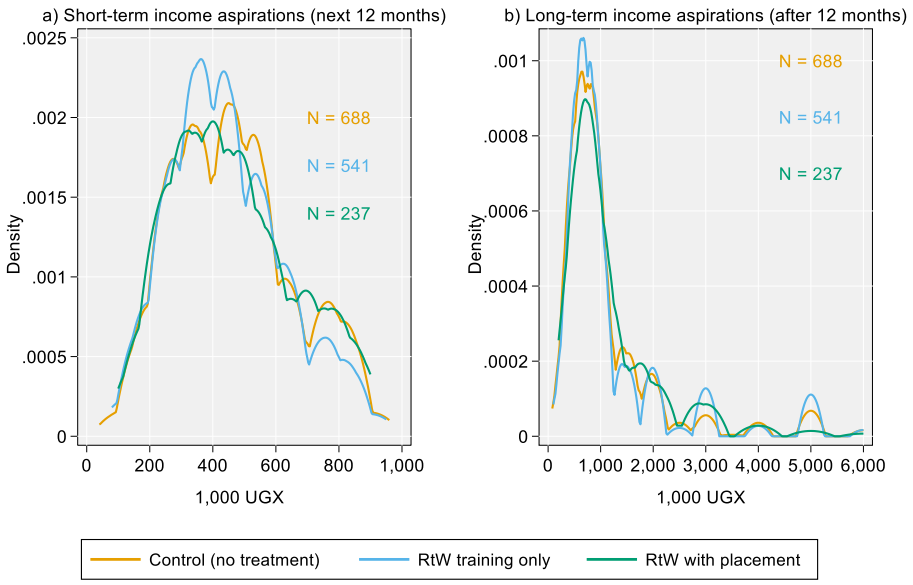


Source: Own calculations based on RtW Survey.

Figure 4.11.1.b describes respondents' confidence in finding a job that satisfies their expectations in terms of tasks and salary in the longer term, i.e., after 12 months. Considerably more respondents are quite or very optimistic to find such work in the long-term (68%) than they are optimistic to find work in the short-term. The confidence in having a job in the long-term is highest in the RtW-with-placement group (74%), followed by the RtW-training-only group (71%), and the no-intervention control group (63%).

Figure 4.11.2 shows the distributions of income aspirations of respondents in the short-term (within the next 12 months) and in the long-term (after 12 months). While in the short-term few participants aspire monthly incomes above 1,000,000 UGX, in the long-term many more do. Yet, even in the long-term, most respondents aspire to earn between 500,000 and 1,300,000 UGX.

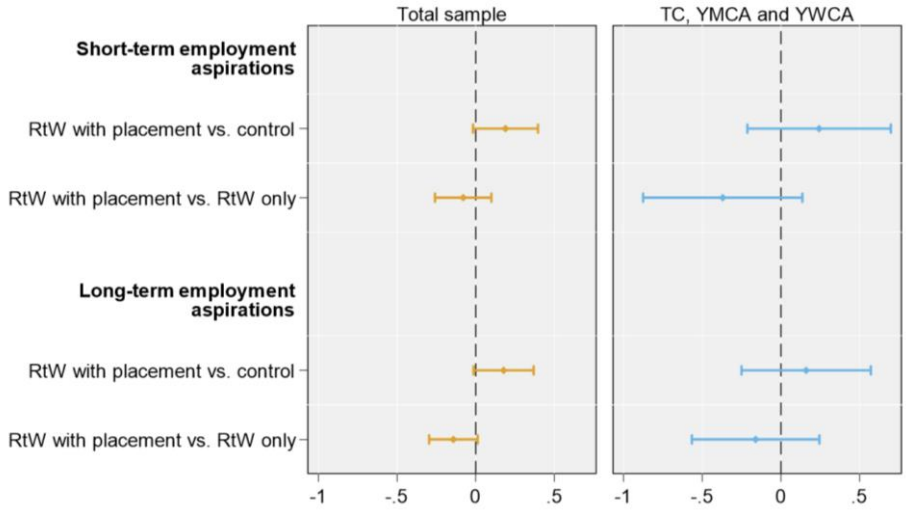
Figure 4.11.2
Monthly income aspirations (in 1,000 UGX) for a job held a) in the next 12 months and b) after 12 months by treatment group



Source: Own calculations based on RtW Survey.

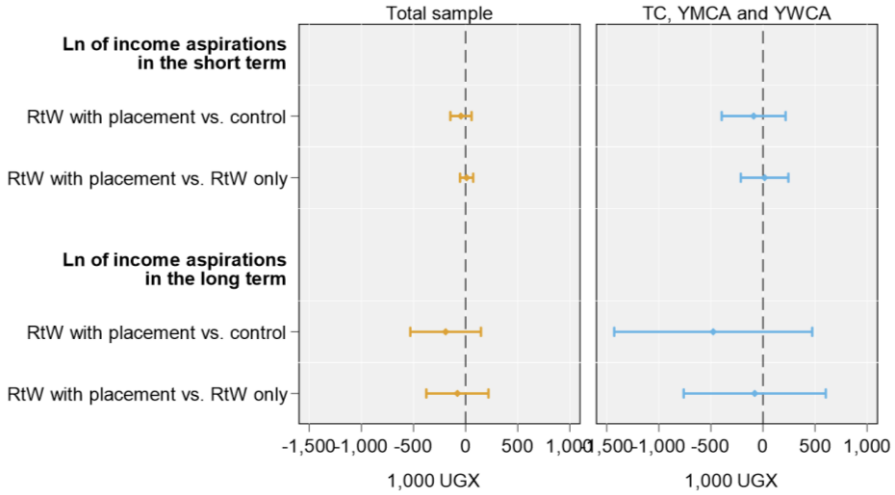
Figures 4.11.3 and 4.11.4 present the estimated treatment effects of the RtW training and placement program on short-term and long-term employment and income aspirations. For employment (Figure 4.11.3), the effect signs differ between whether the RtW-with-placement group is compared to the no-intervention control group (positive) or the RtW-training-only group (negative). However, the confidence intervals are large and suggest that the RtW program had no effect on employment aspirations. Similarly, Figure 4.11.4 shows that the RtW program had no impact on income aspirations in the short- and long-term.

Figure 4.11.3
Estimated treatment effects on short-term (up to 12 months) and long-term (after 12 months) employment aspirations



Source: Own calculations based on RtW Survey.

Figure 4.11.4
Estimated treatment effects on short-term (up to 12 months) and long-term (after 12 months) income aspirations



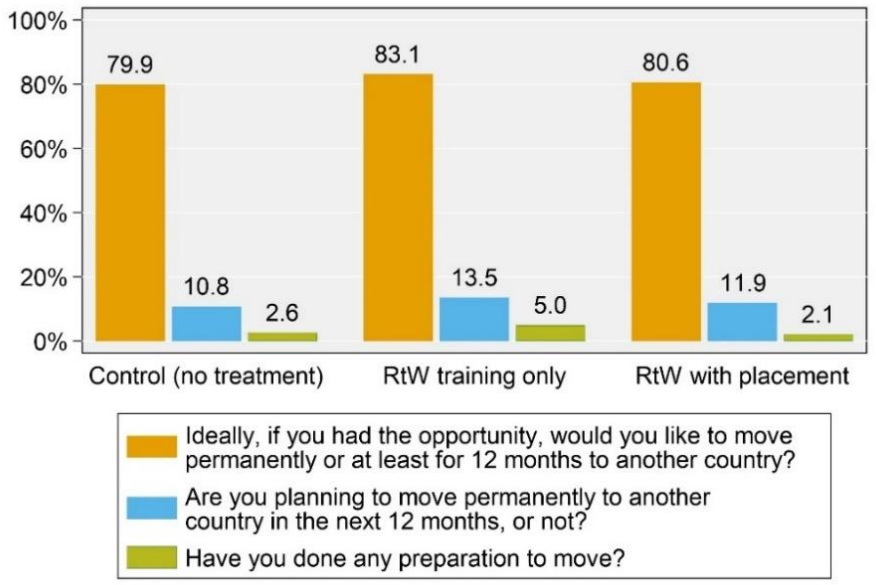
Source: Own calculations based on RtW Survey.

4.11.2 Migration intentions

Figure 4.11.5 presents respondents' attitudes towards international migration. The large majority of respondents (81%) would like to move permanently or at least for 12 months to another country if they had the opportunity. However, much fewer respondents are actually planning to move to another country in the next 12 months (12%) or have made preparations to move (3%). The percentage share of respondents who would move to another country if they had the opportunity is similar across treatment groups (80-83%). However, the percentage share of respondents who have made preparations to move is at least twice as high in the RtW-training-

only group (5%) in comparison to the no-intervention control group (2.6%) and the RtW-with-placement group (2.1%).

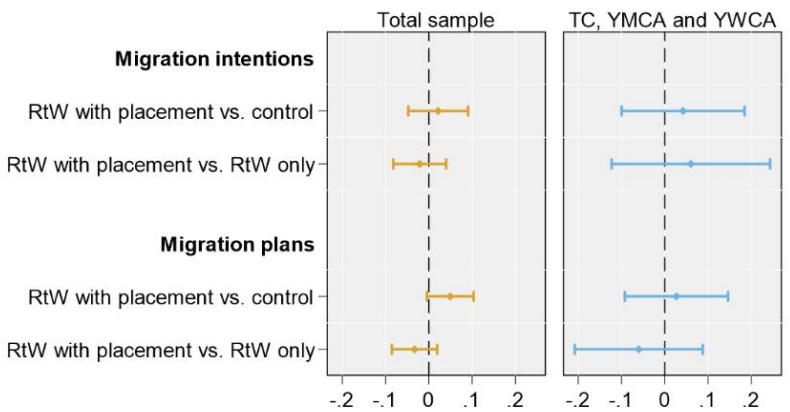
Figure 4.11.5
International migration intentions by treatment group



Note: Responses to the questions described in the figure legend are dichotomous yes/no responses. - Source: Own calculations based on RtW Survey.

Figure 4.11.6 presents the estimated treatment effects on respondents' willingness to migrate and whether they had planned to migrate in the next 12 months. Preparations to migrate were not considered for the treatment effect estimations because too few respondents had made plans (see Figure 4.11.5). Figure 4.11.6 shows that the RtW program had no effect on migration intentions and plans to move in the next 12 months.

Figure 4.11.6
Estimated treatment effects on migration intentions and plans to migrate in the next 12 months



Note: Both are binary (yes/no) response indicators to the questions "Ideally, if you had the opportunity, would you like to move permanently or at least for 12 months to another country" (migration intentions) and "Are you planning to move permanently to another country in the next 12 months, or not?" (migration plans). - Source: Own calculations based on RtW Survey.

4.12 Summary and discussion

The objective of this study was to assess the short-term impact of the RtW program on employment and labor market outcomes among recent vocational training graduates in Uganda. The RtW program included a two-day training in work readiness and a matching to UMA member companies for an internship or job, although, not every participant who was trained in work readiness was also matched to a company. The focus of this impact evaluation was specifically on the effectiveness of the completion of both RtW components as a measure to overcome the common lack of practical experience of vocational training graduates and related difficulties in the school-to-work transition.

Study participants had graduated from their respective vocational training institute in either May or November 2019 and RtW beneficiaries had been trained in work readiness at some point in their final school year. A third of the graduates trained in work readiness was subsequently matched to UMA member companies. In February 2020, RtW beneficiaries as well as graduates who were not part of the RtW program participated in a survey about their current and retrospective employment status and earnings. Initially another survey had been planned for late-2020, which was cancelled due to the COVID-19 pandemic and it's assumed effect on labor markets.

The treatment effect estimations suggest that the RtW training and placement had no impact on employment and earnings in the short term. This result holds for comparisons of the treatment group, i.e., beneficiaries who were trained and placed, with (i) the no-intervention control group and (ii) the RtW-training-only group. The results were further confirmed when alternative estimation methods and subsamples were used.

The heterogeneity of program impacts was analyzed in sub-samples of respondents who (i) were younger than 25 years or 25 years and older, (ii) had been working for one employer for at least six months prior to the study, (iii) had previous work experience in the construction sector, (iv) had previous experience in mechanical or electrical works, and (v) completed either a certificate level or diploma level vocational training. In none of these sub-samples did the RtW program have employment or income effects.

In addition, employment and earnings aspirations as well as migration intentions have been assessed. As for direct labor market outcomes, the RtW program did neither affect aspirations nor migration intentions.

The study is subject to two main limitations. The first and foremost limitation is the point in time at which post-program labor market outcomes were measured. As discussed in section 4.6.2, some respondents in the treatment group were still in internships with UMA companies at the time of the survey and the majority of respondents had only completed their placement in the past three months. This raises major concerns about whether any potential program effects could have existed at all at the time of the survey in February 2020. To tackle and alleviate this concern, all analyses were conducted in a sub-sample of respondents who had attended technical colleges or vocational training institutes with graduation dates in May 2019, rather than in November 2019. The beneficiaries in this subsample had been trained in work readiness and matched with UMA companies considerably earlier and, at the time of the survey in February 2019, essentially all treatment group members had completed their placement at least three months ago. But, also in this sub-sample of early graduates the RtW program did not show positive treatment effects.

However, even for the sub-sample of early graduates, the measured program effects would be considered short-term and some concerns, for example, about beneficiaries' potential disadvantage in the time to search for a job, remain. Thus, the results presented in this study do not exclude the possibility that positive program effects exist, which have not unfolded yet at the time of the survey.

The second limitation of the study is that unobserved or unmeasured differences between treatment and control groups may drive the estimated treatment effects. For example, if the control group had better labor market potential prior to the RtW program and, therefore, decided not to participate in the RtW program, the estimation strategy used in this study may not be able to capture such differences between treatment and control groups and the estimated program impact would be downward biased. To avoid this limitation, randomized controlled trails (RCT) have established as the gold standard of impact evaluation methods. In the RCT approach, study participants cannot select themselves into the program based on their abilities, skills, motivation, and potential. Instead, a random allocation mechanism sorts participants into treatment and control group which ensures that participants of the two groups are similar on average. However, randomized evaluations may not always be practical as they require substantial ex-ante planning (i.e., prior to program implementation) and close coordination between program implementers and researchers.

In contrast, this study was detached from all aspects of program implementation and relied on an ex-post evaluation design. As part of the ex-post design a plethora of participant background characteristics were collected in order to model their decision to join the RtW program. Further, two methods were used to model the selection process, one based on theoretical grounds and one based on a Double Machine Learning algorithm. Yet, these ex-post methods rely on the individual characteristics that were measured, and it cannot be ruled out that unobservable characteristics did downward bias (or up-ward bias) the estimates of treatment effects (also referred to as omitted variable bias), leading to the observed null-effects.

5 Quantitative Evaluation of the S4C program

5.1 The Skills for Construction program

The Skills for Construction (S4C) program is a skill promoting program supported by E4D that aims at improving employment outcomes for Ugandan youths by bridging the skills gap in the natural resource sector. It was implemented by the Solid Rock Group Uganda in partnership with the international consulting firm GOPA.

The S4C program consisted of two training levels and a placement component. Between March 2019 and March 2020, 1,587 disadvantaged young people were trained in eight cohorts on training sites either in Pakwach or in Kampala. The training was designed to meet the needs of upcoming investments in the oil and gas sector, but the skills covered are transferable to other sectors and are highly relevant for the Ugandan labor market in general. The training curriculum strongly emphasized practical training elements, repetition, and consolidation and was implemented by four local training providers, of which two were located in Pakwach and two in Kampala. Before the S4C program started, trainers at these institutions underwent a Continuing Professional Development program, which covered international health and safety standards and instructional methods.

The S4C training consisted of two parts, referred to as S4C level 1 training and S4C level 2 training. In S4C level 1 training, participants underwent a 9-day general work readiness training and a 10-day training in basic construction skills as well as a two-week international certification program in health, safety, and environment standards. S4C level 1 training concluded with an international Engineering Construction Industry Training Board (ECITB) certification level I, called HSE passport. Depending on the cohort and the training site, 35 to 80% of the highest performing students of S4C level 1 training proceeded to S4C level 2 training, which was an ECITB level II training in either rigging or pipe fitting.

Following an outreach campaign to advertise the S4C program, individuals who were interested in participating in the S4C program could register their interest over the phone. At the time of registration, potential participants had to pass a short selection test and a somewhat more comprehensive test in person at a later time in order to be admitted to the program. This test included some very basic questions concerning technical knowledge and the modalities of the training to assess applicants' aptitude and motivation for participation in the program.

In addition to training, the program included a placement component. Upon completion of either S4C level 1 or S4C level 2 training, a subset of beneficiaries was matched to companies for either an internship or a job placement. Placements were facilitated either at career fair events that connected participants who had completed their training with representatives of hiring firms or, alternatively, companies could file individual requests for interns or staff. For the second alternative, companies indicated their preferences for the candidate's skill set and S4C then identified trainees who matched the requested profile and encouraged them to apply for the position. In both matching processes candidates went through a regular selection procedure and could be rejected by the company after all. Companies partnering with the S4C program had committed to receive a certain number of S4C beneficiaries for each quarter of 2020. However, the number of placements actually implemented was lower, which is partly due to the COVID-19 pandemic that hit Uganda in March 2020. The matching process was not structured with respect to cohort, such that a trainee of any cohort could be placed at any time.

The needs-based approach of the matching service had multiple implications for the coverage and selection of beneficiaries. First, not all trainees who completed the S4C training were placed

in internships or jobs because not all of them were offered a position and some trainees who have been offered a position rejected it. Reasons to reject positions included workplaces being located too far away from residences, transport and opportunity costs were perceived as too high, or having a better outside option, such as a regular job offer. Second, and closely related to the first point, placements were based on merit or skills, such that, among trained students, potentially those who had better labor market prospects already, were placed in internships or jobs through the S4C program. Third, placements did not immediately follow the S4C training. Fourth, the length of internships differed across participants. Points three and four suggest considerable heterogeneity in the time of program completion. A more detailed description of the sample and timing of internships follows in sections 5.5 and 5.6.

5.2 Research questions

The main objective of this project is to understand if and how the S4C program has contributed to improved labor market outcomes for participants. The program consisted of three components: (i) the level 1 training, (ii) the level 2 training, and (iii) internship or job placements. However, not all beneficiaries participated in all three components. This report seeks to answer three main research questions by comparing the employment outcomes of different groups of beneficiaries with appropriate control groups.

(RQ#1) What are the overall effects on employment and earnings of the S4C program (includes any combination of program components)?
→ (S4C program) vs. (no treatment)

(RQ#2) What are the gains in employment and earnings from a placement facilitated by S4C in addition to participation in the S4C level 1 or level 1 and 2 training?
→ (S4C training + placement) vs. (S4C training only)

(RQ#3) What are the gains in employment and earnings from participating in the level 2 training in addition to participation in the level 1 training?
→ (S4C level 2 training) vs. (S4C level 1 training)

The first research question assesses the average gains from participation in any component of the S4C program in comparison to not participating at all. Thus, the treatment group contains beneficiaries who participated either only in the level 1 training, with or without a subsequent placement, or in both level 1 and level 2 training, with or without a subsequent placement. The control group consists of respondents who had registered their interest in the program but did not end up participating in any component.

Because the S4C training was quite intensive, lasting either 6 weeks (for level 1 training) or even 10 weeks (for both level 1 and level 2 training), it is expected that the training alone can have a positive impact on job market outcomes even if it was not complemented by a placement. Still, placements may have an additional effect. This is what the second research question addresses by comparing beneficiaries who participated in the S4C training (either only level 1 or level 1 and level 2) and a placement to those who participated only in the training. This comparison identifies the *additional* effect of a placement, but it cannot isolate the effect of a placement alone (not preceded by a S4C training).

Finally, the third research question seeks to isolate the additional effect of the level 2 training. To this end, job market outcomes of beneficiaries who participated in both training levels are compared to those who participated only in the level 1 training. Both groups contain beneficiaries who participated only in the training and others who subsequently obtained an internship placement.

5.3 Research design

Rigorous impact evaluation is the measurement of changes in relevant indicators, e.g., employment, that can be causally attributed to a policy intervention. At the heart of rigorous impact evaluation is the *counterfactual problem,* which is that it is unknown what would have happened to the respective beneficiary group in the absence of the intervention or policy program. The counterfactual problem makes it inherently difficult to causally attribute changes in relevant outcome indicators to the considered program. In order to emulate counterfactual scenarios, specific research designs and estimation methods are used to establish causality of program impacts.

This section describes the estimation methods used to emulate such a counterfactual scenario for the S4C program evaluation. They make use of respondents who (i) registered their interest in S4C but did not end up participating as a comparison group for RQ#1, (ii) beneficiaries who participated only in the training but not in internship placements as a comparison group for RQ#2, and (iii) beneficiaries who participated only in level 1 training as a comparison group for RQ#3.

5.3.1 Estimation methods

To measure the S4C program impacts, a survey of 1,218 respondents who had registered their interest in the program was conducted in November and December 2020. However, not everyone who had registered was actually trained. The S4C training was organized in 8 cohorts that were trained between March 2019 and March 2020. For the analysis, respondents that did not participate in the training were assigned to the cohort for which they had registered. The evaluation sample consisted of five groups of respondents:

1. beneficiaries who participated only in S4C level 1 training,
2. beneficiaries who participated in both S4C level 1 and level 2 training but not in placements,
3. beneficiaries who participated in S4C level 1 training and placements,
4. beneficiaries who participated in S4C level 1 and level 2 training with placement, and
5. respondents who did not take part in any component of the program.

For the estimation of the overall program impacts (RQ#1), the first four groups constitute the treatment group and group five is the control group. For the assessment of the additional impacts of placements (RQ#2), the treatment group comprises the third and fourth group and the control group consists of the first two groups. For the evaluation of the additional effects of the S4C level 2 training, the second and fourth group constitute the treatment group and the control group consists of the first and third group.

As described in section 5.1, admission to the different program components was not random. In order to be considered for program participation, potential beneficiaries had to pass a selection test. Additionally, admission to the level 2 training was based on test results from level 1. Moreover, placements were driven by companies' needs and may have depended on the trainees' skills. Finally, potential beneficiaries could of course always choose not to participate in (a component of) the program even if they were offered a place. This means that for all three analyses, respondents in the treatment group may systematically differ from respondents in the control group.

Because admission to the different components of the program depended, at least to a certain degree, on performance, it is possible that program participants would have had better job market prospects than non-participants even in the absence of the program. Simply comparing job market outcomes of both groups would then result in an overestimation of the program effect. However, it is also possible that potential beneficiaries decided not to participate in (another component of) the program because they had a better outside option, such as a job offer. In this case, the measured program impact would be downward biased.

To restore comparability of treatment and control groups, linear multivariable regression models are used to answer RQ#1 and RQ#2. These models control for observable background characteristics of respondents that might drive participation in the program as well as labor market outcomes. Important background characteristics are demographic and socioeconomic status variables, such as age, gender, region of residence, education, previous work experience, household size and household wealth, education level and job position of the most educated household member, as well as the respective outcome variable measured at registration for program participation. Controlling for pre-intervention outcomes (i.e., the respective outcome at registration) is similar to employing a difference-in-differences method but has the statistical advantage of being more efficient than the classical difference-in-differences interaction model approach.[57] In addition, the estimation strategy does not rely solely on pre-intervention employment outcomes as some difference-in-differences analyses do, but also controls for the sociodemographic background characteristics listed above.

For the evaluation of the additional impact of the S4C level 2 training, a *fuzzy regression discontinuity* design is used. The selection into level 2 of the S4C program was designed to follow a strict decision rule, which was based on the beneficiaries' performance score at the end of the S4C program's level 1 assessment. The decision rule was that the best 35-80% (depending on the respective cohort) of level 1 participants should proceed to level 2. In practice, this rule was not strictly implemented. One reason for not strictly implementing the decision rule might have been that some trainees who had reached the required performance score for admission to S4C level 2 training decided not to take the training, such that the remaining slots were given to less successful S4C level 1 trainees. Still, there was a strong correlation between level 2 participation and having achieved at least the required score in level 1 assessment. The idea of a regression discontinuity design is that beneficiaries who have just missed and beneficiaries who have just reached the required cutoff level essentially have identical abilities and that the minimal differences in their score are due to chance. This implies that, around the cutoff, assignment to treatment is as good as random. Thus, outcomes for beneficiaries sufficiently close to both sides of the cutoff can be compared in order to obtain the causal treatment effect of training participation. In our analysis, we additionally control for demographic and socioeconomic background characteristics to increase the precision of the estimates and to account for potential other selection mechanisms into the S4C level 2 training.

[57] *Difference-in-differences is an econometric method to measure the impacts of an intervention or policy program by comparing the average change over time in the outcome variable, e.g., employment status, for the treatment group with the average change over time in the outcome variable for the control group. The key assumption of this method is that the change over time in the outcome variable for the treatment group would have been the same as the change over time in the outcome variable for the control group in the absence of the intervention or policy program. This method thus relies on the observed outcome variables and, in the case of the S4C impact evaluation, assumes that pre-intervention outcomes reflect participants labor market potential.*

5.3.2 COVID-19 and the timing of data collection

Data collection was originally planned to be carried out separately for each cohort nine months after the respective end of S4C level 2 training. This would have meant to start data collection for cohort 1 in January 2020 and complete data collection for cohort 8 in December 2020. The original design was not feasible due to two reasons. First, a placement component was added to the program. This implied that there would have been very little time for employment benefits to materialize for beneficiaries who completed an internship after the end of the training if data collection were carried out according to the original schedule. Second, the COVID-19 pandemic hit Uganda in late March 2020.

Therefore, it was decided to jointly collect data for all eight cohorts in November and December 2020. By November 2020, some of the most severe government measures for the containment of the COVID-19 pandemic (e.g., curfews and public transport restrictions) had been relaxed. Nevertheless, the labor market was still affected by the pandemic and corresponding government measures.[58] It is, thus, expected that program evaluation based on respondents' labor market outcomes in November and December 2020 will be affected by the impact of the pandemic and may not reflect the full potential benefits from program participation. Therefore, the survey also elicited pre-pandemic labor market outcomes via a set of recall questions.

The analyses in section 5.10 evaluate program impacts with respect to two different endlines: a) February/ March 2020 (pre-pandemic), b) November/ December 2020. In a robustness check, outcomes were measured in November/ December 2020, but respondents who indicated to have lost their job or income due to the pandemic were considered as still employed or earning their pre-pandemic income, respectively. The intention was to circumvent the potential impact on outcomes resulting from either the pandemic (if measuring them in November/ December) or from not having had enough time to search for a job (if measuring them in February/ March).

5.3.3 Outcome indicators

Primary outcomes

The S4C program seeks to improve labor market prospects and outcomes of young people in Uganda. Therefore, the primary outcomes of this impact evaluation constitute direct employment benefits, which include different forms of measuring employment status and earnings.

The different employment measures are dichotomous variables (1 if employed, 0 if unemployed) and include the following indicators:

1. **Employment status,** which is defined as any paid work. It includes self-employment and excludes paid internships.
2. **Decent employment,** which is defined as paid work of at least 20 hours per week that earns at least 148,632.54 UGX per month. It includes self-employment and excludes paid internships.
3. **Self-employment,** which is defined as any own-account work or own business.
4. **Job with contract,** which relies on the respondents' answer to whether he or she had a contract with the employer (excluding internships).

[58] *Workplace closures and stay-at-home restrictions were still in place. Public transport restrictions, however, had been lifted (https://ourworldindata.org/policy-responses-covid).*

5. **Formal employment,** which relies on guidelines of the United Nation's International Labor Organization on measuring informality based on the 15th and 17th International Conference of Labour Statisticians (International Labour Organization 2009, 2013). A person who works for someone else with pay and receives at least one of thirteen benefits (e.g., paid sick leave, insurance, social security etc.) or pays taxes is considered to have a formal job (excluding internships). Self-employment is considered formal if the enterprise is registered or taxes on earnings were paid.

The income measures assessed include:

6. **Total income** in UGX**,** which is the sum of earnings from all paid work for participants who have paid work and zero for participants who do not have paid work (internships are not considered work).
7. **Total income among employed** in UGX**,** which is the sum of earnings from all paid work for participants who have paid work (internships are not considered work; estimation sample excludes those without paid work).
8. **Hourly wage among employed** in UGX**,** which is the average hourly wage rate or earnings from all paid work for those who have paid work (internships are not considered work; estimation sample excludes those without paid work).

In addition to employment status, income, and hourly wages, also working hours were assessed in order to fully understand the relationship between total income and wages. Specifically, the following outcome indicator was used:

Total number of hours worked among employed, which is the sum of hours worked per month in all paid jobs for participants who have paid work (estimation sample excludes those without paid work).

Secondary outcomes

In addition to employment and earnings, the following secondary outcomes were assessed:

Employment aspirations in the **long term**:

- Employment aspirations in the long term, which is the response to the question "How optimistic are you that you will have work in the long term, i.e., 5 years from now?" on a scale from 1 (not optimistic at all) to 5 (very optimistic).
- Income aspirations in the long term, which is the response to the question "How much would **you earn monthly from that work?"**

Migration intentions:

- Internal migration intentions, which is the response to the question "To what extent, if at all, would you like to move to **another region within Uganda** to live?" on a scale from 0 (not at all) to 3 (a lot).
- External migration intentions, which is the response to the question "To what extent, if at all, would you like to move to **another country** to live?" on a scale from 0 (not at all) to 3 (a lot).

5.4 Data collection

5.4.1 Sample size and response rates

The survey took place between November 12 and December 29, 2020. The targeted sample size was 2293, however, only 1,218 interviews were conducted. Of these 1,218 interviews

- **233** respondents were beneficiaries who participated **only in the S4C level 1 training**,
- **357** respondents were beneficiaries who participated in **both S4C level 1 and level 2 training** but not in placements,
- **20** respondents were beneficiaries who participated in the **S4C level 1 training and a placement**,
- **49** respondents were beneficiaries who participated in the **S4C level 1 and level 2 training with a placement**, and
- **559** respondents did not take part in any component of the program.

Contact details were available for 3,581 potential respondents, but it was not possible to reach out to all potential respondents due to time and budget constraints. The interviewing process started with the first cohort and then moved on to the next until reaching cohort 8. In total, 2,753 potential respondents were attempted to be called. 1,218 completed interviews, thus, corresponds to a response rate of 44%. The main reason for attrition were unanswered phone calls (46%) and wrong phone numbers (24%).

The distribution of treatment groups across cohorts with different end dates of training is shown in Table 5.4.1 for control group members, Table 5.4.1 reports hypothetical end of training dates, which are based on the end date of level 1 training of the cohort for which they had registered.

Table 5.4.1
Number of respondents by cohort and treatment group

	(Hypothetical) end of training[59]			Participation in	
Cohort	Level 1	Level 2	No training	Level 1	Level 2
1	March 2019	April 2019	205	47	45
2	May 2019	June 2019	147	43	55
3	June 2019	July 2019	64	48	60
4	July 2019/ August 2019	August 2019	35	25	50
5	September 2019	October 2019	38	22	65
6	October 2019	November 2019	31	22	65
7	November2019/ December 2019	January 2020	20	9	52
8	January 2020	February 2020/ March 2020	19	11	31

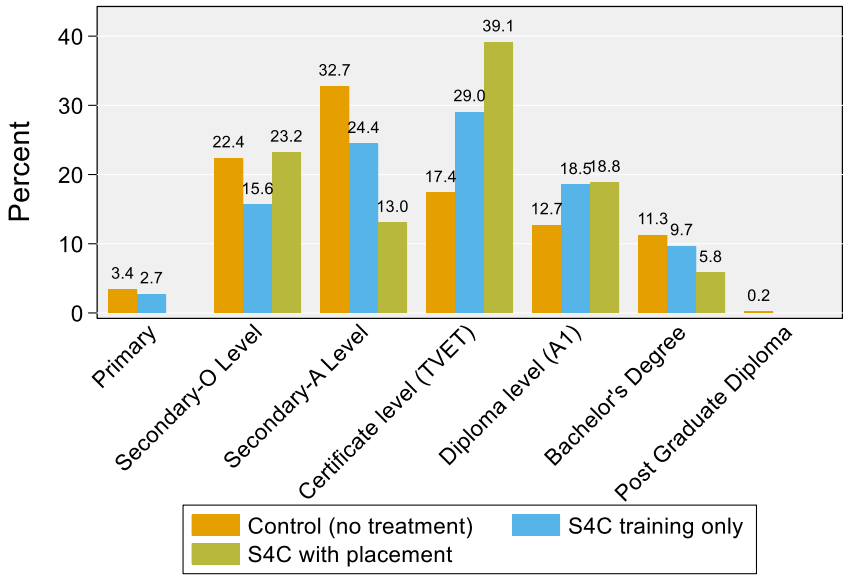
Source: Own calculations based on S4C survey.

[59] *If two dates are indicated, the first one refers to the training site in Pakwach and the second one to the site in Kampala. If only one date is indicated, they were the same for both sites.*

5.4.2 Study participants by treatment group, region of residence, and education

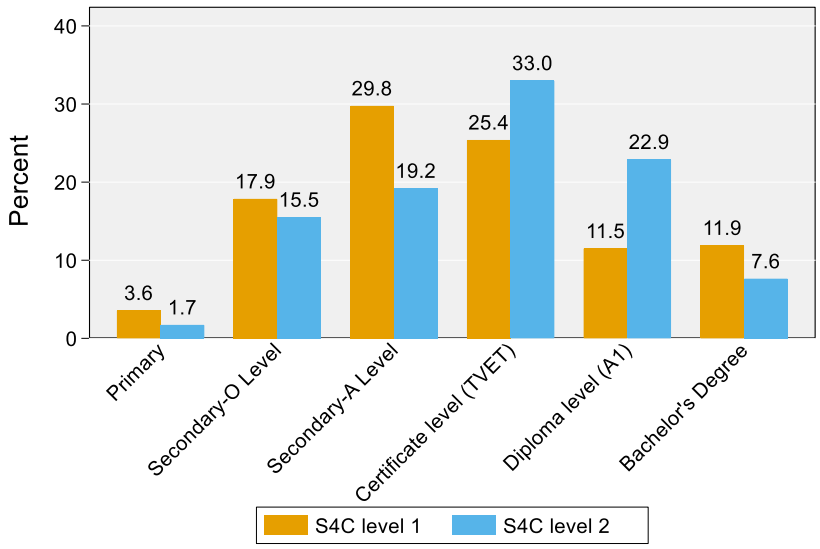
Figure 5.4.1 shows the distribution of respondents in the different treatment groups no training, training only and training with placement by their highest level of education completed. Figure 5.4.2 shows the distribution of training levels (no training, S4C level 1, S4C level 2) by education. While all educational levels except the highest (post graduate diploma) and the lowest (primary education) are represented in all treatment groups and training levels, it is evident that a higher proportion of respondents who were more intensely treated have received vocational training as their highest level of education. 39% of respondents who were trained and subsequently placed had completed a certificate level and 19% a diploma level. About 50% of respondents who received only the S4C training had vocational training (29% certificate level and 19% diploma level). More than half of the respondents in the no-treatment control group, however, had completed only secondary education (22% secondary-O level and 33% secondary-A level). The picture that emerges for the different training levels (Figure 5.4.2) is very similar: More than 60% of those who were trained in S4C level 2 have completed vocational training (33% certificate level and 30% diploma level), whereas the shares are much lower for the S4C level 1 training. Vocational training was not a prerequisite for admission to the S4C program. Nevertheless, the data suggest that it was helpful for (successful) participation. This is plausible since the S4C program was very practice oriented and trained technical skills that some beneficiaries had studied already during their vocational training.

Figure 5.4.1
Distribution of respondents in different treatment groups by level of education



Source: Own calculations based on S4C survey.

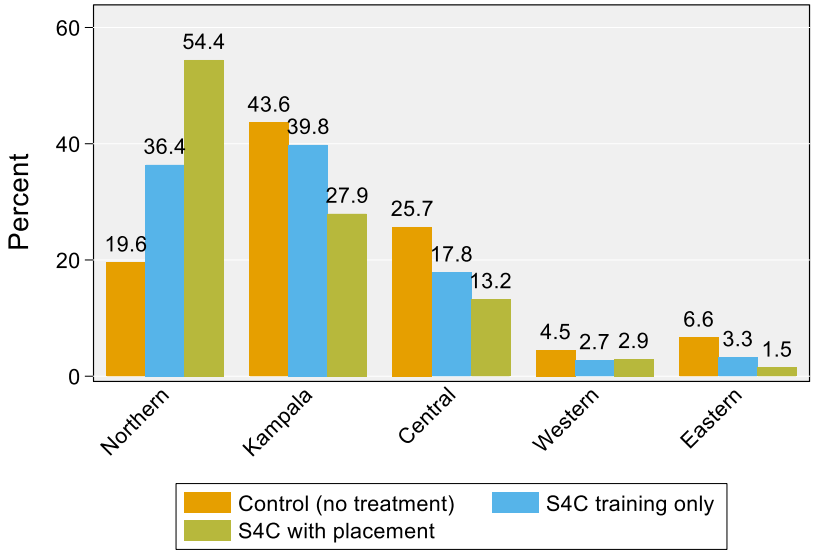
Figure 5.4.2
Distribution of respondents in different training levels by level of education



Source: Own calculations based on S4C survey.

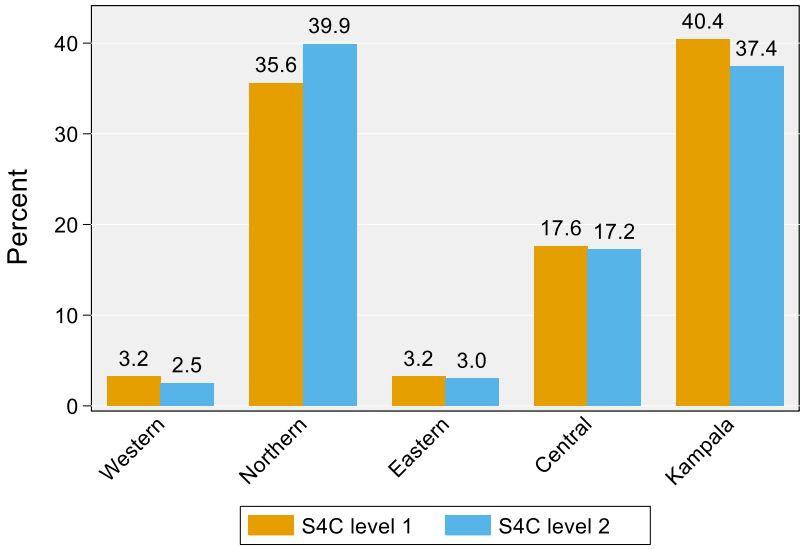
Figure 5.4.3 shows the distribution of treatment groups by region of residence at the time of registration and Figure 5.4.4 shows the distribution of participants in the different training levels by their region of residence. It is evident that the majority of respondents lived in the Northern region and in Kampala. This is not surprising since this is where the training sites were located. Potential beneficiaries who lived near the training sites were more likely to learn about the S4C program and, thus, to register their interest. Further, participation was less costly for those living closer to the training sites because it was easier and cheaper for them to commute.

Figure 5.4.3
Distribution of respondents by treatment group and region of residence



Source: Own calculations based on S4C survey.

Figure 5.4.4
Distribution of respondents by training level and region of residence



Source: Own calculations based on S4C survey.

The distributions shown in Figures 5.4.1 to 5.4.4 illustrate the importance of controlling for respondents' education and region of residence at the time of registration for the S4C program when evaluating the effectiveness of the S4C program because these background characteristics potentially influenced selection into the program.

5.4.3 Timing of survey and program roll-out

The time that had passed between program participation and the survey in late 2020 was determined by the cohort, the training level attended and the matching process of companies and participants. Figure 5.4.5 illustrates the timing of the survey relative to these relevant events and program activities by treatment group and training level, considering the following two measures:

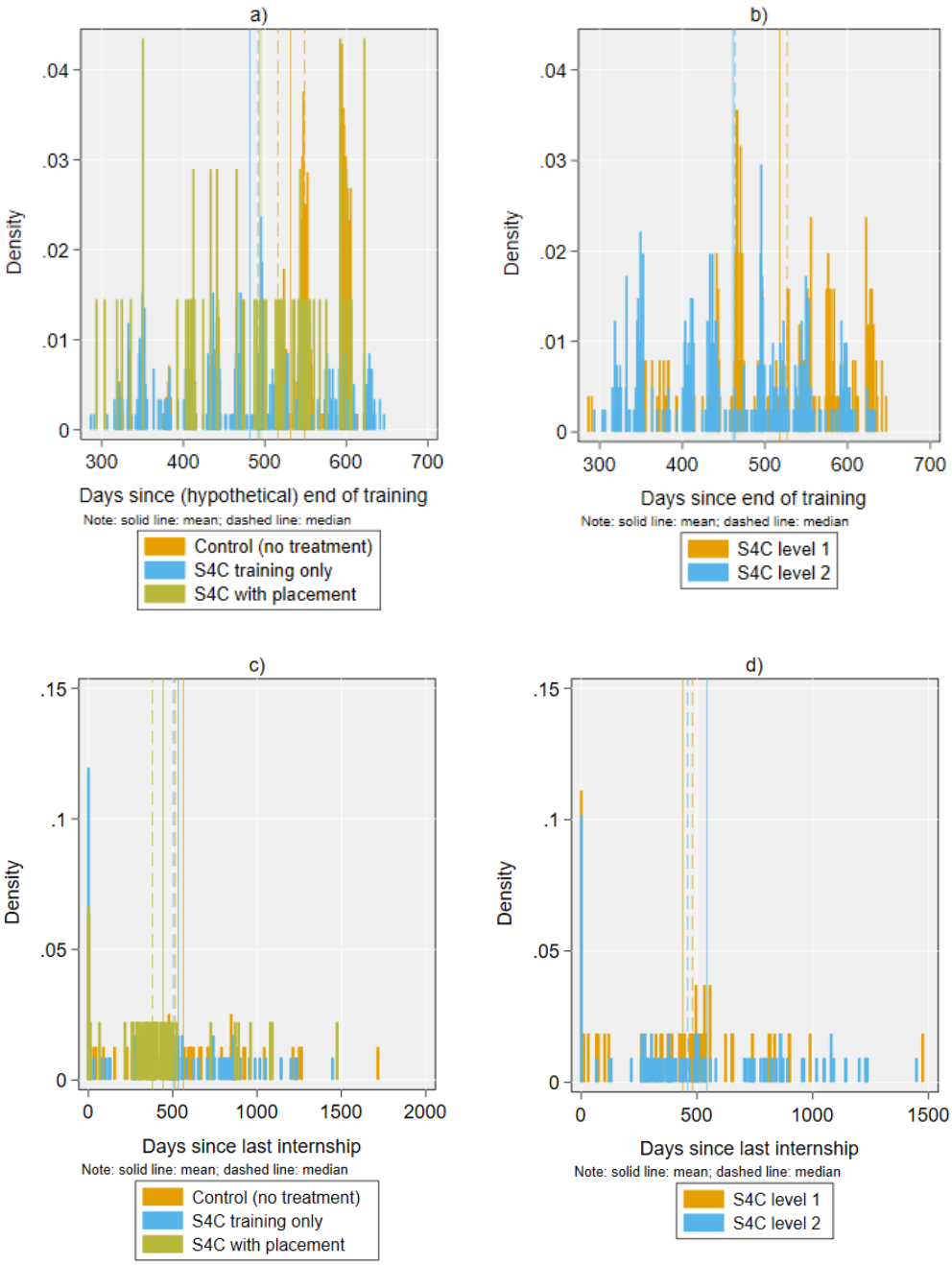
i. the number of days between the last day of the training and the survey (Figures 5.4.5.a-b),

ii. the number of days between completion of the last internship and the survey (Figures 5.4.5.c-d).

Overall, the number of days between program completion and the survey differs by treatment group and training level. The survey was conducted at the same time for all respondents. Therefore, the time period between training completion and the survey is shorter for respondents who participated in higher training levels. The same applies for days since last internship and participation in the S4C placement component. Theoretically, respondents from the control group and from the S4C training only group could have participated in internships independently of the S4C program. These individually sought internships are also included in Figures 5.4.5.c-d. Still, the time period between completion of the last internship and the survey is shorter for the S4C with placement group.

It is reassuring that even for those groups that completed the program most recently, the last day of the program was on average 460 days (last day of training for participants in S4C level 2 training; Figure 5.4.5.b) to 480 days (last day of training for S4C training only participants; Figure 5.4.5.a) before the survey. This means that the analyses can recover short to medium term effects of the program. Some internships were still ongoing when the survey was conducted, but on average the last internship ended 380 to 515 days before the interview for all study participants.

Figure 5.4.5
Time between relevant program activities and date of the survey



Source: Own calculations based on S4C survey.

5.5 Background characteristics of study participants

This section describes individual level characteristics of respondents across treatment groups and training levels with respect to participants'

a) sociodemographic background characteristics and

b) pre-intervention employment characteristics.

The pre-intervention employment indicators were inquired via a set of recall questions and date back to the time when study participants had registered their interest in the S4C program. The registration date depends on the cohort for which respondents registered and ranges between January and December 2019. Similarly, some of the sociodemographic background indicators, such as region of residence and household wealth, are based on recall questions in order to represent the situation prior to the S4C program and, thus, unaffected by the program.

The examination of individual level background characteristics serves two purposes. First, it describes the study population for whom treatment effects are measured. And second, it allows a detailed investigation of the comparability of the different groups.

5.5.1 Sociodemographic characteristics

Table 5.5.1 presents sociodemographic background characteristics of respondents in the different treatment groups (columns (1), (2), (4), and (5)) and the two training levels (columns (7) and (8)). Columns (3), (6), and (9) compare background characteristics of S4C participants vs. pure control group, S4C training plus placement beneficiaries vs. training only, and S4C level 2 vs. S4C level 1 trainees, respectively.

Sample characteristics – Columns (1), (2), (4), (5), (7), and (8)

Column (1) presents mean background characteristics (or percentage shares for binary indicators) of the pure control group that did not participate in the S4C program at all. Column (2) refers to all S4C beneficiaries. Column (4) shows mean characteristics of beneficiaries that participated only in the training, while column (5) relates to those who participated in both the training and the placement component. Columns (7) and (8) refer to level 1 and level 2 trainees, respectively.

The share of females among study participants is very low (9-19%) and the average age is 25 years. Most respondents have participated in a TVET program at some point during their education. The share of those with TVET education is lowest in the pure control group (43%) and highest among those who participated in both the training and the placement (73%). Between 12 and 24% of respondents were attending school at the time when they were interviewed. A large part (50-63%) had some work experience in the construction sector and a continued work experience of at least 6 months when they registered for the S4C program. About 80% were single and the average number of dependents was 2 (1 child and 1 adult).

Table 5.5.1
Sociodemographic background characteristics of study participants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Pure control	S4C	Std. diff. (1)-(2)	S4C training only	S4C + placement	Std. diff. (4)-(5)	S4C level 1	S4C level 2	Std. Diff. (7)-(8)
	% or Mean	% or Mean		% or Mean	% or Mean		% or Mean	% or Mean	
Respondent characteristics									
Age	24.99	25.33	-0.09	25.39	24.88	0.16	25.12	25.47	-0.10
Female	19.32	15.17	0.11	15.93	8.70	0.22	14.62	15.52	-0.02
Ever participated in TVET program	43.47	62.22	-0.38	61.02	72.46	-0.24	52.57	68.23	-0.32
Still in school	24.69	17.15	0.19	17.80	11.59	0.18	22.53	13.79	0.23
Construction experience	50.09	59.42	-0.19	59.08	62.32	-0.07	52.57	63.70	-0.23
Job ≥ 6 months	50.81	54.17	-0.07	54.24	53.62	0.01	49.01	57.39	-0.17
Single	79.43	79.36	0.00	78.81	84.06	-0.13	77.87	80.30	-0.06
Child dependents	0.91	1.06	-0.09	0.96	1.94	-0.44	1.05	1.07	-0.01
Adult dependents	0.78	0.90	-0.07	0.81	1.62	-0.42	1.04	0.81	0.12
Region of residence									
Kampala	43.63	38.56	0.10	39.79	27.94	0.25	40.40	37.41	0.06
Central	25.67	17.36	0.20	17.84	13.24	0.13	17.60	17.21	0.01
Eastern	6.64	3.07	0.17	3.26	1.47	0.12	3.20	2.99	0.01
Northern	19.57	38.25	-0.42	36.36	54.41	-0.37	35.60	39.90	-0.09
Western	4.49	2.76	0.09	2.74	2.94	-0.01	3.20	2.49	0.04
Area of residence									
City	11.29	10.75	0.02	9.43	22.06	-0.35	8.40	12.22	-0.13
Town	59.50	59.45	0.00	59.69	57.35	0.05	58.80	59.85	-0.02
Peri-urban area	15.95	16.44	-0.01	16.98	11.76	0.15	19.20	14.71	0.12
Rural area/village	13.26	13.36	-0.00	13.89	8.82	0.16	13.60	13.22	0.01
Household characteristics									
PPI	63.10	61.05	0.14	61.89	53.81	0.55	61.56	60.74	0.06
Household member with highest education: at least TVET	55.28	64.64	-0.19	64.58	65.22	-0.01	59.29	67.98	-0.18
Position of household member with highest education									
Did not work	18.46	23.14	-0.12	23.94	16.18	0.19	24.60	22.22	0.06
Office worker	22.04	21.77	0.01	21.56	23.53	-0.05	20.24	22.72	-0.06
Skilled tradesman	19.53	22.98	-0.08	22.92	23.53	-0.01	18.65	25.68	-0.17
Unskilled tradesman	26.52	22.37	0.10	21.22	32.35	-0.25	22.62	22.22	0.01
Still studying	12.19	9.13	0.10	9.68	4.41	0.21	13.10	6.67	0.22
Don't know	1.25	0.61	0.07	0.68	0.00	0.12	0.79	0.49	0.04
Test of joint orthogonality									
F-statistic			5.72			2.77			2.43
p-value			0.00			0.00			0.00
Observations	559	659		590	69		253	406	

Source: Own calculations based on S4C survey.

More than two thirds of the sample lived either in Kampala or in the Northern region at registration and the area of residence was mostly urban (roughly 10% city and 60% town). The average Poverty Probability Index (PPI)[60], which measures the likelihood of a household being poor based on a short set of questions about the household's asset ownership and socioeconomic characteristics, was about 60% for all households in the sample. In 60% of households, the household member with the highest level of education had completed at least a TVET certificate. Occupations were categorized into three groups, including office workers (e.g., managerial workers, engineers, professionals, and teachers), skilled tradesmen (e.g., technicians, soldiers, and shop assistants), and unskilled tradesmen (unskilled laborers, farmers, assemblers, cleaners, helpers, security guards, drivers, street sellers, etc.) and the occupations of the household member with the highest educational status were equally divided between these three categories.

Differences in background characteristics across groups – Columns (3), (6), and (9)

Columns (3), (6), and (9) present statistics that allow to assess the similarity and, thus, comparability of the groups. Column (3) compares S4C trainees to the pure control group. Column (6) contrasts beneficiaries who received a placement in addition to the training with those who participated only in the training and column (9) assesses the difference between level 2 and level 1 trainees. Columns (3), (6), and (9) show standardized differences (Std. Diff. in Table 5.5.1) as measures of similarity for each characteristic separately. A standardized difference takes the difference in means of the two groups and weights it by the variance of the indicators (the variance is a measure of how much the indicator spreads around its mean). Standardized differences of 0.2 (or -0.2) or less are considered small. Thus, when the similarity of two groups is evaluated, everything that is larger than the 0.2 threshold (or smaller than –0.2) would be considered as not similar.

The last three rows of Table 5.5.1 in columns (3), (6), and (9) present the number of observations in each group and measures that evaluate the similarity across treatment and control groups considering all presented characteristics jointly. A p-value below 0.1 would imply that the listed background characteristics of the two considered groups are significantly different from each other and, therefore, the two groups would not be comparable without applying econometric techniques that restore comparability.

The test statistic at the bottom of column (3) indicates that S4C program beneficiaries and non-beneficiaries significantly differ from each other taking into account all characteristics jointly. Considering the individual standardized differences, there are two characteristics – namely, participation in a TVET program before registration and region of residence at registration – that exceed the threshold of 0.2 or -0.2. Participation in a TVET program and construction sector experience are less prevalent among the pure control group.

[60] *The Poverty Probability Index (PPI) is a tool to measure the likelihood that a household is living below the poverty line. Higher values of the PPI imply a higher probability of living below the poverty line. The index is assessed via 10 standardized questions where each answer option is linked to a country-specific score (https://www.povertyindex.org/about-ppi).*

Comparing beneficiaries who received a placement in addition to the training to those who only participated in the training, the p-value at the bottom of column (6) reveals large differences. The individual standardized differences can be considered large for almost all background characteristics. In general, beneficiaries who received a placement had a better starting point than those who did not.

Turning to the comparison between S4C level 1 and level 2 trainees, the p-value at the bottom of column (9) indicates marked differences as well. The background characteristics with the largest differences are again TVET education and construction sector experience, as well as current school attendance of the respondent and school attendance of the household member with the highest level of education.

5.5.2 Employment status and earnings prior to the S4C program

Table 5.5.2 presents mean employment and income indicators at registration for study participants across the different groups. The table reveals that a substantial part (37-44%) of respondents were in paid employment when they registered for the S4C program. Between 29% (S4C with placement) and 38% (pure control group) even were in decent employment. The share of study participants in formal employment ranged between 23% (S4C training only) and 30% (pure control group). Self-employment and jobs with contract, by contrast, were only held by about 10% of respondents across all groups. Average monthly income among employed amounted to roughly 550,000 UGX.

The p-values at the bottom of columns (3), (6), and (9) reveal that only the groups “S4C training only” and “S4C with placement” differ significantly when considering all employment and income indicators jointly. However, none of the individual indicators has a standardized difference of more than 0.2 (or less than -0.2) for any two groups.

Table 5.5.2
Employment status and earnings at registration

	1 Pure control	2 S4C	3 Std. diff. (1) - (2)	4 S4C training only	5 S4C with placement	6 Std. diff. (4) - (5)	7 S4C level 1	8 S4C level 2	9 Std. Diff. (7) - (8)
	% or Mean	% or Mean		% or Mean	% or Mean		% or Mean	% or Mean	
Employment status									
Employed	44.19	39.45	0.10	39.49	39.13	0.01	37.15	40.89	-0.08
Decent employment	38.10	35.51	0.05	36.27	28.99	0.16	32.81	37.19	-0.09
Self-employment	7.59	7.87	0.01	7.25	7.63	0.01	8.37	6.32	-0.08
Job with contract	9.30	7.13	0.08	6.61	11.59	-0.17	8.70	6.16	0.10
Formal employment	29.52	23.98	0.13	23.39	28.99	-0.13	23.32	24.38	-0.02
Income									
Total income	255,762.08	216,945.37	0.06	219,200.00	197,666.66	0.05	201,094.86	226,822.66	-0.06
Total income (employed)	578,829.96	549,873.08	0.03	555,055.79	505,148.15	0.10	541,244.68	554,759.04	-0.02
Hourly wage (employed)	1,256.69	1,074.17	0.06	1,077.94	1,041.90	0.02	952.66	1,149.89	-0.09
Test of joint orthogonality									
F-statistic			1.05			2.82			1.00
p-value			0.39			0.02			0.42
Observations	559	659		590	69		253	406	

Source: Own calculations based on S4C survey.

5.5.3 Restoring the comparability of treatment and control groups

The similarity and thus comparability of treatment and control groups is important in order to attribute measured employment effects to the impact of the S4C program and its components. If treatment and control groups are not comparable, the measured treatment effects could be due to participants' different background characteristics and labor market potential. For example, if S4C program beneficiaries were better educated than non-beneficiaries and better education leads to better labor market opportunities, then estimated program effects may be due to respondents' education if the differences in education across the treatment groups were ignored in the estimation approach. In order to restore the comparability of treatment and control groups, we apply the estimation strategies outlined in section 5.3.1. That is, we control for relevant background characteristics of participants that correlate with treatment group membership and labor market outcomes and exploit the decision rule based on level 1 test scores for admission to level 2 training. Still, there may remain differences in characteristics that could not be measured, such as ability or motivation.

5.6 Descriptive statistics on S4C placements

This section presents descriptive statistics about the placement component of the S4C program. Specifically, it explores who was placed, who was still expecting to be placed when interviewed, and when placements took place.

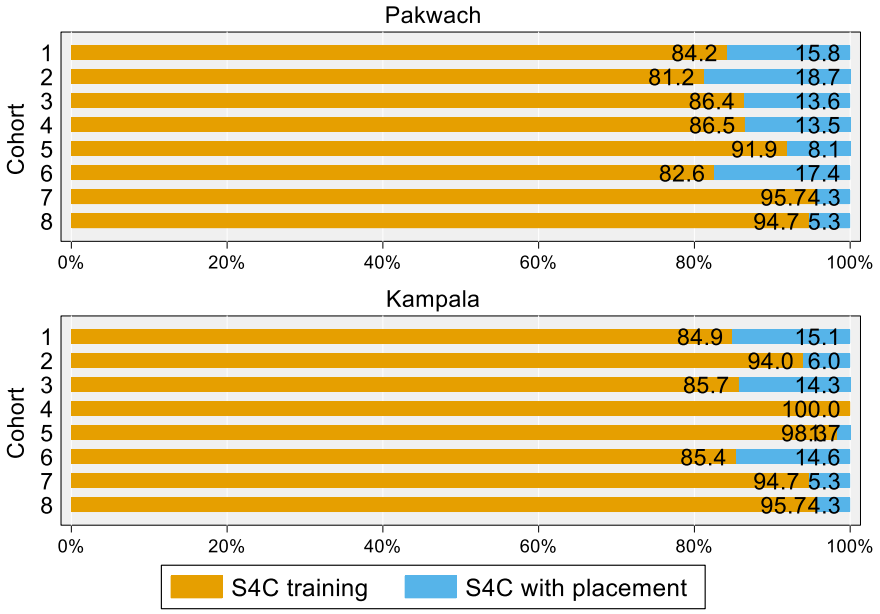
5.6.1 Who was placed?

Of the 659 respondents who participated in the S4C training, 69 (10%) received a placement through the S4C program. Two types of placements were possible, either an internship placement or a job placement. Detailed information about the placements, such as type of placement and dates, is available for 50 study participants. Out of those, 15 indicated that the placement was an internship with or without pay, respectively. 18 specified to have worked for someone else with pay and 2 worked for someone else without pay.

This section describes the percentage shares of placed beneficiaries by sociodemographic background characteristics. This means that the sample is restricted to respondents with a certain characteristic, for example, being female, and the share of placed beneficiaries among respondents with that characteristic is presented. All figures in this section refer to the sample of 69 beneficiaries who were placed through the S4C program.

Figure 5.6.1 presents the percentage share of placed beneficiaries from each cohort and training site. Relatively more respondents trained in Pakwach (between 4 and 19% per cohort) than respondents trained in Kampala (0 to 15% per cohort) received a placement.

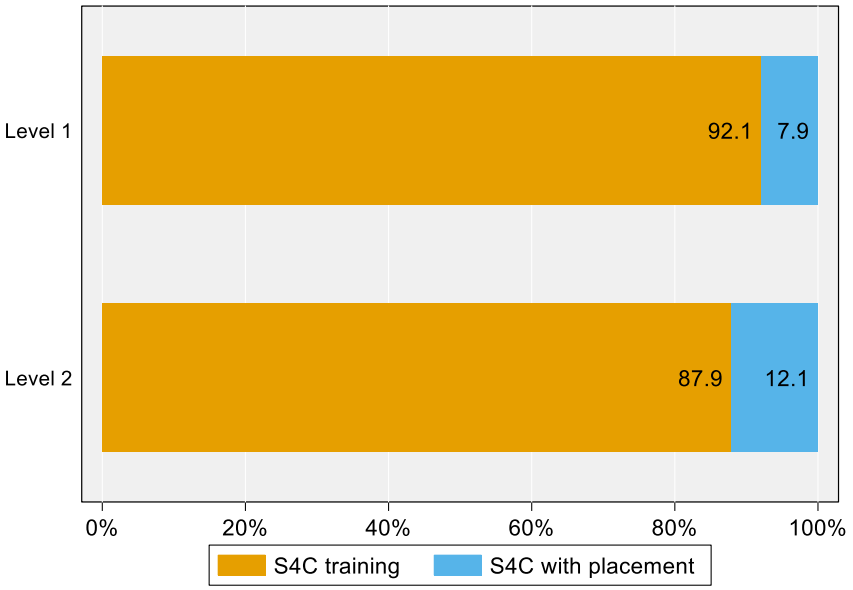
Figure 5.6.1
Percentage share of beneficiaries placed by cohort and training site



Source: Own calculations based on S4C survey.

Figure 5.6.2 reveals that 4 percentage points more beneficiaries of level 2 training received a placement compared to those who participated only in level 1.

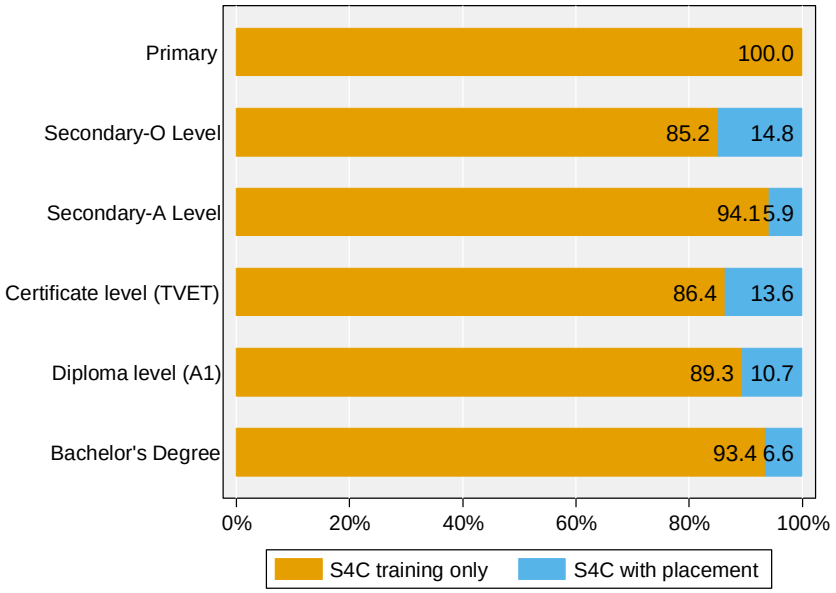
Figure 5.6.2
S4C placements by training level



Source: Own calculations based on S4C survey.

Figure 5.6.3 shows that the likelihood of receiving a placement differed by level of education as well. Relatively more beneficiaries whose highest level of education was TVET received a placement (14% for certificate level and 11% for diploma level), followed by those with secondary education (15% for Secondary-O Level and 6% for Secondary-A Level). This is well in line with the observation from section 5.4.2 that respondents with these educational levels are more likely to be more intensely treated.

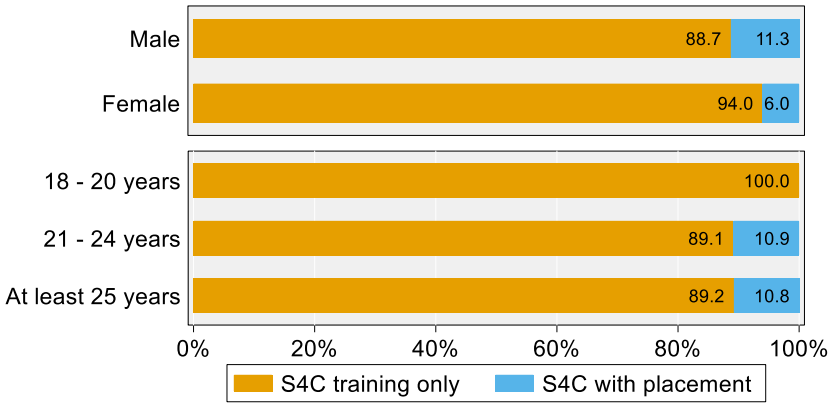
Figure 5.6.3
Percentage share of beneficiaries placed by level of education



Source: Own calculations based on S4C survey.

Figure 5.6.4 presents the shares of beneficiaries placed by their gender and age. A slightly higher share of male than female program participants received a placement (11% for males compared to 6% for females). With regard to age groups, the proportion who received a placement was higher for older beneficiaries (11% for those aged 21-24 and those aged at least 25, respectively, compared to 0% for beneficiaries younger than 21 years).

Figure 5.6.4
Percentage share of beneficiaries placed by gender and age



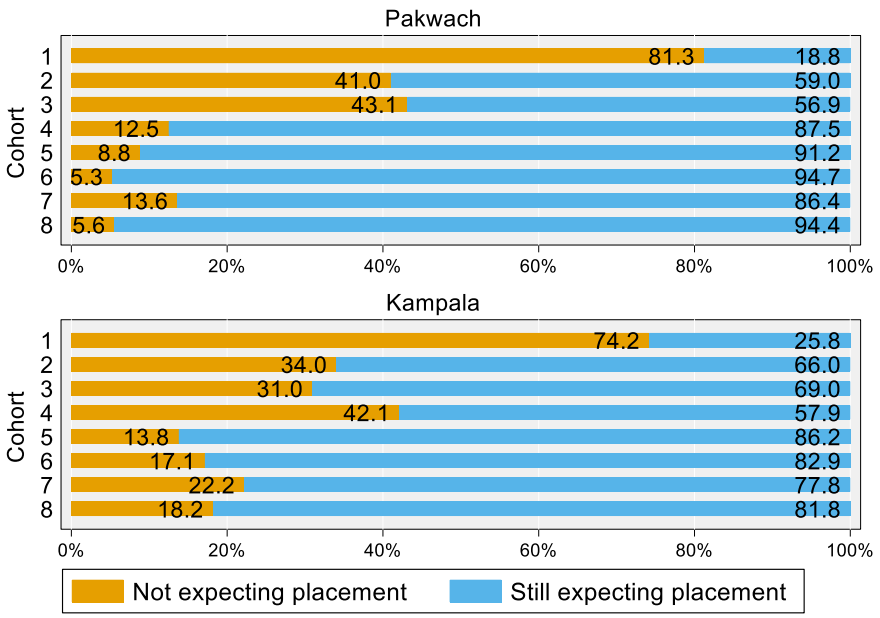
Source: Own calculations based on S4C survey.

5.6.2 Who is still expecting to be placed?

At the time when the survey was conducted, only a small share of beneficiaries had received a placement through the S4C program. However, companies had committed to about 900 placements in total. Thus, the majority of participants who had not yet received a placement were still expecting to be placed at a later point in time. It is plausible that these individuals invested less effort in their job search because they aimed to start a (permanent) job only after completing their placement. In this case, the analysis in this report chapter would overestimate the effectiveness of the program's placement component in improving employment outcomes. The impact of the whole S4C program, by contrast, would be underestimated. The reverse logic could also apply: Those who are already employed are not expecting a placement anymore.

Figures 5.6.5, 5.6.6, and 5.6.7 show the share of beneficiaries who are expecting to be placed at a later point in time by cohort and training site, education, and gender and age group, respectively. The proportion of trainees who are not expecting to receive a placement anymore is larger for earlier cohorts (Figure 5.6.5).

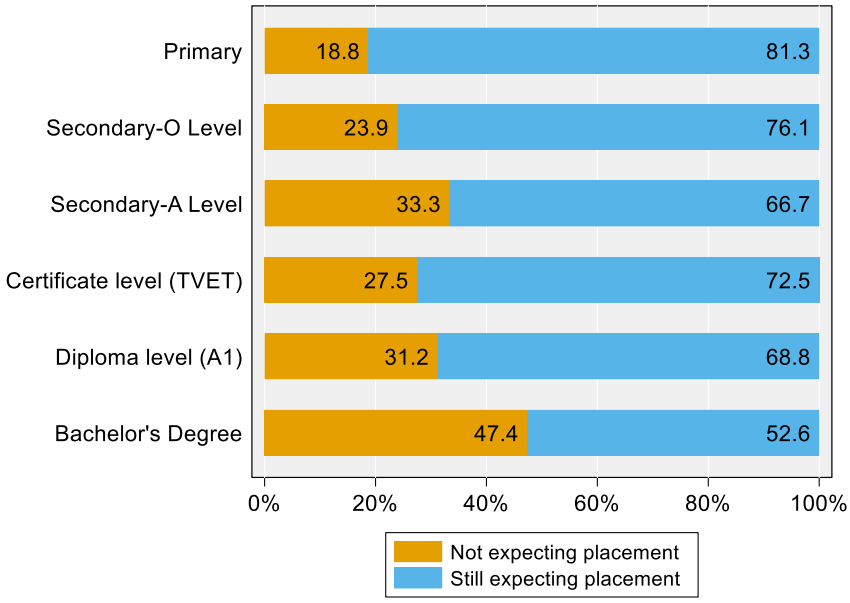
Figure 5.6.5
Percentage share of beneficiaries who are still expecting to receive a placement by cohort and training site



Source: Own calculations based on S4C survey.

Figure 5.6.6 shows that the share of beneficiaries who is not expecting a placement is lower among higher educated beneficiaries than lower educated ones. Potentially, this may be because higher educated participants have a higher overall employment probability and are less dependent on the placement.

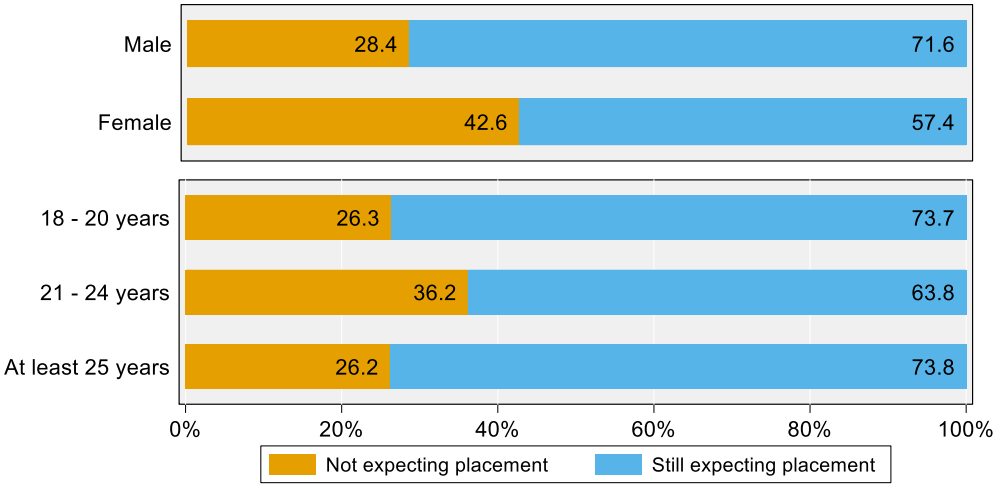
Figure 5.6.6
Percentage share of beneficiaries who are still expecting to receive a placement by level of education



Source: Own calculations based on S4C survey.

Figure 5.6.7 further shows that the share of beneficiaries who is not expecting a placement is higher among females than males and among 21- to 24-year-olds than among 18- to 20-year-olds and 25-year-olds and older ones.

Figure 5.6.7
Percentage share of beneficiaries who are still expecting to receive a placement by gender and age



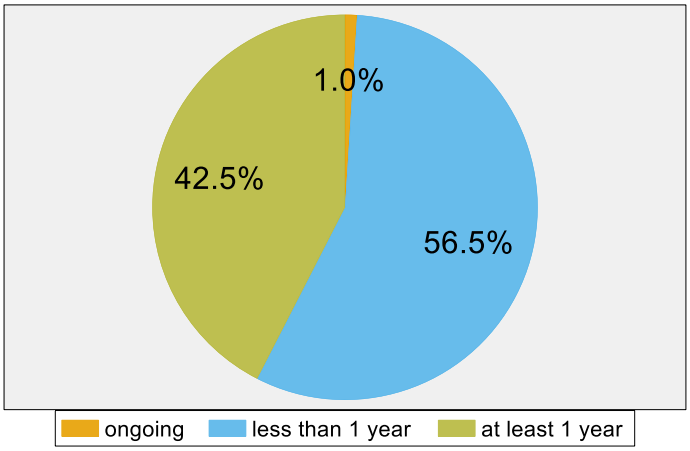
Source: Own calculations based on S4C survey.

5.6.3 When were beneficiaries placed in internships?

This section refers only to placements into internships, but not job placements. Figure 5.6.8 describes how long before the survey the S4C internship placements had ended. This is important because it determines how much time there was for the employment benefits of the S4C

internship placements to materialize. While 1% of internships were still ongoing at the time of the survey, for the remaining participants the internship had terminated at least 6 months ago and for 43% it had even ended more than one year prior to the survey. This implies that the analysis measures the short- to medium-term effects of the placement component.

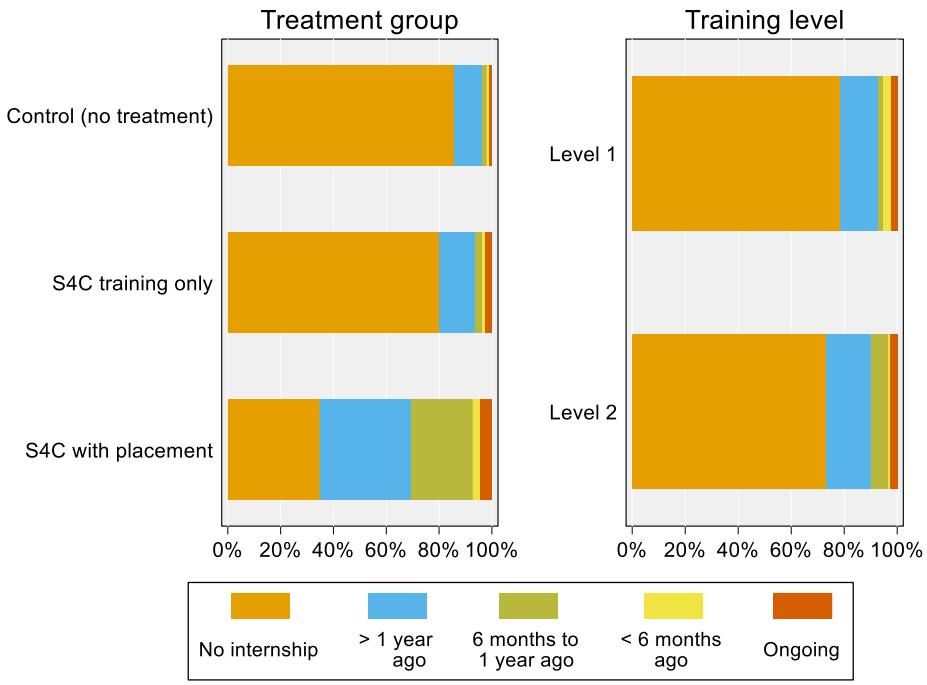
Figure 5.6.8
Timing of S4C internship placement



Source: Own calculations based on S4C survey.

Figure 5.6.9 shows the share of respondents from each treatment group and training level that completed any internship and the timing of those internships. The share of respondents who did not complete any internship is slightly smaller among "S4C training only" beneficiaries and substantially smaller in the group that additionally participated in the placement component.

Figure 5.6.9
Timing of any internship



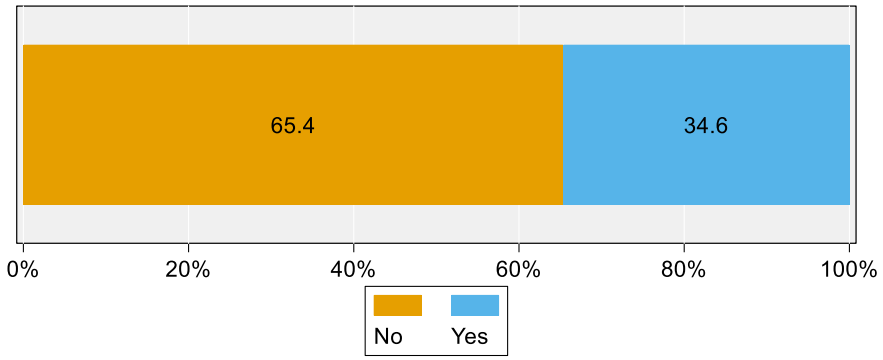
Source: Own calculations based on S4C survey.

To some extent, this reflects internships facilitated by the S4C placement component. The figure shows that not everyone in the "S4C with placement" group completed an internship because some beneficiaries were directly placed into jobs or no information was available on whether the type of placement was an internship or a job. In addition, internships ended somewhat more recently in this group. The same pattern emerges when comparing S4C level 1 to level 2 participants. The differences, however, are very small.

5.7 Participation in alternative employment promoting programs

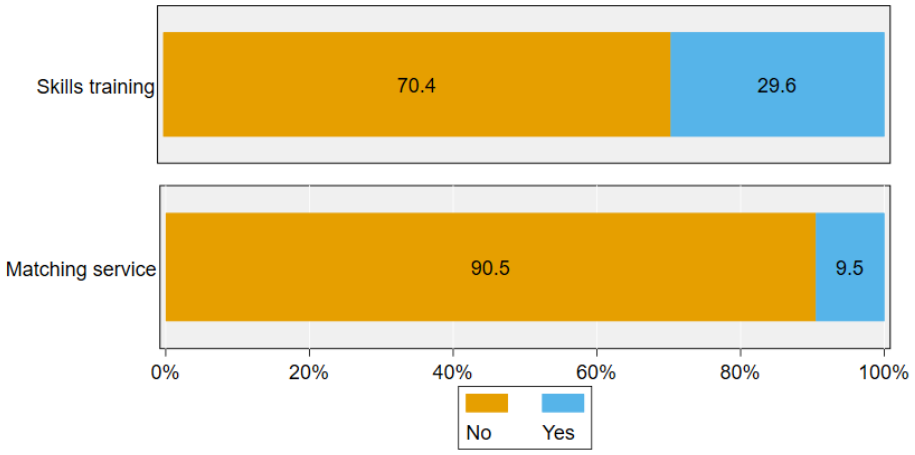
Respondents of the pure control group may have participated in alternative employment promoting programs other than S4C. Figure 5.7.1 reveals that this is indeed the case for 35% of non-beneficiaries. 30% completed a skills training program and 10% participated in a matching service (Figure 5.7.2).

Figure 5.7.1
Participation in any employment promoting program among pure control group



Source: Own calculations based on S4C survey.

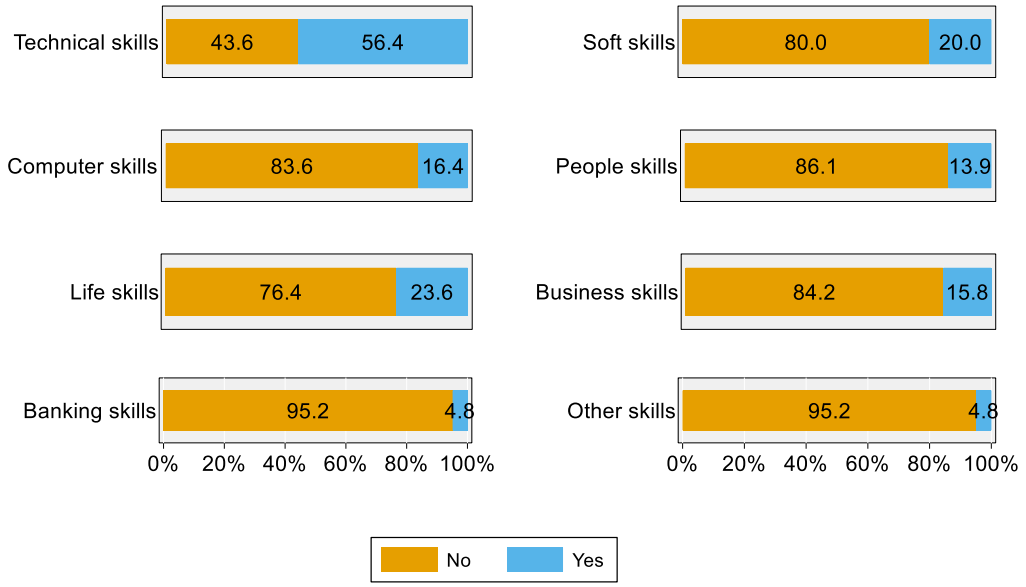
Figure 5.7.2
Type of employment promoting program attended by control group



Source: Own calculations based on S4C survey.

Figure 5.7.3 depicts the types of skills trainings attended by control group members. This was not a single choice question because a training program can cover different types of skills trainings. 56% of those who participated in a skills training program received training in technical skills, followed by 24% trained in life skills and 20% in soft skills. Thus, in terms of content, the trainings attended by respondents from the control group were roughly comparable to the S4C training.

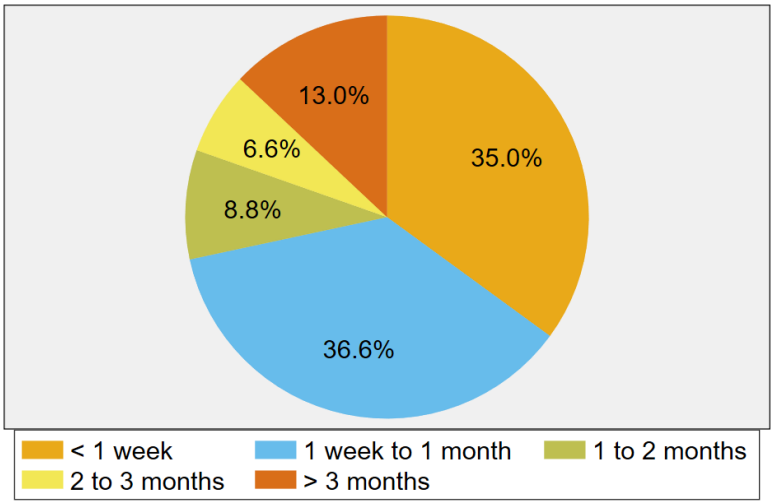
Figure 5.7.3
Types of skills trainings attended by control group



Source: Own calculations based on S4C survey.

The duration of those training programs is depicted in Figure 5.7.4. More than 70% of the trainings had a duration of less than one month and 35% even less than one week, implying they were shorter than S4C level 1 training.

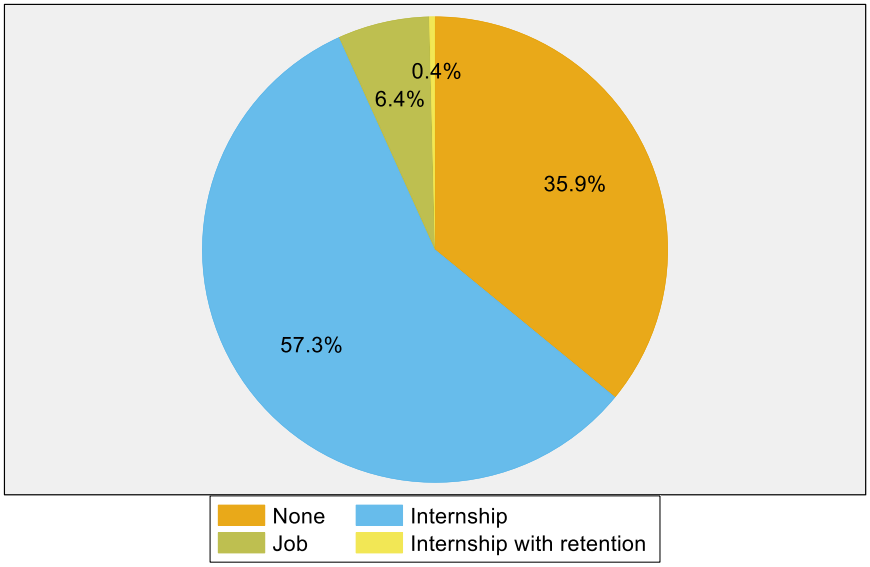
Figure 5.7.4
Duration of trainings attended by control group



Source: Own calculations based on S4C survey.

Figure 5.7.5 sheds more light on the type of work experience that resulted from matching services in which control group members participated. 36% did not lead to any work experience, 57% led to an internship, and only 7% to a job or an internship with subsequent retention.

Figure 5.7.5
Success of matching services



Source: Own calculations based on S4C survey.

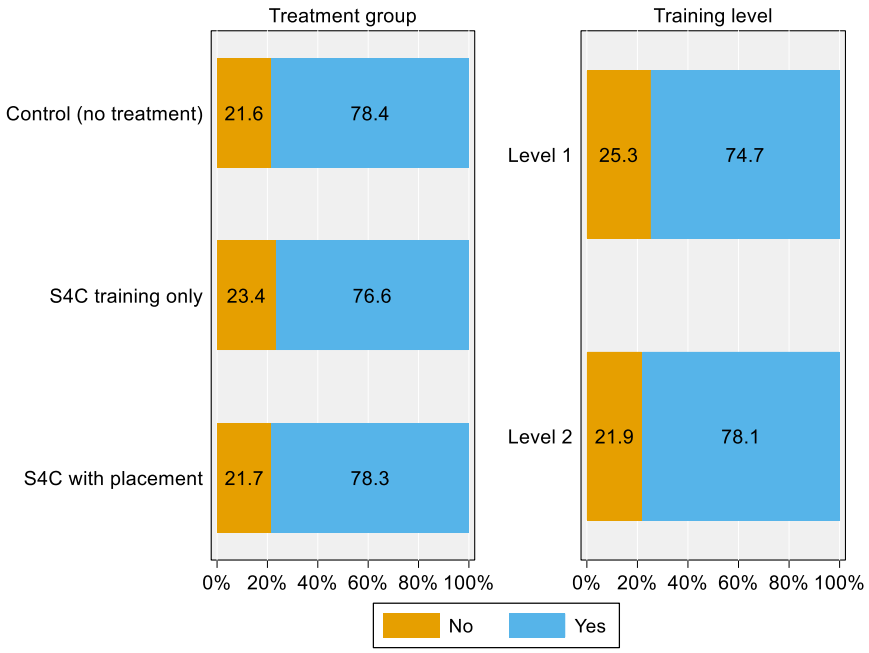
5.8 Descriptive statistics on respondents' job search

This section provides some insights into respondents' job search. It describes the success of job interviews, the duration and methods of job search and the impact the COVID-19 pandemic on the search.

5.8.1 Interviews

Figure 5.8.1 shows the percentage share from the three treatment groups and two training levels that received an invitation to at least one job interview. Respondents who were gainfully employed in March or November are coded as having had an interview in this section. The share of beneficiaries that were invited to at least one interview is slightly higher among level 2 trainees compared to level 1 but is virtually the same for the three groups "pure control", "S4C training only", and "S4C with placement".

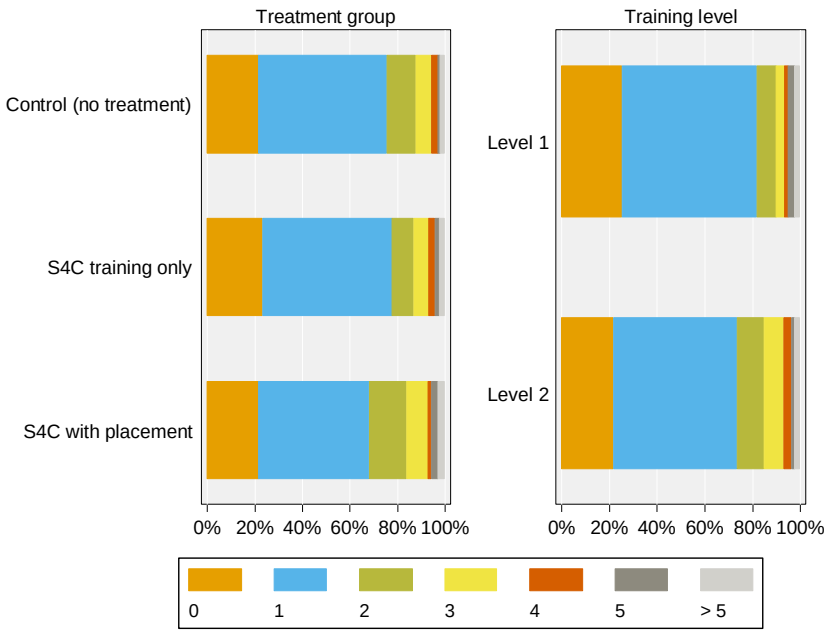
Figure 5.8.1
Invitation to job interviews



Source: Own calculations based on S4C survey.

Figure 5.8.2 compares the number of interview invitations received by members of the three treatment groups and two training levels. Beneficiaries in the "S4C with placement" group were invited to a slightly higher number of interviews than members of the other two treatment groups. This is also the case for level 2 trainees compared to level 1 trainees.

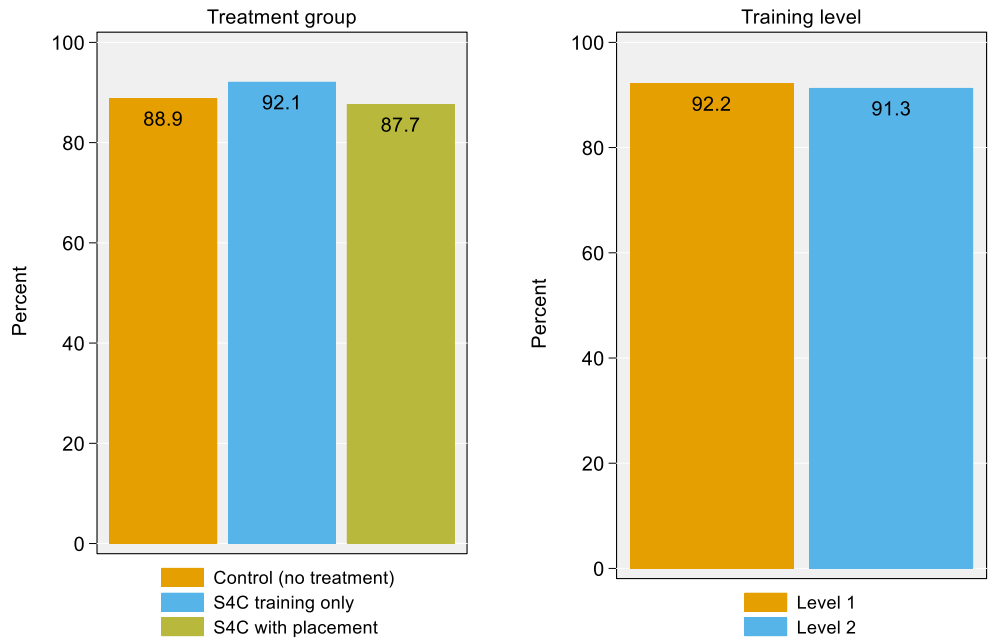
Figure 5.8.2
Number of interview invitations



Source: Own calculations based on S4C survey.

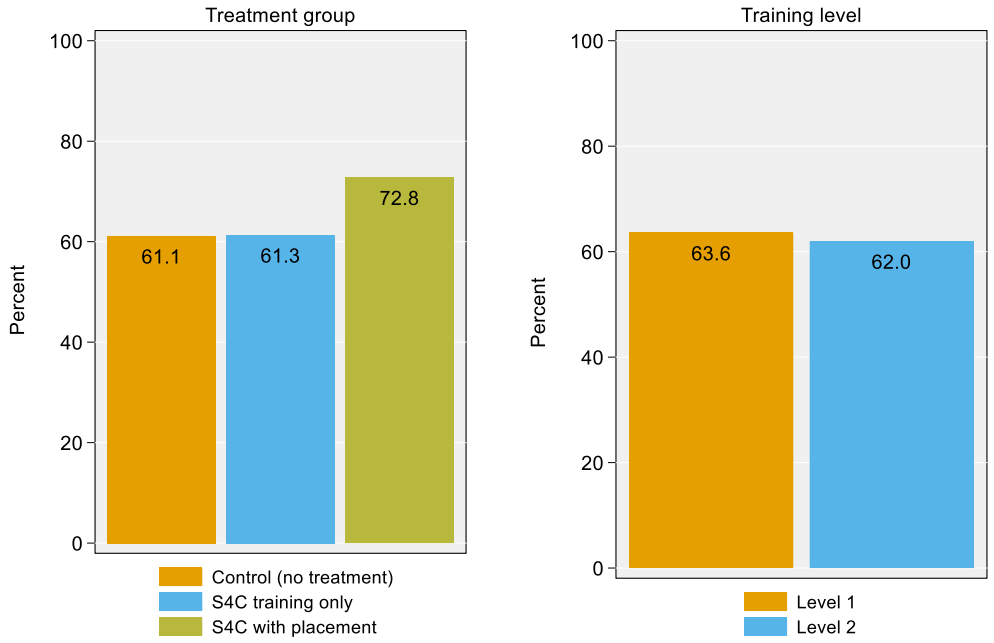
Figures 5.8.3 and 5.8.4 show the percentage shares of interviews that respondents attended and the proportion of attended interviews that led to a job. The share of attended interviews is similar across all groups. With 73% of attended interviews that led to a job, the success rate was highest in the group that received a placement in addition to training, compared to only 61% in the other groups (Figure 5.8.4). This difference, however, is partly due to the fact that it was necessary to complete an interview in order to receive a placement.

Figure 5.8.3
Percentage share of attended interviews



Source: Own calculations based on S4C survey.

Figure 5.8.4
Percentage share of attended interviews that led to a job

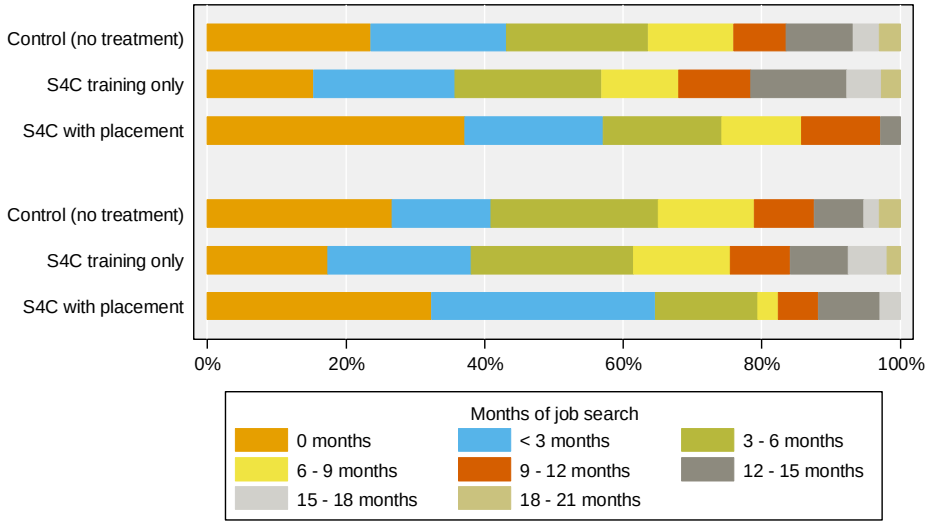


Source: Own calculations based on S4C survey.

5.8.2 Length of job search

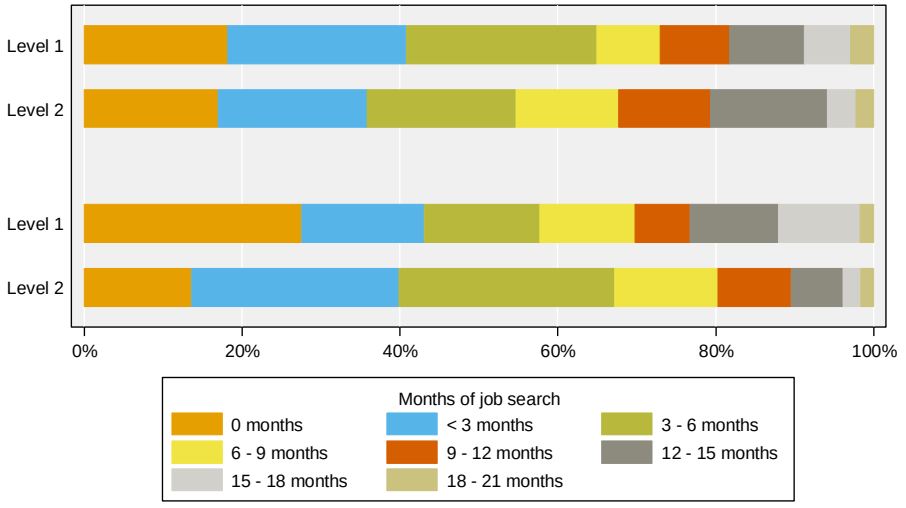
To gain further insights into the process of searching for a job, the length of job search is another interesting variable to look at. On the one hand, a longer duration of job search can increase the probability to find employment. On the other hand, more successful respondents will need less time to find a job. Figure 5.8.5 shows for how long the different treatment groups searched for a job after the (hypothetical) end of their S4C training by their employment status. Figure 5.8.6 presents the same relationship for the two training levels. The proportion of respondents who did not search at all is largest among those who received a placement after having completed the S4C training. Partly, this may reflect the fact that beneficiaries invested less (or no) time in their job search while completing the placement.

Figure 5.8.5
Length of job search by treatment group and employment status in November/December



Source: Own calculations based on S4C survey.

Figure 5.8.6
Length of job search by training level and employment status in November/December



Source: Own calculations based on S4C survey.

Table 5.8.1 reveals the average length of job search by employment status in November/ December for those who searched at all in all groups. In general, the mean search duration is shorter among those who were employed. With 4.7 months, the average duration is shortest for beneficiaries who completed an S4C placement and were employed at the time when the survey was conducted.

Table 5.8.1
Average duration of job search for those who searched at all

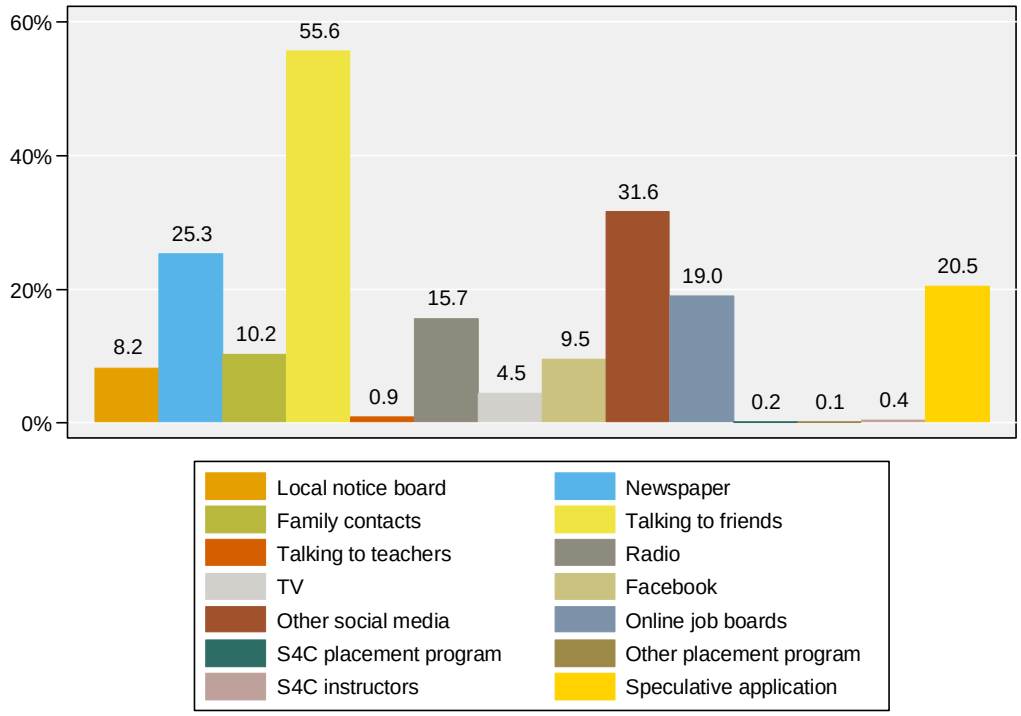
	Control (not treatment)	S4C training only	S4C with placement	S4C level 1	S4C level 2
Employed	6.4 months	6.4 months	4.7 months	7.6 months	5.5 months
Not employed	6.6 months	6.9 months	5.2 months	6.4 months	7.1 months

Source: Own calculations based on S4C survey.

5.8.3 How did respondents search for a job?

To provide optimal support for the job search, it is crucial to understand how youths in Uganda search for jobs and what particular challenges they face. Because there were no significant differences between the groups of respondents, the graphs in this section are only shown for the complete sample. According to Figure 5.8.7, the most prominent method of job search is talking to friends (56%). Moreover, the internet plays an important role with 32% drawing on social media and 19% making use of online job boards. Yet, 25% still read newspapers for information about job offerings. Another important channel are speculative applications (21%). Note that these channels are not mutually exclusive and, thus, percentages do not add up to 100.

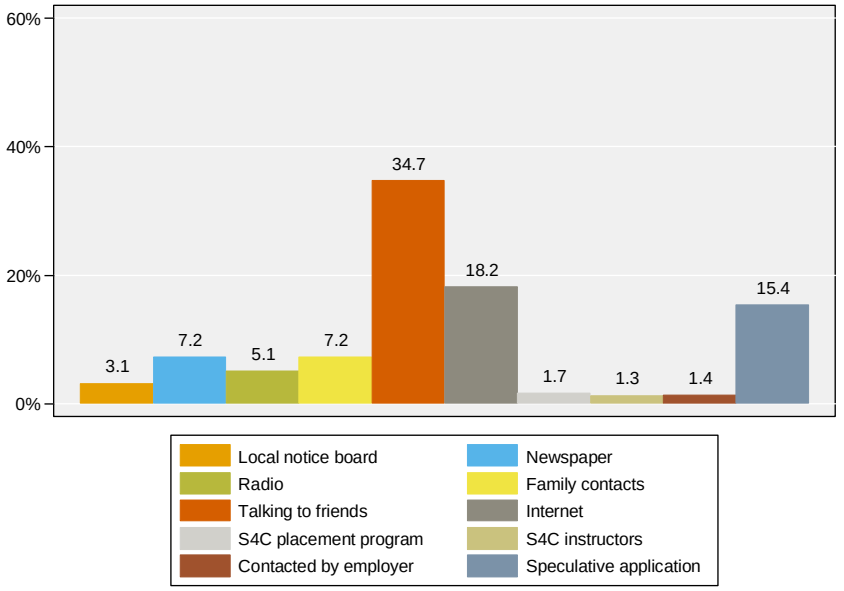
Figure 5.8.7
Methods of job search



Source: Own calculations based on S4C survey.

Figure 5.8.8 reveals that, overall, the most popular methods are also the most successful ones. Newspapers, however, do not appear to be a very promising channel.

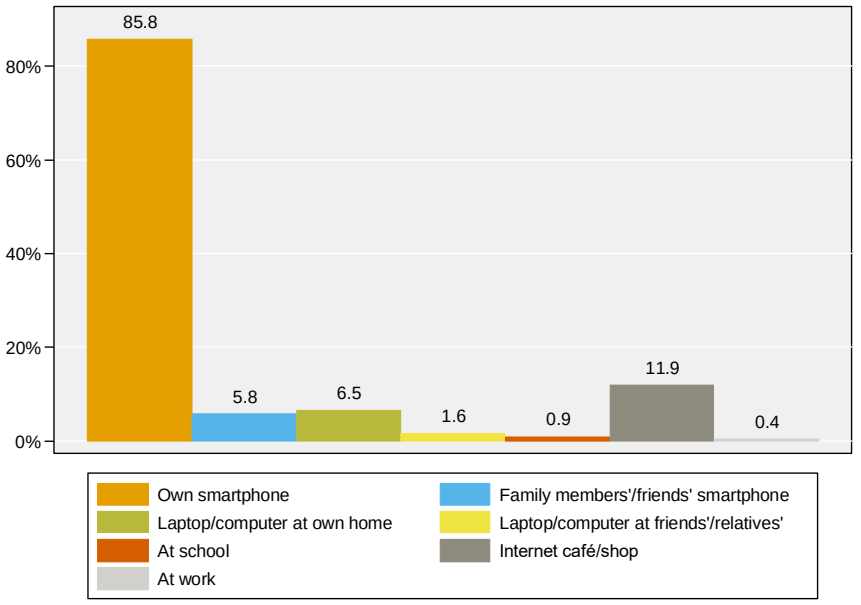
Figure 5.8.8
Most successful methods of job search



Source: Own calculations based on S4C survey.

The vast majority (86%) of respondents who search the internet for job opportunities do so using their own smartphone (Figure 5.8.9). Only 12% rely on internet cafés. Accordingly, a mere 2% mentioned lack of internet access as an important challenge for job search (Figure 5.8.10).

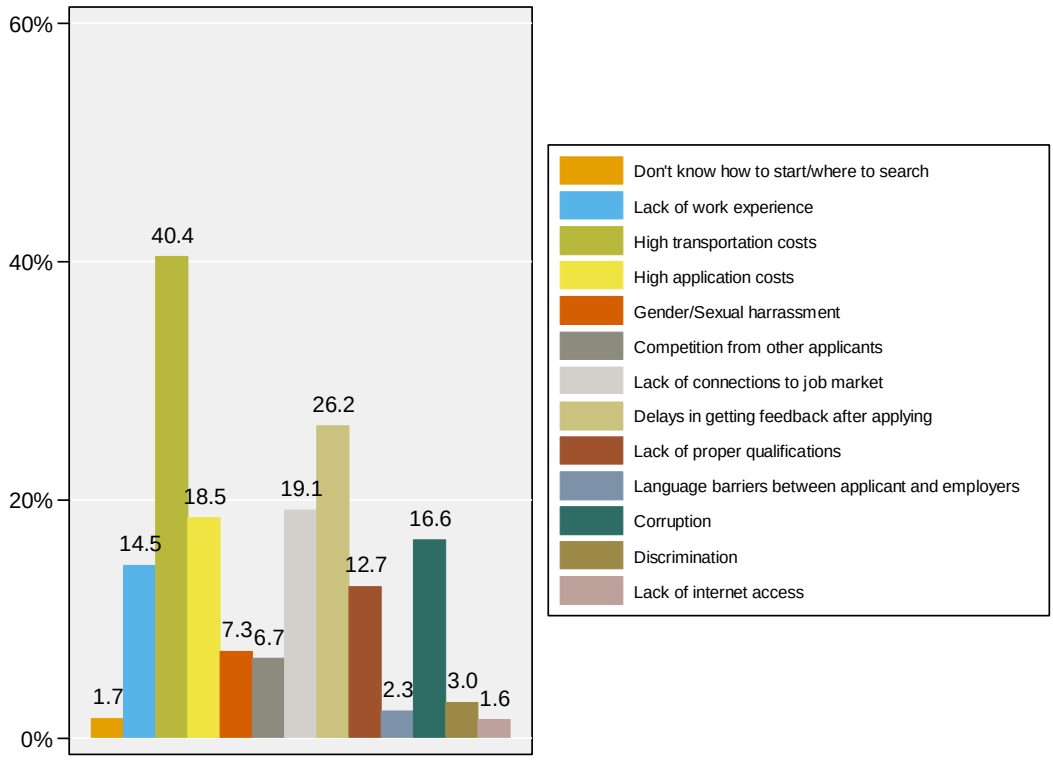
Figure 5.8.9
Internet access for job search



Source: Own calculations based on S4C survey.

The most frequently mentioned challenges were high transportation costs (40%), followed by delays in getting feedback after applying (26%), lack of connections to the job market and high application costs (both 19%). Application costs include, for example, costs for airtime and photocopies or administration costs. Lack of proper qualifications was mentioned by 13% and lack of work experience by 15%. This shows that skills training and placement programs such as S4C are very important. That these issues were also mentioned by program participants does not necessarily imply that the program was not effective but could as well mean that respondents participated in the program because they were facing these particular challenges. It is rather worrying that corruption was brought up by 17% of respondents and 7% reported gender-related discrimination and sexual harassment, which occurred predominantly among female participants (14%) but also among male ones (6%).

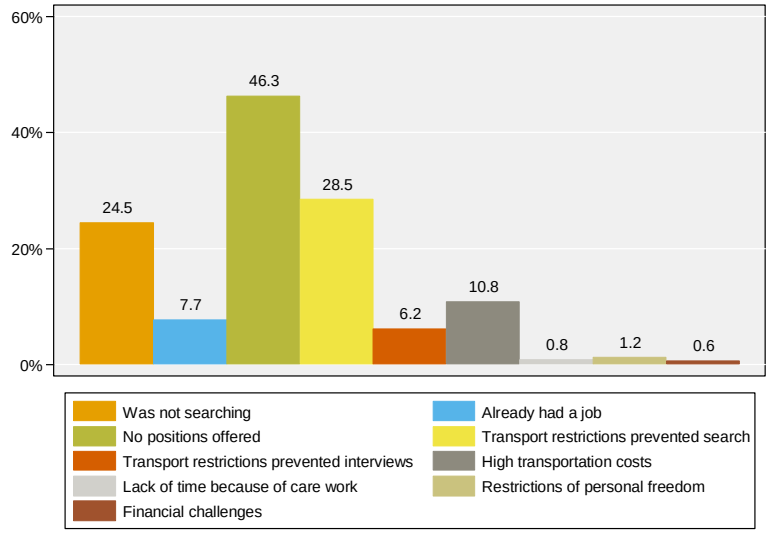
Figure 5.8.10
Challenges in job search



Source: Own calculations based on S4C survey.

In addition to the factors outlined in Figure 5.8.10, the COVID-19 pandemic may have adversely affected job search in various ways. This is depicted in Figure 5.8.11. 46% reported that no positions were offered due to the pandemic. Moreover, transport restrictions heavily impacted the search process. 29% mentioned that transport restrictions prevented their search, 6% said that the transport restrictions prevented them from attending interviews and 11% complained about increasing costs of transportation.

Figure 5.8.11
Impacts of the COVID-19 pandemic on job search



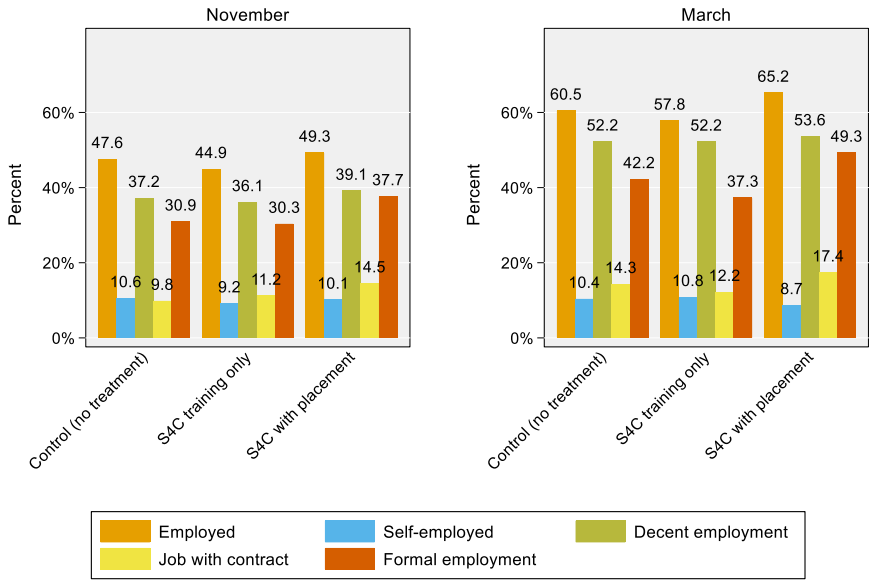
Source: Own calculations based on S4C survey.

5.9 Descriptive statistics on labor market outcomes

5.9.1 Labor market characteristics at endline

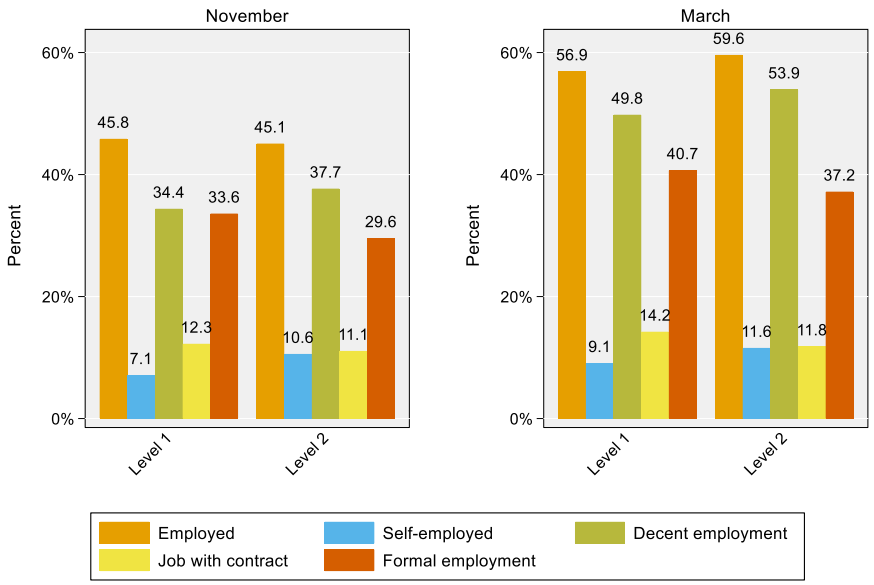
Figures 5.9.1 and 5.9.2 show descriptive statistics on employment outcomes in November (when the survey was conducted) and March (pre-COVID) by treatment group and training level, respectively. Across all groups, employment outcomes were better in March than in November. This is not surprising as the labor market was affected by the COVID-19 pandemic in November, but not in March. The distribution of employment outcomes is roughly similar across all groups of respondents.

Figure 5.9.1
Employment outcomes by treatment group



Source: Own calculations based on S4C survey.

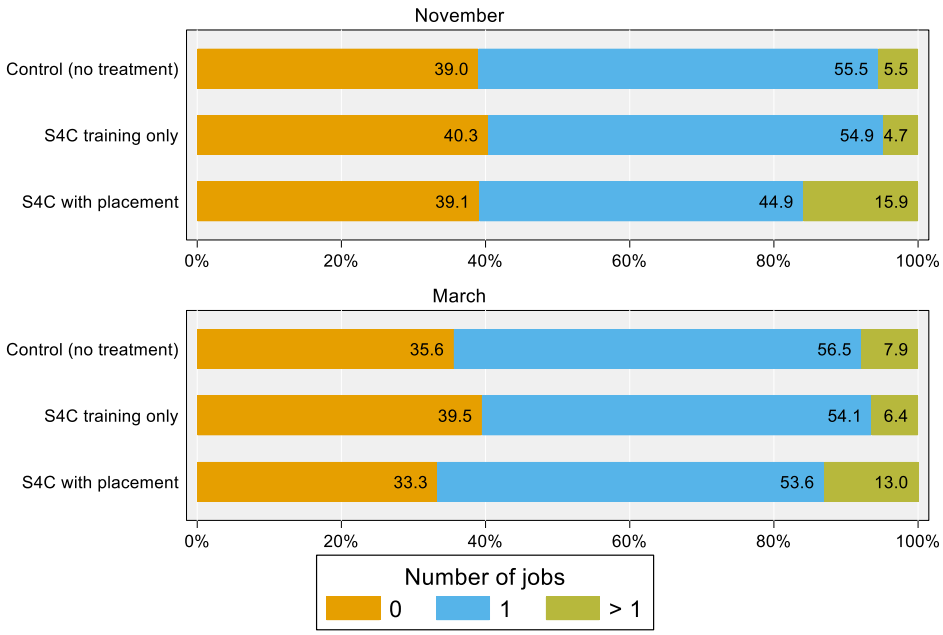
Figure 5.9.2
Employment outcomes by training level



Source: Own calculations based on S4C survey.

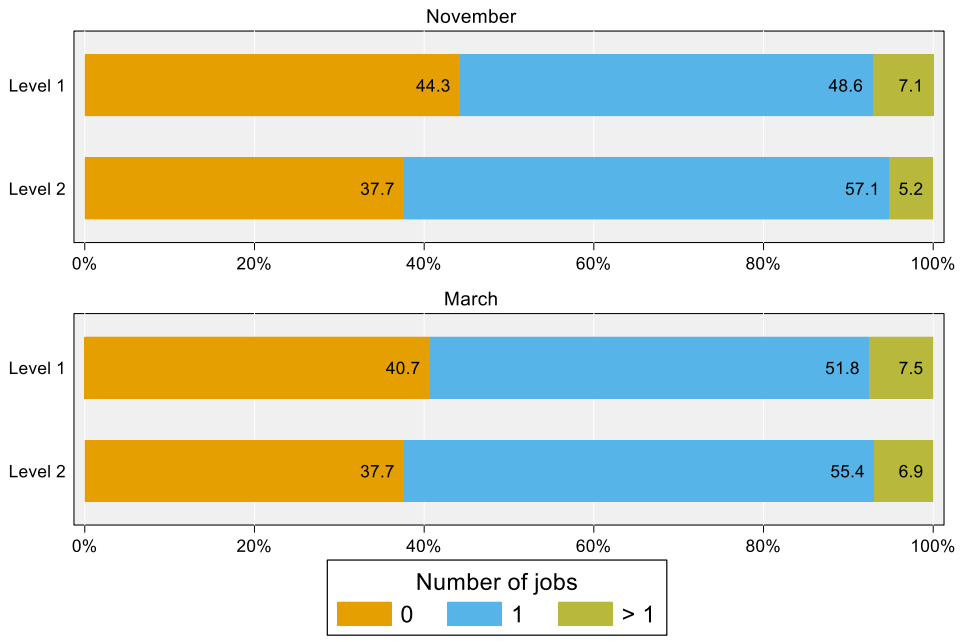
The fact the respondents may have had multiple jobs is explored in Figures 5.9.3 and 5.9.4 by treatment group and S4C level, respectively. They reveal that the great majority had either none or one job in both November and March. The share of respondents with more than one job is largest in November for the group that completed the S4C training and placement (16%).

Figure 5.9.3
Number of jobs by treatment group



Source: Own calculations based on S4C survey.

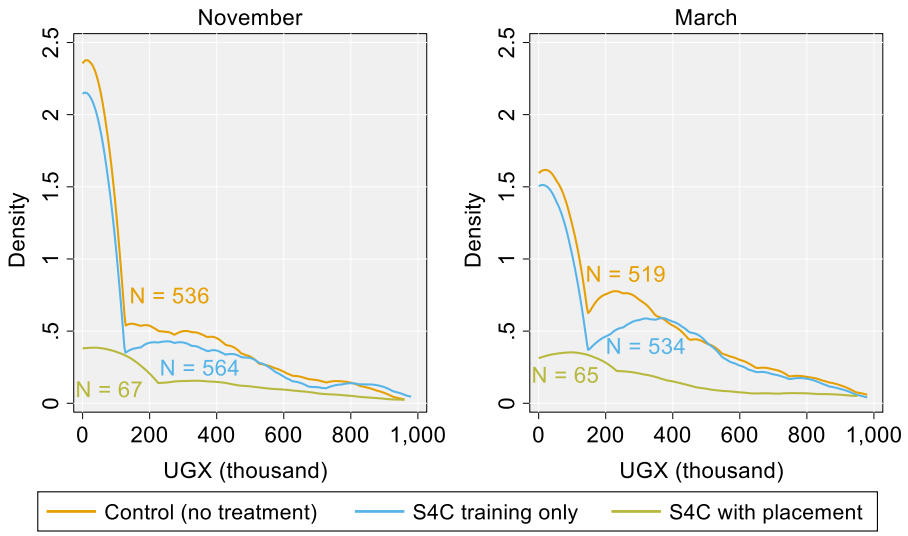
Figure 5.9.4
Number of jobs by training level



Source: Own calculations based on S4C survey.

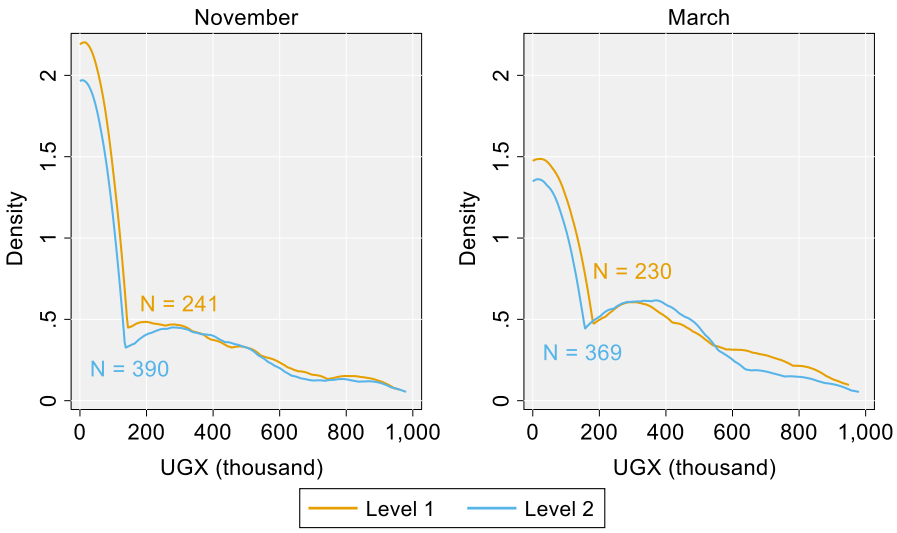
For the assessment of total income, the monthly incomes from all jobs held at the respective time were added up. Figures 5.9.5 and 5.9.6 show the distributions of total monthly income by treatment group and training level. Not surprisingly, the share of zero earnings is higher in November than in March. Total income varied substantially from no income at all to monthly earnings of more than 1,000,000 UGX. The flat distribution of monthly incomes in the S4C plus placement groups indicates that the variation was largest in this group (Figure 5.9.5).

Figure 5.9.5
Total monthly income by treatment group



Source: Own calculations based on S4C survey.

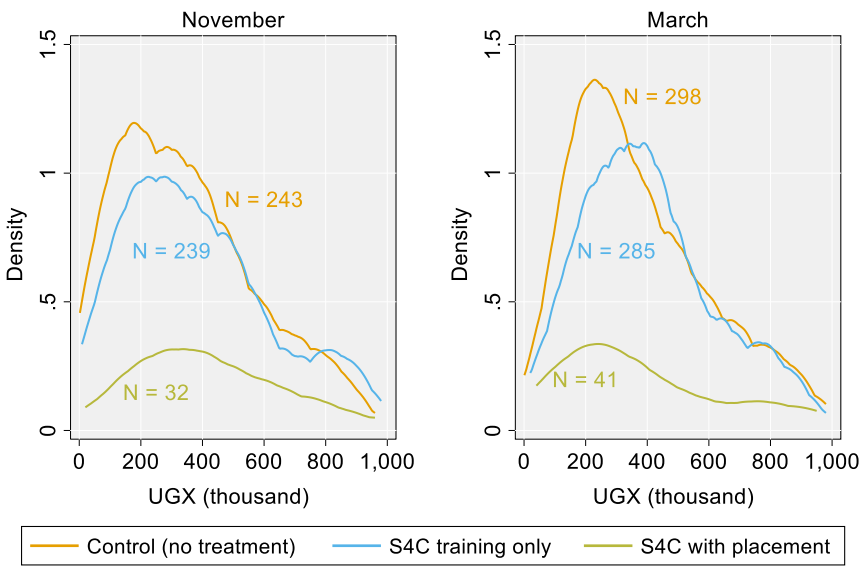
Figure 5.9.6
Total monthly income by training level



Source: Own calculations based on S4C survey.

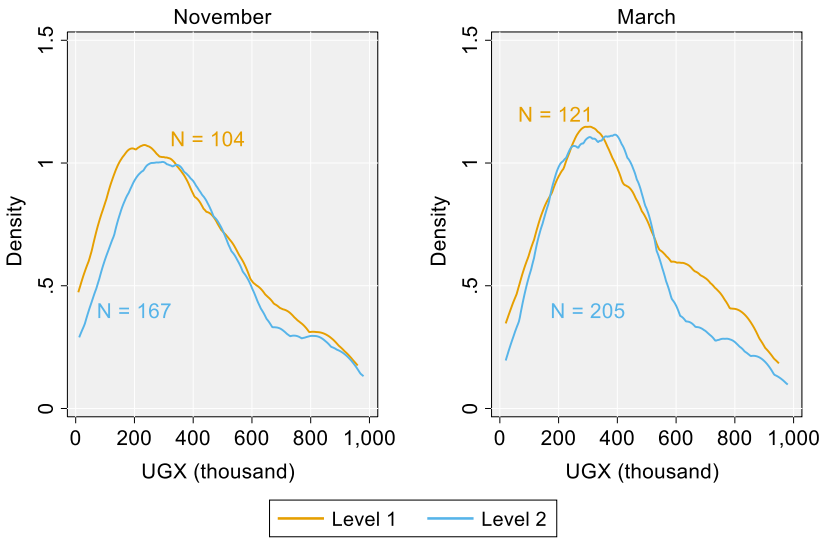
Figures 5.9.7 and 5.9.8 show the distributions of only positive incomes in November and March by treatment group and S4C training level, respectively. Compared to the pure control group, the distribution of income among employed for S4C trainees is slightly shifted to the right, which means they had somewhat higher incomes. Most employed respondents earned roughly between 150,000 and 300,000 UGX per month. Again, the largest variation can be found in the S4C plus placement group (Figure 5.9.7).

Figure 5.9.7
Total monthly income among employed by treatment group



Source: Own calculations based on S4C survey.

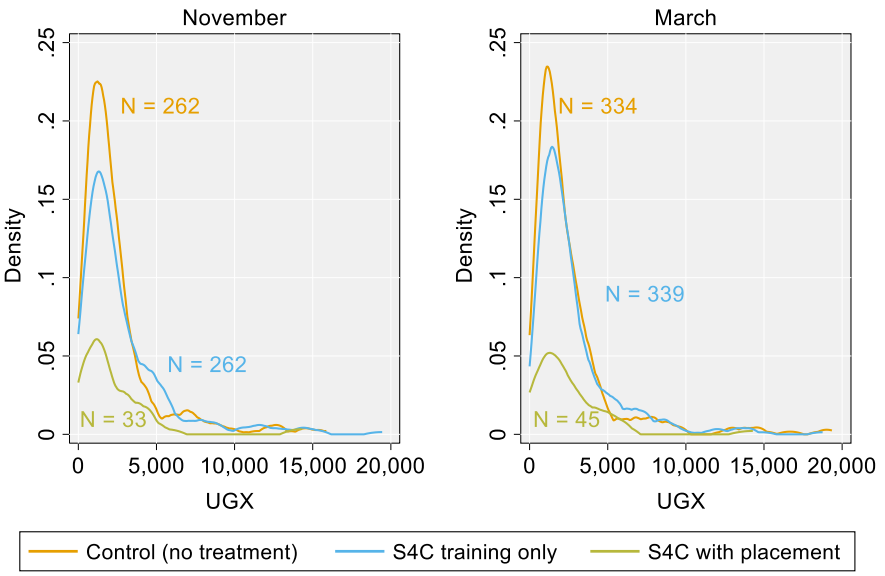
Figure 5.9.8
Total monthly income among employed by training level



Source: Own calculations based on S4C survey.

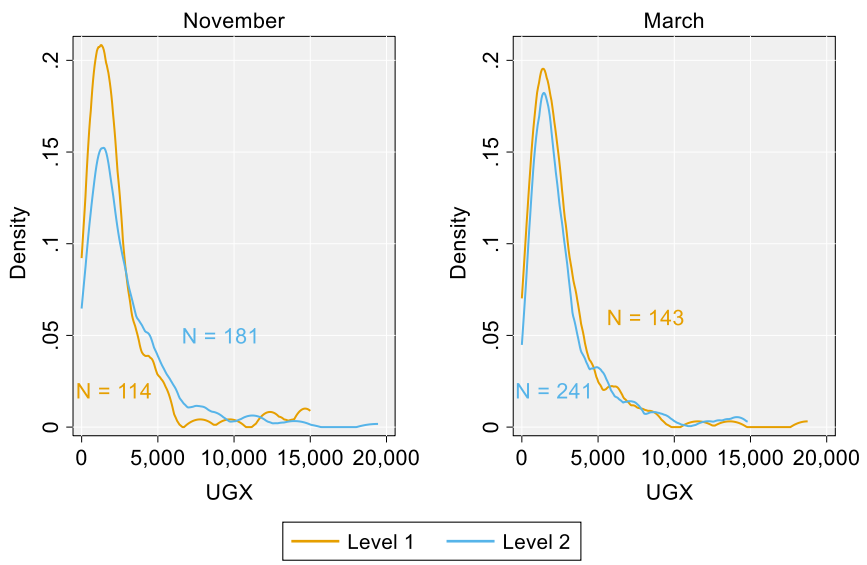
The distribution of hourly wages among employed is depicted in Figures 5.9.9 and 5.9.10 by treatment group and S4C training level, respectively. Most respondents earned about 1,000 to 3,000 UGX per hour.

Figure 5.9.9
Hourly wage among employed by treatment group



Source: Own calculations based on S4C survey.

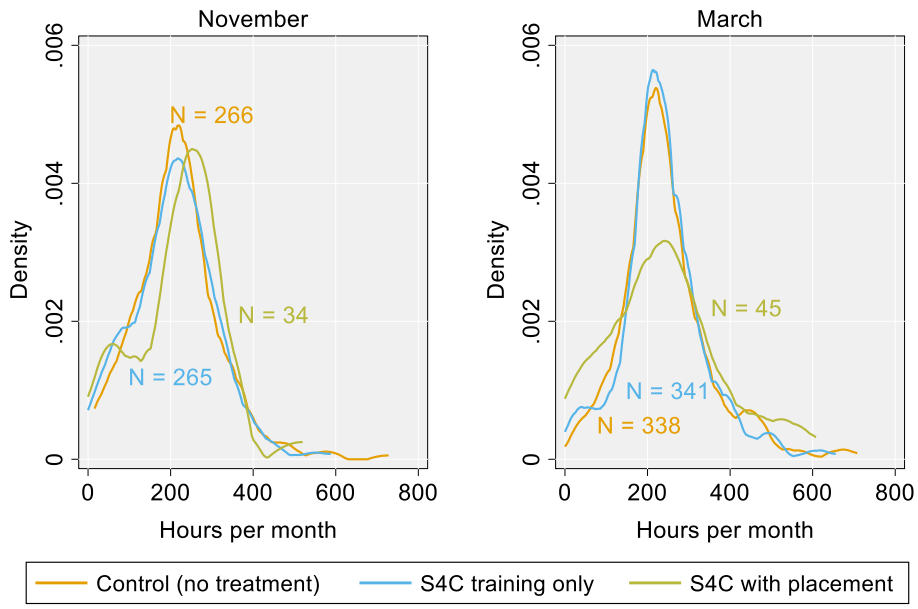
Figure 5.9.10
Hourly wage among employed by training level



Source: Own calculations based on S4C survey.

The distribution of monthly working hours by treatment group and S4C training level is presented in Figures 5.9.11 and 5.9.12. The left-hand side of Figure 5.9.11 shows that, in November, beneficiaries of the S4C plus training group worked longer hours than the other groups.

Figure 5.9.11
Working hours among employed by treatment group

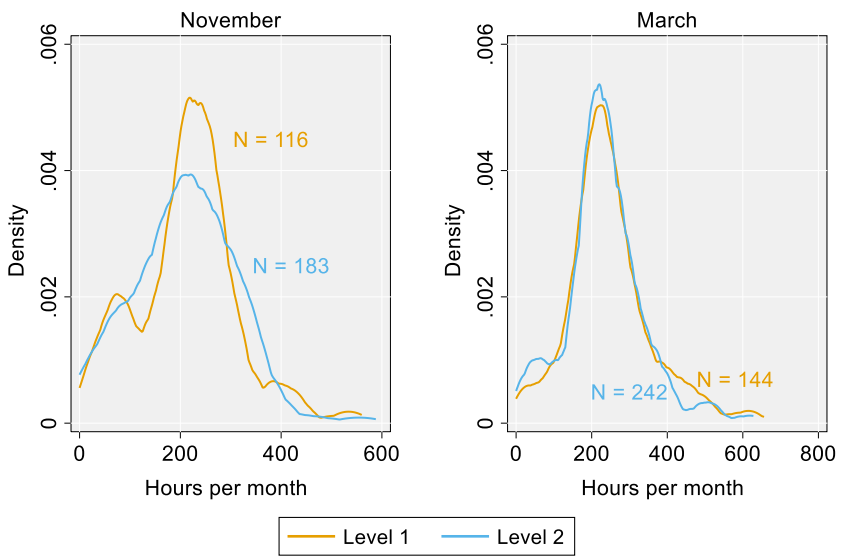


Source: Own calculations based on S4C survey.

In general, working hours varied from just a few to more than 500 hours, with the majority working between 150 and 400 hours per month. Considering fulltime employment as working 40 hours per week, the majority of respondents worked more than fulltime. The assessment of

working hours considers all jobs held by the respondents, but as shown in Figures 5.9.3 and 5.9.4, most of them had only one job. Interestingly, working hours were much more similar in March than in November between the pure control group and S4C training only beneficiaries and between the two training levels.

Figure 5.9.12
Working hours among employed by training level



Source: Own calculations based on S4C survey.

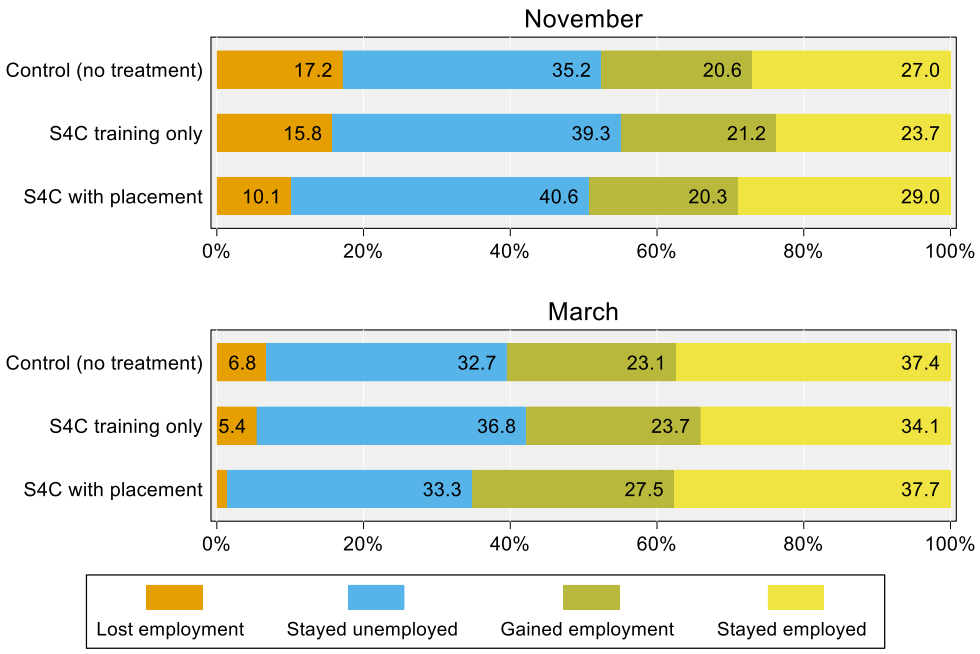
5.9.2 Changes in labor market characteristics between baseline and follow-up

In the survey, information on jobs held and incomes earned in the previous years was elicited. On this basis, baseline outcomes were defined as outcomes at the time when respondents had registered their interest in the S4C program. These outcomes are not yet affected by the program and may have influenced the selection into the training and the placement components. The registration date depended on the cohort and ranged between January and December 2019.

Changes in employment status

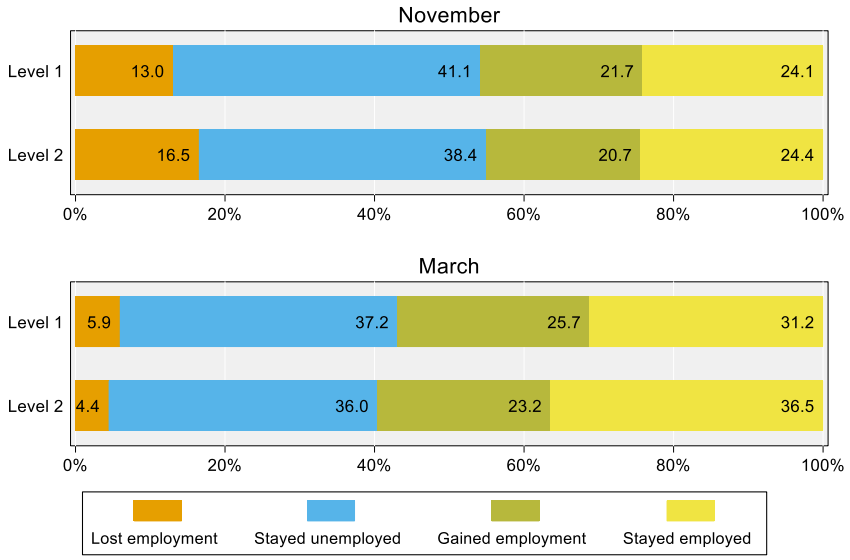
In November 2020, between 10 and 17% of respondents in the different groups had lost their employment held at baseline (Figures 5.9.13 and 5.9.14). The share is largest in the pure control group and smallest in the S4C plus placement group. However, in comparison to the pure control group, 5 percentage points more respondents from the S4C plus placement group stayed unemployed. Across all groups, about 20% had gained employment in November. The general pattern of employment transitions is similar for November and March, but unemployment was more prevalent in November. The overall differences between treatment groups and training levels are rather small.

Figure 5.9.13
Changes in employment status by treatment group



Source: Own calculations based on S4C survey.

Figure 5.9.14
Changes in employment status by training level



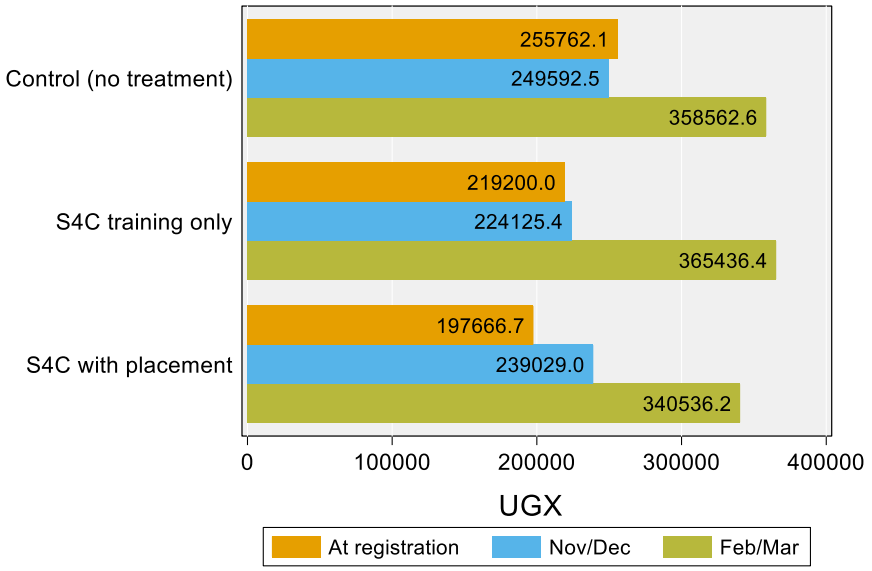
Source: Own calculations based on S4C survey.

Changes in income

Figures 5.9.15 and 5.9.16 show average total income at registration, in November, and in March by treatment group and training level, respectively. Across all groups total income was highest in March. At registration, average total income was highest in the pure control group (about 256,000 UGX) and lowest in the S4C plus placement group. However, beneficiaries in the S4C

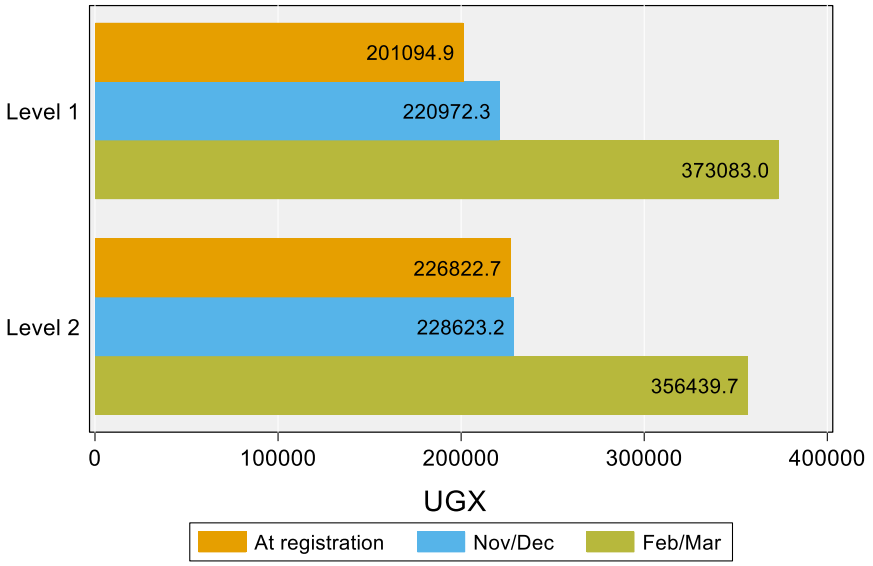
plus placement group could, on average, improve their incomes in November by 41,000 UGX, while they remained unchanged for the other groups.

Figure 5.9.15
Changes in income by treatment group



Source: Own calculations based on S4C survey.

Figure 5.9.16
Changes in income by training level

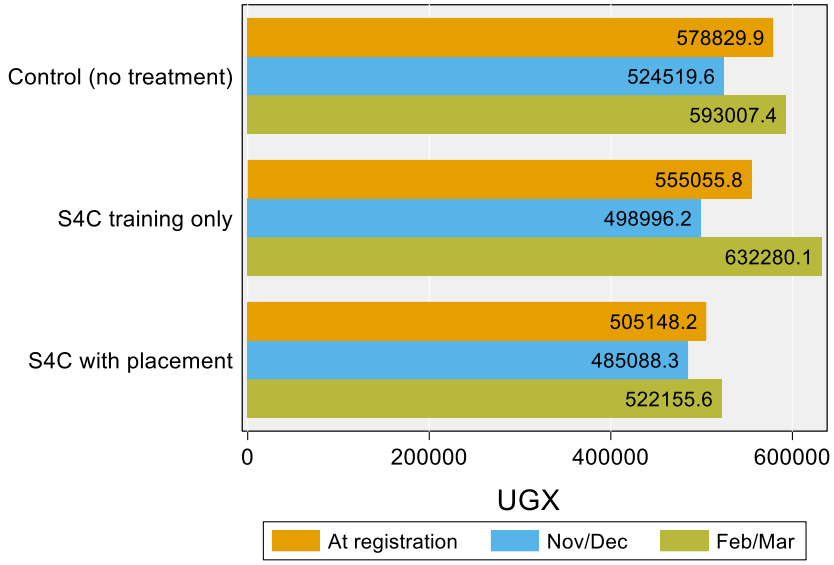


Source: Own calculations based on S4C survey.

Figures 5.9.17 and 5.9.18 present average income among those with paid employment at registration, in November, and in March by treatment group and training level, respectively. Differences between registration and the two endlines are small for all groups (Figures 5.9.17 and 5.9.18). This suggests that the differences in Figures 5.9.15 and 5.9.16 are mainly driven by high

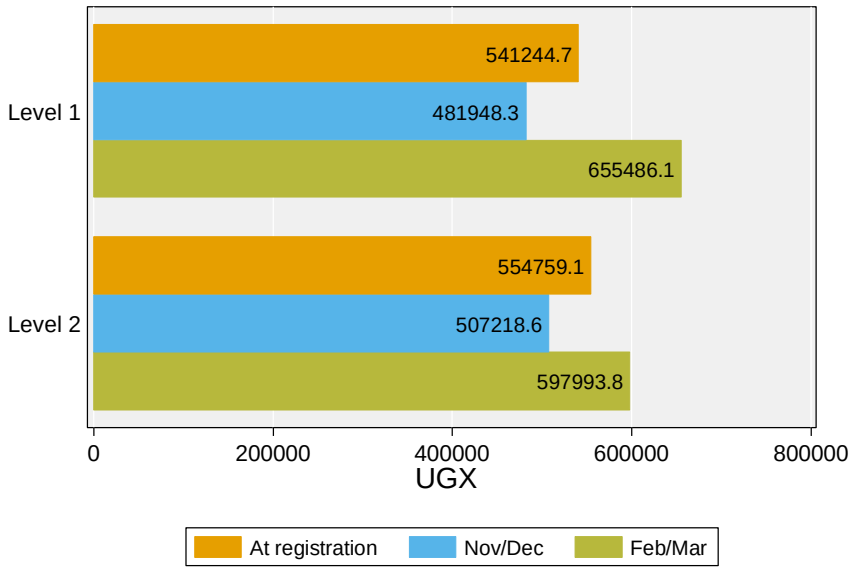
levels of unemployment at registration and in November. Compared to the baseline, income among employed increased in March and decreased in November, implying that the COVID-19 pandemic did not only lead to higher unemployment but also to loss of income for those who managed to keep their job.

Figure 5.9.17
Changes in total income among employed by treatment group



Source: Own calculations based on S4C survey.

Figure 5.9.18
Changes in total income among employed by training level

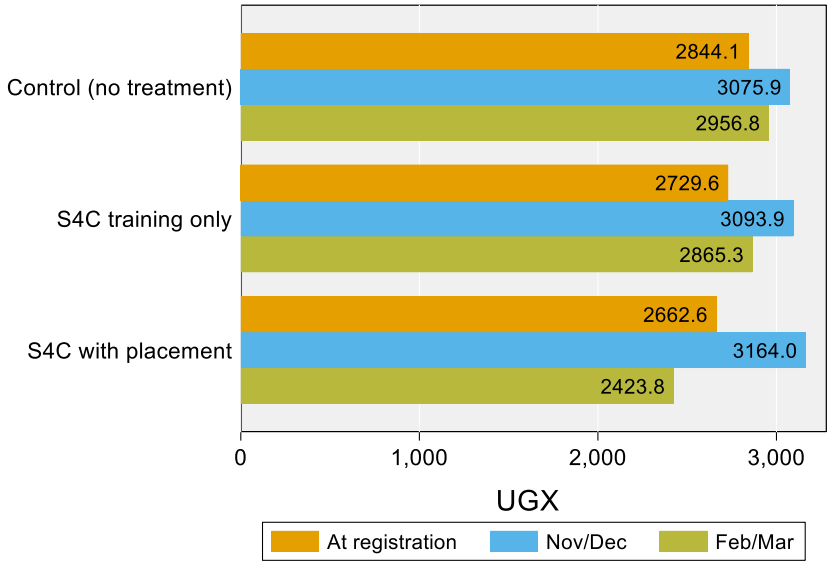


Source: Own calculations based on S4C survey.

Changes in hourly wages and working hours

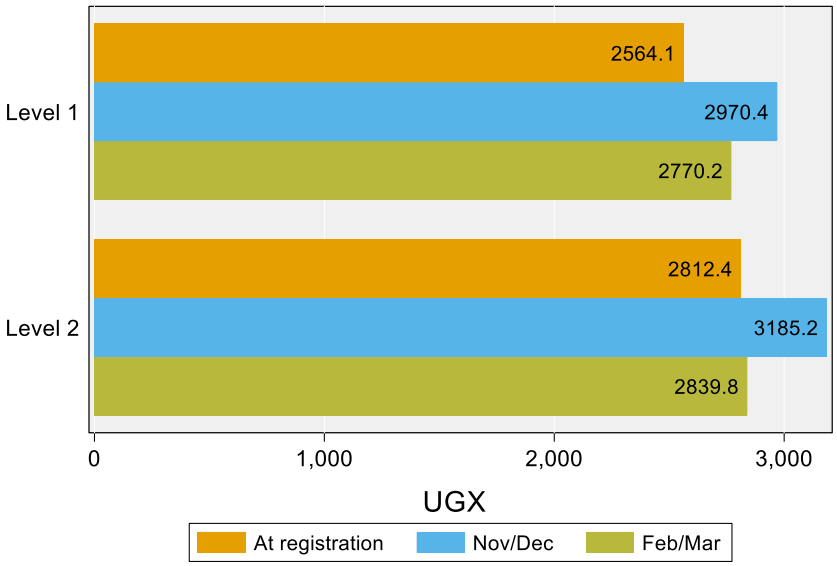
Figures 5.9.19 and 5.9.20 show mean hourly wages among employed at registration, in November, and in March by treatment group and training level, respectively. Interestingly, hourly wages for all groups were highest in November (Figures 5.9.19 and 5.9.20). Comparing wages at registration and in March, there was an increase for all groups except S4C plus placement beneficiaries.

Figure 5.9.19
Changes in hourly wage among employed by treatment group



Source: Own calculations based on S4C survey.

Figure 5.9.20
Changes in hourly wage among employed by training level

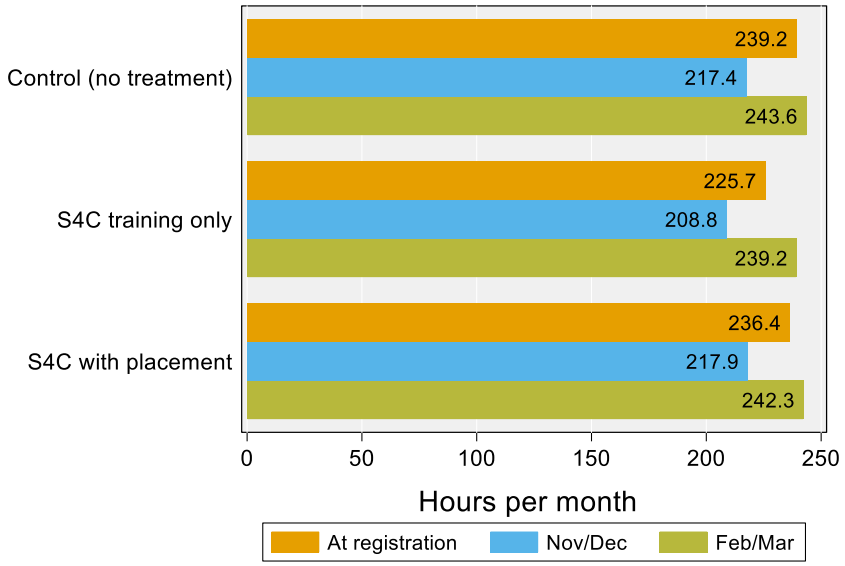


Source: Own calculations based on S4C survey.

By looking at the change in total working hours per month, Figures 5.9.21 and 5.9.22 shed more light on the relationship between total income and hourly wages. Across all groups, working

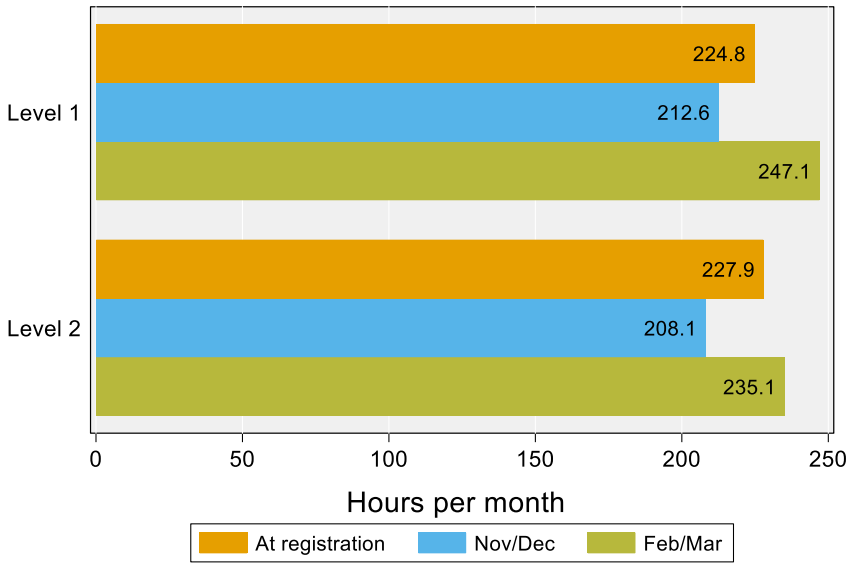
hours among employed were highest in March and lowest in November. The increase in hourly wages in November depicted in Figures 5.9.19 and 5.9.20, thus, seems to be driven by a decrease in working hours.

Figure 5.9.21
Change in working hours among employed by treatment group



Source: Own calculations based on S4C survey.

Figure 5.9.22
Change in working hours among employed by training level

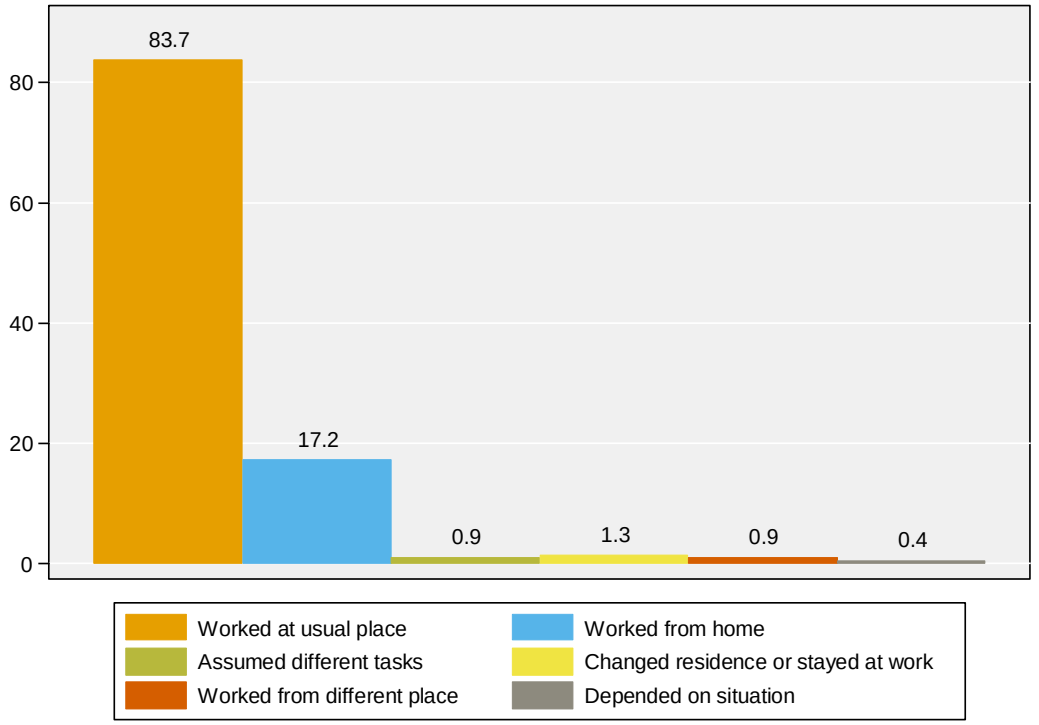


Source: Own calculations based on S4C survey.

Figure 5.9.23 explores how work in Uganda was organized during the COVID-19 pandemic. 84% of study participants who were working worked at their usual place and 17% worked from home.

Other answers to this question were assuming different tasks, changing residence to be closer to the workplace or staying at work, and working from a different place (1% each).

Figure 5.9.23
Organization of work during the pandemic



Source: Own calculations based on S4C survey.

5.10 Estimated program effects on labor market outcomes

This section presents the impact estimations of the S4C program on labor market outcomes based on the estimation methods discussed in section 5.3.1. Specifically, it presents the estimated impact of a) having participated in any component of the S4C program compared to no participation at all (RQ#1), b) having participated in both the S4C training and the placement component compared to the S4C training only (RQ#2), and c) having participated in both S4C level 1 and level 2 training compared to only level 1 training (RQ#3).

The figures in this section show estimated treatment effects, also referred to as coefficients or point estimates, as well as 95% confidence intervals of the estimated effects. Confidence intervals are a measure of precision of the effect estimates and provide a range of effect estimates in which the true parameter, i.e., the true program effect, might be. Here, we use 95% intervals, which means that the estimated confidence interval will contain the true value of the treatment effect with a probability of 95%. If a confidence interval does not include zero, the estimated treatment effect is said to be significantly different from zero and, thus, considered to have a significant positive (or negative) effect on the respective outcome.

All presented employment measures are binary indicators which only have two categories or levels. Effect estimations on employment outcomes can be interpreted as *percentage point* changes in the respective employment indicator. For example, a coefficient of 0.1 can be interpreted as a 10-percentage point increase in the employment rate due to the intervention, i.e., S4C training (RQ#1), S4C placement in addition to the training (RQ#2), or S4C level 2 training in addition to level 1 training (RQ#3).

Income and wage outcomes were measured in Ugandan shilling (UGX) and logarithmized using the natural logarithm (referred to as "Ln of [income indicator]" in the figures that follow). Using logarithms of income and wage variables has the advantage that the estimated program impacts can be interpreted as *percent changes* and compared across different contexts and currencies. An estimated effect of 0.1, for example, indicates a 10% increase in income due to the intervention, i.e., S4C training (RQ#1), S4C placement in addition to the training (RQ#2), or S4C level 2 training in addition to level 1 training (RQ#3).

Effects greater than 0.3 (or smaller than –0.3) cannot be directly interpreted as percentage changes (although they are still a close approximation) but require conversions that are detailed in the respective figure notes, if applicable.[61]

Working hours are measured as continuous variables and their coefficients can be interpreted as a change in hours worked per month due to the intervention, i.e., S4C training (RQ#1), S4C placement in addition to the training (RQ#2), or S4C level 2 training in addition to level 1 training (RQ#3).

Six effect estimates are presented for each outcome, which results from the two endlines we consider – one in November/ December 2020, when the survey was conducted, and one pre-COVID endline in February/ March 2020 – and the three comparisons we make based on the three research questions. Effect estimates for RQ#1 and RQ#2 are presented together in the same graphs, whereas results for RQ#3 are presented subsequently in a separate section. The reason is that confidence intervals in the estimations related to RQ#3 were much larger such that estimates for RQ#1 and RQ#2 would have been hard to distinguish if all three were presented in the same graph.

Two additional analyses were conducted as robustness checks. First, respondents who dropped out of S4C level 1 training were added to the pure control group and those who did not complete S4C level 2 training were considered to have received only level 1 training. The results of this analysis are presented in Appendix A4.1. Second, an additional endline was defined, in which outcomes were measured in November/ December 2020, but respondents who indicated to have lost their job or income due to the pandemic were still counted as employed or earning their pre-pandemic income, respectively. The results of this additional analysis are discussed in section 5.10.3.

5.10.1 Program impact on employment status

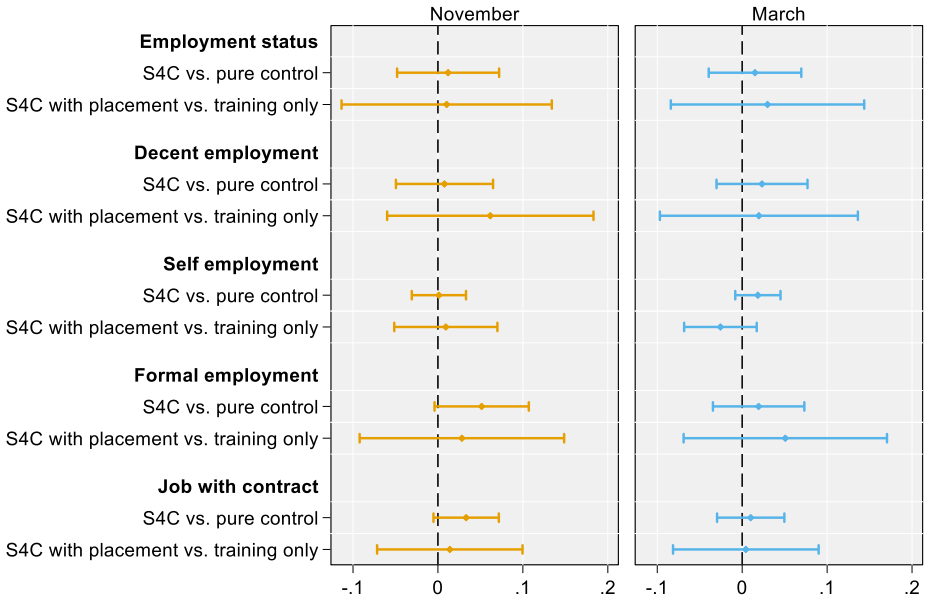
Impact of the S4C training and the placement component on employment status

Figure 5.10.1 shows the estimated impacts of the S4C training and the additional effect of the placement on (i) having paid work (employment status), (ii) having decent employment, (iii) being self-employed, (iv) being in formal employment, and (v) having a job with contract in November/ December 2020 and in February/ March 2020.

For both endlines, most point estimates are positive, but all confidence intervals include zero. Overall, Figure 5.10.1 suggests that the S4C training had no significant short- to medium-term impact on employment outcomes and the S4C placement had no significant additional effect.

[61] *The conversion follows the following formular:* $(e^{coef} - 1) * 100 = program\ effect.$

Figure 5.10.1
Estimated treatment effects of S4C training and placement on binary employment indicators

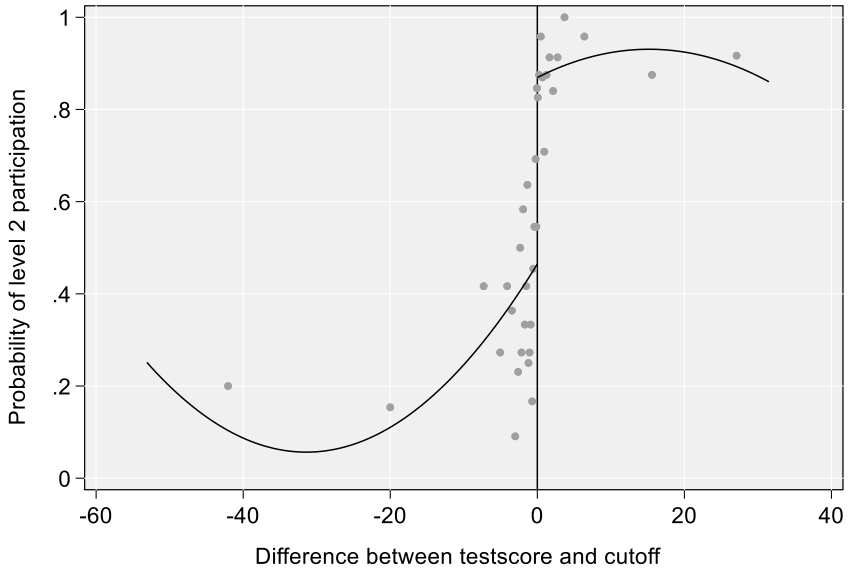


Source: Own calculations based on S4C survey.

Additional impact of S4C level 2 training

Figure 5.10.2 explores whether the decision rule for level 2 participation was indeed implemented and can, thus, be exploited to assess the additional effect of level 2 training on labor market outcomes. It plots the probability of level 2 participation against the difference in respondents' level 1 test score to the cutoff score required to proceed to level 2.

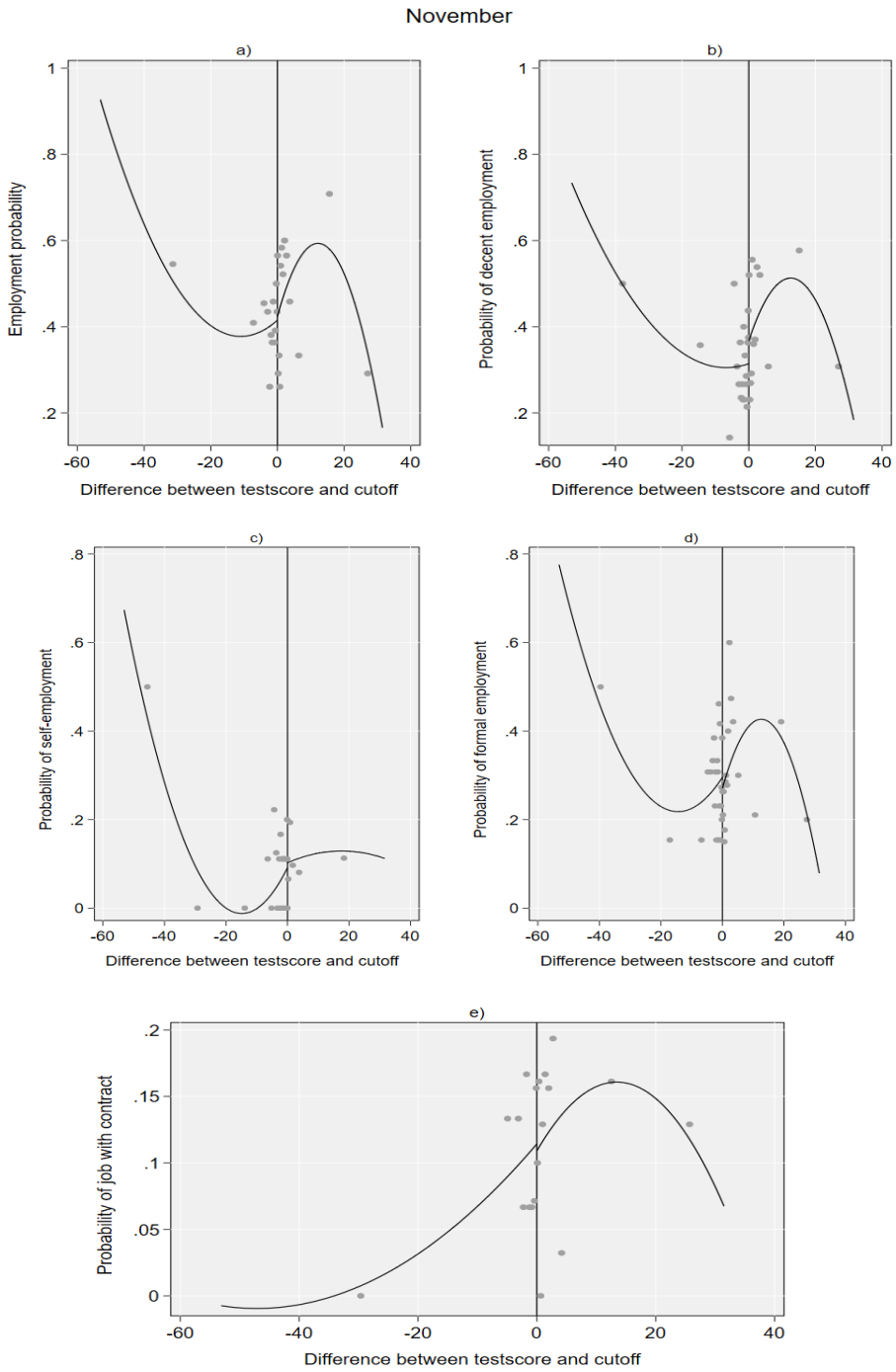
Figure 5.10.2
Level 1 test scores and probability of level 2 participation



Source: Own calculations based on S4C survey.

Additionally, it shows a quadratic fit of the relationship between level 2 participation and tests scores below and above the cutoff. If the decision rule had been strictly implemented, there would be a jump from 0 to 1 in the probability of level 2 participation at the cutoff. The figure reveals that, in practice, admission to S4C level 2 training deviated from the decision rule. But still, the probability of level 2 participation increases by 40 percentage points at the cutoff. This means that the rule was partially implemented and can, therefore, be used to identify the effect of level 2 training in a fuzzy regression discontinuity design.

Figure 5.10.3
S4C level 1 test scores and employment outcomes in November/December

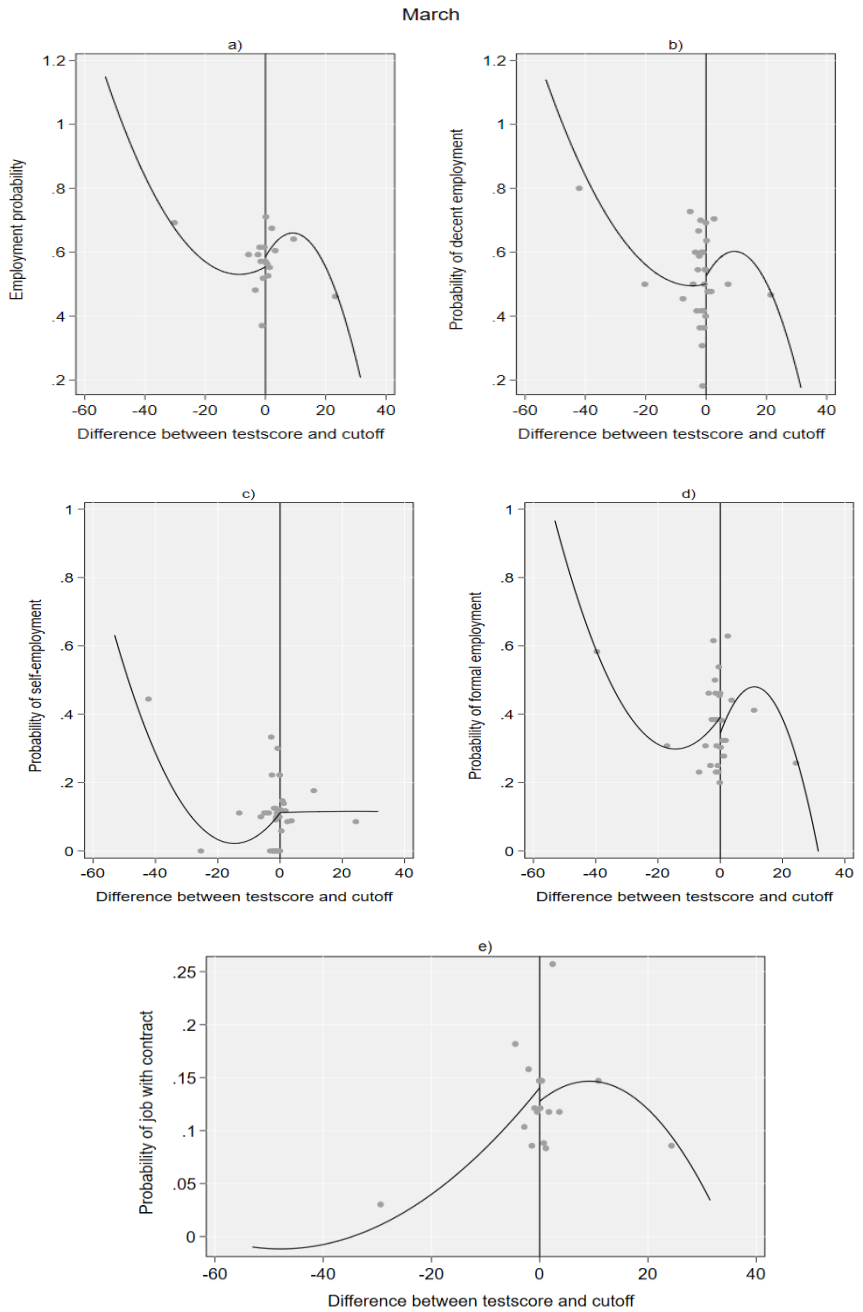


Source: Own calculations based on S4C survey.

Figures 5.10.3 and 5.10.4 show plots for the relationship between level 1 test scores and employment outcomes. A jump in this relationship at the cutoff hints at an impact of level 2 participation on the outcome. Whether this impact is significant, however, must be assessed by regression analyses.

Figure 5.10.3 shows that, in November, S4C trainees who had just passed the threshold of level 1 test scores required for S4C level 2 participation were about 7 percentage points more likely to be in decent employment than trainees who had just missed the threshold (panel b)). For the remaining employment outcomes, the differences between respondents just above and just below the threshold are very small.

Figure 5.10.4
S4C level 1 test scores and employment outcomes in February/March



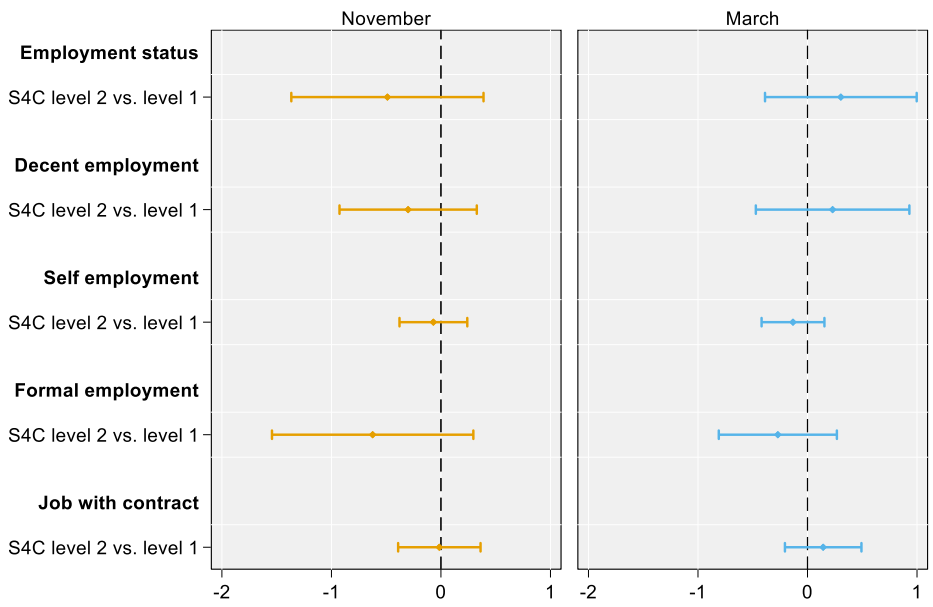
Source: Own calculations based on S4C survey.

In March, as shown by Figure 5.10.4, trainees just above the cutoff of S4C level 1 test scores were about 2 percentage points more likely to be employed and to have decent employment than those just below the cutoff (panels a) and b)). The probabilities of formal employment and having a job with contract, by contrast, were slightly smaller for respondents above the threshold.

The results of the regression analyses for the additional impact of S4C level 2 training are presented in Figure 5.10.5. Estimates of the additional impact of S4C level 2 training on employment outcomes are very imprecise, as evidenced by the large confidence intervals. This is partly due to the small sample size used in this analysis. Compared to the previous section, the sample size for this analysis is reduced for two reasons: (i) the sample comprises only S4C trainees – level 1 only trainees in the control group and level 2 trainees in the treatment group, (ii) the estimation only relies on trainees with level 1 test scores close to the cutoff as only these are comparable (see section 5.3.1 for a more extensive explanation).

For the November-endline, all point estimates are negative. When considering the pre-COVID endline, the coefficients for employment status, decent employment and job with contract become positive. However, all confidence intervals include zero and Figure 5.10.5, thus, indicates that the S4C level 2 training, in comparison to the S4C level 1 training only, had no additional impact on employment outcomes in the short to medium term.

Figure 5.10.5
Estimated treatment effects of S4C level 2 training on binary outcome indicators



Source: Own calculations based on S4C survey.

In some cases, the jumps in the relationship between test scores and outcome variables in Figures 5.10.3 and 5.10.4 and the point estimates in Figure 5.10.5 have opposing directions, e.g., the jump shows an increase in the outcome whereas the coefficient shows a decrease in the outcome. This is because Figures 5.10.3 and 5.10.4 do not account for the possibility that beneficiaries in the two groups may not be directly comparable due to differences in important background characteristics, such as education or household wealth. Figure 5.10.5, by contrast, corrects for these differences and, thus, restores the comparability between the two groups.

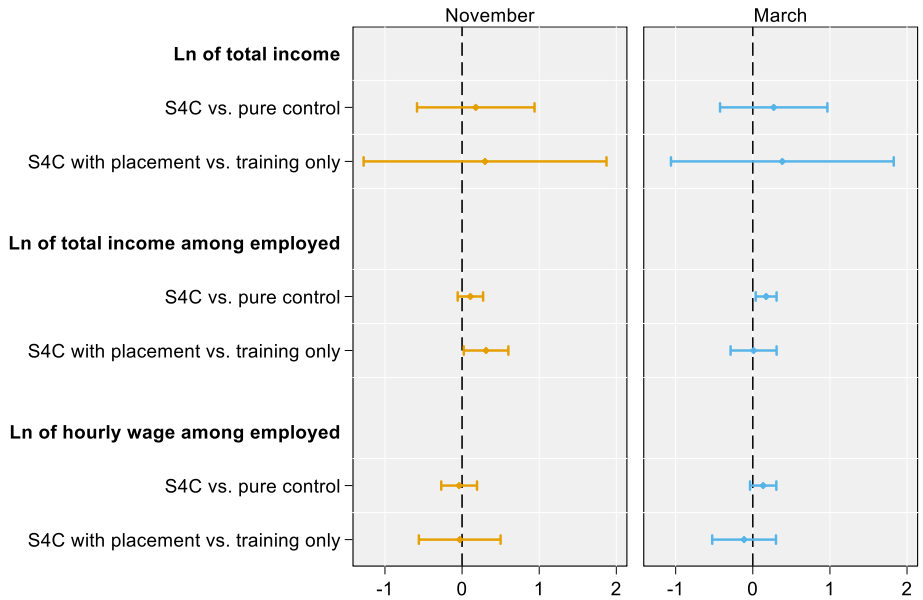
5.10.2 Program impact on incomes and wages

Impact of the S4C training and the placement component on incomes and wages

Figure 5.10.6 presents the effect estimates of the S4C training and the S4C placement component on income earned from all jobs among all respondents and among employed respondents as well as the average hourly wage from all jobs among employed respondents in November/ December 2020 and February/ March 2020.

Most confidence intervals in Figure 5.10.6 include zero. However, there was a significantly positive impact of the S4C placement (compared to S4C training only) on total income among employed in November and a significantly positive effect of the S4C program (compared to the pure control group) on total income among employed in March. Figure 5.10.6, thus, shows that participation in the S4C program led to 17% higher income among those who were gainfully employed in March 2020 and completion of an S4C placement in addition to the training increased income for those who were employed in November 2020 by 36%.

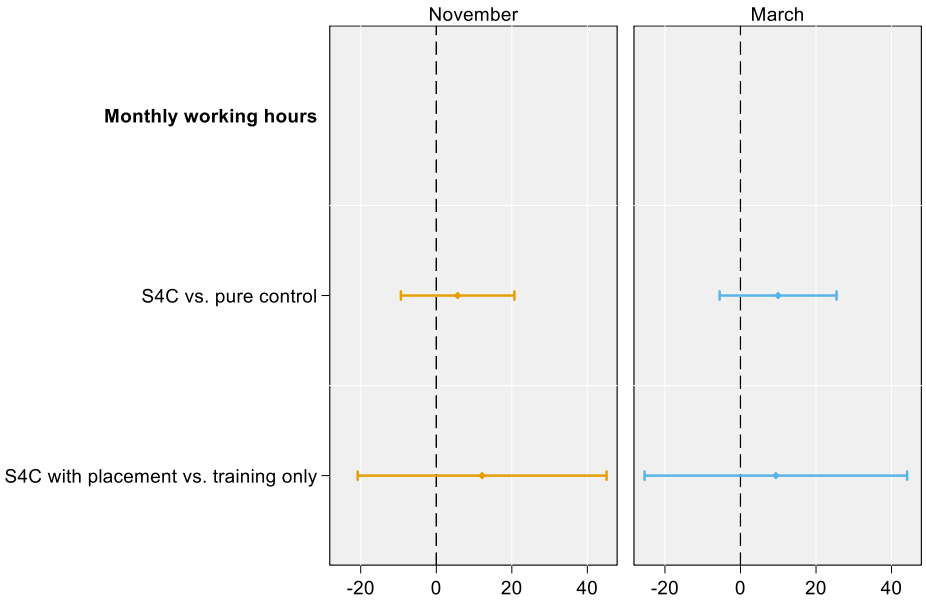
Figure 5.10.6
Estimated treatment effects of S4C training and placement on total income, income among employed and wages among employed



*Notes: The coefficient for "S4C with placement vs. training only" on ln of total income among employed in November translates to an income increase of (exp(0.31) – 1)*100 = 36%. - Source: Own calculations based on S4C survey.*

Figure 5.10.7 presents the estimated effects of the S4C training and the S4C placement component on working hours among employed respondents in November/ December 2020 and February/ March 2020. The treatment estimates vary between 5 and 12 hours per month and all confidence intervals include zero, suggesting that the S4C training and placement component did not significantly impact working hours.

Figure 5.10.7
Estimated treatment effects of S4C training and placement on hours worked among employed



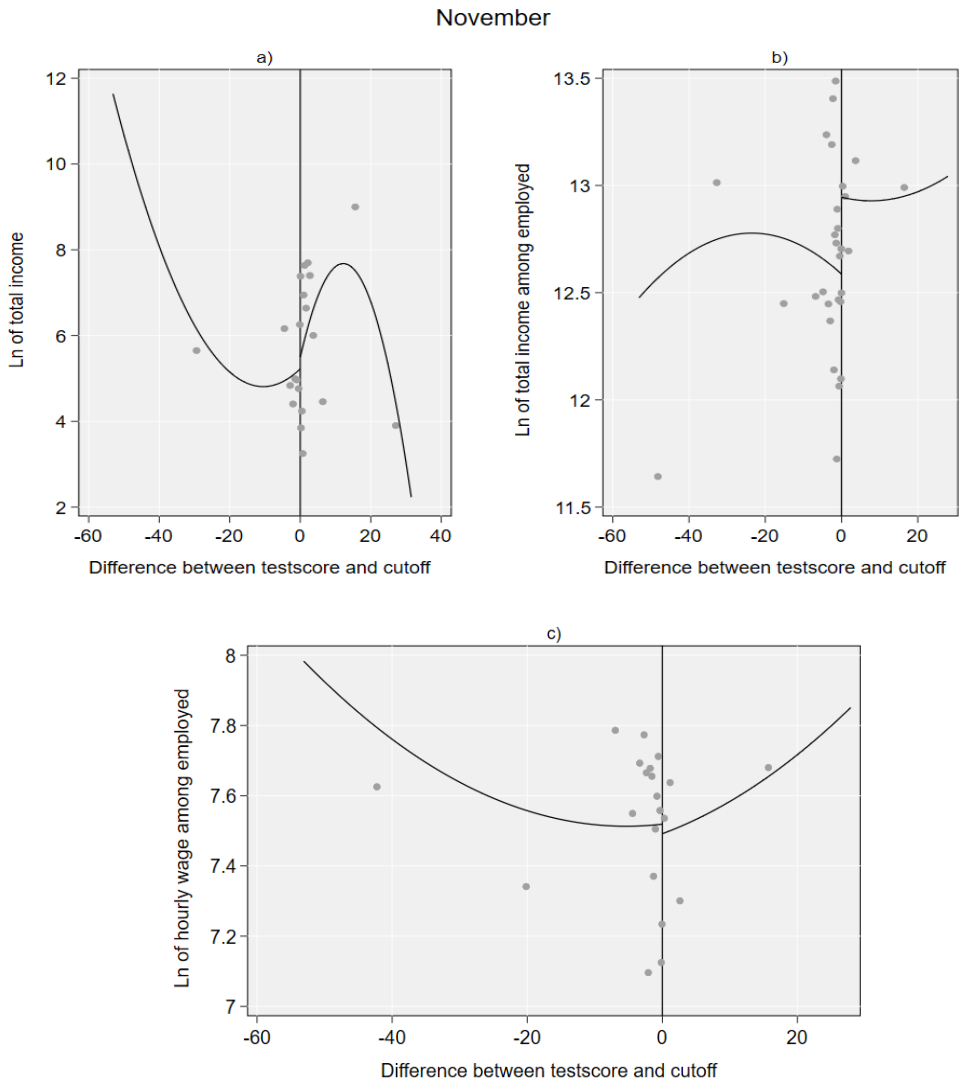
Source: Own calculations based on S4C survey.

Although no significant impacts on working hours could be detected, there is a tendency for treated individuals to work longer hours per month. Hourly wages did not change significantly either. This suggests that the increase in income among employed is driven by a mixture of higher wages and longer working hours. For the placement component, the impact of working hours seems to be larger, whereas in the assessment of the S4C program as a whole, wages tend to increase as well.

Additional impact of S4C level 2 training on incomes and wages

Figures 5.10.8 and 5.10.9 plot total income, total income among employed, and mean hourly wage among employed against level 1 test scores. Figure 5.10.8 shows that, in November, trainees with test scores just above the cutoff had higher income, higher income among employed, and higher hourly wages than trainees just below the cutoff.

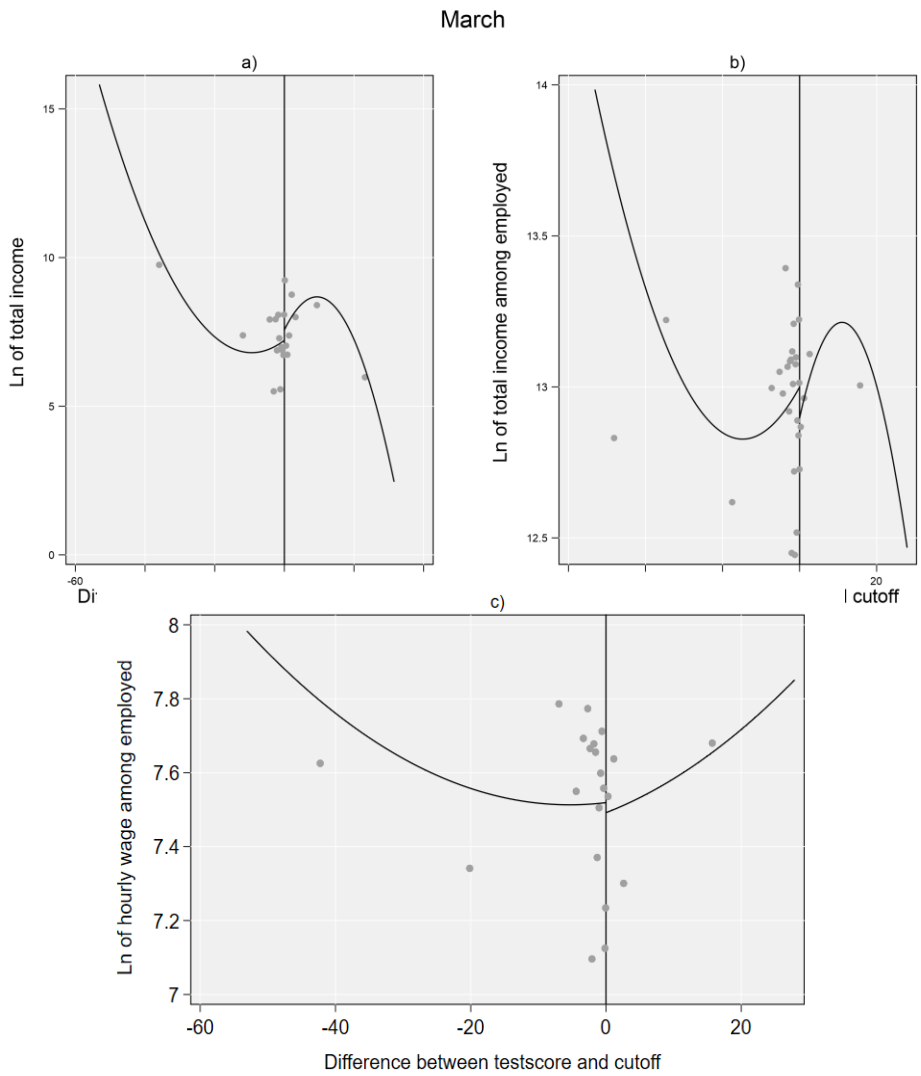
Figure 5.10.8
S4C level 1 test score, incomes, and wages in November



Source: Own calculations based on S4C survey.

For March 2020, Figure 5.10.9 shows that, overall, trainees whose S4C level 1 test scores were just above the cutoff value earned slightly higher incomes than those who scored just below this value. Total income among employed and hourly wages, by contrast, were somewhat lower for those above the threshold.

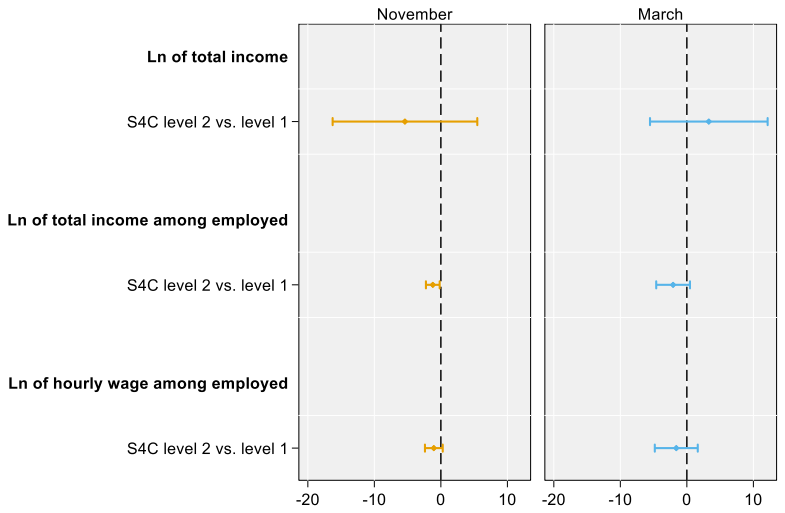
Figure 5.10.9
Level 1 test score, incomes, and wages in March



Source: Own calculations based on S4C survey.

Figure 5.10.10 shows the respective estimated treatment effects on incomes and wages. Because the effects are very imprecisely estimated, point estimates and confidence intervals, particularly for total income, are huge. There is a significantly negative effect of S4C level 2 training on total income among employed. Potentially, this is due to the timing of the training end relative to the endline. S4C level 2 trainees completed their training about one month after S4C level 1 trainees. This means that S4C level 1 trainees had been longer in their job by November and, thus, may have had higher job security, resulting potentially in fewer lay-offs, wage cuts or reductions in working hours. All other confidence intervals include zero, implying that S4C level 2 training had no additional impact on incomes and wages in the short to medium term.

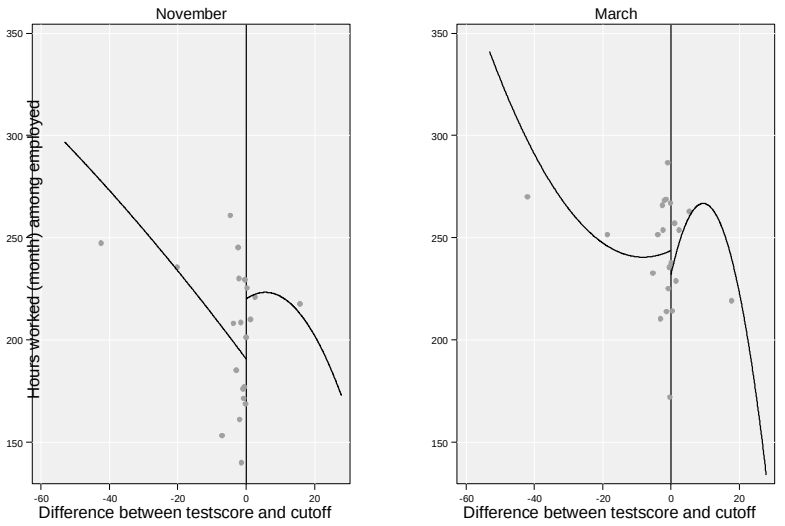
Figure 5.10.10
Estimated treatment effects of S4C level 2 training on incomes and wages



*Notes: The coefficient on ln of total income in November translates to a decrease of (exp(-5.4) – 1)*100=-100%. The coefficient on ln of total income among employed in November translates to a decrease of (exp(-1,215) – 1)*100=-70%. The coefficient on ln of hourly wage among employed in November translates to a decrease of (exp(-1.06) – 1)*100=-65%. The coefficient on ln of total income in March translates to an increase of (exp(3.302) – 1)*100)=2617%. The coefficient on ln of total income among employed in March translates to a decrease of (exp(-2.069) – 1)*100=-87%. The coefficient on ln of hourly wage among employed in March translates to a decrease of (exp(-1.589) – 1)*100=-80%. - Source: Own calculations based on S4C survey.*

Figures 5.10.11 and 5.10.12 explore the relationship between S4C level 2 participation and monthly working hours. In November, trainees who scored just above the cutoff level in their final test of S4C level 1 training worked about 10 hours more per month than those who just missed the cutoff level. In March, however, trainees above the threshold worked slightly shorter hours.

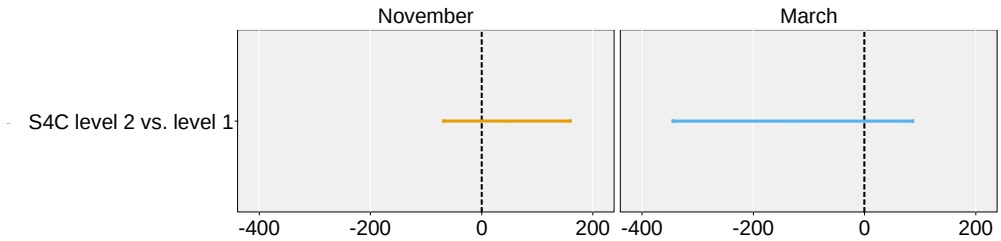
Figure 5.10.11
Level 1 test scores and monthly working hours



Source: Own calculations based on S4C survey.

Point estimates for monthly working hours (Figure 5.10.12) are large and negative for March (-134 hours) and positive for November (45 hours), but imprecisely estimated. The analysis, thus, suggests that, compared to S4C level 1 training only, S4C level 2 training had no significant additional effect on trainees' monthly working hours.

Figure 5.10.12
Estimated treatment effects of S4C level 2 training on monthly working hours



Source: Own calculations based on S4C survey.

5.10.3 Robustness checks

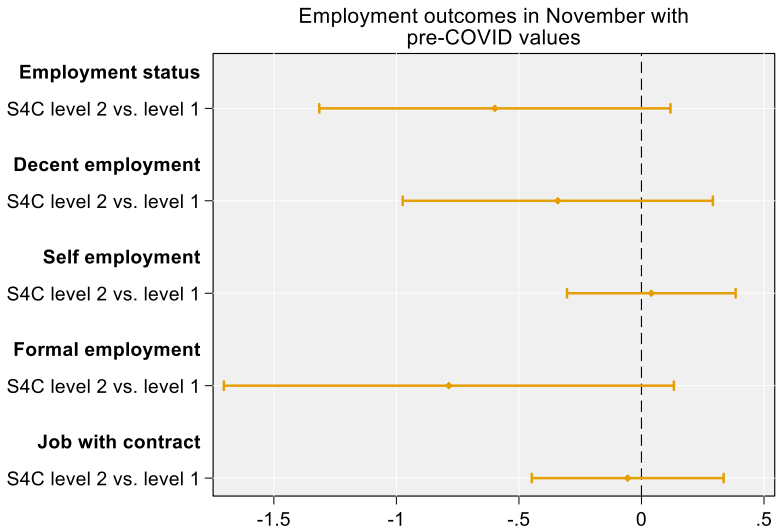
To assess the robustness of the results presented in sections 5.10.1 and 5.10.2, two additional analyses were conducted. First, those respondents who did not complete S4C level 1 training were added to the pure control group and those who dropped out of S4C level 2 training were considered to have received only S4C level 1 training. The results are very similar to those of the main analysis and are presented in Appendix A4.1. Second, labor market outcomes were evaluated with respect to an alternative endline: Outcomes were measured in November/ December 2020, but the pre-pandemic outcome was used for study participants whose labor market outcomes had been negatively impacted by the COVID-19 pandemic. The intention of this approach was to address both the potential impact on outcomes resulting from the pandemic (if measuring them in November/ December) and from not having had enough time to search for a job (if measuring them in February/ March).

Figure 5.10.13
Estimated treatment effects of S4C training and placement on binary outcomes for alternative November endline

Source: Own calculations based on S4C survey.

Estimation results for binary employment indicators are presented in Figures 5.10.13 and 5.10.14. Figure 5.10 13 shows coefficient estimates for the S4C program and its placement component and Figure 5.10.14 for S4C level 2 training. Estimated effects are very similar to those found when considering actual employment outcomes measured in November. In comparison to the results for the November endline, the point estimates for the S4C program and the placement component in Figure 5.10.13 are somewhat larger. However, all estimates fail to reach significance.

Figure 5.10.14
Estimated treatment effects of S4C level 2 training on binary indicators for alternative November endline



Source: Own calculations based on S4C survey.

Figures 5.10.15 and 5.10.16 present estimated treatment effects on incomes and wages. Figure 5.10.15 refers to the impact of the S4C program and its placement component and Figure 5.10.16 to S4C level 2 training. Similarly to binary employment indicators, the pattern of results very much resembles the regular November endline.

Figure 5.10.15
Estimated treatment effects of S4C training and placement on incomes and wages for alternative November endline

*Notes: The coefficient of "S4C with placement vs. training only" on Ln of total income translates to an increase of (exp(0.562) – 1)*100=75%. - Source: Own calculations based on S4C survey.*

However, the S4C program as a whole (compared to the pure control group) had a significantly positive impact of 16% on total income among employed, which is remarkably close to the result found for the pre-COVID endline in March. The additional impact of the S4C placement, by contrast fails to reach significance. With a point estimate of 11%, it lies between the estimated effects for the regular November endline (36%) and March (1%).

For the impact of S4C level 2 training on total income among employed (Figure 5.10.16), the significantly negative effect already detected for the regular November endline (Figure 5.10.10) remains. As explained in section 5.10.2, this is most likely due to the relative timing of the end of training and the pandemic.

Figure 5.10.16
Estimated treatment effects of S4C level 2 training on incomes and wages for alternative November endline

*Notes: The coefficient on ln of total income translates to a decrease of (exp(-4.957) – 1)*100= -99%. The coefficient on ln of total income among employed translates to a decrease of (exp(-1.772) – 1)*100= -83%. The coefficient on ln of hourly wage among employed translates to a decrease of (exp(-0.794) – 1)*100= -55%.* - Source: Own calculations based on S4C survey.

In short, the robustness checks confirm the results presented in sections 5.10.1 and 5.10.2.

5.11 Program impact for sub-groups of S4C respondents

The estimated treatment effects discussed in section 5.10 present average treatment effects for the whole study population. However, the treatment effects for specific sub-populations may differ from those for the whole sample. In this section, the sample is divided into sub-samples of participants with specific background characteristics and the impact of the S4C training program and the placement component is estimated for those sub-populations. The additional impact of the S4C level 2 training is not analyzed in this section because sample sizes would become too small to obtain meaningful estimates. The background characteristics under consideration are:

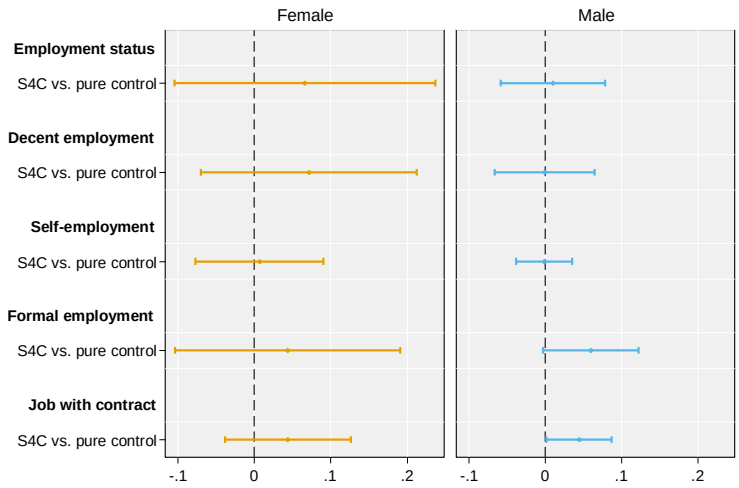
1. **female** respondents versus **male** respondents,
2. respondents aged **less than 25 years** versus respondents **25 years and older**,
3. respondents who have **participated in a TVET program** versus respondents who were **never inscribed in a TVET program**,
4. respondents who have a **sustained work experience of at least 6 months** versus respondents who do not have a sustained work experience, and
5. respondents who have any **previous work experience in the construction sector** (could be a paid or unpaid job or internship) versus respondents that do not have construction sector experience.

The figures in this section only include treatment effects on employment indicators, incomes, and wages for November. The results for March are presented in Appendix A4.2.

5.11.1 Program effect heterogeneity by respondents' gender

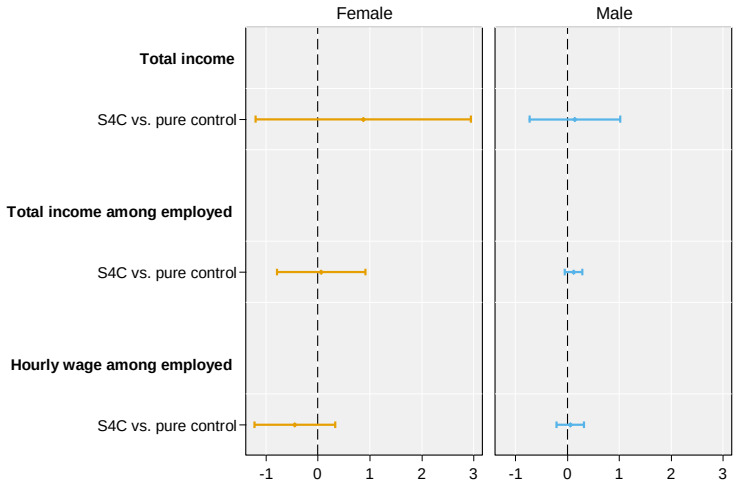
This section explores the differential impact of the S4C program by respondents' gender. The additional impact of the placement component cannot be evaluated as only 6 female study participants obtained a placement. Figure 5.11.1 presents the estimated impact of the S4C program on employment outcomes for female respondents on the left-hand side and for male respondents on the right-hand side. Figure 5.11.2 shows the impact on incomes and wages for the same subsamples. While there are no significant effects for female study participants, participation in the program increased the likelihood of having a job with contract by 4 percentage points for male respondents. However, overall, the effect sizes for females and males are very similar and, except for the difference in significance for formal employment, are also insignificant. These results do not suggest differential impacts by respondents' gender.

Figure 5.11.1
Program effect heterogeneity on binary outcomes by respondents' gender in November



Source: Own calculations based on S4C survey.

Figure 5.11.2
Program effect heterogeneity on incomes and wages by respondents' gender in November

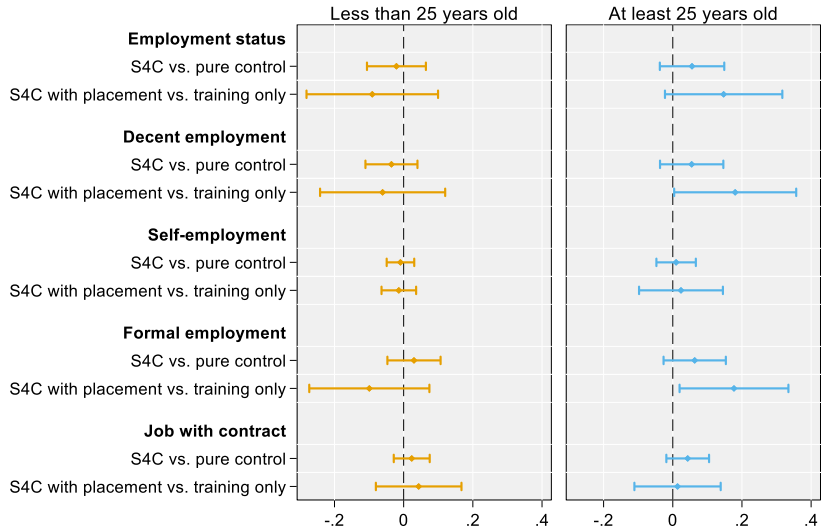


Source: Own calculations based on S4C survey.

5.11.2 Program effect heterogeneity by respondents' age

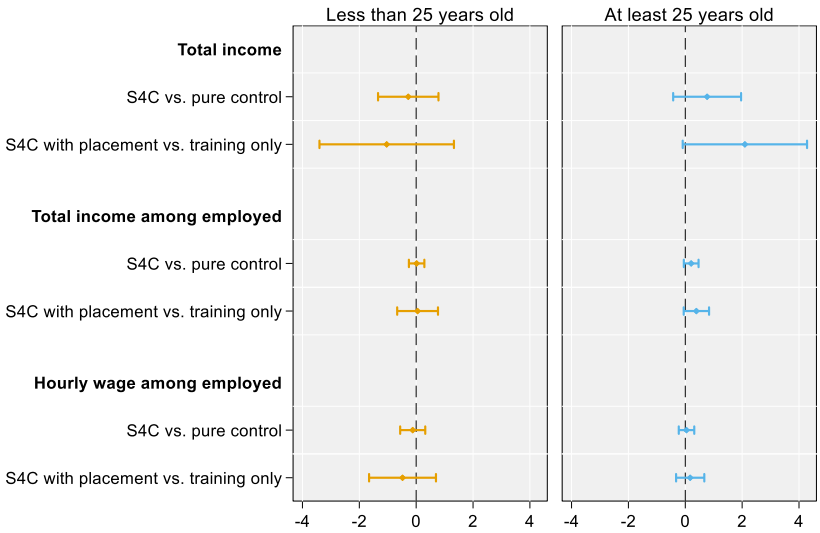
Figure 5.11.3 presents estimated treatment effects of the S4C training and the S4C placement component for respondents younger than 25 years on the left-hand side and for those aged 25 years and older on the right-hand side. Figure 5.11.4 explores treatment effects on incomes and wages by respondents' age group. In general, estimated effects are larger for older than for younger respondents. Compared to training only, additional participation in the S4C placement component leads to a significant increase of 17 percentage points in the probability of having decent employment and a significant increase of 16 percentage points the probability of being formally employed among study participants aged 25 years and older. Other than that, no significant effects can be found for either age group.

Figure 5.11.3
Program effect heterogeneity on binary outcomes by respondents' age in November



Source: Own calculations based on S4C survey.

Figure 5.11.4
Program effect heterogeneity on incomes and wages by respondents' age in November

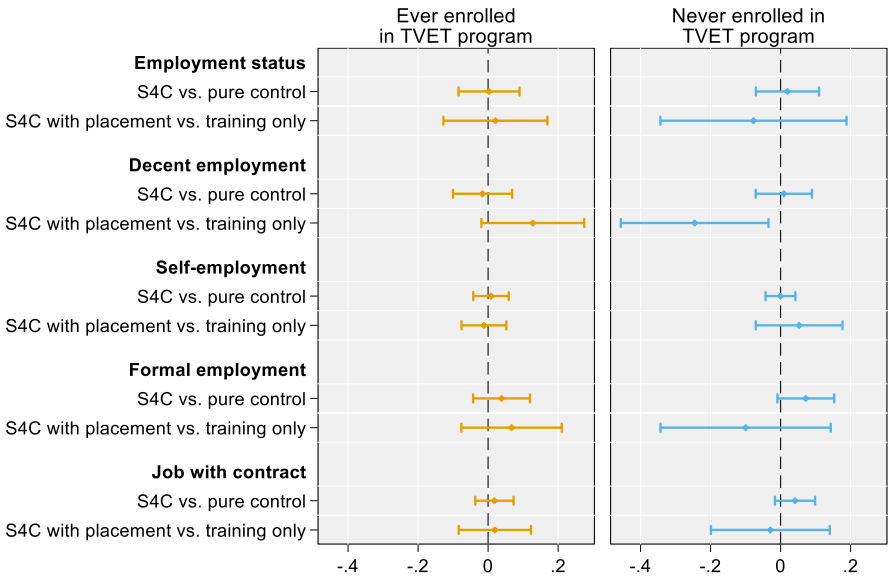


Source: Own calculations based on S4C survey.

5.11.3 Program effect heterogeneity by respondents' vocational training level

Figure 5.11.5 presents treatment effect estimates of the impact of the S4C program and the S4C placement component on employment outcomes for a subsample of study participants who participated in a TVET program at some point before they registered for the S4C program on the left-hand side and for the subsample of respondents who were never inscribed in a TVET program on the right-hand side.

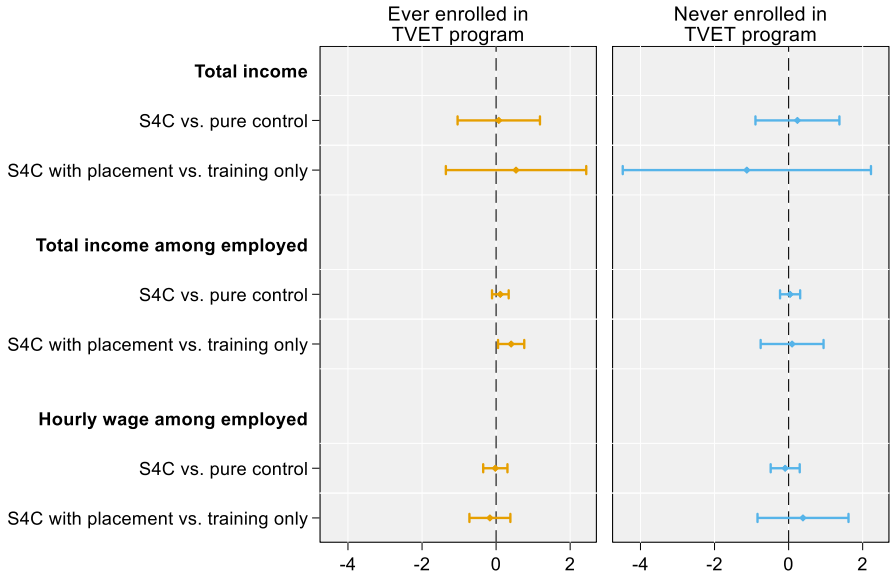
Figure 5.11.5
Program effect heterogeneity on binary outcomes by respondents' vocational training experience in November



Source: Own calculations based on S4C survey.

Figure 5.11.6 shows estimated impacts on income and wages for the same subsamples. The estimated impact of the S4C placement component on the probability of having decent employment is significantly negative for respondents who never participated in a TVET program (Figure 5.11.5). However, this estimate is based on only 18 respondents in the treatment group, of which 3 were in decent employment, and should, thus, not be overrated. For respondents with TVET education, completion of the S4C placement in addition to training leads to a significant increase in income among employed of 50%.

Figure 5.11.6
Program effect heterogeneity on incomes and wages by respondents' TVET experience in November

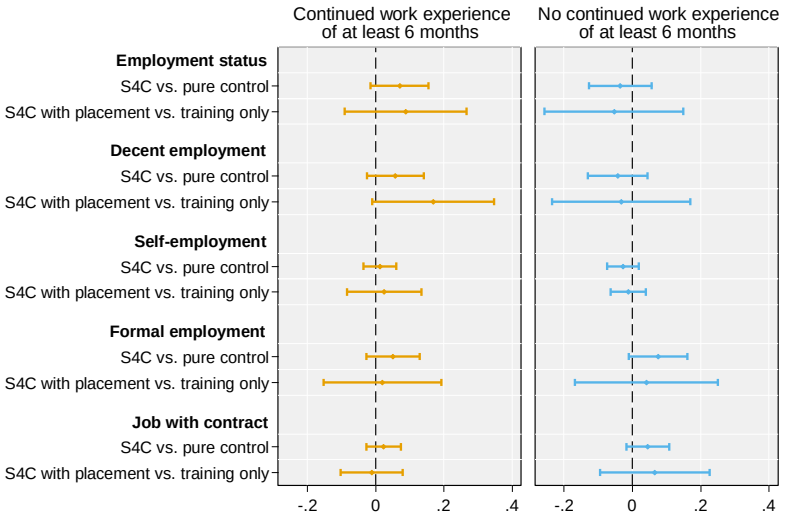


*Notes: The coefficient for "S4C with placement vs. training only" on total income for respondents with TVET education translates to an increase of (exp(0.547) – 1)*100=73%. The coefficient for "S4C with placement vs. training only" on ln of total income among employed for respondents with TVET education translates to an income increase of (exp(0.406) – 1)*100 = 50%. The coefficient for "S4C with placement vs. training only" on ln of total income for respondents without TVET education translates to a decrease of (exp(-1.192) – 1)*100=-70%. The coefficient for "S4C with placement vs. training only" on ln of hourly wage among employed for respondents without TVET education translates to an increase of (exp(0.314) – 1)*100=37%. - Source: Own calculations based on S4C survey.*

5.11.4 Program effect heterogeneity by respondents' previous work experience

Figures 5.11.7 and 5.11.8 show estimated treatment effects of the S4C training and the S4C placement component for respondents with a continued work experience of at least six months prior to their registration for the S4C program on the left-hand side and for respondents without such a work experience on the right-hand side. Overall, point estimates tend to be larger for study participants who had worked at least six months for one employer before participating in the training. For this particular group, participation in the S4C placement component led to a significant and large increase in income among employed of 71%. All other estimates fail to reach significance.

Figure 5.11.7
Program effect heterogeneity on binary outcomes by respondents' work experience in November



Source: Own calculations based on S4C survey.

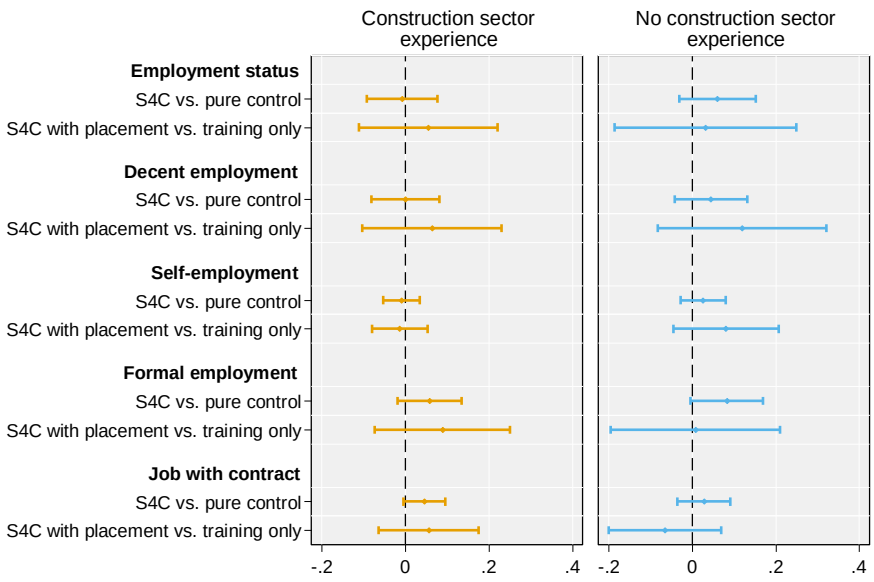
Figure 5.11.8
Program effect heterogeneity on incomes and wages by respondents' work experience in November

*Notes: The coefficient for "S4C vs. pure control" on ln of total income for respondents with a continued work experience of at least 6 months translates to an increase of (exp(0.889) – 1)*100 =143%. The coefficient for "S4C with placement vs. training only" on ln of total income for respondents with a continued work experience of at least 6 months translates to an increase of (exp(1.424) – 1)*100=315%. The coefficient for "S4C with placement vs. training only" on ln of total income among employed for respondents with a continued work experience of at least 6 months translates to an increase of (exp(0.537) – 1)*100=71%. The coefficient for "S4C vs. pure control" on ln of total income for respondents without a continued work experience of at least 6 months translates to a decrease of (exp(-0.425) – 1)*100= -35%. The coefficient for "S4C with placement vs. training only" on ln of total income for respondents without a continued work experience of at least 6 months translates to a decrease of (exp(-0.761) – 1)*100= -53%. The coefficient for "S4C with placement vs. training only" on ln of hourly wage among employed translates to a decrease of (exp(-0.557) – 1)*100= -43%. - Source: Own calculations based on S4C survey.*

5.11.5 Program effect heterogeneity by respondents' construction sector experience

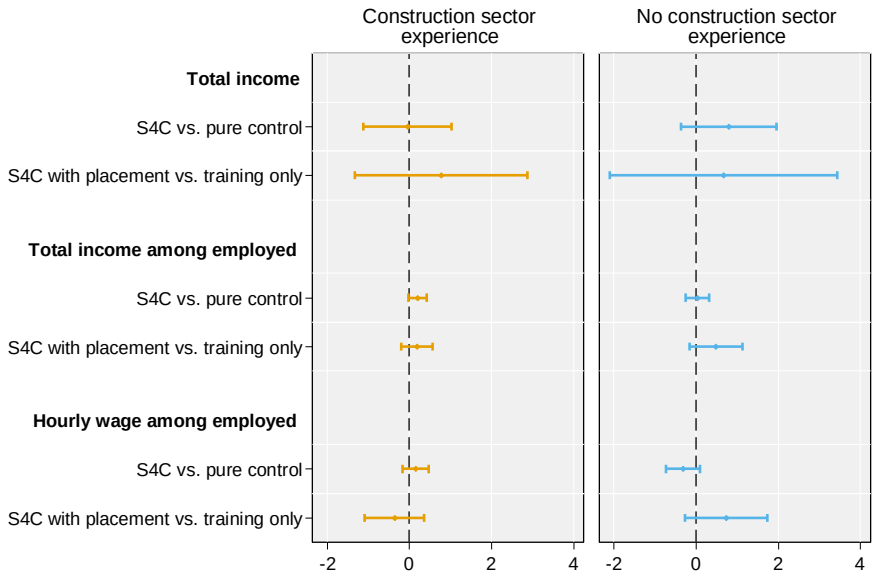
Figure 5.11.9 presents the impact of the S4C program and the S4C placement component on employment outcomes for the subsample of respondents who had prior work experience in the construction sector on the left-hand side and for the subsample of respondents who had never worked in the construction sector before on the right-hand side.

Figure 5.11.9
Program effect heterogeneity on binary outcomes by respondents' construction sector experience in November



Source: Own calculations based on S4C survey.

Figure 5.11.10
Program effect heterogeneity on incomes and wages by respondents' construction sector experience in November



Source: Own calculations based on S4C survey.

Figure 5.11.10 shows the impact on incomes and wages for the same sub-groups. In general, point estimates tend to be larger for respondents without previous construction sector experience. This stands to reason as the S4C program comprises training in basic construction skills and might, thus, be particularly helpful for participants who want to work in the construction sector but have no previous experience. However, no coefficient is significantly different from zero, so we cannot conclude that the S4C program and its placement component had a significant positive medium-term impact for study participants with or without construction sector experience.

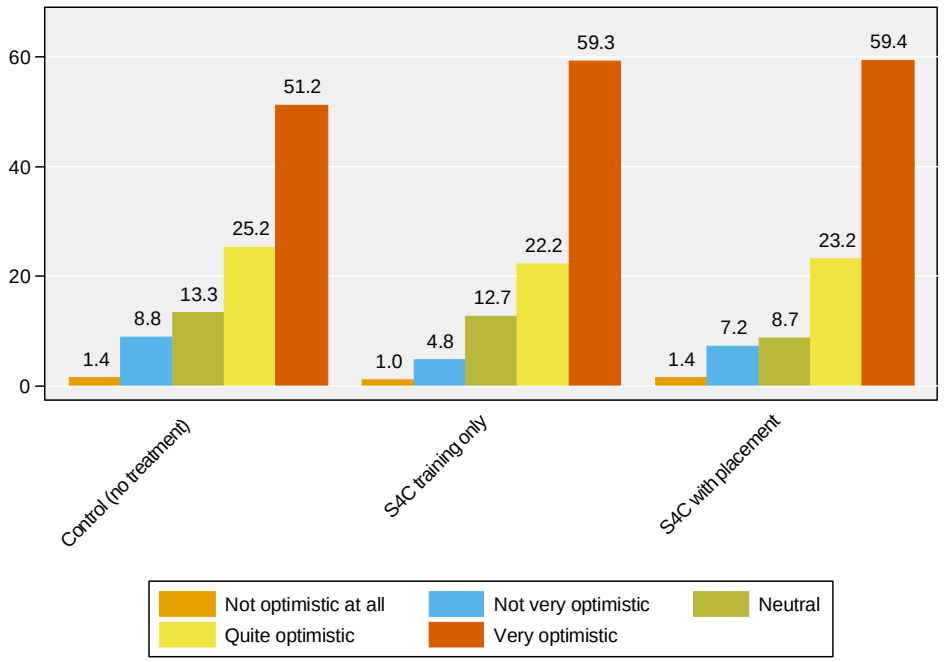
5.12 Program impact on secondary outcomes

Besides employment and earnings, the S4C program may have affected participants' lives in multiple other ways. This section explores the impact of the S4C program as a whole, the additional impact of the placement component, and the additional impact of S4C level 2 training on participants' aspirations with respect to future labor market outcomes as well as their migration intentions.

5.12.1 Employment and earnings aspirations

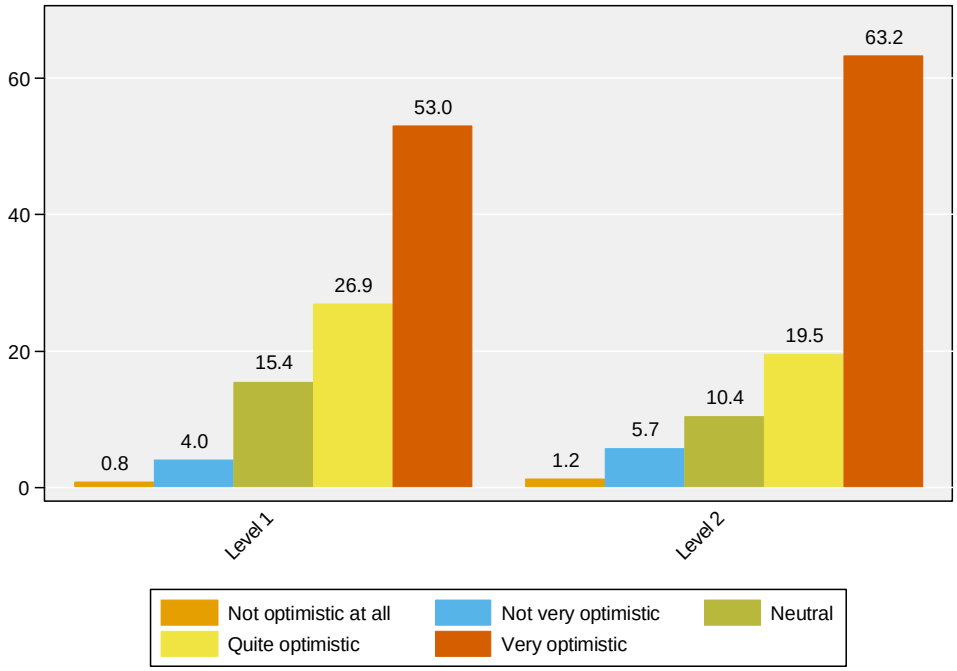
Figures 5.12.1 and 5.12.2 describe respondents' optimism about being employed in the long term, i.e., 5 years after the survey, by treatment group and training level, respectively. S4C trainees are more optimistic about their employment in the long run than the pure control group. While 51% of respondents in the pure control group indicated to be "very optimistic", this share amounts to 58% in both the "S4C training only" and "S4C with placement" groups (Figure 5.12.1). 53% of level 1 trainees and even 63% of level 2 trainees are very optimistic about being employed in the long term (Figure 5.12.2). The share of respondents who are not optimistic is very small across all groups.

Figure 5.12.1
Optimism about employment in the long term by treatment group



Source: Own calculations based on S4C survey.

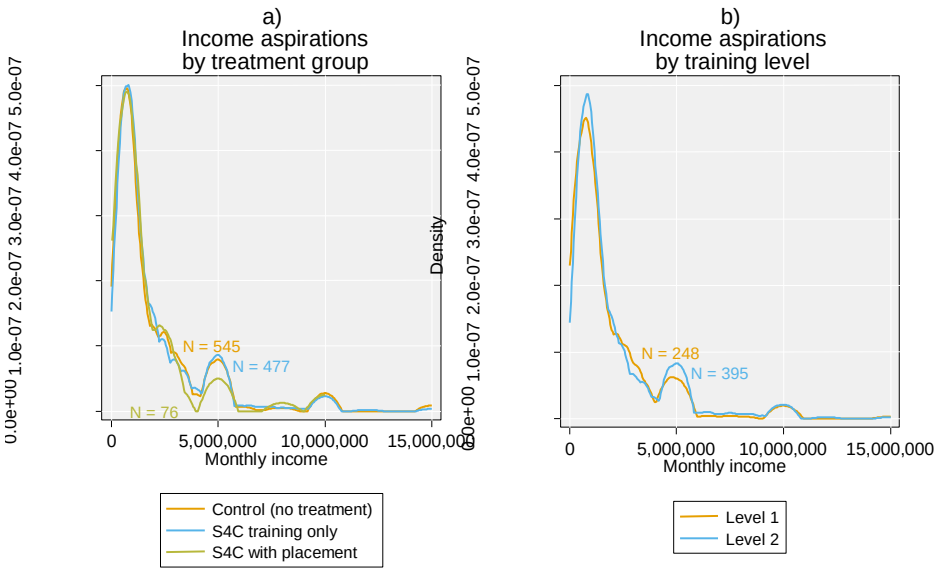
Figure 5.12.2
Optimism about employment in the long term by training level



Source: Own calculations based on S4C survey.

Figure 5.12.3 shows that long-term income aspirations are very similar across all treatment groups and training levels. For all groups, income aspirations range between 2,370,000 UGX and 2,820,000 UGX per month.

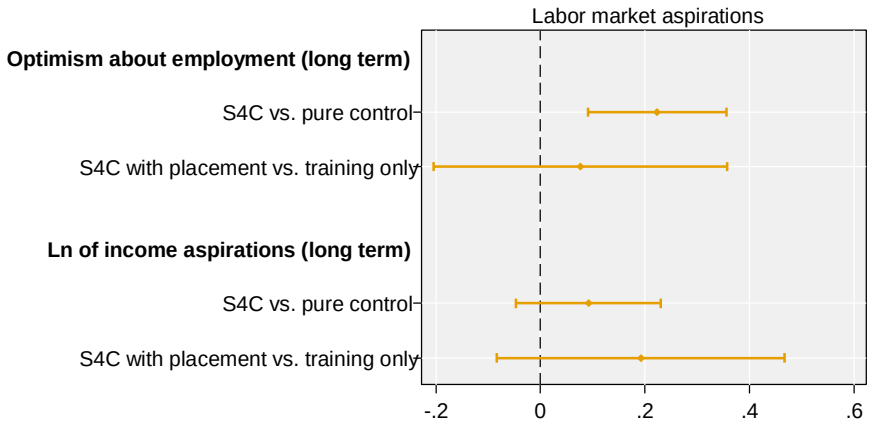
Figure 5.12.3
Income aspirations in the long term



Source: Own calculations based on S4C survey.

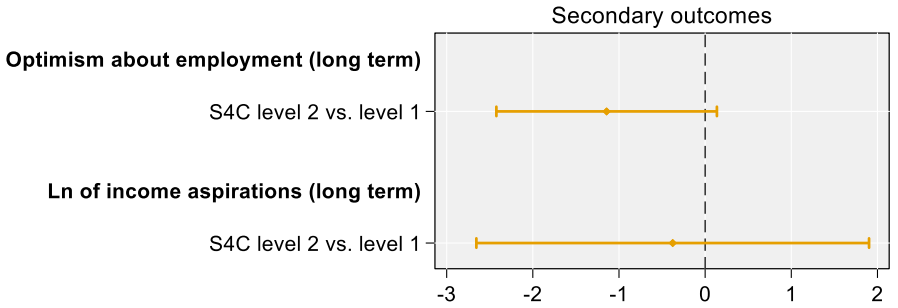
Figures 5.12.4 and 5.12.5 present estimated treatment effects on long-term labor market aspirations. Figure 5.12.4 shows that participation in the S4C program indeed has a significantly positive impact on study participants' optimism about being employed in the long term, i.e., 5 years from the survey. However, the S4C placement did not have an additional effect on employment aspirations in comparison to the S4C training. Similarly, the S4C level 2 training (see Figure 5.12.5) has no additional impact on long-term labor market aspirations. Further, neither the program nor any of its components do significantly affect income aspirations in the long term.

Figure 5.12.4
Estimated treatment effects of the S4C program and placement component on labor market aspirations in the long run



Source: Own calculations based on S4C survey.

Figure 5.12.5
Estimated treatment effects of S4C level 2 training on labor market aspirations in the long run

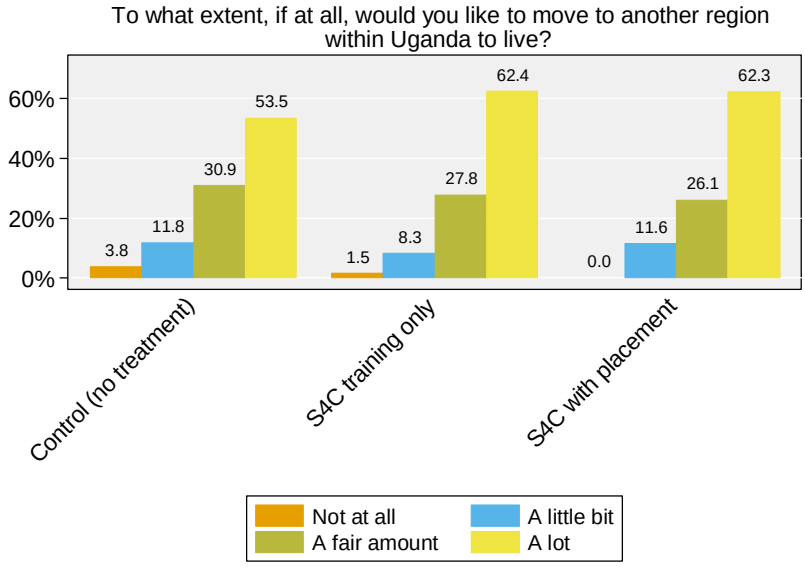


Source: Own calculations based on S4C survey.

5.12.2 Migration intentions

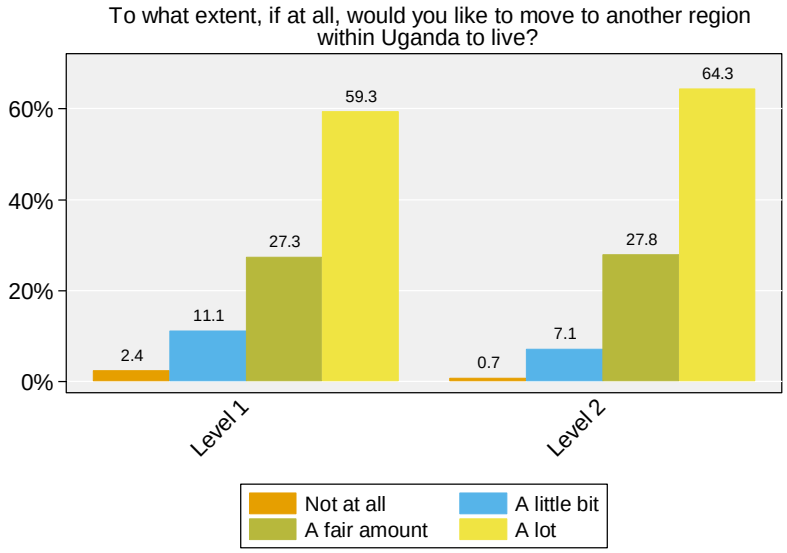
Figures 5.12.6 and 5.12.7 present study participants' intentions to move to another region *within* Uganda by treatment group and training level, respectively. More S4C program participants than non-beneficiaries have high internal migration intentions (Figure 5.12.6). The same applies for S4C level 2 trainees compared to S4C level 1 trainees (Figure 5.12.7). The difference between training-only beneficiaries and those who additionally completed a placement, however, is minimal (Figure 5.12.6).

Figure 5.12.6
Internal migration intentions by treatment group



Source: Own calculations based on S4C survey.

Figure 5.12.7
Internal migration intentions by training level

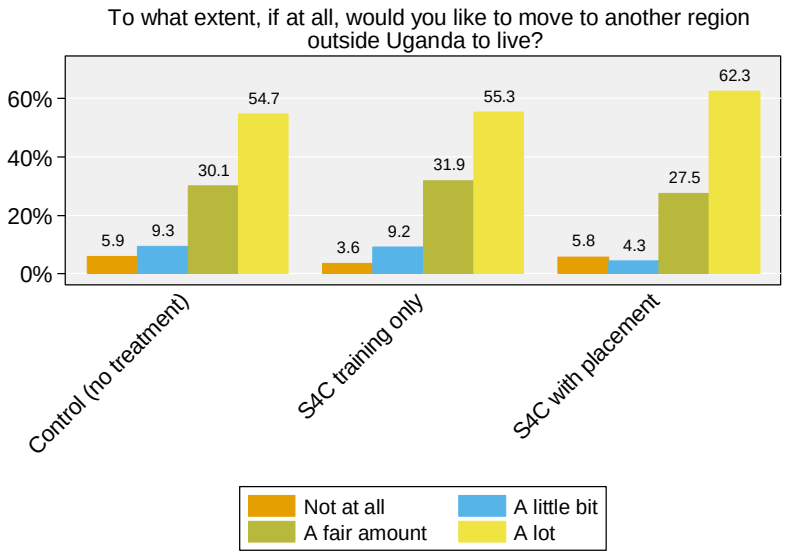


Source: Own calculations based on S4C survey.

Figures 5.12.8 and 5.12.9 show intentions to move to another region *outside* Uganda by treatment group and training level, respectively. External migration intentions are similar for the pure control group and beneficiaries that participated only in the S4C training (Figure 5.12.8). Respondents who also completed the placement component, by contrast, were more likely to state that they would very much like to migrate to another country (Figure 5.12.9). Comparing S4C level 2 trainees to S4C level 1 trainees (Figure 5.12.9), level 2 trainees have higher external

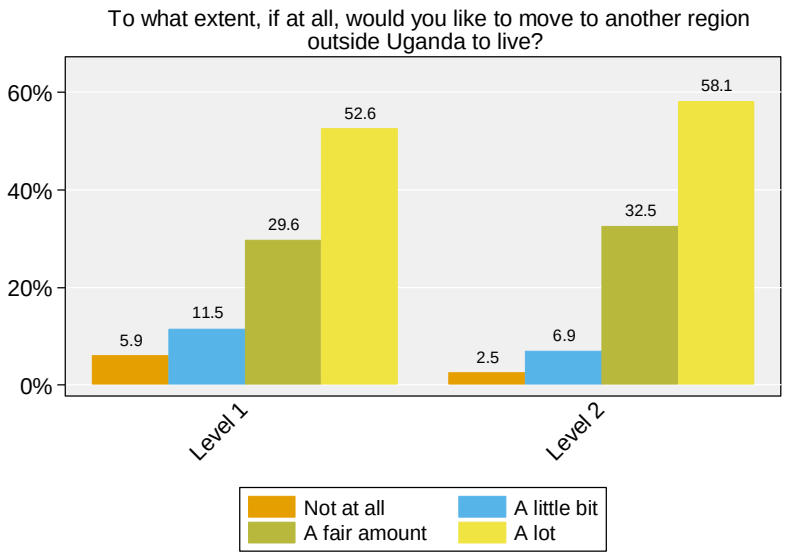
migration intentions. The differences, however, are not as marked as for internal migration intentions.

Figure 5.12.8
External migration intentions by treatment group



Source: Own calculations based on S4C survey.

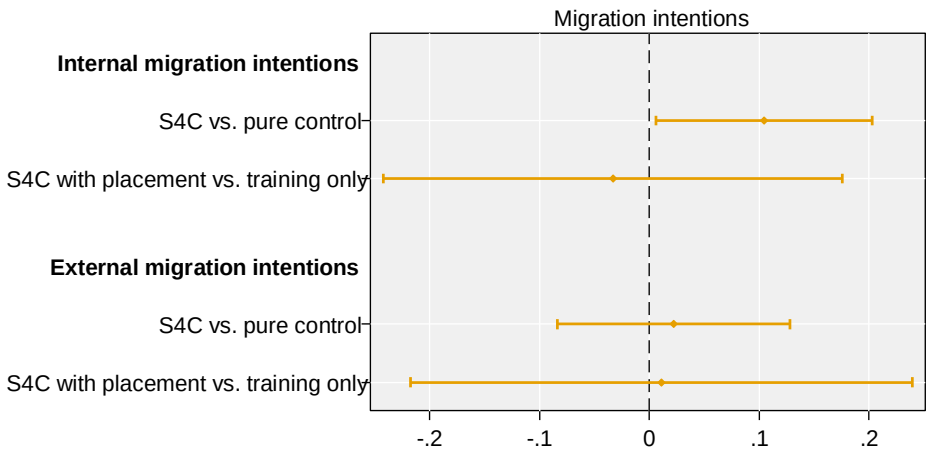
Figure 5.12.9
External migration intentions by training level



Source: Own calculations based on S4C survey.

Figures 5.12.10 and 5.12.11 present estimated treatment effects on migration intentions. Compared to the pure control group, S4C beneficiaries have significantly higher internal migration intentions (Figure 5.12.10). The point estimate on external migrations intentions is positive as well, but not significantly different from zero. Figure 5.12.10 further shows that the placement component did not have additional significant effects on neither internal nor external migration intentions.

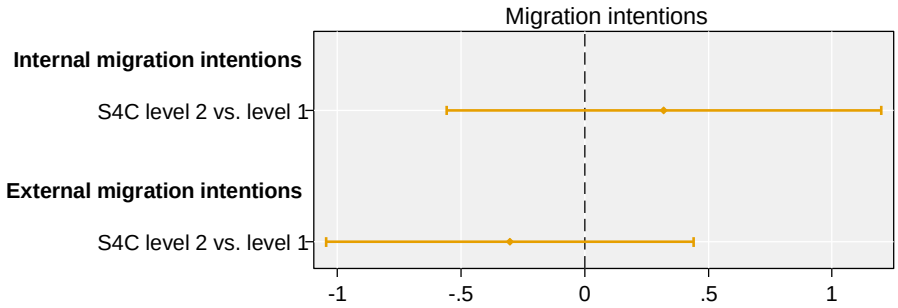
Figure 5.12.10
Estimated treatment effects of the S4C program and placement component on migration intentions



Source: Own calculations based on S4C survey.

Figure 5.12.11 presents the effect estimates of the S4C level 2 training on internal and external migration. While the coefficient on internal migration is positive, it is negative for external migration. However, both coefficients are not significant and also not statistically different from each other.

Figure 5.12.11
Estimated treatment effects of S4C level 2 training on migration intentions



Source: Own calculations based on S4C survey.

5.13 Summary and discussion

This study aimed to evaluate the short- to medium-term impact of the S4C program on employment and labor market outcomes among young people in Uganda. The S4C program comprised a two-stage training and a matching to companies for an internship or job. However, not every beneficiary participated in all components of the program. The S4C level 1 training consisted of a six-week training course in work readiness, basic construction skills as well as health, safety and environment standards. A certain share of level 1 trainees then proceeded to the S4C level 2 training, which was a one-month training in either rigging or pipe fitting. Admission to the level 2 training was mainly based on trainees' performance during level 1, measured by their test scores

in the final exam. After completion of the training component, a subset of level 1 and level 2 trainees were matched to companies for an internship or job placement.

This impact evaluation focused on (i) the effectiveness of the S4C program as a whole, (ii) the additional impact of the S4C placements compared to the trainings only, (iii) the additional impact of the level 2 training compared to only the level 1 training.

Depending on the cohort and level in which participants were trained, beneficiaries had completed the S4C training between March 2019 and March 2020. In November and December 2020, S4C beneficiaries and young people, who had registered their interest in the program but did not end up participating, were interviewed about their current and retrospective labor market outcomes.

Because the Ugandan labor market was affected by the COVID-19 pandemic and corresponding government measures at the time when the survey was conducted, labor market outcomes were evaluated with respect to two different endlines: a) November/ December 2020, and b) February/ March 2020. The earlier endline bears the advantage that labor market outcomes were still unaffected by the pandemic. However, the early endline is very close to the end of the S4C training for some beneficiaries, such that employment benefits may not yet have unfolded.

The effect estimates that include all study participants show that the S4C program as a whole as well as the placement component and the level 2 training specifically had no significant impact on binary employment outcomes neither in November nor in March. However, a significantly positive impact on total income among employed was detected for the S4C program as a whole in March (17%) and for the placement component in November (36%). This increase appears to be driven by a mixture of higher wages and longer working hours.

The program effects for specific subgroups of beneficiaries were analyzed by (i) respondents' gender, (ii) respondents' age, (iii) respondents' participation in a TVET program, (iv) respondents' sustained work experience of at least 6 months, and (v) respondents' previous work experience in the construction sector. Participation in the S4C program led to a small increase in the likelihood of having a job with a contract in November for male respondents, for female respondents the effect sizes were similar but not significant. The S4C placement had an additional positive and large effect on the probability of being in decent and formal employment for study participants aged 25 years and older and large positive impacts of 50% and 71% on income among employed for respondents with TVET education and a previous sustained work experience of at least six months, respectively.

In a recent meta-study, Kluve et al. (2019) systematically reviewed and compared 113 impact evaluations of youth employment interventions worldwide. The estimated program effects of the S4C program are smaller than the standardized average effect sizes found in Kluve et al. (2019) of 0.04 standard deviations in income and 0.05 standard deviations in employment outcomes. For the S4C program, effect sizes on income among employed range between 0.001 and 0.043 standard deviations. The effects of the S4C program on employment indicators lie between 0.002 to 0.012 standard deviations.

An additional analysis focused on the impact of the S4C program as well as the S4C placement and level 2 training component more specifically on labor market aspirations and migration intentions. The analysis showed that S4C beneficiaries are significantly more optimistic about being employed five years after the survey and have significantly higher internal migration intentions than non-beneficiaries.

The results presented in this study suggest that, even though the S4C program did not affect the overall employment probability, it managed to significantly improve the income among employed. Further, for specific sub-groups – older, TVET educated and experienced participants – the placement component improved the job quality measured by decent employment, formal employment, and employment with contract.

All evaluations that revealed a positive impact of the S4C program as a whole also showed positive impacts of the placement component. Moreover, point estimates for the S4C program as a whole and the placement component are not significantly different from each other whenever a positive impact was detected for one of the two. Jointly, these results, thus, suggest that the S4C placement was the program component that is most effective and drives the program's overall impact. Interestingly, Kluve et al. (2019) found larger impacts for skills training programs (0.05 standard deviations) and no impact for employment services. One explanation might be that not all employment services considered by Kluve et al. (2019) were preceded by a skills training. Against the background that only a small share of designated S4C placements had been realized at the time when the survey was conducted, the evaluation results suggest that program implementers should follow up on firms that had committed to a larger number of placements than actually implemented. This is also important in light of the evidence that the majority of unplaced beneficiaries are still expecting to receive a placement.

The present study has two main limitations. The first and foremost limitation is the point in time at which post-program labor market outcomes were measured. As described in section 5.3.2, labor market outcomes in November/ December 2020 were impacted by the COVID-19 pandemic and the alternative endline in February/ March 2020 was very close to (or even before) the program end for some beneficiaries, such that there was very little time for potential employment benefits to unfold. Kluve et al (2019) and Card et al (2017) suggest that the impact of youth employment interventions increases in the long-term. Remarkably, for the S4C program, positive effects could be detected even in the presence of the pandemic.

The second limitation of the study is that unobserved or unmeasured differences between treatment and control groups may drive the estimated treatment effects. To address this concern, a plethora of participant background characteristics were collected in order to model the participants' decision to join the S4C program. Further, the decision rule assigning S4C level 1 trainees to the S4C level 2 training based on beneficiaries' test scores has been used to identify the additional impact of the S4C level 2 training on labor market outcomes. Unfortunately, for this estimation method the obtained sample size was not sufficient to obtain reliable and precise estimates. The analyses mostly rely on the individual characteristics that were measured, and it cannot be ruled out that unobservable characteristics did downward bias or up-ward bias the estimates of treatment effects (also referred to as omitted variable bias). For example, if only trainees who did not have an outside option, such as a job offer, decided to participate in S4C level 2 training the estimated impact of level 2 training would be downward biased. In future evaluations, it is important to overcome these methodological limitations and to corroborate the effectiveness of the S4C program and, in particular, the S4C placement component.

6 Labor Mobility and the Relevance of Regional Income Differentials among S4C Participants

6.1 Introduction

Labor mobility constitutes a key assumption of E4D's theory of change. Many E4D projects offer trainings in locations and regions where potential participants are not residing and potential employers in the resource sector often cluster in places where natural resources can be quarried but which are remote and offer few other job opportunities. The programs studied in section 2 to 5 of this report offer skills trainings with subsequent internship placements. One of the eligibility criteria to participate in the respective program was commitment to travel or move to the training and placement locations. In many cases the locations of the firm or candidate determined whether a match took place as firms feared potential interns who were not living in proximity to drop out and the participant's decision to work at the respective firm's location hinged on their willingness to relocate or travel. Mobility is essential for potential participants to enroll in and complete E4D programs and, hence, is important for the targeting of E4D interventions. In addition, mobility is key to actually benefit from the intervention as envisaged by E4D's theory of change by acquiring a job in locations where resource sector firms are based.

In order to gain a deeper understanding of participants' labor mobility, this chapter provides descriptive evidence on Skills for Construction (S4C) study participants' intentions to migrate internally, i.e., within Uganda, and internationally as well as the reasons that prevent them from migrating. Further, we investigate how information on wage differentials across regions affects S4C study participants' intention to migrate as misperceptions about wage differentials may distort migration intentions and behavior. To study the relevance of wage information, we informed a random subset of study participants about the median monthly wages of each region of Uganda.

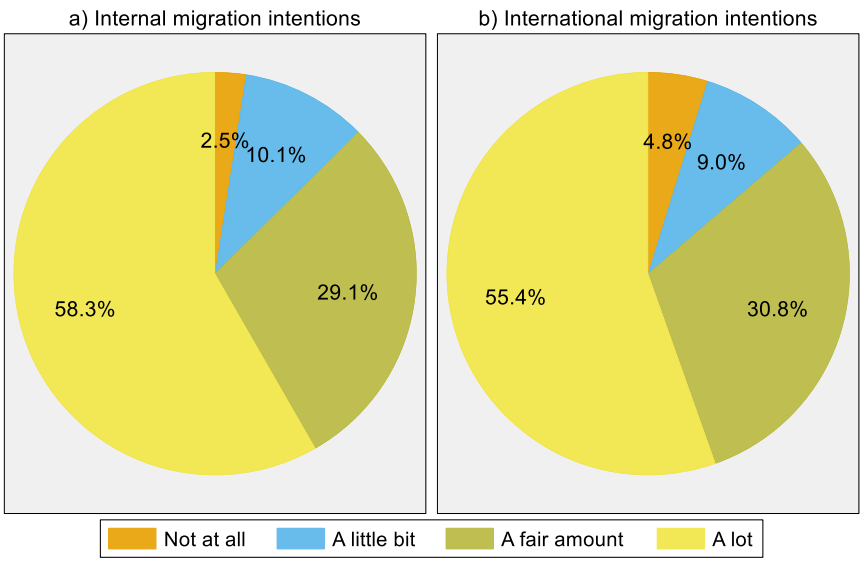
6.2 Descriptive evidence on migration intentions

In this section we describe S4C study participants' intentions and preparations made to migrate within and outside of Uganda, irrespective of whether they benefitted from the S4C program or received the regional wage information treatment. In addition, this section reflects on why participants are not willing to relocate within Uganda for work.

6.2.1 Internal and international migration intentions and preparations

Figure 6.2.1 shows the extent to which participants are willing to migrate internally, i.e., to another region of Uganda, and internationally. Almost 90% of participants are willing to migrate within Uganda or outside of Uganda by a fair amount or a lot. Only few participants are not willing to migrate at all. The differences in migration intentions by internal and international migration are very small. 3% of participants are not willing to migrate internally at all and 5% are not willing to migrate internationally at all. These numbers suggest that there is overall a very high disposition of study participants to relocate. However, Figures 6.2.2 and 6.2.3 - which describe the study participants' preparations for a move - suggest that about 70% of the sample has not made any preparations at all.

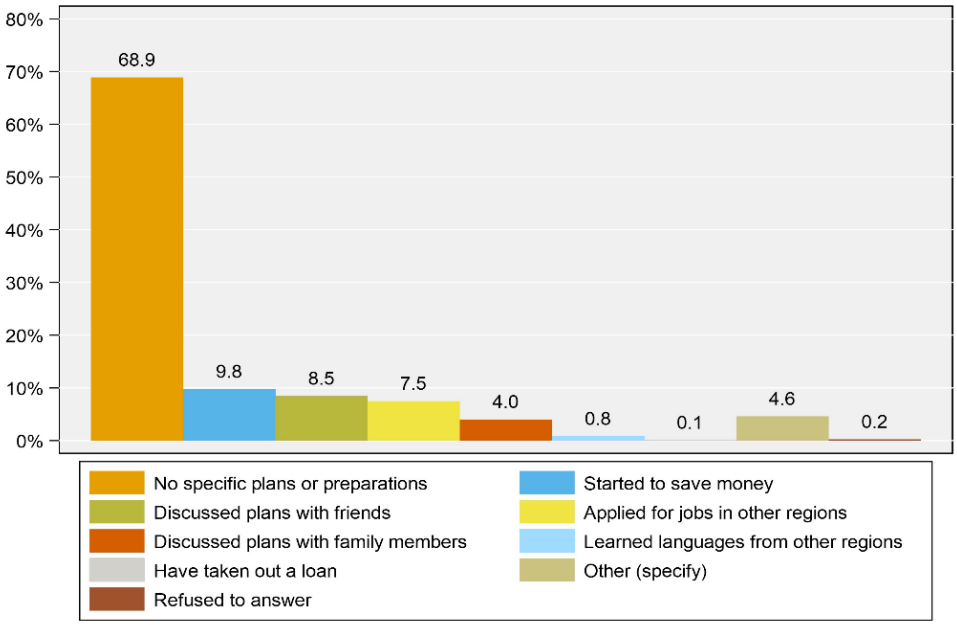
Figure 6.2.1
Responses to the questions "To what extent, if at all, would you like to move to…" a) "… another region within Uganda to live?" and b) "… another country to live?"



Source: Own calculations based on S4C survey.

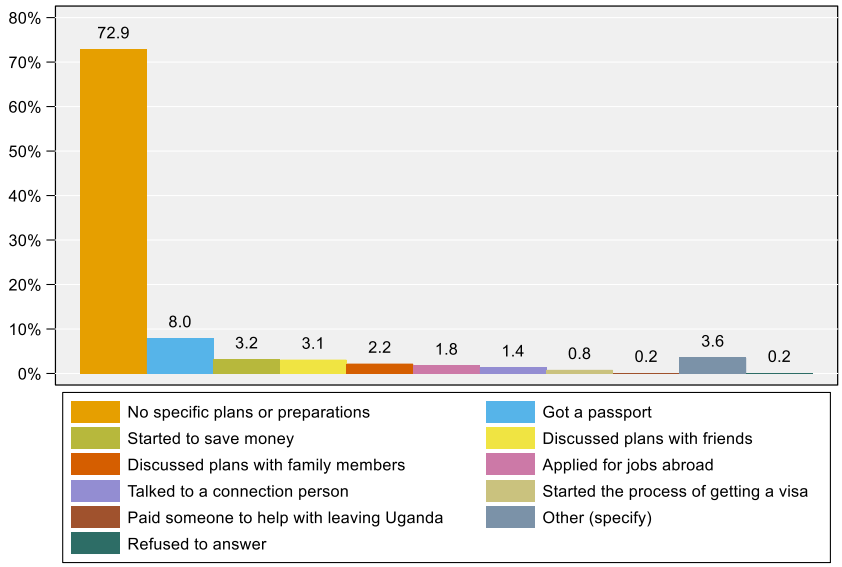
Specifically, Figure 6.2.2 describes the preparations made by study participants for internal migration. The most common preparation (10%) was to save money, which was closely followed by having discussed plans to move with friends (9%), having applied for jobs in other regions (8%), and having discussed plans with family members (4%). Another 1% of participants started to learn languages spoken in other regions of Uganda and 5% mentioned other unspecified plans.

Figure 6.2.2
Response to "What kind of planning or preparation have you done to move to another region within Uganda to live?" (multiple answers were possible)



Source: Own calculations based on S4C survey.

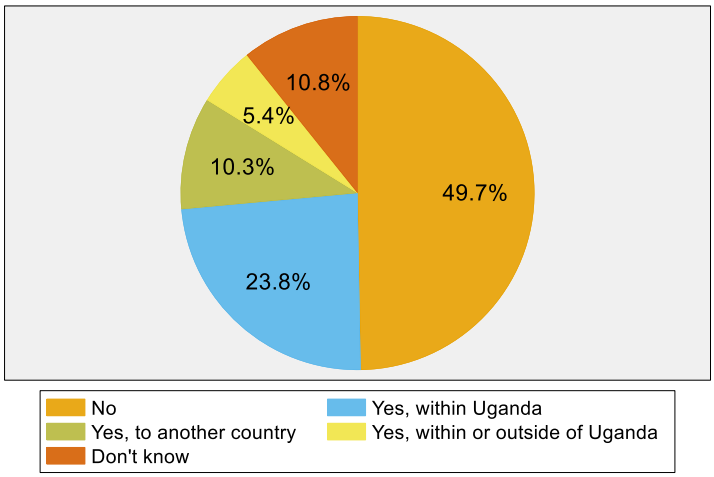
Figure 6.2.3
Response to "What kind of planning or preparation have you done to move to another country to live?" (multiple answers were possible)



Source: Own calculations based on S4C survey.

Figure 6.2.3 describes the preparations made by study participants for international migration. The most common preparation (16%) was to get a passport, whereas all other preparations including saving money, discussed plans with friends or family, applied for jobs abroad, talked to a connection person, started the visa process and paid someone to help leaving Uganda were mentioned by 3% of respondents or less. In addition to enquiring about the respondents' own migration intentions and preparations, they were asked about migration plans of their household members. Figure 6.2.4 shows that half of the respondents stated that no other member of their household have made plans to migrate. 24% and 10% of respondents indicated that members of their household made plans to move within Uganda or to move to another country, respectively, and 5% stated that members plan to move either within or outside of Uganda.

Figure 6.2.4
Response to „As far as you know, does any member of your household have plans to move to another region, district or country?"

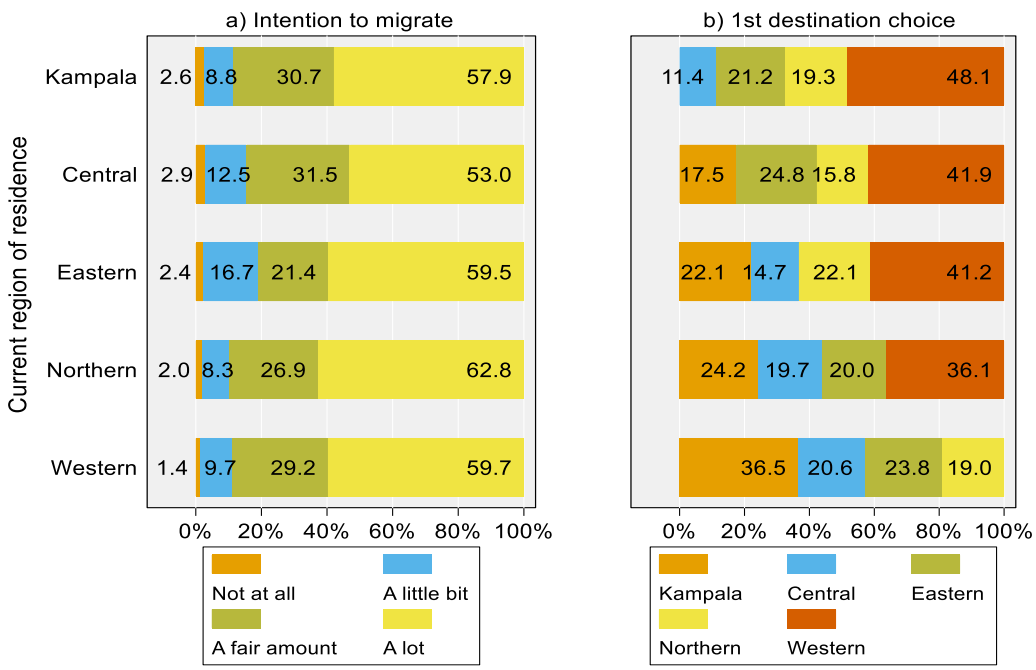


Source: Own calculations based on S4C survey.

6.2.2 Destination choices

This section presents evidence on the preferred migration destinations of respondents. Figure 6.2.5 presents the preferred destination regions for migration within Uganda by respondents' region of residence on the right-hand side (Figure 6.2.5.b). The left-hand side of Figure 6.2.5 presents respondents willingness to move by region of residence (Figure 6.2.5.a). Figure 6.2.5.a shows that respondents' willingness to migrate is fairly similar across regions of residence and that the percentage share of respondents who are interested in moving is large in all regions (81 to 90% are willing to move a fair amount or a lot). Figure 6.2.5.b shows that there are considerable differences in the preferred destination by current region of residence. The largest differences in the percentage shares of respondents by region of residence appear for Kampala and Western Uganda as destination choices, whereas the percentage shares of respondents who would like to move to Central, Eastern or Northern Uganda are somewhat similar across regions of residence (11 to 25% across the three regions). The share of respondents who prefer to move to Kampala, is 37% for residents residing in Western Uganda, 24% in Northern Uganda, 22% in Eastern Uganda and 18% in Central Uganda. The share of respondents who prefer to move to Western Uganda, is 48% for respondents residing in Kampala, 42% in Central Uganda, 41% in Eastern Uganda and 36% in Northern Uganda. These results suggest that there is not one region of interest, e.g., one that is particularly economically strong, to which most respondents would move to. However, section 6.3.3.1 discusses respondents' wage expectations for each region of Uganda and shows that wages are expected to be highest in Kampala, followed by Western Uganda. Thus, it seems that respondents prefer to relocate to regions where they expect to earn the highest incomes.

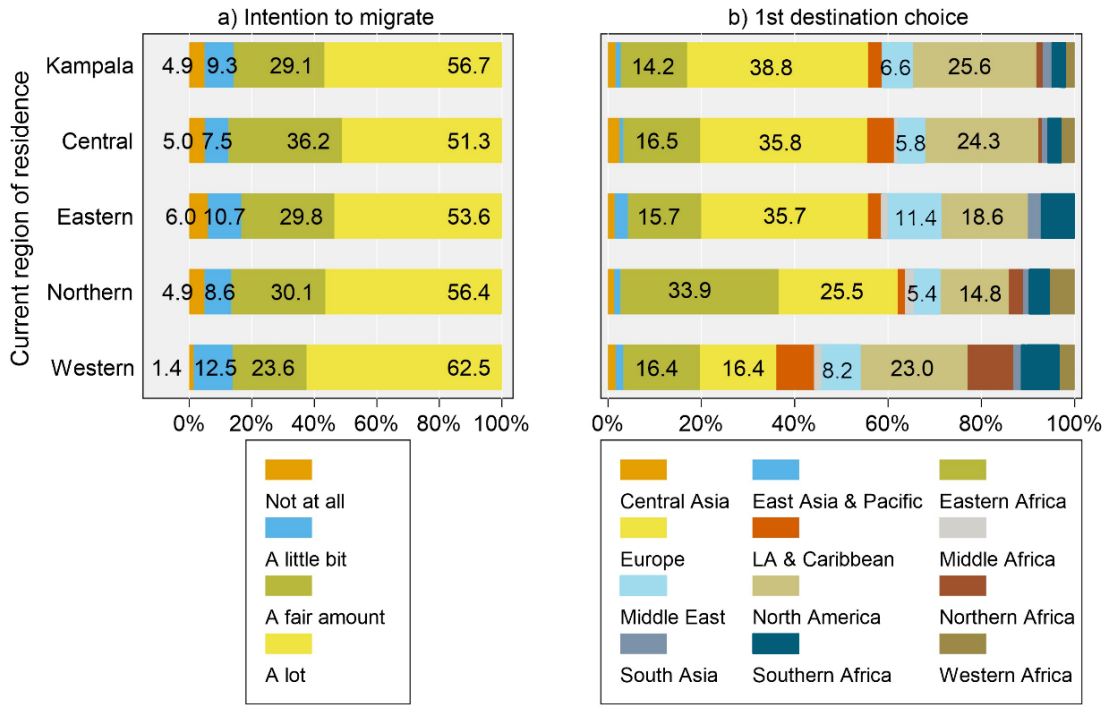
Figure 6.2.5
Current region of residence (subfigure a) and first national destination choice (subfigure b)



Notes: Subfigure b) presents the response to the question: "If you were to move to another region, where would you most want to live?" - Source: Own calculations based on S4C survey.

Figure 6.2.6 presents the preferred destination regions for migration intentions outside of Uganda by the respondents' region of residence on the right-hand side (Figure 6.2.6.b). The left-hand side of Figure 6.2.6 further presents respondents' willingness to move outside of Uganda by region of residence (Figure 6.2.6.a). Figure 6.2.6.a suggests that respondents' willingness to migrate internationally is fairly similar across regions of residence in Uganda and that the percentage share of respondents who are interested in moving to another country is large in all regions (83 to 88% are willing to move a fair amount or a lot). Figure 6.2.6.b shows that there are considerable differences in the preferred destination country by current region of residence. For respondents who reside in Kampala, Central Uganda, and Eastern Uganda, the most preferred region to move to outside of Uganda is Europe (36 to 39%), followed by North America (19 to 26%), Eastern Africa (14 to 16%), and the Middle East (6 to 11%). For respondents residing in Northern Uganda, the most preferred region to move to outside of Uganda is Eastern Africa (34%), followed by Europe (26%), North America (15%), the Middle East (5%), and Western Africa (5%). For respondents residing in Western Uganda, the most preferred region to move to outside of Uganda is North America (23%), followed by Eastern Africa (16%), Europe (16%), Northern Africa (10%), Latin America and the Caribbean ((%), the Middle East (8%), and Southern Africa (8%). Considering all respondents, irrespective of their current region of residence in Uganda, and summing all African destination regions together, about one third of respondents stated a country in Africa as their preferred destination choice.

Figure 1.2.6
Current region of residence (subfigure a) and first international destinations choice (subfigure b)



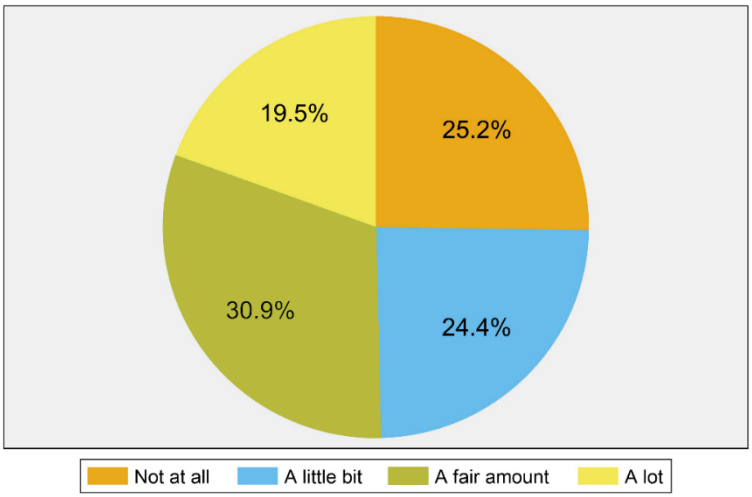
Notes: Subfigure b) is presents the response to the question: "If you were to move to another country, where would you most want to live?" - Source: Own calculations based on S4C survey.

6.2.3 Internal migration for internships and jobs opportunities in the oil and gas sector

This section provides evidence on respondents' willingness to migrate for an internship or job as well as their willingness to migrate to the Albertine region in Western Uganda where large scale investments in the oil and gas sector were expected. Because the E4D program is set out to improve employment opportunities and incomes of beneficiaries through interventions targeted at labor demands resulting from the expected foreign direct investments in the resource sector, the Albertine region is of particular interest and beneficiaries' willingness to move to the Albertine region a key assumption for E4D interventions to achieve the targeted effects.

Figure 6.2.7 presents respondents' willingness to move to another region or district within Uganda for work if they have not secured a job yet in the place they are moving to. A quarter of respondents state that they are not interested in moving at all if they had not secured a job. This figure is substantially different to respondents' overall willingness to move within Uganda as presented in Figure 6.2.1. Figure 6.2.1 shows that only 2.5% of respondents are not interested in moving at all when the question is not linked to their working status at the destination. The difference in the two statistics highlights the relevance of work as a driving factor of migration or at least as a necessary condition for being interested in moving.

Figure 6.2.7
Response to "To what extent, if at all, would you move to another region or district within Uganda for work without having secured a job in that region or district yet?"

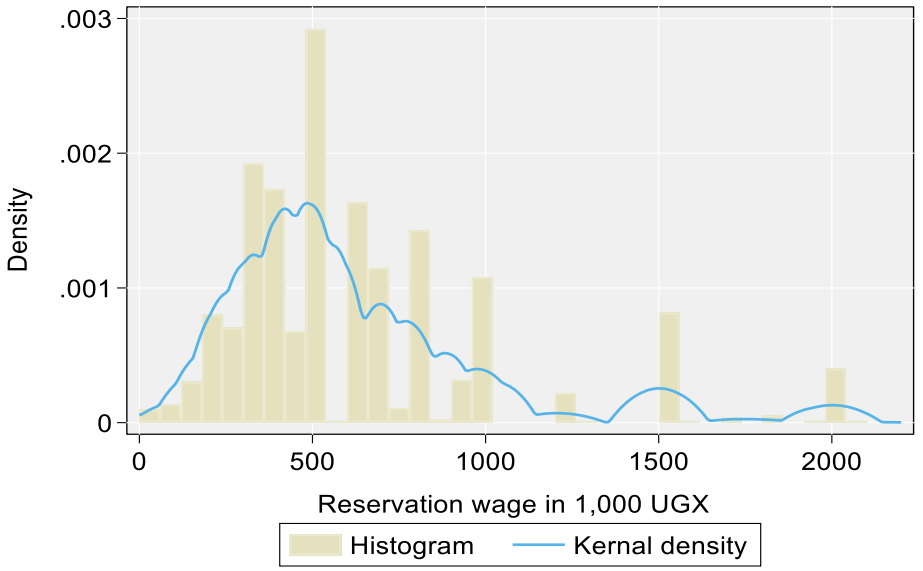


Source: Own calculations based on S4C survey.

Figure 6.2.8 presents respondents' reservation wage to migrate within Uganda, i.e., the minimum wage respondents would need to earn per month in order to move to another region or district. For 90% of respondents the reservation wage lies between 100,000 and 1,000,000 UGX. The mean reservation wage is 660,000 UGX and the median reservation wage is 500,000 UGX, which was also the most common reservation wage. These wages are higher than the median wages in the overall Ugandan population as measured in the Uganda National Household Survey 2016/2017 (Uganda Bureau of Statistics, 2018), which range across regions between 150,000 UGX (Western Uganda) and 400,000 UGX (Kampala) for men and 80,000 UGX (Northern Uganda) and 200,000 UGX (Kampala) for women. However, wages from the Uganda National Household Survey 2016/ 17 refer to earnings only from the main job, whereas respondents of

the S4C survey may expect to hold multiple jobs with higher total monthly earnings. Yet, respondents' earnings at the time of the S4C survey from all jobs for those respondents that had paid work as well as their earnings prior to the pandemic in March 2020 were also lower than their indicated reservation wages. Prior to the pandemic, respondents earned 600,000 UGX on average and the median total monthly income was 400,000 UGX. The monthly earnings during the pandemic were lower than before the pandemic. Overall, only 20% of respondents indicated a reservation wage that was less than their total monthly earnings in March. The mean and median difference between their reservation wage and actual earnings in March was 300,000 UGX, respectively, suggesting that a wage premium of 300,000 UGX is on average required for respondents to move within Uganda for work. This is a substantial wage premium equivalent to 50% of the pre-pandemic average monthly earnings or even 75% of the pre-pandemic median monthly earnings.

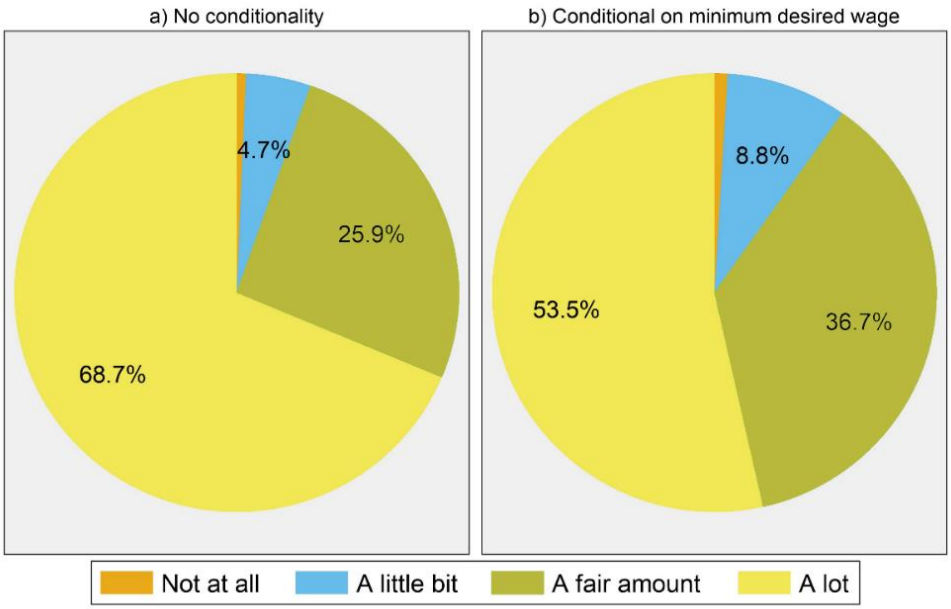
Figure 6.2.8
Response to "What is the smallest amount of money you would need to earn per month to move to another region or district within Uganda just for work?"



Source: Own calculations based on S4C survey.

Figure 6.2.9.a presents respondents' willingness to move to the Albertine region to work in the oil and gas sector value chain. Figure 6.2.9.b also shows respondents' willingness to move to the Albertine region for work but provided that they would earn their reservation wage in that job. Considerably fewer respondents are willing to move a lot provided they earn the desired reservation wage than when they are asked the same question without the conditionality on the reservation wage. This may indicate that the wage premium for migration might even be higher when the destination is the Albertine region than for other parts of the country.

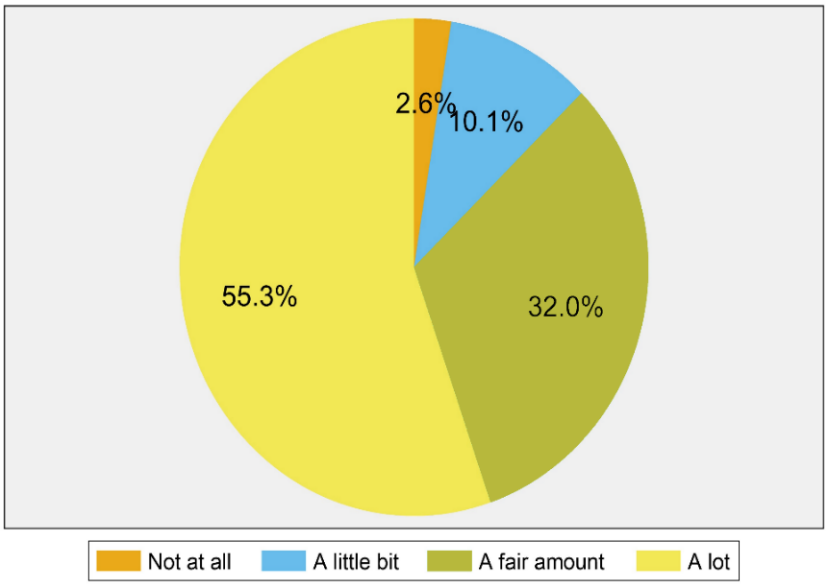
Figure 6.2.9
Extent to which respondents would like to move to the Albertine region to work in the oil and gas sector value chain by conditionality on whether they would earn the reservation wage to move to another region that was indicated in Figure 6.2.8



Notes: Subfigure 3.a) presents the response to the question "In the Albertine region, large scale investments in the oil and gas sector are being made. To what extent, if at all, would you like to move to the Albertine region to work in the oil and gas sector value chain?" Subfigure 3.b) presents the response to the question "In the Albertine region, large scale investments in the oil and gas sector are being made. To what extent, if at all, would you like to move to the Albertine region to work in the oil and gas sector value chain if you were paid the monthly amount of [RESERVATION WAGE] UGX you mentioned in the previous question?" - Source: Own calculations based on S4C survey.

Figure 6.2.10 presents the reasons why respondents are not interested in moving to the Albertine region for work at all. Note that because few respondents answered that they are not interested at all Figure 6.2.10 relies on responses from only 10 people. Four respondents indicated that they would not move because living conditions are not good in the Albertine region, 3 stated that it is expensive to live there, two indicated that the Albertine region is far from their families and friends or that they do not know anyone in the Albertine region, and one person, respectively, stated that he or she does not like the Albertine region or does not want to move anywhere.

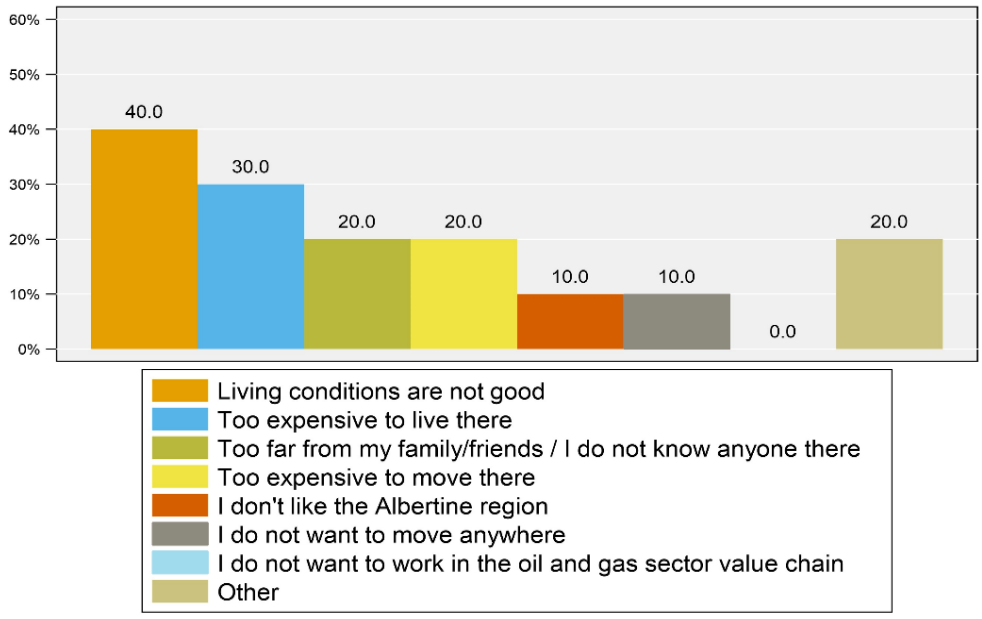
Figure 6.2.10
Response to question „ Why would you not like to move to the Albertine region to work in the oil and gas sector value chain? " (multiple responses were possible)



Source: Own calculations based on S4C survey.

Figure 6.2.11 presents respondents' willingness to move within Uganda for a 3 to 6 months internship which pays transport allowance, free meals, and work equipment. The respondents' willingness to temporarily move for an internship is similar to the overall migration intentions of respondents. 55% of respondents are interested in moving a lot, 32% are interest a fair amount and 10% are interested a little bit.

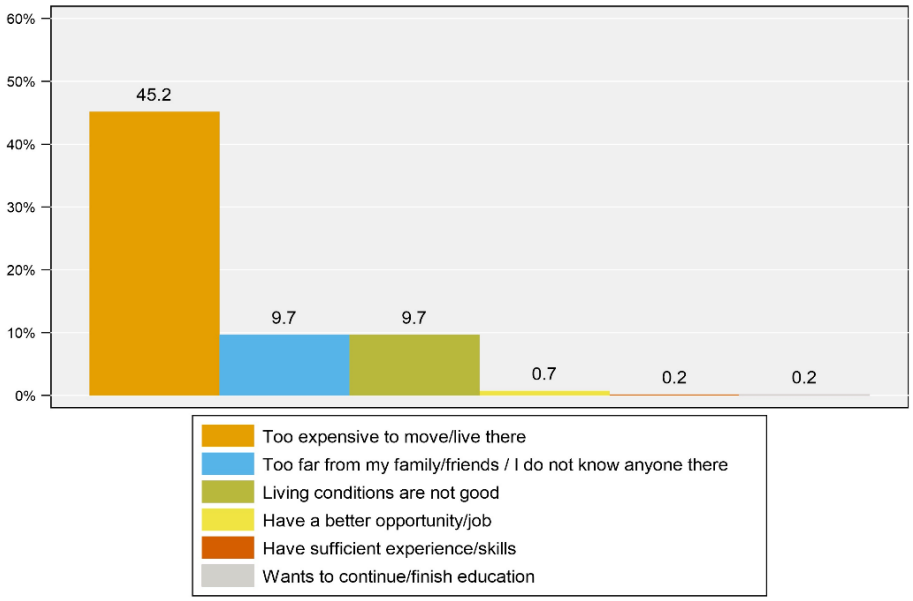
Figure 6.2.11
Response to question „To what extent, if at all, would you like to move to another district for a 3 to 6 months internship where you would get transport allowance, free meals and equipment?"



Source: Own calculations based on S4C survey.

Figure 6.2.12 presents respondents' reasons why they do not want to move for an internship. 45% of the 31 respondents who do not want to move at all for an internship stated that moving or living at the destination would be too expensive and 10% of respondents indicated that the internship would be too far from friends and family or that living conditions would not be good, respectively. Only less 1% of respondents indicated that they have better job opportunities, have sufficient experience or skills or want to continue or finish education, respectively.

Figure 6.2.12
Response to question "Why would you not like to move to another district for an internship?"



Source: Own calculations based on S4C survey.

6.3 The importance of information about wage differentials

This section investigates how information on wage differentials across regions affects S4C study participants' intentions to migrate within Uganda. Misperceptions about wage differentials may distort migration intentions and behavior. Thus, providing participants with correct wage information might influence their intentions and destination choices.

6.3.1 The information treatment

In order to study the importance of knowledge about wage differentials across regions for migration intentions, a random subset of S4C study participants received information about wages in each region of Uganda. The provided information was based on data collected by the Uganda Bureau of Statistics in 2016/ 2017 and as published in the Uganda National Household Survey 2016/ 17 Report (Uganda Bureau of Statistics, 2018). The report provided the median monthly wage in the main job for each region of Uganda, i.e., Central, Eastern, Northern, and Western, as well as Kampala. As regional median wages differ by sex, the regional wage information was specific to the respondent's sex. Further, the information was put in relation to the respondent's current location of residence. Specifically, the information provided, based on the sample of a male respondent residing in Kampala, was:

Here is how much a median man earned per month from his main job in each region, in Ugandan Shilling. This data is from when it was last collected by the Uganda Bureau of Statistics in 2016/2017.

- In Kampala, where you currently reside, the median wage in the main job is 400,000 Ugandan Shilling.
- In Central Uganda, the median wage in the main job is 250,000 Ugandan Shilling, that is 0.6 times as much as the median wage in Kampala, where you reside.
- In Eastern Uganda, the median wage in the main job is 154,000 Ugandan Shilling, that is 0.4 times as much as the median wage in Kampala, where you reside.
- In Northern Uganda, the median wage in the main job is 160,000 Ugandan Shilling, that is 0.4 times as much as the median wage in Kampala, where you reside.
- In Western Uganda, the median wage in the main job is 150,000 Ugandan Shilling, that is 0.4 times as much as the median wage in Kampala, where you reside.

For all other regions and females, the information was conveyed analogously, i.e., the respondent's region of residence was named and, subsequently, the wages in UGX and relative to the region of residence were listed.

Unfortunately, the information treatment was erroneously delivered in the first days of data collection and, therefore, a third of the observations in the sample had to be dropped for the evaluation of the information treatment.[62] This reduced the sample size from 1,218 observations to 862 observations.

The information on wage differentials was preceded by questions about participants' intentions to migrate internally, their preferred destination and their best guess of median monthly wages in each region of Uganda. The respondents best guess on median wages by region enables to understand the extent to which participants were mis-informed about wage levels and differentials and, hence, in what ways the treatment contributed to knowledge updates towards actual wage levels and differentials.

Because participants were potentially not familiar with the concept of medians, the interviewer explained to the respondent that the median is the middle value based on an example and probed the respondents' understanding prior to asking them for their best guess for the median monthly wage of each region and providing the information treatment to a random subset of participants.

After the information of median wages by region was provided, participants were asked once more a set of questions about their internal migration intentions and preferred destinations. Further, respondents were asked about their intentions to move (i) for an internship, (ii) if they had not yet a job secured, (iii) to the Albertine region, and (iv) outside of Uganda.

At the end of the survey module on migration intentions and after the information treatment, respondents were de-briefed about the fact that not only wages differ across regions and districts but that also costs of living differ substantially and that it is important to consider these factors before moving.

[62] *The survey instrument was programmed to display the information on wage differentials to randomly selected individuals only. However, in the first few days of the survey interviewers used paper-based survey instruments, although differently instructed, to read the information treatment to all study participants.*

6.3.2 Estimation strategy and randomization

The information treatment was set up as a randomized controlled trial. The information was provided to half of the study participants by random chance. Therefore, treatment and control group participants are in expectation similar and, thus, comparable. Given this experimental set-up, the treatment effect estimation essentially relies on a simple comparison of outcomes of the treatment and control group.

To check whether randomization worked out as intended and treatment and control group participants are indeed similar, Table 6.3.1 provides means of background characteristics of the treatment and control group in columns (1) and (3). Column (5) presents the difference in background characteristics, i.e., the difference between column (1) and (3). If the difference in column (5) shows one or more stars (*), then this difference is said to be statistically significant and, therefore, the two groups cannot be considered to be similar with respect to the considered background characteristic. Column (6) presents the standardized differences, which is an additional measure for for testing the two groups' similarity. The standardized difference weighs the difference presented in column (5) by the variance of the considered indicators (the variance is a measure of how much the indicator spreads around its mean). Standardized differences with an absolute value of less than 0.2 are considered small. Thus, when the similarity of the treatment and control group is evaluated, everything that is larger than the 0.2 threshold (or smaller than –0.2) would be considered as not similar.

Although treatment and control group are expected to be similar based on the random allocation to treatment and control groups, it possible that individual background characteristics differ from each other by chance. Column (5) of Table 6.3.1 shows that there are some significant differences between the treatment and control group with respect to respondents' region of residence, their education level, their employment status in March, and whether respondents had moved within Uganda since they registered their interest in participating in the S4C program. However, the standardized differences in column (6) do not exceed 0.2 (or -0.2), suggesting that the individual significant differences in column (5) are moderate in size.

The results of a statistical test, called F-test, which examines whether the mean characteristics in Table 6.3.1 jointly differ across treatment and control group, are displayed at the bottom of Table 6.3.1. The p-value of the F-test suggests that the two groups do not differ significantly from each other when all characteristics are considered jointly. Overall, the treatment and control group can therefore be regarded comparable. Yet, to ensure that the significant individual differences identified in column (5) do not affect the estimated information treatment impacts, the regression models used for the effect estimation will control for the respective significant characteristics.

Table 6.3.1
Comparison of background characteristics of treatment and control group

	(1)	(2)	(3)	(4)	(5)	(6)
	Control		Treatment		Difference (1)-(3)	Standardized difference[1]
	Mean/SD	N	Mean/SD	N		
Female	0.14	430	0.14	432	-0.006	-0.018
Age	25.86	427	25.55	431	0.312	0.078
Region of residence						
Kampala	0.36	429	0.37	432	-0.009	-0.019
Central	0.24	429	0.18	432	0.060**	0.147
Eastern	0.06	429	0.06	432	0.000	0.002
Northern	0.28	429	0.34	432	-0.058*	-0.126
Western	0.07	429	0.06	432	0.007	0.031
Education Level						
None	0.03	430	0.02	432	0.007	0.044
Primary	0.15	430	0.20	432	-0.050*	-0.132
Secondary-O Level	0.27	430	0.26	432	0.008	0.019
Secondary-A Level	0.27	430	0.27	432	-0.001	-0.002
Certificate level (TVET)	0.19	430	0.17	432	0.022	0.057
Diploma level (TVET)	0.09	430	0.08	432	0.017	0.060
Poverty probability index	63.12	428	61.55	430	1.568	0.109
Employment indicators						
Employed in Nov. (during COVID-19)	0.46	430	0.49	432	-0.037	-0.075
Employed in March (before COVID-19)	0.58	430	0.66	432	-0.078**	-0.161
Self-employed in Nov. (during COVID-19)	0.11	430	0.10	432	0.010	0.032
Self-employed in March (before COVID-19)	0.13	430	0.09	432	0.033	0.106
Migration preferences and behavior before the information treatment						
Risk willingness score	8.26	430	8.06	432	0.200	0.086
Respondent moved since registration	0.19	430	0.15	432	0.043*	0.114
Respondent moved since birth	0.52	430	0.50	432	0.026	0.051
Migration intentions	2.69	430	2.48	432	0.207	0.061
Preferred migration destination						
Kampala	0.15	430	0.17	432	-0.018	-0.049
Central	0.11	430	0.12	432	-0.011	-0.035
Eastern	0.22	430	0.22	432	0.001	0.002
Northern	0.13	430	0.12	432	0.010	0.030
Western	0.39	430	0.38	432	0.016	0.032
F-test of joint significance (F-stat)					1.197	
p-value of F-test					0.219	
F-test, number of observations					853	

Notes: [1]A standardized difference takes the difference in means of the two groups and weights it by the variance of the indicators (the variance is a measure of how much the indicator spreads around its mean). - Source: Own calculations based on S4C survey.

6.3.3 Results

6.3.3.1 Descriptive evidence on respondents' expected regional wages

Figure 6.3.1 depicts the distribution of expected median monthly wages in the main job in each region of Uganda by respondents' sex. The expected wages range between 10,000 to 1,000,000 UGX, however, the distributions are not particularly flat, indicating that many respondents had similar wage expectations and that the concept of median wages was generally understood.

Figure 6.3.1
Distribution of expected median monthly wage in main job by region of residence and sex of respondent

Source: Own calculations based on S4C survey.

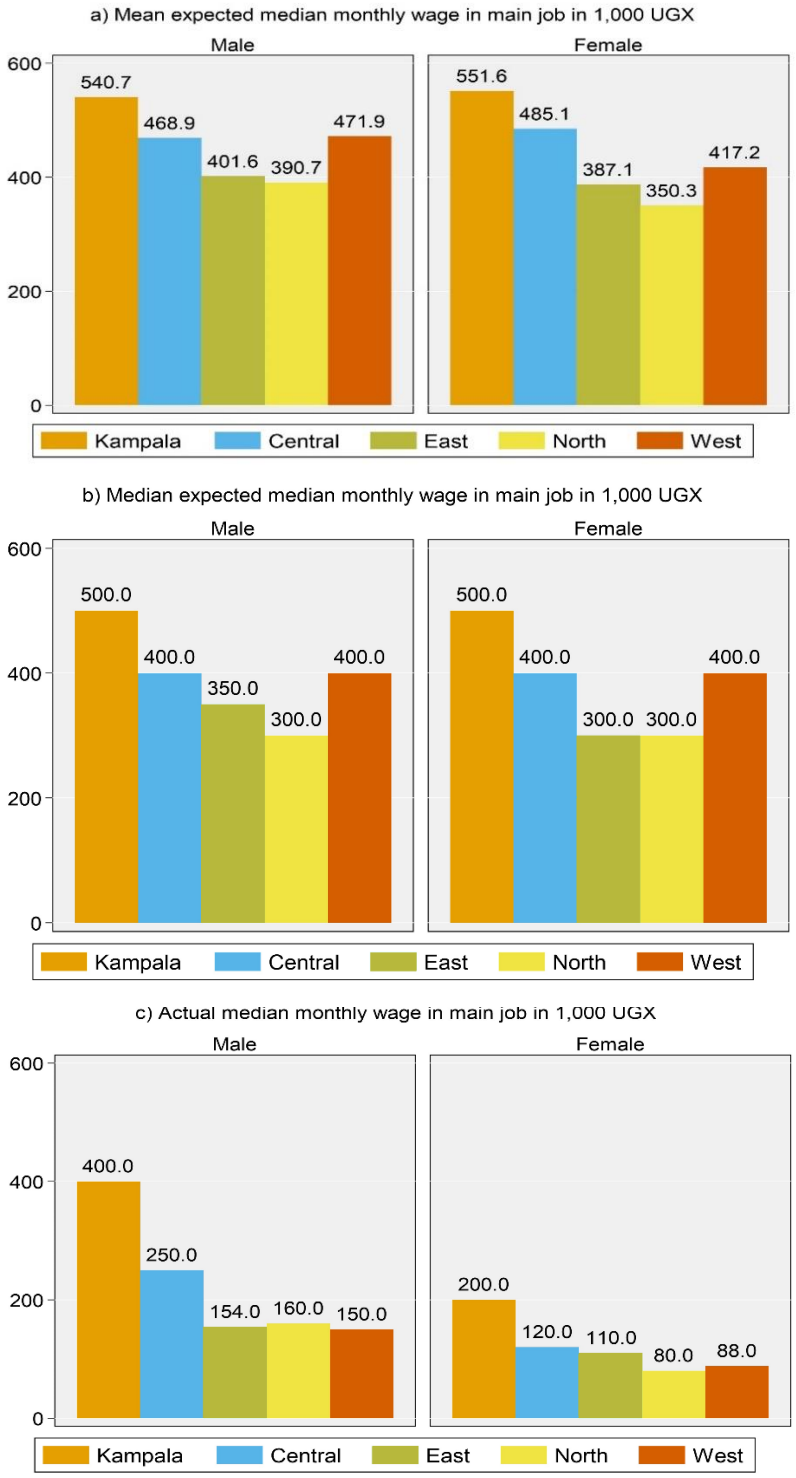
Figure 6.3.2 presents further evidence on respondents' expected wages for each region by sex. Specifically, Figure 6.3.2.a shows the mean expected median monthly wage and Figure 6.3.2.b the median expected median monthly wage. The median and mean expected monthly wages are quite similar and provide similar ranking of regions by wage. Among males, the highest wages are expected in Kampala, followed by Western Uganda, Central Uganda, Eastern Uganda, and Northern Uganda. For females, the highest wages are expected in Kampala, followed by Central Uganda, Western Uganda, Eastern Uganda, and Northern Uganda. Note, however, that for males as well as for females the *median* expected wages in Central and Western Uganda are the same.

For comparison, Figure 6.3.2.c shows the actual median monthly wages in the main job for each region by sex. The actual wages differ substantially from the expected wages. First, the actual wages are considerably lower than the expected ones, especially for females. Second, the ranking of regions by their actual median wage differs from the ranking based on expected wages. Respondents were correct to assume that the highest wages can be earned in Kampala. However, they were wrong to assume that the Kampala wage is closely followed by wages in Central and Western Uganda. Central Uganda, which was ranked third based on mean expected wages of males and second for females, is on rank two of actual wages for both sexes. Further, for males, the actual difference in wages between Kampala and Central Uganda is much larger

(150,000 UGX) than expected (72,000 UGX for mean expected wages). Western Uganda, which is ranked third based on mean expected wages of males and third for females, actually has the lowest monthly wage for males (150,000 UGX) and the second lowest for females (88,000 UGX).

Figure 6.3.2

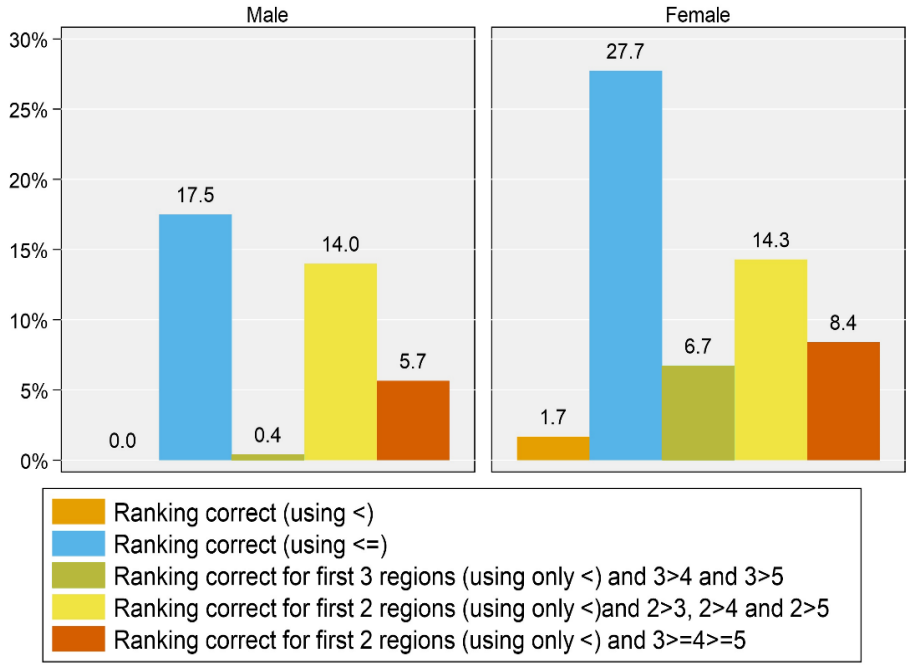
Expected and actual median monthly wage in main job by region of residence and sex of respondent



Source: Own calculations based on S4C survey.

Figure 6.3.3 presents the percentage share of respondents who ranked the regions correctly by their wages by respondents' sex. The ranking uses the mean expected median monthly wage and applies different ranking criteria. The first ranking criteria (orange) is the exact ranking. Figure 6.3.3 shows that none of the males and only 1.7% of females stated expected wages that are consistent with the actual ranking of regions. If more loose definitions of correct rankings are used, for example, considering the respondents ranking as correct when only parts of the ranking were correct, then the share of correct rankings increases to 0.4 up to 17.5% for males and 6.7 up to 27.7% for females. Overall, it seems that females are better informed about regional wage differences. Yet, Figures 6.3.2 and 6.3.3 imply that many study participants were not aware of regional wage differentials measured by the median monthly wage in the main job of the Ugandan population. This suggests that the information treatment potentially has leverage to correct wage expectations and, thereby, change regional migration intentions.

Figure 6.3.3
Percentage share of respondents who ranked regions by monthly median wage correctly



Source: Own calculations based on S4C survey.

6.3.3.2 Impacts of the information treatment

This section presents results of the impact estimation of the regional wage information treatment on migration intentions. Migration intentions were measured on a categorical scale ranging between 0 "would not like to migrate at all" to 3 "would like to migrate a lot". In some estimations, this categorical migration intentions variable is used as a continuous outcome variable in an Ordinary Least Squares (OLS) estimation. The disadvantage of this procedure is that the coefficients are difficult to interpret as the outcome is unit-free and a coefficient of, for example, 0.1 on a scale from 0 to 4 does not allow an intuitive quantitative interpretation of the results.[63] However, such a coefficient provides an overall tendency whether migration intentions reduced or increased on average. A more appropriate way to deal with categorical outcome variables is

[63] *Another disadvantage is that the difference in willingness to migrate between the categories must not be the same.*

the use of Ordered Logit estimations, which provide insights about how much the percentage share of respondents that affirm a specific answer category changed when they were provided with regional wage information. Specifically, in this section marginal effects of ordered logit coefficients are presented and a coefficient of 0.1 for the category "A lot" means that the share of respondents who answered "would like to migrate a lot" increased by 10 percentage points due to the information treatment.

Table 6.3.2 presents the main effect of the information treatment on internal migration intentions. Column (1) shows the average effect of the information treatment using migration intentions as a continuous outcome variable in an OLS estimation. Overall, migration intentions were significantly reduced due to the information treatment. Based on the evidence from the previous section that expected wages were considerably higher than actual wages, the reduction in migration intentions seems plausible.

In column (2) of Table 6.3.2, marginal effects of the information treatment on the different internal migration intention responses based on an ordered logit estimation are presented. The results in column (2) show that the interest to migrate a lot has significantly decreased by 7.4 percentage points, whereas the interest to migrate a fair amount, a little bit and not at all increased significantly by 4.2, 2.7 and 0.6 percentage points, respectively. The coefficients of the ordered logit estimation in column (2) provides additional insights on the composition of the significant average effect in column (1).

Columns (3) to (7) of Table 6.3.2 present the effect of the information treatment on migration intentions in subpopulations based on respondents' current region of residence. Splitting the sample into regional subpopulations reduces the sample sizes and, therefore, the precision of estimates is lower in columns (3) to (7) than in columns (1) and (2). In Northern Uganda, displayed in column (7), too few respondents answered "a lot" and, therefore, this response category could not be included in the estimation.

Table 6.3.2
Information treatment effects on internal migration intentions

	OLS	Ordered Logit (marginal effects)					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All regions	All regions	Kampala	Central	East	West	North
	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE
Treatment	-0.138**						
	(0.045)						
Not at all		0.006*	0.006	0.013	0.024	0.000	0.099
		(0.003)	(0.005)	(0.010)	(0.021)	(0.002)	(0.065)
A little bit		0.027*	0.029	0.040	0.061	0.003	0.065
		(0.011)	(0.019)	(0.026)	(0.037)	(0.018)	(0.043)
A fair amount		0.042**	0.039	0.047	0.128*	0.005	-0.164
		(0.016)	(0.025)	(0.029)	(0.054)	(0.035)	(0.099)
A lot		-0.074**	-0.074	-0.101	-0.213**	-0.008	
		(0.029)	(0.047)	(0.061)	(0.080)	(0.055)	
N	859	859	312	179	52	263	53

Source: Own calculations based on S4C survey.

The percentage share of respondents who would like to migrate a lot reduced by 7.4 percentage points in Kampala, by 10.1 percentage points in Central Uganda, and even 21.3 percentage points in Eastern Uganda, whereas the largest gains in these regions were in the answer category "a fair amount" (3.9, 4.7 and 7.8 percentage points in Kampala, Central, and Eastern Uganda, respectively) and "a little bit" (2.9, 4.0 and 6.1 percentage points in Kampala, Central, and Eastern Uganda, respectively). However, only the large effects for Eastern Uganda were statistically significant.

In Northern Uganda, for which the category "a lot" is omitted, responses to the answer category "a fair amount" decreased insignificantly by 16.4 percentage points, whereas the responses "a little bit" and "not at all" increased by 6.5 and 9.9 percentage points, respectively. Northern Uganda is also the region where the response "not at all" increased considerably (although insignificantly) in size (9.9 percentage points). In Western Uganda no changes in migration intentions due to the information treatment could be observed.

Table 6.3.3 presents the impact of the information treatment on destination choices. Specifically, columns (1) and (2) show the effect of the information treatment on whether respondents improved their destination choice compared to their pre-treatment destination choice towards destinations with higher actual median monthly wages that can potentially be earned at their first and second destination choice, respectively. Similarly, columns (3) and (4) show the effect of the information treatment on whether respondents worsened their destination choice compared to their pre-treatment destination choice towards destinations with lower actual median monthly wages for their first and second destination. Columns (5) and (6) show the treatment effect on potential monthly wages at their first and second destination choice, respectively. Columns (7) and (8) show the effect of the information treatment on whether respondents chose a destination that is higher up in the wage ranking compared to their pre-treatment destination choice for destination choices one and two, respectively. A negative coefficient implies that respondents improved the ranking (rank 1 refers to the highest wage region and rank 5 to the lowest wage region).

Columns (1) and (2) of Table 6.3.3 show that the information treatment significantly increased the probability to choose a destination with higher actual wages compared to their pre-treatment destination choice for the first destination choice by 9.7 percentage points. However, the information treatment also increased the probability to select a second destination with lower wages than the pretreatment destination by 9.3 percentage points as shown in columns (3) and (4). Columns (5) and (6) present the information treatment's impact on potential wages at their choice of destination. The coefficient in column (5) shows that there was a significant increase in potential wages of about 12,000 UGX at the first destination choice. According to columns (7) and (8) there was also a significant improvement in destination choices based on their ranking by wages. Overall, these results suggest that respondents understood the provided information and changed their destination choices in accordance with where the highest wages can be earned.

Table 6.3.3
Information treatment effects on destination choices measured by potential wage improvements

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Improved wage (binary)		Decreased wage (binary)		Monthly wage (1,000 UGX)		Ranking	
	Choice 1	Choice 2	Choice 1	Choice 2	Choice 1	Choice 2	Choice 1	Choice 2
	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE
Treatment	0.097***	0.053	0.052	0.093***	11.952*	-0.314	-0.521*	-0.387
	(0.026)	(0.028)	(0.027)	(0.027)	(5.110)	(4.522)	(0.207)	(0.740)
N	861	861	861	861	858	836	861	861

Source: Own calculations based on S4C survey.

Table 6.3.4 presents the impact of the information treatment on international migration intentions (column (1)) as well as on internal migration intentions if the respondent had not secured a job at the destination (column (2)), temporary migration for an internship (column (3)), migration to the Albertine region for work (column (4)) and migration to the Albertine region for work that pays the respondent's reservation age (column (5)). Overall, the coefficients in Table 6.3.4 are small in size and insignificant. The information treatment does not seem to have had an impact on international migration intentions or migration intentions for more specific purposes.

Table 6.3.4
Information treatment effects on other migration intentions outcomes

	(1)	(2)	(3)	(4)	(5)
	International	No job (internal)	Internship (internal)	Albertine, no condition (internal)	Albertine, conditional (internal)
	β / SE	β / SE	β / SE	β / SE	β / SE
Not at all	0.004	-0.008	-0.002	-0.000	0.002
	(0.006)	(0.022)	(0.003)	(0.001)	(0.002)
A little bit	0.008	-0.003	-0.010	-0.000	0.011
	(0.010)	(0.008)	(0.012)	(0.006)	(0.013)
A fair amount	0.013	0.004	-0.015	-0.003	0.024
	(0.017)	(0.010)	(0.018)	(0.039)	(0.030)
A lot	-0.025	0.007	0.027	0.004	-0.036
	(0.033)	(0.020)	(0.033)	(0.045)	(0.046)
N	861	858	859	417	444

Source: Own calculations based on S4C survey.

6.3.3.3 Impacts of the information treatment by S4C beneficiary status

This section examines whether the information treatment had different impacts in the subsamples of S4C beneficiaries and the no intervention control group. The sample of S4C beneficiaries includes all respondents who had participated in any of the S4C program components, i.e., the level 1 training, the level 2 training, or the internship placement. Respondents of the control group did not receive any S4C benefits.

Table 6.3.5 presents the effects of the information treatment on migration intentions by S4C treatment group. The overall impact based on the OLS estimation is similar across the two subsamples, but it is more precisely estimated and, thus, significant in the subsample of S4C participants. In both groups the information treatment reduced migration intentions and in the S4C subsample significantly so by 0.16. The ordered logit estimations suggest that in both groups the information treatment reduced respondents' interest to migrate a lot and increased the share of respondents who answered with one of the other response categories. The effects are overall larger and significant in the subsample of S4C beneficiaries. Among S4C beneficiaries the percentage share of respondents who are willing to migrate a lot reduced by 7.6 percentage points, whereas the response categories "a fair amount" and "a little bit" increased significantly by 4.4 and 2.6 percentage points.

Table 6.3.5

Information treatment effects on internal migration intentions by S4C treatment group

	(1)	(2)	(3)	(4)
	OLS		Ordered Logit (marginal effects)	
	No S4C	S4C	No S4C	S4C
	β / SE	β / SE	β / SE	β / SE
Treatment	-0.113	-0.155**		
	(0.074)	(0.059)		
Not at all			0.003	0.006
			(0.003)	(0.003)
A little bit			0.019	0.026*
			(0.018)	(0.013)
A fair amount			0.025	0.044*
			(0.025)	(0.021)
A lot			-0.047	-0.076*
			(0.046)	(0.037)
N	338	521	339	521

Source: Own calculations based on S4C survey.

Table 6.3.6 shows the improvements in destination choices based on actual median monthly wages at respondents' preferred destinations. Columns (1) to (4) suggest that control group respondents had higher potential wage increases due to the information treatment than S4C beneficiaries, whereas improvements in the ranking were larger among S4C beneficiaries than among control group members. However, none of the coefficients in Table 6.3.6 are significantly different from zero and, thus, we cannot conclude that the information treatment was effective in one subsample but not in the other.

Table 6.3.6
Information treatment effects on destination choices measured by potential wage improvements by S4C treatment group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Monthly wage (1,000 UGX)				Ranking			
	No S4C		S4C		No S4C		S4C	
	Choice 1	Choice 2	Choice 1	Choice 2	Choice 1	Choice 2	Choice 1	Choice 2
	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE
Treatment	10.423	12.886	1.347	-1.025	-0.086	-0.220	-2.506	0.457
	(8.156)	(7.344)	(8.497)	(5.746)	(0.153)	(0.121)	(1.756)	(0.888)
N	337	512	328	509	339	522	339	522

Source: Own calculations based on S4C survey.

6.4 Summary and conclusion

Labor mobility constitutes a key assumption of E4D's theory of change. Many E4D projects offer trainings in locations and regions where potential participants are not residing and potential employers in the resource sector often cluster in places where natural resources can be quarried but which are remote and offer few other job opportunities. Thus, mobility of potential beneficiaries is key for participants to actually benefit from E4D interventions as envisaged by its theory of change, i.e., by acquiring jobs in the resource sector after completing an E4D intervention.

In this chapter, respondents' overall willingness to migrate within Uganda and outside of Uganda were analyzed as well as respondents' preferred destination choices. In addition, it was investigated how information on wage differentials across regions affects S4C study participants' intention to migrate as misperceptions about wage differentials may distort migration intentions and behavior. To study the relevance of wage information, we informed a random subset of study participants about the median monthly wages of each region of Uganda.

The results show that overall study participants have considerable interest in moving to another region of Uganda and also outside of Uganda. Almost 90% of respondents would like to move "a lot" or "a fair amount" within Uganda or outside of Uganda. However, only 30% of respondents have made any preparations for their move yet. Thus, their willingness to migrate potentially reflects respondents' overall attitude towards migration rather than having made concrete plans and being ready to move.

The most popular destination choices within Uganda are Kampala and Western Uganda, which are also the two regions where respondents expect the highest monthly wages. While indeed the highest wages can be earned in Kampala, the actual regional median monthly wage is lowest in Western Uganda for men and second lowest for women. The most common international migration choices are countries in Europe and Northern Africa. When all African regions are considered together, one third of respondents indicate an African country as their preferred destination choice.

Respondents' intentions to migrate when they were enquired more specifically with respect to moving for work, for work to the Albertine region, or for an internship are similar to the overall migration intentions. However, respondents' willingness to migrate was much lower in a hypothetical scenario in which participants had not secured a job at their destination yet. Under this

circumstance only 50% of respondents would like to move “a lot” or “a fair amount”. While this number is substantially lower than the overall migration intentions, it is still very large.

Prior to the regional wage information treatment, participants were asked about their expected median wages in each region of Uganda. Overall, participants expected wages were higher than the actual wages measured by the median monthly wage earned in the main job in the respective region based on the Uganda National Household Survey 2016/ 2017 Report (Uganda Bureau of Statistics, 2018). Further, participants ranked Kampala correctly as the highest wage region but Western Uganda erroneously as the second highest wage region. In fact, the median monthly wage is lowest in Western Uganda among males and second lowest among females.

In line with respondents’ high wage expectations, the information treatment led to an overall decline in respondents’ willingness to migrate “a lot”. This effect was largest for respondents currently residing in Central Uganda where the treatment led to a significant reduction in the response category “a lot” by 21.3 percentage point. In contrast, there were no effects detected for Western Uganda. Respondents also changed their destination choices in line with the theory towards regions where higher wages could be earned. For example, due to the information treatment, the first destination choice was changed to regions where 12,000 UGX higher wages could potentially be earned on average.

In addition, the information treatment effects were investigated by S4C treatment group affiliation. For migration intentions, the results were similar across treatment and control groups, but the coefficients were more precisely estimated in the larger sample of S4C beneficiaries and, therefore, were only significant in the sample of S4C beneficiaries. With respect to destination choices, higher gains in potential wages due to the treatment appeared in the control group, whereas the ranking of regions by wages improved more in the S4C treatment group. However, none of the effects regarding wages at destination choices are significant.

Overall, these results suggest that mobility and migration is a topic of high relevance to many study participants. The information treatment showed that many participants respond to the provided information by updating their wage expectations, migration intentions, and preferred destinations. The results further show that while participants are generally willing to migrate, they will not move under all circumstances. The decision to move heavily depends on the wages they can earn at the destination location and whether they have a job secured already at their destination. For the E4D program the results suggest that providing information to potential beneficiaries about mobility requirements to participate in E4D programs, about locations or regions of the country where the trained skills are likely in demand, and about sector specific earnings in locations where the trained skills are in demand could enable potential beneficiaries to better gauge the benefits from participating in the E4D program. Providing such information about the benefits and risks of migration to potential beneficiaries, may result in a more efficient targeting of E4D interventions and an increased intervention effectiveness. Not only could such information campaigns take place in the beneficiary selection process but could also be integrated into the skills trainings. For E4D programs in Uganda, this information could be geared towards resource sector specific employment aspects such as region, sector, and position specific information on wages, time to search for jobs, or working conditions.

7 Value for Money

7.1 Value for Money as a concept

Value for Money is a concept used in international development that describes the relationship between the use of resources and the impact of the invested resources. Value for Money is not a tool but rather offers a framework for considerations that can be made in balancing resources and impact (Jackson, 2012). It can be used to guide, steer, and justify individual development projects, specific country programs or whole agency portfolios. The concept of Value for Money is not uniquely defined and, therefore, the definitions of and approaches to Value for Money differ across development agencies. Table 7.1.1 summarizes some of the definitions of Value for Money of different development agencies.

While these definitions leave open the exact approaches to Value for Money, they highlight that the objective of Value for Money is to:

(1) maximize the impact given a unit of resources spent, or

(2) minimize resources given a certain impact goal.

The different definitions further indicate important criteria for the balancing of resources and impact. These include scale, depth, inclusion, effectiveness, efficiency, and economy.

It is important to note that Value for Money is not an ad-hoc evaluation of costs and impact. Value for Money is *implemented* in daily activities of project-level agency staff as well as at the agency management level. This is important because Value for Money is often confused with the *assessment* of Value for Money through cost-effectiveness analyses.

The subsequent sections discuss why Value for Money considerations are important (section 7.2), the caveats of Value for Money (section 7.3), a conceptual framework of Value for Money that includes the most common criteria mentioned in the various definitions (section 7.4), tools that can be used to implement and evaluate Value for Money (section 7.5), a detailed description of cost-effectiveness analyses as one tool to assess Value for Money (section 7.6), and an application of the cost-effectiveness analysis to the E4D example (section 7.7).

Table 7.1.1
Value for Money Definitions

No.	Organization/ Institution	Definition	Reference
1	DFID	Maximize the impact of each pound spent to improve poor people's lives.	DFID (2011, p.3)
2	Christian Aid	Christian Aid's approach to Value for Money is about achieving the best results we can with the money and resources we have. In defining the 'best' results, we are concerned with scale (numbers of people benefiting), depth (intensity and sustainability of change) and inclusion (in other words, a change has greater impact if it benefits people who are more excluded and marginalized).	Christian Aid (2012, p.1)
3	Oxfam Australia	The best use of resources to contribute to positive significant change in the most vulnerable people's lives. Significant change includes: Consideration of scale (the number of people benefitting); Depth (the intensity and sustainability of change); and Inclusion (the change benefits people who are vulnerable and marginalized).	Besley (2016, p.5)
4	Penny Jackson, OECD	The optimum combination of whole-life cost and quality (or fitness for purpose) to meet the user's requirement. It can be assessed using the criteria of economy, efficiency and effectiveness.	Jackson (2012, p.2)
5	World Bank/ UK treasury	Value for Money is the optimum combination of whole-of-life costs and quality (or fitness for purpose) of the good or service to meet the user's requirements.	World Bank Institute and Public-Private Infrastructure Advisory Facility (2013, p.9)
6	ITAD	Value for money examines the optimal relationship between costs/resources and benefits/outcomes - delivered through processes that transform inputs through activities to outputs which are necessary and sufficient to trigger outcomes	Barnett, Barr, Christie, Duff, & Hext (2010, p.9)
7	King, OPM	Value for Money is an evaluative question about how well resources are used, and whether the resource use is justified.	King (2018, p.7)
8	UK National Audit Office	The optimal use of resources to achieve intended outcomes.	National Audit Office (n.d.)

Source: See rightmost column of table.

7.2 Why evaluate Value for Money?

The concept of Value for Money and, more specifically, the implementation of Value for Money concepts in international development serves multiple purposes. These purposes can be grouped into three categories of (i) optimal resource allocation ex-ante, (ii) learning, and (iii) communication and justification.

1. **Optimal resource allocation ex-ante:**
   During the planning phase of a project, program or portfolio, the costs and benefits should be evaluated and compared against alternative uses of the resources. Such comparisons can help to reduce opportunity costs. A good allocation would be one which compares favorably in comparison to its next best alternative (King, 2018). In the case that the next best alternative is not measurable, value for money can still fulfil the purpose of informing stakeholders if the limited resources are being used legitimately or not (King, 2018; ICAI, 2011; ITAD, 2010).

2. **Learning:**
   Upon conclusion of a project or program, Value for Money assessments can facilitate learning about what successful projects and programs are in terms of their implementation, the achieved benefits given the invested resources and how resources could be reduced by holding the benefits constant (King, 2018). A continuous process of assessing Value for Money and identifying projects, programs, and portfolios with better or worse Value for Money will contribute to building a knowledge base that can support and guide the optimal resource allocation ex-ante and allows to channel funds most effectively in order to achieve the best possible impact.

3. **Communication and justification:**
   Development projects are often funded by taxpayers, who do not benefit from development activities directly. The effective use of tax revenues is, therefore, particularly important. Value for Money assessments can help to communicate more clearly and, thus, effectively how tax revenues are used, that funds are well and efficiently managed, and can demonstrate the benefits that are achieved with tax revenues (King, 2018; Jackson, 2012). Further, Value for Money can produce a strong evidence to defend the validity of aid and can demonstrate that those in charge are consistently working to produce better results.

7.3 Caveats of Value for Money

Although Value for Money is conceptually compelling, its implementation is challenging and involves the following considerations:

1. **Risk of a shift in policy focus towards easily assessable programs and portfolios:**
   Value for Money could lead to a shift in policy focus away from issues or topics that potentially require policy attention and towards issues and topics that can be more easily assessed. For example, projects, programs, or portfolios that target complex social, behavioral, or institutional changes are more difficult to assess than interventions that target more easily measurable outcomes such as income (King, 2018).

2. **Risk of shift in policy focus towards target groups that tend to have higher Value for Money:**
   Value for Money could lead to a shift in policy focus away from target groups that potentially require policy attention and towards target groups for which Value for Money tends to be higher (DFID 2011). For example, if interventions are more costly to

implement in marginalized or remote groups because they are more difficult to reach, then the Value for Money for this target group may be lower than for a more accessible and less disadvantaged group (assuming the same value). But the consequence of this should not be that interventions should not be implemented among marginalized groups. Similarly, simply because an intervention is cheaper to implement in country A than in country B, the consequence should not be that all project funds should be channeled towards country A.

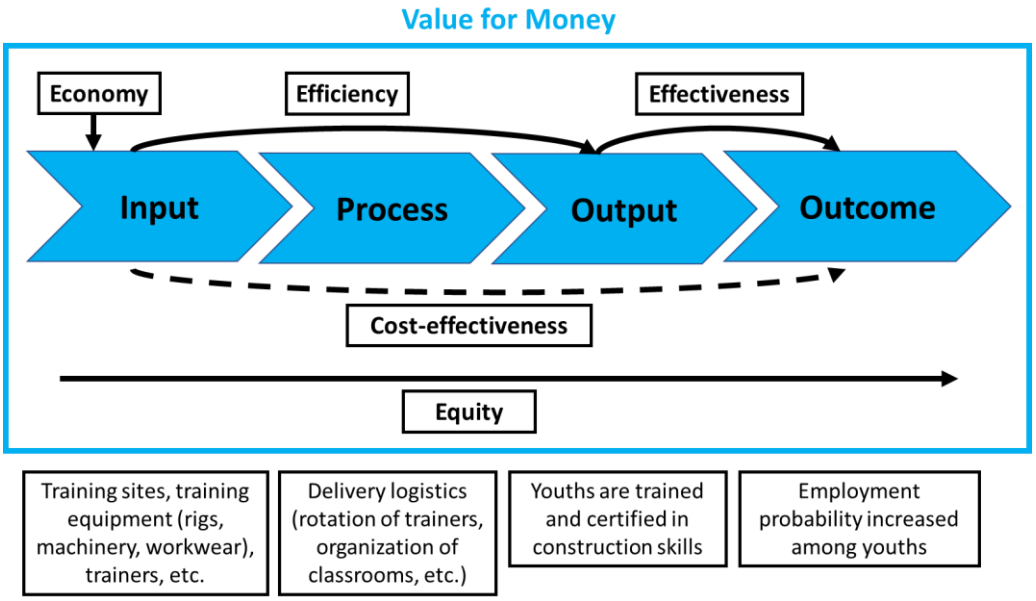
3. **Limited comparability of Value for Money in different contexts, interventions, and outcomes:**
   It is difficult to compare Value for Money across projects, programs, or portfolios as both the costs and the benefits across interventions are often not comparable owing to the difference in contexts, interventions, and outcomes. For example, the benefits of a vocational training program in Uganda measured in additional jobs created is difficult to compare with a deworming program among school-aged children in India.

The potential issues associated with the concept of Value for Money mirror the characteristics of good Value for Money. It is important to carefully consider the target group, context, type of intervention, and outcomes. Importantly, good Value for Money should target the most disadvantaged groups and not merely those for whom high benefits can be achieved at low costs.

7.4 Conceptual framework of Value for Money: *The 4E approach*

Most approaches to Value for Money follow, in one way or another, a framework by the Department for International Development (DFID) referred to as the *4Es* which assesses the following dimensions of a project, program or portfolio: (i) its economy, (ii) efficiency, (iii) effectiveness, and (iv) equity (Figure 7.4.1). Figure 7.4.1 illustrates how the 4Es come into play at a different stage of the result chain - either at the result chain's input stage, its outputs stage or as an outcome.

Figure 7.4.1
VfM Conceptual Framework (DFID, 2011)



Source: Own illustration based on DFID (2011.

In the following, we will introduce each of the 4Es in more detail and provide E4D-specific examples using the S4C program in Table 7.4.1.

Economy regards the acquisition of inputs - for example, raw materials, services, and staff - and asks whether input costs could be reduced while holding the quality of outputs constant. The evaluation of `Economy' includes the following considerations (Jackson, 2012; DFID, 2011):

- Can the same or equivalent inputs be bought for less money?
- Would the use of cheaper inputs be causing costs elsewhere?
- Would the use of cheaper inputs affect the effectiveness of the intervention?

Efficiency regards the use of inputs in their conversion to outputs. It is about the management and logistics of delivering or implementing an activity. The evaluation of `Efficiency' includes the following considerations (Jackson, 2012; DFID, 2011):

- Can the inputs be used or managed differently so that the same results are achieved but fewer inputs used?
- Would the different use of inputs and different management of activities cause costs elsewhere?
- Would the different use of inputs and different management of activities affect the effectiveness of the intervention?

Effectiveness regards the translation of outputs into outcomes and includes the following considerations (Jackson, 2012; DFID, 2011):

- Are the outcomes intended to result from the activities and outputs achieved?

Equity regards the targeting of activities, outputs, and outcomes. It ensures that target groups of interest are reached and that not only those benefit from the activities and outputs for whom high Value for Money can be achieved. The evaluation of `Equity' includes the following considerations (ICAI, 2011; Jackson, 2012; DFID, 2011):

- Were the poorest, most disadvantaged, or most marginalized population groups reached?

Some Value for Money frameworks omit the concept of 'Equity' and, therefore, refer to the framework as *The three E's*, i.e. Economy, Efficiency, and Effectiveness. However, the concept of 'Equity' is important to address the second caveat discussed in section 7.3. If equity is not considered and if program selection would strictly follow Value for Money, then the poorest, most disadvantaged, or most marginalized groups may not be targeted at all. Clearly, this contradicts the purpose of development aid.

Table 7.4.1
The 4Es: concept and examples

Concept	Considerations	S4C Example
Economy	Can the same or equivalent inputs be bought for less money? Would the use of cheaper inputs be causing costs elsewhere? Would the use of cheaper inputs affect the effectiveness of the intervention?	*Could rigs used in the level 2 training be bought cheaper?* *Do the cheaper rigs need fixing or maintenance more often?* *Are cheaper rigs more difficult to handle, which could slow down the training such that fewer exercises can be conducted and learning outcomes would be lower?*
Efficiency	Can the inputs be used or managed differently so that the same results are achieved but fewer inputs used? Would the different use of inputs and different management of activities cause costs elsewhere? Would the different use of inputs and different management of activities affect the effectiveness of the intervention?	*Could trainers be working at one training site only rather than switching sites every other training cycle?* *Although travel time and costs may be reduced if trainers do not regularly switch sites, trainers at less liked sites may stay absent more often and the organization of replacement may be difficult and costly* *If trainers do not regularly switch sites, trainers at less liked sites may stay absent more often which would result in less training hours and lower learning outcomes.*
Equity	Were the poorest, most disadvantaged, or most marginalized population groups reached?	*Did the reach-out and recruitment of participants target the poorest?* *Did the poorest participate in and complete the S4C program?* *Did the poorest improve their construction skills, employability, and employment probability?*

Source: Own illustration.

The concepts of the 4Es illustrate that Value for Money is about weighing and balancing different alternatives and aspects of how to acquire, organize and manage resources, inputs, activities, and outputs. Reducing the resources used to implement an activity may increase Value for Money. However, when resources are reduced typically also effectiveness is affected. If effectiveness is reduced due to the cost reduction, then Value for Money may not increase but decrease. Thus, the net result of the reduction in costs is relevant.

All of the 4Es are important to consider separately and in unison. An activity may be economically and efficiently implemented but not effective. Hence, cutting costs further will not improve Value for Money. Reversely, Effectiveness is important but only in joint consideration with Economy and Efficiency. It can make sense to focus on and to invest in an activity with relatively low effectiveness if the resources required to implement the activity and to generate the output are comparatively cheaper than for the high effectiveness activities.

7.5 Approaches to Value for Money

The implementation of Value for Money concepts such as the 4Es into daily activities is a challenging task in practice. Objective and universally applicable guidelines or implementation procedures are lacking as Value for Money is highly case specific due to differences in contexts, interventions, outcomes and target groups across programs, projects, portfolios, and agencies.

The Economy concept requires the understanding of costs of one's agency's own spending and also the spending of partners. It requires information about prices and the quality of inputs that are used and of those that pose alternative inputs. Further, it probes to understand how actual and alternative inputs translate into outputs and outcomes. Similar challenges arise in the evaluation of Efficiency.

DFID's Value for Money framework, which is guided by the concepts of Economy, Efficiency, Effectiveness, and Equity comprises "A [Value for Money] cycle which ensures that we allocate and manage resources so that we achieve the maximum development impact; and […] six [Value for Money] enablers that help drive the quality of our work within this [Value for Money] cycle" (DFID 2011, p. 9). DFID enablers are actors at all levels of the agency, ranging from the Management Board to program staff, each fulfilling different tasks of planning, scrutinizing, monitoring, measuring, and reporting. The Value for Money cycle is equally comprehensive, including (1) business plans, (2) spending reviews, (3) operational plans, (4) indicator targets, (5) procurement, (6) financial management, (7) program management and monitoring, and (8) evaluations and learning. A summary of Value for Money enablers and the cycle can be found in Table 7.5.1.

Table 7.5.1
DFID's Value for Money enablers and the Value for Money cycle

VfM Enablers	Process/ Work Stream/ Behavior	Value for Money Purpose
Our Skills and Behaviors	**Diligence, Good sense and skills of DFID staff**	– Staff behaviors are crucial to enabling Value for Money. – For the Value for Money agenda to succeed we all need to show diligence in what we decide to do and how we implement (knowing what works and being sure that we are delivering development results in the most effective and efficient way). – This involves a high level of skill and diligence from our staff.
Transparency	**UK Aid Transparency Guarantee and International Aid Transparency initiative**	– Publishing strategy documents, project data and documentation, and individual payments opens our decisions to public scrutiny and enables the public to judge whether we are delivering Value for Money.
	Existing Management decision-taking process	– The first line of scrutiny comes internally through our management chain. – Managers take responsibility that all aspects of their work represent good Value for Money.
Internal Scrutiny	**Internal Audit**	– Provides internal scrutiny and informs the Accounting Officer (the Permanent Secretary) and DFID managers on the whole system of risk management, control and corporate governance.
	Quality Assurance Unit	– The quality of DFID appraisals and the Business Case more generally is assured through the Quality Assurance Unit and Chief Economist.
	Special Measures for Multilaterals	– Improve the performance of multilaterals offering poor Value for Money - signal to these multilaterals that DFID wants an urgent improvement in its performance and will monitor progress carefully. Progress will be reviewed after two years.
External Scrutiny	**Independent Commission for Aid Impact (ICAI)**	– Independent evaluations that focus on whether value for money has been achieved for UK taxpayer and impact of aid has been maximized. – Systematic process of follow up on recommendations resulting from these evaluations fosters lesson learning.
	National Audit Office/ Public Accounts Committee	– Continue to carry out Value for Money studies on DFID to scrutinize whether we are delivering good Value for Money.

Table 7.5.1 continued

VfM Enablers	Process/ Work Stream/ Behavior	Value for Money Purpose
Results & Value for Money tools	Evidence	– We need to understand what works best to achieve our results and ensure investment decisions represent good Value for Money; RED is leading on research to establish what works and what doesn't. – If evidence is flawed or assumptions don't hold, we risk not achieving our results. – DFID will open in July 2011 a quality assessed internet-based evidence databank, which includes evidence papers, systematic reviews and access to a wide range of research and e journals, so staff know what type of interventions are likely to work in particular situations.
	Results & Value for Money Indicators	– The results agenda is about being clear what outputs and outcomes we can realistically expect from an intervention · If we don't know what we're buying we can't assess whether it represents good Value for Money. – Indicators help us measure results achieved. – Value for Money indicators are useful tools to help signal whether interventions represent good Value for Money. – Key international institutions have knowledge leadership roles (e.g., World Bank) and roles in benchmarking performance and unit cost data. We need them to do that work so we in turn have better data on our results & Value for Money.
	Sector Value for Money Framework, guidance andcapability	Strengthen the quality of our sector Value for Money work: – Develop a set of core indicators in sectoral results and VFM frameworks and embed at each level of DFID Business Planning. – Undertake greater internal benchmarking on unit costs and comparators of sectoral investments. – Provide relevant sectoral evidence on what works to achieve results at good value for money. – Monitoring of aggregate portfolio performance. – Build DFID wide sectoral results and Value for Money capability. – Engage international system and institutions to get better results and Value for Money from global sectoral investments - and strengthen international benchmarking of country results and unit costs.
	Commercial Strategy	– Focusing strategically on getting Value for Money from all commercial activities.
System Development	Innovating and Improving	– Bringing in improvements to Value for Money cycle e.g., BAR and MAR for allocation process, new Business Case, Logframes, Annual Reviews etc.
Influencing Partners	Ability to affect others	– Being aware that DFID is often only one of many funders in each organization, and so may only hold a small fraction of the political and financial capital – Nonetheless, DFID must effectively use its financial and human resources to influence third parties to implement Value for Money reforms at a more holistic level (as outlined in the MAR) and engage other partners and donors in support of this agenda. Coordination at both HQ and country level helps to strengthen our ability to do this.

Table 7.5.1 continued

VfM Enablers	Process/ Work Stream/ Behavior	Value for Money Purpose
Value for Money Cycle	**DFID Business Plan**	– Sets-out our areas of focus. – Indicators on our impact and costs reported publicly. – Performance against the indicators enables the public to judge whether we are delivering good value for money at the corporate level.
	SR & Allocations (BAR, MAR, HERR etc.)	– Spending Review sets-out resources available to achieve the Business Plan. – BAR, MAR and other reviews help us to choose how to allocate our resources on the basis of results in our areas of focus (from Business Plan) and Value for Money considerations.
	Operational Plans	– Sets-out our plans to deliver results from BAR, MAR etc. – Getting the right staffing mix to achieve the results. – Teams create their own strategy to ensure Value for Money in their program.
	Corporate performance framework	– Indicators and targets enable Management Board to assess whether DFID is on track to deliver agreed results in line with planned costs
	Business Case	– Embeds evidence and consideration of value for money early in the decision making and design stages of all projects and programs. – Multilateral business case sets out the reforms and results we want to see to improve value for money from core financing; · Strengthens DFID's commercial awareness to ensure intervention design yields better value for money in procurement. – Clarifies our approach to evaluation from the outset and in the design of the intervention; and – The quality of appraisal and the Business Case more generally is assured through the Quality Assurance Unit and Chief Economist.
	Procurement	– Procurement best principles drive the commercial aspect of Value for Money. Involves: – Being clear what we need. – Having a sound procurement strategy which identifies most appropriate funding method and route to market to deliver maximum Value for Money. – Developing well focused terms of reference that enable bidders or other Partners such as CSOs and Multilaterals using DFID funds to be clear what they are being asked to deliver and by when. – Having appropriate evaluation criteria and assessment, conducting good negotiations. – Maintaining appropriate relationships with the supplier, carrying out effective monitoring. – Payment by results. – Learning lessons.

Table 7.5.1 continued

VfM Enablers	Process/ Work Stream/ Behavior	Value for Money Purpose
	Financial Management	– Ensures that all spending decisions are guided by Value for Money principles. – Ensures that resources are allocated to programs where they can make the most impact. – Provides clarity on how much interventions cost and from which budget. – Provides a clear framework to reallocate funds where appropriate. – Ensures good monitoring, reporting and accountability of expenditure. – Enables the organization to continually drive down costs and promote efficiency across the organization.
	Monitoring and Management	– Ongoing program management to keep interventions on track to achieve intended results. – Annual reviews to place emphasis on scoring based on a comparison of the results actually achieved against those expected at project design stage. – Sectoral Portfolio reviews to consider sector wide Value for Money issues and lesson learning; and – Approvers regularly assess whether intervention still represents good Value for Money.
	Evaluations & Lesson Learning	– Evaluating the results and effectiveness achieved from DFID's policies, programs and partnerships plays a key role in demonstrating whether specific interventions achieved good Value for Money. – Going forward – embedding evaluation through a decentralized system with greater ownership from teams enables us to measure results, value for money, and build evidence base about what works. – Teams will be encouraged to go for studies focusing on impact and cost effectiveness, with a strong underlying theory of change.

Source: DFID (2011), pp.12-15.

Not only is the integration of Value for Money into daily activities challenging, but also the evaluation of Value for Money. In particular, Economy and Efficiency require qualitative and to some extent subjective assessments. The Independent Commission for Aid Impact, an organization tasked to scrutinize UK aid spending, suggests a review for the purpose to assess Value for Money based on four guiding criteria – objectives, delivery, impact, and learning – and propose a number of guiding questions alongside with these four criteria (ICAI, 2011, pp. 11-13). For example, one of the guiding questions as part of the `Delivery' criterion is "Is there a clear view of costs throughout the delivery chain?" (ICAI, 2011, p.13). For each review conducted based on these criteria they propose to use a traffic light system to judge Value for Money. However, the assessment tools and methods required to apply the criteria and to decide on a traffic light judgement are diverse and case specific. ICAI (2011) states: "We will need to develop the right approach to answering the questions we set ourselves for each review. To do this, we will use a wide range of assessment tools, drawing on the most appropriate for the task in hand." (ICAI 2011, p.14).

The evaluation of `Effectiveness' is somewhat more clearly defined. Through a toolkit of econometric methods, the impact of an intervention, i.e., the change in the outcome that is causally attributable to the intervention, can be measured. Although methodologically well understood, the implementation of such econometric methods is challenging in practice as data requirements

are high and the methods cannot be applied in all contexts. Examples of effectiveness measurements pose chapters 2, 4, 5, and 6 of this report, i.e., the quantitative evaluations of the KAM, RtW, and S4C programs and of the income information treatment.

These strategies and frameworks highlight that there is no one-fits-all solution to Value for Money and that it requires a comprehensive, time consuming and joint effort of GIZ management, E4D program management, E4D program teams (locally and centrally), local project teams, and project partners. Central elements as part of this effort constitute transparency, scrutiny, monitoring and reporting. As such, the existing planning and monitoring of result matrices and logframes as well as procurement strategies already constitute important processes to enable Value for Money. However, this could be systematically improved if agency-wide guidelines, scrutiny mechanisms, and staff trainings were offered.

7.6 Tools to evaluate Value for Money

While Value for Money is taking place as part of day-to-day activities of aid workers or agency staff, a range of tools exist that attempt to quantitatively evaluate Value for Money and, thus, enable the comparison of Value for Money across programs, projects, and portfolios. Fleming (2013) proposes six tools that can be used for Value for Money evaluations:

1. Cost effectiveness analysis
2. Cost utility analysis
3. Cost benefit analysis
4. Social return on investment
5. Rank correlation of cost vs. impact
6. Basic efficiency resource analysis

Table 7.6.1 summarizes these six tools, including a short description, their advantages and disadvantages, and their similarities and differences. The most commonly used tools are the cost-effectiveness analysis and the cost-benefit analysis. Both compare costs with impact. However, the cost-effectiveness analysis uses the outcome of interest to describe the impact, i.e., the number of beneficiaries that found employment, whereas the cost-benefit analysis attaches a monetary value to the outcome. The cost-benefit analysis, thus, directly measures whether benefits exceed costs and enables to compare interventions with different outcomes, for example, the impact of HIV treatment on health versus the impact of a technical skills training on employment probability. The main disadvantage of the cost-benefit analysis is that it is immensely difficult to attach monetary values to outcomes. For example, how much is it worth to save a life or improve life quality through antiretroviral therapy? Or, how much is it worth to improve female empowerment through financial literacy training of women?

The cost-effectiveness analysis is often considered the more viable option. It also summarizes a complex intervention in a ratio and, hence, allows to compare interventions easily. It indicates, for example, how much it costs to create one job or how many jobs can be created per monetary unit such one EUR or USD. The comparability is straightforward for interventions with the same outcome, e.g., employment probability, but is also possible for interventions with different outcomes, e.g., health and employment probability. When cost-effectiveness ratios of different outcomes are compared the evaluator herself can weigh the two outcomes against each other and may have good reasons to come to a subjective conclusion. Further, many programs and portfolios target specific outcomes, such as E4D targets to improve employment rates, incomes, and job quality. Thus, in many cases comparisons within the same outcome classes are being made naturally.

Table 7.6.1
Tools for evaluating Value for Money

Method	Description	Advantages	Disadvantages	Similarities and Differences
Cost-Effectiveness Analysis	The evaluation of two or more alternatives, based on the relative costs and outcomes (effects), in reaching a particular goal. This method can be used when comparing programs that aim to achieve the same goal.	This method is well suited to the comparison of alternatives that are being considered for reaching the same goal.	– This method cannot compare alternatives with different goals – It cannot make an overall determination of whether a program is worthwhile in an absolute sense—this method will not help determine whether total benefits exceed total costs, only whether an alternative is a relatively more cost-effective solution.	Cost Effectiveness and Cost Utility analyses are useful for evaluating programs that aim to reach the same goal in non-monetary terms. For education programs, that might mean a goal of increased school enrolment, attendance, completion, or cognitive development. The main difference between the two methods is that CU takes beneficiary perspectives into account. Well known applications of CU analysis are in the health sector, with the use of Quality Adjusted Life Years (QALYs). The QALY allows each potential program to be measured according to the extent to which it extends life expectancy while also improving the quality of each year lived. Developing this indicator involves determining satisfaction derived from different health states.
Cost-Utility Analysis	The evaluation of two or more alternatives by comparing their costs to their utility or value (a measure of effectiveness developed from the preferences of individuals). This method can be used where monetizing outcomes is not possible or appropriate. This method is most commonly used in health through quality-adjusted life years (QALY). The QALY allows the comparison of medical interventions by the number of years that they extend life.	– Makes careful attempts to consider individual preferences – A large number of potential outcomes can be included in the evaluation Can contribute to consensus building and participatory decision-making as stakeholders are called upon to assess their preferences for diverse outcomes.	– Results are often difficult to reproduce among different evaluators because of the numerous and sometimes conflicting methodologies that are used to estimate importance weights.	

Table 7.6.1 continued

Method	Description	Advantages	Disadvantages	Similarities and Differences
Cost-Benefit Analysis	The evaluation of alternatives by identifying the costs and benefits of each alternative in money terms and adjusting for time. This method can be used to identify if a course of action is worthwhile in an absolute sense - whether the costs outweigh the benefits - and allows for comparison among alternatives that do not share the same objective or the same sector.	This method can help determine: – A comparison between alternatives with different objectives – That any particular alternative has benefits that exceed its costs – Which of a set of alternatives within a given sector has a higher ratio of benefits to costs.	Benefits and costs must be assessed in money terms. For this reason, this method is best used when the majority of benefits can be converted to monetary values or when those that cannot be converted are unimportant or are similar among the alternatives considered.	Cost Benefit Analysis and Social Return on Investment evaluate whether a program is beneficial in an absolute sense. They both monetize outcomes. Both methods allow for comparison of programs with different objectives or from different sectors. The difference between them is that Social Return on Investment measures social, environmental, and economic costs and benefits.
Social Return on Investment	Measures social, environmental, and economic costs and benefits. Like Cost Benefit analysis, Social Return on Investment analysis can be used when comparing programs with different goals or in different sectors.	Can contribute to participatory decision making as stakeholders are called upon to identify and value programs outcomes • The Social Return on Investment Network provides an assurance process that ensures the analysis has been completed to a consistent standard.	Cost data can be disputed as different evaluators use numerous and sometimes conflicting methodologies to derive value.	
Rank Correlation of Cost vs Impact	Allows for the relative measurement of Value for Money across a portfolio of initiatives.	This method can help determine a comparison between alternatives with different objectives – Can be useful for multi-unit programs. It shows the impact and performance of each unit relative to other units. – Can contribute to participatory decision making as stakeholders are called upon to identify and value program outcomes.		Rank correlation of cost vs impact and Basic Efficiency Resource Analysis both evaluate the relative costs and benefits of many programs. The first method ranks and correlates costs and impact while the second examines relative value by plotting programs on a four-quadrant graph based on costs and impacts.

Table 7.6.1 continued

Method	Description	Advantages	Disadvantages	Similarities and Differences
Basic Efficiency Resource Analysis	Provides a framework for evaluating complex programs by comparing impact to resources and offering a relative perspective on performance where units analyzed are judged in comparison to other peer units.	This method can help determine a comparison between alternatives with different objectives – Can be useful for multi-unit programs. It shows the impact and performance of each unit relative to other units. – Can be used to aid discussions into the performance of units, their challenges, opportunities, and operating environment.	Simplifies complex information and should not be relied on alone. It should be used in conjunction with other data, and never as the only analytical approach.	

Source: Fleming (2013), pp. 6, 8, 9 (Tables 2 and 4).

7.7 Cost-effectiveness analysis

7.7.1 Ingredients and procedure

A cost-effectiveness analysis can be carried out in eight stages. The first four stages regard the effectiveness measurement, the next two stages the cost measurement, and the last two stages regard combining effectiveness and costs in a ratio.

Part I: Effectiveness measurement

1. **Identify competing interventions:**
   Identify interventions for which Value for Money should be evaluated. This could be interventions that are competing for resources. For example, when a new project portfolio is set-up and different projects from a previous project phase are considered, projects from that previous phase could be selected.

2. **Identify the appropriate outcome:**
   Identify the best outcome to determine the effectiveness of the programs. This step is necessary to enable comparability across interventions and to ensure Value for Money is measured for an outcome that is of most relevance for the considered types of interventions. For example, for skills development training the best outcome could be employability – i.e., abilities that enable participants to gain and maintain a job – or it could be employment – i.e., the number of people in jobs.

3. **Estimate the impact of the interventions:**
   The impact of the interventions, i.e., the changes in the identified outcome that can be causally attributed to the interventions, needs to be estimated in order to establish the effectiveness of the interventions. Note the importance of the interventions in question causing changes in outcomes. If causality is not established and, for example, before-after comparisons are used instead, the cost-effectiveness ratio becomes vague in some cases it may even be misleading - and would not pose reliable grounds for

decision making and learning.
While establishing the causal impact, the following points should be kept in mind:

a) In some impact evaluations the treatment group consists of individuals who are offered the treatment and participate in it and people who are offered the treatment but decide not participate in it. The impact estimate resulting from this composition of the treatment group is called intention-to-treat effect (ITT). In other evaluations the treatment groups only consists of individuals who participated in the treatment. The impact estimates resulting from this set-ups is called treatment effect on the treated (ToT). These two impact estimates cannot be compared directly and require a small adjustment to restore comparability (Dhaliwal, Iqbal et. al., 2013).
The ITT is the aggregate impact of the ToT among eligible treated individuals and the zero effect among eligible untreated individuals. Therefore, information on the share of treated individuals can be used to derive the ToT from the ITT. For example, in the KAM program not all participants were placed. If the treatment group was defined as those who were in the pool of trainees that KAM companies considered for internships, then for the estimation of ITTs, this pool of trainees would form the treatment group. If the IIT was 0.1 (e.g., 10 percentage points in employment probability) and 70% of the treatment group was placed in an internship, then the ToT would be 0.14=0.1/0.7. Reversely, if the ToT is 0.14 and 70% of eligible people were treated, the ITT is 0.1=0.14*0.7. The ITT is typically considered the policy relevant measure as policy programs are often offered to people but participation is not mandatory.

b) Cost-effectiveness calculations should only consider interventions as effective that have a “statistically significant” impact. An estimated effect is said to have a “significant” impact when its p-value is less than 0.1. The p-value is a measure of confidence about effect estimates and is important because differences in outcomes between treatment and control group can occur at random chance rather than due to the intervention. To better understand the concept of p-values, imagine that a technical skills training program improved the employment probability by 5 percentage points (0.05) and that the effect estimate has a p-value of 0.2. The p-value of 0.2 implies that there is a 20% probability that the intervention in fact did not improve employment. However, if the p-value was 0.02, this would imply that there is only a 2% probability that the intervention did not improve employment. Evidently, the lower the p-value, the greater the confidence that the effect is larger than zero. A p-value of 0.1 (10%) or less is commonly accepted as a threshold for which the probability of a zero effect is sufficiently low and one can, thus, be sufficiently confident that there is indeed a positive (or negative) treatment effect. In order to check for the sensitivity of a cost-effectiveness ratio to the level of significance, one can compare the intervention’s cost-effectiveness ratio based on its point estimate to its cost-effectiveness ratios based on the effect’s upper and lower bound of the confidence interval (Dhaliwal et. al., 2013).

c) It can be interesting to consider the impact of the intervention for different groups of the population. The effectiveness of interventions in subgroups is not required to calculate cost-effectiveness, but it can provide insightful details as to how the characteristics of participants could impact the intervention and how future interventions could be optimally targeted.

4. **The total impact of the intervention:**
Impact estimates typically provide average and per person effect estimates. For example, the technical skills training program improved participants' incomes by 100 EUR. However, as total costs are typically reported rather than costs per person, also the total impact must be calculated based on the causal effect estimate, the number of beneficiaries and the duration of the program. The result of the following calculation can be used to calculate the cost-effectiveness ratio:

$$TotalImpactOfProgram = Impact * SampleSize$$

Part II: Cost measurement

5. **Gathering the cost data:**
Cost estimates of each intervention need to be calculated. It is important to include all costs of an intervention as the omission of resources will distort the cost-effectiveness ratio and may lead to wrong conclusions. Consider two soft skills training interventions, A and B, that were selected to be evaluated and compared. For both interventions only the costs that incur with project partners are available, but intervention A consumed much more time effort from GIZ staff than intervention B. Ignoring the GIZ staff costs will reduce the costs of intervention A relative to intervention B and, thus, erroneously give an advantage to intervention A in terms of Value for Money.
A method to ensure that all costs are gathered is the Ingredients Method. The Ingredients Methods lists all ingredients to an intervention and their prices and allocates them to broader categories. Each cost category must be included in both interventions, if applicable, and the ingredients to each category should be complete. Some common categories are as follows (McEwan, 2012 & J-PAL, n.d.):

a) **Program administration:** This comprises expenditures that incur, for example, in the hiring and paying of staff involved in the implementation of the intervention, costs of office facilities as well as overhead costs.

b) **Targeting costs:** This comprises resources that are spent in the planning and implementation of the outreach phase of an intervention, for example, when identifying and targeting beneficiaries, raising awareness about the intervention or enrolling beneficiaries.

c) **Staff training:** This comprises expenditures on training the staff responsible for implementing the intervention such as costs for training facilities, personnel that trains the trainers, or training equipment.

d) **User training:** This consists of the costs incurred in training beneficiaries such as costs of training facilities, personnel that trains beneficiaries, or training equipment.

e) **Implementation costs:** This comprises expenditures on implementation costs that were not covered by the previous categories - such as personnel that manages daily activities and logistics, equipment required such as laptops or mobile phones as well as phone credit or transportation costs.

f) **User costs:** This comprises costs incurred by the beneficiaries to participate in the intervention such as the opportunity cost of their time or expenses to travel to the training facility. Often these costs are not available, but, in theory, pose an important cost category.

g) **Averted costs:** This comprises costs that were successfully avoided due to the implementation of the program and should be subtracted from the final costs of the intervention. Often these costs are not available, but, in theory, pose an important cost category. For example, if an intervention enrolls participants into health insurance and, in consequence, participants have better health, then costs for the treatment of illnesses, cost from absenteeism at work and income losses are averted.

h) **Monitoring costs:** This comprises expenditures for monitoring the progress and outcomes of an intervention and may include data collection costs, personnel costs, or equipment costs (laptop, software etc.).

If cost data are not available at all, one could alternatively look for information on costs of comparable interventions based in similar contexts. An example for this poses Kremer and Miguel (2004) who evaluated the effectiveness of the distribution of deworming drugs in primary schools in Kenya. It is important to note that this approach should be used if and only if, there is no way of gathering costs using the ingredients method. Using a cost estimate from another intervention requires the interventions to be extremely similar. In order to ensure the similarity, a rigorous check of the key assumptions from the parent study needs to be done. This task will be very challenging and time-consuming. Therefore, great care needs to be taken before resorting to this approach of estimating cost data.

For both approaches, the final amount of total costs must be presented together with a list of categories and ingredients included to reach the total. Further, cost calculations should be accompanied by an appropriate discussion of caveats and the implications that the omission or inclusion of certain ingredients or categories have for the cost-effectiveness ratio of each intervention. In addition, if interventions of different contexts, regions or countries are compared, different price levels that the two interventions are facing should be critically reviewed and commented on.

Note that in case the intervention has multiple impacts, the evaluator could consider the program as a whole and divide the cost of the whole program by the number of outcomes. For example, a skills training program to improve participants' employability can positively affect both their technical skills and their employment status. The evaluator can then take the entire cost of the intervention, divide it into two, and calculate the cost-effectiveness ratio for the two sets of outcomes (i.e., number of people who improved technical skills and who found employment) separately.

6. **Converting the costs into common units:**

The costs gathered through the Ingredients Methods may not be directly comparable across interventions, but also even within one intervention itself. Costs may incur over

a longer period of time during which inflation can occur. Further, often different currencies are used to pay for ingredients. Therefore, the cost ingredients need to be inflation and exchange rate adjusted in order to be comparable (Dhaliwal, Iqbal et. al., 2013).

Part III: Combining effectiveness and costs

7. Calculating the cost-effectiveness ratio:

Using the total costs calculated in steps 5 and 6 and the impact of the interventions as derived in steps 1 to 4, the cost-effectiveness ratio can be calculated using the following formula:

$$Cost - effectiveness\ ratio = \frac{Total\ impact\ of\ program}{Total\ costs\ in\ common\ units}$$

The obtained ratio indicates the increase or decrease in the outcome for a unit amount of money spent. For example, it could indicate how many people were able to find employment for every Euro spent on providing technical training to people. Alternatively, one could reverse the numerator and denominator such that the ratio indicates the amount of money needed to generate a unit increase in the outcome – e.g., how much money needs to be invested in a technical training program to enable one person to find employment.

8. Sensitivity analysis:

Finally, the sensitivity of the cost-effectiveness ratio can be checked using the bounds of the confidence interval as suggested in step 3. To do so, the cost-effectiveness ratios based on the interventions' estimated point estimates, i.e., the results from the main analysis, need to be compared to cost-effectiveness ratios that use the upper and lower bound of the effect estimates' confidence intervals. For example, consider intervention A, which has a point estimate of 0.1 and a confidence interval of 0.15 (upper bound) and 0.05 (lower bound), and intervention B, which has a point estimate of 0.08 and a confidence interval of 0.12 (upper bound) and 0.04 (lower bound). The main analysis would have used the point estimates of 0.1 (intervention A) and 0.08 (intervention B). In the sensitivity analysis, the calculation of the cost-effectiveness ratios would then be repeated for the upper and lower bonds for each intervention.

If the ranking of cost effectiveness ratios is the same across the different effect sizes applied (i.e., point estimates and upper and lower bounds), then the sensitivity analysis confirms the results of the main analysis based on point estimates and, hence, one intervention is more cost-effective than the other. However, if the cost-effectiveness rankings are different in the sensitivity analysis, the results of the main analysis are called into question and evoke a more detailed discussion. In the above example, intervention A has a higher point estimate than intervention B, but the point estimate of intervention B is more precisely estimated, such that the lower bound effect size of intervention B is higher than of intervention A. The cost-effectiveness ratio of course also depends on the total costs, but in such scenarios the ranking of the cost-

effectiveness ratios in the main analysis and the sensitivity analysis may not coincide. In addition to the bound analysis, the sensitivity analysis should critically review cost assumptions as discussed in step 5 and any other assumptions or aspects of the cost-effectiveness calculations that can have implications for the obtained ratio and, thus, the conclusions to be drawn from the analysis.

7.7.2 Example calculation for the KAM and S4C program

This section follows the steps outlined in section 7.7.1 to calculate the cost-effectiveness ratio for E4D programs.

1. **Identify competing interventions:**
   The interventions for which the cost-effectiveness ratios should be calculated are the RtW program and the S4C program in Uganda as well as the KAM program in Kenya. The programs were chosen because all three programs constitute large skills training and internship placement programs, which are typical for the E4D portfolio. Importantly, all three programs were planned to be continued during E4D's second program phase.
   However, the evaluation of the RtW program in section 4 did not show significant impacts. Further, the S4C program only showed significant employment effects for subgroups of beneficiaries, but not for the sample on average. Therefore, the example calculation in this section focusses on the KAM program only.
   Furthermore, the example calculation focusses on the comparison of beneficiaries that participated in the whole program.

2. **Identify the appropriate outcome:**
   The outcome of interest is decent employment as this is the main employment indicator of E4D which combines having paid work of at least 20 hours per week and a minimum income from that work of at least 6,209.93 KES per month.

3. **Estimate the impact of the interventions:**
   Section 2 quantitatively evaluated the effect of the KAM program on decent employment. The results suggest that the KAM program, i.e., having completed the skills training and an internship placement, improved decent employment by 28 percentage points. If only the effect of the placement in addition to the skills training is considered, the estimated point estimate is 13 percentage points (0.127). As the group of KAM beneficiaries consist of individuals who were trained and placed as well as only trained, also the treatment effect of completing the KAM skills training only is required, which is unknown. It is assumed that the difference between the effect of completing both components and of only the placement is the effect for the training only: 0.283 – 0.127 = 0.156, i.e., 16 percentage points. Note, that this is a strong assumption because interaction effects between the training and the placement may exist, i.e., the training alone may not be as effective as in combination with the placement. The sensitivity analyses in step 8 will use a different assumption about the effectiveness of the KAM training to calculate the cost-effectiveness ratio. Further, the sensitivity analysis will

use the confidence interval bounds to calculate the cost-effectiveness ratios.
The two relevant effect estimates are thus 28 percentage points (0.283) for beneficiaries who completed both KAM program components and 16 percentage points (0.156) for beneficiaries who completed only the KAM skills training. Both effects present treatment effects on the treated (ToT). As there is no information available about the number of people who were offered the KAM program, the intention-to-treat effect (ITT) cannot be calculated and the ToT is used.

4. **The total impact of the intervention:**
In total, 1,019 individuals received the skills training and a subset of those also received a placement. Unfortunately, the number of placed individuals is not known and needs to be deduced from the sample of study participants. 803 of the study participants were trained and 504 additionally completed an internship. Thus, 63% (=504/803) of study participants received both program components. Applying this ratio to the total number of trained participants, i.e., 0.63*1,019, would suggest that roughly 642 individuals were trained and placed.
The number of 642 beneficiaries can be used to calculate the total impact of completing the skills training and a placement, which is the average treatment effect times the number of beneficiaries: 0.283 * 642 = 182. Hence, 182 individuals are in decent employment due to participating in both KAM program components. Note, that this calculation assumes that the treatment effect among study participants is the same as among individuals who were beneficiaries but not part of the quantitative evaluation.
In addition to the roughly 642 individuals who were trained and placed there are roughly 377 individuals (=1,019-642) who were only trained. Given the assumed treatment effect of 16 percentage points as discussed in step 3, the total impact of the training is: 0.156 * 377 = 59, i.e., 59 beneficiaries are in decent employment due to participating in the KAM training.
The total impact is thus 182 + 59 = 241, i.e., 241 individuals are in decent employment due to participating in the KAM program.

5. **Gathering cost data:**
The available cost data are limited. Available are an estimate of the in-kind partner contribution, which is 180,480 EUR, and a GIZ contribution of 583,201.94 EUR. These figures include all costs incurred in the implementation of the KAM program; however, it is not known how these costs split over specific implementation cost categories. The figures include salaries of GIZ's project management staff, but do not include costs for M&E activities, other project administration by GIZ, user costs or averted costs. Thus, the resulting cost-benefit ratio will be upward biased to a substantial extent and should not be used for comparisons across projects.

6. **Converting the costs into common units:**
The costs gathered under step 5 are indicated in one currency (EUR) and, therefore, do not require exchange-rate adjustment. The KAM program ran from October 2016 to October 2019 and inflation likely occurred during these three years. However, it is

unknown when during this period the costs incurred exactly. For illustration purposes, we assume that a third of the total budget (254,560.65=763,681/3) was spent each year (2016, 2017, and 2018) in October and convert it to January 2021 prices using inflation-tool.com:

Table 7.7.1
Example calculation for inflation adjustment

Year	Costs at time of Expenditure in EUR	Inflation factor	Costs in January 2021 in EUR
2016	254,560.65	1.04	264,743.076
2017	254,560.65	1.05	267,288.683
2018	254,560.65	1.04	264,743.076
		Total	797,774.835

Source: Own calculations.

The total costs of the program adjusted for inflation amount to 797,774.835 EUR as of January 2021. If the resulting cost-effectiveness ratio was to be compared to the cost-effectiveness ratio of other interventions, the cost of the other intervention would also need to be expressed in January 2021 EUR.

7. **Calculating the cost-effectiveness ratio:**

To calculate the cost effectiveness ratio, the total impact is divided by total costs:
241 jobs / 797,774.835 EUR = 0.000302 jobs/EUR
The cost-effectiveness ratio suggests that one EUR spend creates 0.000302 decent jobs. Alternatively, one could reverse the ratio and divide the costs by the number of decent jobs (797,774.835 EUR / 241 jobs = 3,310.27 EUR/job), which would suggest that one decent job costs 3,310.27 EUR.

8. **Sensitivity analysis:**

Table 7.7.2 shows the results of a number of sensitivity analyses. The sensitivity analyses include the calculation of the cost-effectiveness ratio based on (i) the assumption that the implementation costs make up two-thirds of total costs, (ii) the assumption that the KAM training alone did not have an effect on decent employment (i.e., effect size of zero), (iii) using the upper bound of the confidence interval of the KAM training and placement effect for the impact calculation, and (iv) using the lower bound of the confidence interval of the KAM training and placement effect for the impact calculation. The resulting cost-effectiveness ratios based on alterations (i) to (iii) range between 0.00021 and 0.00024 jobs per EUR invested and are thus relatively similar, suggesting that these are quite reasonable alterations. As expected, the cost-effectiveness ratio based on the upper bound of the confidence interval is much larger, amounting to 0.00036 jobs per EUR invested.

Table 7.7.2
Sensitivity analyses of the cost-effectiveness ratio

Topic	Changes	Cost-effectiveness ratio:
Cost categories	The total costs under step 5 (763.681,94 UR) only include implementation costs. In the sensitivity analysis, it is assumed that implementation costs make up 2/3 of the total costs. This would result in total costs of 763.681,94 EUR * (3/2) = 1,145,522.91 EUR	241 jobs / 1,145,522.91 EUR = 0.00021 jobs/EUR
Treatment effect of KAM training	A treatment effect of 16 percentage points was assumed among beneficiaries who only participated in the KAM training but did not complete an internship placement. This is a strong assumption, and a more conservative assumption would be a KAM training effect of zero. In this case the total impact would only amount to 182 decent jobs created: (0.283 * 642) + (0 * 377) = 182 jobs	182 jobs * 797,774.84 EUR = 0.00023 jobs/EUR
Upper bound of confidence interval	The upper bounds of the confidence interval of the KAM training and placement effect is 0.3552445. The confidence interval of the training only effect is unknown, as it was calculated from the difference in point estimates from two other effects, and sensitivity analysis of the training only effect is conducted above. Therefore, this sensitivity analysis will only vary the effect the KAM training and placement had. The calculation of the total impact is: (0.355 * 642) + (0.156 * 377) = 287 jobs	287 jobs / 797,774.84 EUR = 0.00036 jobs/EUR
Lower bound of confidence interval	The lower bound of the confidence intervals of the KAM training and placement is 0.211227. The confidence interval of the training only effect is unknown, as it was calculated from the difference in point estimates from two other effects, and sensitivity analysis of the training only effect is conducted above. Therefore, this sensitivity analysis will only vary the effect the KAM training and placement had. The calculation of the total impact is: (0.211227 * 642) + (0.156 * 377) = 194 jobs	194 jobs / 797,774.84 EUR = 0.00024 jobs/EUR

Source: Own calculations.

7.8 Conclusion

Value for Money is about the balancing of resources and impact of the invested resources. Value for Money is not an ad-hoc evaluation of costs and impact but is *implemented* in daily activities of project-level agency staff and agency management. This is important because Value for Money is often confused with the *assessment* of Value for Money through cost-effectiveness analyses. Value for Money can be used to guide, steer, and justify individual development projects, specific country programs or whole agency portfolios.

The most common conceptualizations of Value for Money follow DFID's 4Es framework – Economy, Efficiency, Effectiveness, and Equity -, which consists of considerations that regard the economic spending on resources, the efficient conversion of inputs to outputs, the effectiveness of outputs in reaching outcomes (i.e., the impact of an intervention), and the equitable targeting of interventions in order to reach the poorest or most marginalized population groups (DFID, 2011).

There are no one-fits-all guidelines on how to approach and implement the 4Es framework as Value for Money depends on many complex program-, context- and agency-specific conditions and procedures. To guide the implementation of Value for Money, DFID (2011) suggests the use of a Value for Money-cycle that touches upon various project-related aspects and organizational levels, including (1) business plans, (2) spending reviews, (3) operational plans, (4) indicator targets, (5) procurement, (6) financial management, (7) program management and monitoring, and (8) evaluations and learning. Applying this to the GIZ or, more specifically, E4D-context implies that Value for Money requires a comprehensive and joint effort of GIZ management, E4D program management, E4D program teams (locally and centrally), local project teams, and project partners. Central elements as part of this effort constitute transparency, scrutiny, monitoring and reporting. As such, the existing planning and monitoring of result matrices and logframes as well as procurement strategies already constitute important processes to enable Value for Money. However, this could be systematically improved if agency-wide guidelines, scrutiny mechanisms, and staff trainings were provided.

A common tool to *evaluate* (not to implement) Value for Money is the cost effectiveness analysis. It summarizes a complex intervention in a ratio of total impact to total costs and, hence, allows comparisons of interventions easily. It indicates, for example, how many jobs can be created per monetary unit such as one EUR spent or how much it costs to create one job. The computation of cost-effectiveness ratios itself is simple. However, the inputs for its computation require elaborate efforts because (i) the causal impact of the intervention needs to be assessed and before-after comparisons that are part of standard M&E procedures are not sufficient and (ii) total costs of the intervention's implementation, administration (locally and centrally), and M&E that incur at the agency and the partner as well as the time when the cost incurred need to be known in order to convert the costs to common units that are inflation- and exchange rate-adjusted.

The causal impact of an intervention and total costs are often not known. The required information on costs, i.e., all cost items, the time when the costs were incurred and the currency they incurred in, should, in theory, be known through accounting procedures and could be made available for such exercises. This assumes that GIZ staff who are involved with the program in question, such as local program managers, administrative staff, and staff in GIZ's headquarter, are clocking the time they are spending on specific programs. However, an intervention's causal impact cannot be drawn from existing agency structure. As rigorous impact evaluations are costly, it is important to carefully choose the interventions for which they are conducted. Important selection criteria could be whether the intervention is a 'typical' intervention of the program portfolio and whether it is an intervention that is considered to be continued in future program phases.

To illustrate the computation of the cost-effectiveness ratio, the results of the quantitative KAM program evaluation were used. The quantitative evaluation suggested that beneficiaries who participated in the KAM skills training and who were subsequently placed in an internship improved their decent employment probability by 28 percentage points. However, not every participant of the KAM program was placed in internships and for those who only attended the KAM training an effect size of 16 percentage points in decent employment was assumed, which resulted from the difference between the effects from participating in the training and the placement (0.28) and participating in the placement in addition to the training (0.12). In a sensitivity analysis, an effect size of zero is assumed for beneficiaries who only participated in the training.

For the cost data, information on total implementation costs in EUR were available. As the specific timing of when the costs were incurred was unknown, inflation adjustments could not be appropriately conducted and, for the purpose of illustration, it was assumed that a third of the costs occurred in each of the three project years. Because the costs were indicated in EUR already, no exchange rate adjustments were needed. The total costs do not include costs that regard the management, administration, monitoring and evaluation of the project, which results in an upward bias of the Value for Money assessment. If the calculated cost-effectiveness ratio is compared to other cost-effectiveness ratios, the omission of the cost categories of the KAM program may result in misleading conclusions. Thus, a careful consideration of the included cost categories is indispensable. In a sensitivity analysis it is assumed that implementation costs account for two-thirds of all costs, however, this may still be an optimistic assumption.

Overall, the implementation of Value for Money and the assessment of Value for Money require considerable efforts at various levels of the agency and project stages. A number of monitoring and procurement mechanisms are already in place that enable Value for Money and such mechanisms could be more systematically extended. The implementation further requires the internalization of Value for Money principles into daily activities. All agency staff could be purposefully trained to implement Value for Money into their specific tasks. The assessment of Value for Money requires more detailed cost reporting, which could potentially build upon existing accounting structures. Further, rigorous evaluations would need to be conducted in order to learn about the causal impact of an intervention. All together, these requirements seem challenging, but they are not prohibitive and have the potential to contribute to more efficient and effective development aid.

8 Indirect and Induced Employment Effects

8.1 Concept

Employment effects refer to economy-wide changes in employment due to an intervention or investment. Job creation as a result of an intervention occurs through different channels and at different levels and, therefore, does not only affect beneficiaries, i.e., participants of an intervention, directly but also triggers a large variety of effects on individuals and companies outside the primary group of beneficiaries. These, sometimes unintended, employment effects can be seen as ripple effects of the initial intervention effect. Positive ripple effects include jobs that are created due to additional local consumption as a result of higher earnings of beneficiaries, whereas negative ripple effects include substitutions of workers outside the group of beneficiaries with beneficiaries.

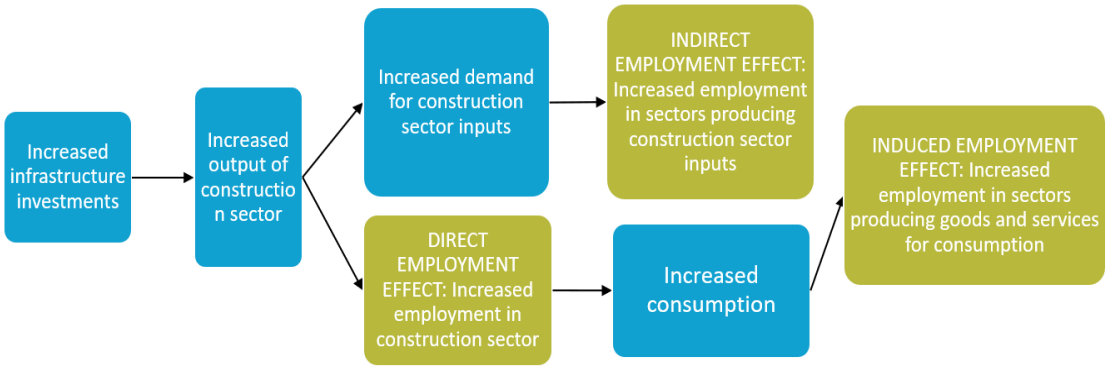
In gauging the economy-wide job-creation effect of an intervention an assessment and subsequent cumulation of three individual effects is necessary: direct, indirect, and induced job creation (IFC, 2013). While direct employment effects are typically the primary goal of an intervention, indirect and induced employment effects are triggered by the direct effect of an intervention on its beneficiaries through ripple effects (Kluve and Stöterau, 2014). Measuring only direct effects inaccurately reflects the overall benefits of an intervention and involves the risk to design future program portfolios towards interventions with large direct effects but low economy-wide effects.

Although the benefits of measuring economy-wide effects are salient in theory, in practice its measurement poses challenges that are difficult to overcome. This is particularly the case for human capital interventions, such as the skills trainings that are the subject of this report's impact evaluations, and, more generally, development cooperation projects that do not entail large infrastructure investments. The subsequent sections summarize the state of the literature in measuring economy-wide employment effects (henceforth employment effects) and illustrate the challenges that arise in the context of human capital focused interventions. Acknowledging these challenges, a conceptualization of employment effects for human capital interventions put forward by Kluve and Stöterau (2014) is introduced and an exemplary calculation of employment effects based on multipliers of human capital interventions from the literature is applied to the interventions that are studied in chapters 2 to 5.

8.2 Literature and Implementation of Theory

In the literature, the concept and measurement of employment effects are predominantly discussed in the context of infrastructure projects and large financial investments (Bacon and Kojima, 2011; IFC, 2013; MacGillivray et al., 2017; ILO, 2020b). In these contexts, the theory of change of employment effects is based on the idea that infrastructure increases production or that financial investments allow a business to grow. The additional output that is or will be generated through these interventions requires more labor input - and, thus, creates *direct employment effects* - and non-labor input. The higher demand for non-labor input supports and creates further jobs along the companies' or sector's supply chain, which creates *indirect employment effects*. Finally, households benefitting from additional income through the direct and indirect employment effects increase their consumption of goods and services and, thereby, further increase output and in consequence jobs. This increase in employment through increased consumer demand is referred to as *induced employment effects*. The impacts triggered by infrastructure interventions and the associated employment effects are illustrated in Figure 8.2.1.

Figure 8.2.1
Employment effects of infrastructure investments



Source: Own representation based on ILO (2020b).

A cost-effective tool for the calculation of employment effects of an intervention is the use of employment multipliers. A multiplier adjusts estimates of created direct jobs for the direct jobs' ripple effects on indirect and induced employment. A multiplier is greater than one if it is expected that the intervention generates benefits to the economy that exceed their immediate impact on direct employment (Kluve and Stöterau, 2014). The magnitude of an employment multiplier depends highly on the characteristics of the intervention and the context in general. Generally, estimates of employment effects through multipliers can only be interpreted as an approximate benchmark (USAID, 2016).

Employment multipliers can be measured at two levels – at the microeconomic and at the macroeconomic level (Kluve and Stöterau, 2014; Bacon and Kojima, 2011). At the microeconomic level a bottom-up approach is used and the analysis focuses on either single persons or companies on the market. While at the person-level the interventions' effect on a participant's employment probability can be analyzed, at the company-level the effect on the company's employment and competitiveness is measured. At the macroeconomic level a top-down-approach is used and the employment multipliers of interventions are measured for more aggregate levels such as for a region, sector or country. The first approach is typically used to estimate an aggregate employment impact for a given portfolio, whereas the latter is typically used to estimate changes in aggregate indicators on the basis of input-output tables for a specific sector or an entire country (Kluve and Stöterau, 2014).

The estimation of employment multipliers is particularly challenging because the counterfactual scenario, i.e., how economy-wide employment would have changed in the absence of the intervention, is difficult to determine and data requirements are substantial (Kluve and Stöterau, 2014). Often only a partial analysis of employment effects is possible which captures indirect effects under specific circumstances but usually does not include effects outside the target group (Kluve and Stöterau, 2014). Ideally, micro-founded macro models are used, which are complicated and not viable in practice.

Because most of the literature on employment multipliers refers to interventions with external output shocks (portfolio investment, etc.), a practical top-down approach to derive multipliers is offered by so-called Input-Output Tables (IOTs) and Social Accounting Matrices (SAMs). IOTs and SAMs provide links between inputs and outputs for several sectors in the economy based on historical data (ILO, 2020a). In detail, IOTs describe the sale and purchase relationships between producers and consumers within an economy and show the flows of intermediate and final goods

and services defined according to industry outputs for a given year. However, IOTs do not contain data on the expenditure patterns of economic actors (government, enterprises, and households). Here the SAM enters the picture, which is an expansion of the IOT and adds other actors in the economy: government, savings and investments, households, and the rest of the world (imports and exports). Since it includes all actors and agents in the economy, a SAM can also be described as a tabular form of the circular flow of income in the economy. Neither SAMs nor IOTs contain any labor market information as such and, therefore, the estimated employment effects are based on the extent of the changes in output (USAID, 2016). Since IOTs and SAMs only contain financial information, i.e., the transactions between sectors and actors in the economy, they are linked to sectoral-level employment data to provide macro-level employment multipliers and, thus, overall employment effects of certain interventions in the economy.

Although the use of IOTs or SAMs is a comparatively simple method for the estimation of employment multipliers, the implementation of such top-down approaches is still methodologically demanding and the required data are typically not readily available. Elaborate IOTs exist only for some countries - mainly European countries and only for a few African countries (MacGillivray et al., 2017). Existing IOTs are often outdated and only include highly aggregated sectors, in some case sectors are not separated at all (Bacon and Kojima, 2011; ILO, 2020a). However, sector-specific results most likely deviate widely from results based on sector averages. A major methodological disadvantage of IOTs and SAMs is their static approach. The IOT or SAM analysis assumes that input-output relationships and prices do not change dynamically, as they would in reality, and, further, assumes no supply constraints (ILO, 2020a). Both the IOT's static approach and the unconstrained supply can lead to under- or overestimations of employment effects and, thus, the resulting employment multipliers can be misleading (IFC, 2013).

Like other human capital interventions, E4D projects are targeted at individual beneficiaries (e.g., TVET students or HR staff at a company) and, thus, at a different level of implementation than vast investment programs, which are targeted at large companies or sectors and are expected to have macroeconomic impacts. Therefore, approaches and definitions established in the literature often do not apply to the case of human capital interventions. In the context of human capital-focused portfolios with individual-level interventions – such as skills trainings and job placements or entrepreneurial trainings of staff at selected small and medium enterprises (SMEs) - IOTs and SAMs are not suitable methods to derive employment multipliers. For interventions that are targeted at individuals, there are no supply chains that can be analyzed. For firm level interventions, supply chains could theoretically be analyzed, but the IOT or SAM method postulates an increase in output and this increase would have to be measured and causally attributed to the human capital intervention first. Further, firms would need to be willing to share information about their supply chain in order to calculate indirect and induced employment effects occurring through the increased input-demand of the beneficiary firm as a result of the output increase caused by the intervention.

Because the main literature on employment effects and employment multipliers does not offer a feasible approach for interventions included in the E4D portfolio, the subsequent sections discuss a conceptualization of economy-wide employment effects in the context of individual-level human capital-focused interventions put forward by Kluve and Stöterau (2014).

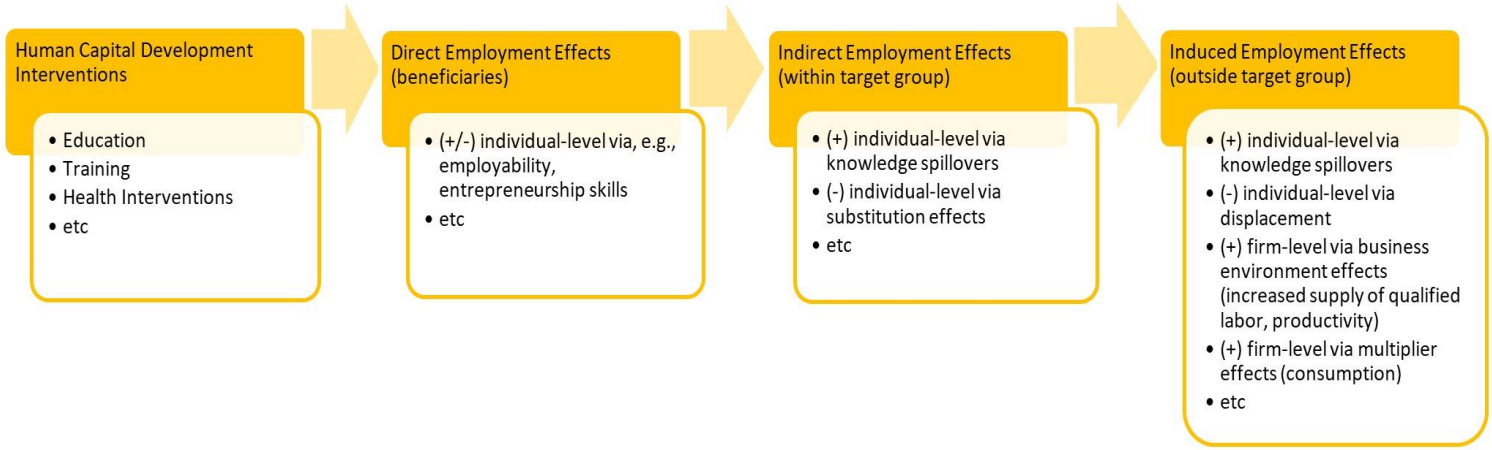
8.3 The case of human capital interventions

8.3.1 Conceptualization of employment effects of human capital interventions

Kluve and Stöterau (2014) provide a conceptualization of employment effects for individual-level interventions at the example of human capital-focused interventions. They suggest an adapted theory of change which is based on the identification and differentiation of three groups: (i) beneficiaries, (ii) the target group, and (iii) the non-target group. Every intervention has a target group for which it aims to reach a specific outcome; for example, an intervention is set out to improve the employment probability (= *outcome*) of adolescent girls in region A (= *target group*). Beneficiaries pose a subgroup of the target group, they are actually participating in the intervention and, therefore, are affected by the intervention outputs; for example, adolescent girls in region A who participate (= *beneficiaries*) in the training (= *output*). The non-target group includes all actors, who are not part of the target group, i.e., older females or males in region A or females and males of any age outside of region A (= *non-target group*).

The identification of these three groups allows to conceptualize the theory of change of employment effects of human capital interventions, which is visualized in Figure 8.3.1. A skills training intervention, for example, may lead to an increased employment probability of beneficiaries through improved employability or signaling of a certificate – these are *direct employment effects*. Individuals of the target group who are outside of the beneficiary group may also improve their employment probability, for example, through knowledge spillovers from beneficiaries, resulting in *indirect employment effects.* However, indirect employment effects also include substitution effects and, therefore, can also be negative. Substitution effects occur when open positions are filled with beneficiaries which in the absence of the skills training would have been filled with a non-beneficiary of the target group.

Figure 8.3.1
Multistage job-creation effect of human capital development interventions



Source: Own representation based on Kluve and Stöterau (2014), p. 17.

In a final step, *induced employment effects* may occur for individuals outside of the target group due to knowledge spillover effects or the increased overall consumption of employees who benefited from the direct or indirect employment effects. The latter leads to increased hiring in sectors producing goods and services for consumption. However, also at this stage negative induced employment effects are possible through displacement effects. Displacement effects are similar to substitution effects but occur outside of the target group. At the individual level, displacement effects occur when a person of the target group is hired instead of a person outside of the target group because, for example, the training intervention established a positive image of the target group. Box 8.3.1 summarizes the definition and logic of employment effects of human capital-focused interventions based on Kluve and Stöterau (2014).

Direct employment effects: Typically, the primary goal of an intervention. All changes in employment outcomes that are directly caused by the outputs of the intervention *among its beneficiaries.*

Indirect employment effects (also referred to as supply multiplier): Triggered by the direct effect of an intervention on its beneficiaries. Changes in employment outcomes *among the target population* which are caused by direct (employment and non-employment) effects of the intervention. Indirect effects comprise (e.g.) multiplier effects or substitution effects. Typically, indirect effects are more difficult to evaluate than direct effects since they cannot be measured among the beneficiaries
*** Substitution effects:** Occur within the target population, meaning that participants in a program take jobs that individuals who did not participate would otherwise have held.

Induced employment (also called consumption or income multiplier): Changes in employment outcomes among individuals and firms which are not part of *(i.e., outside) the target group* of an intervention. The effects are induced through the entire initial economic effect of an intervention – including direct and indirect effects. These include employment effects along the value chain (multiplier effects), effects of an altered business environment, or displacement. Generally, not a goal of an intervention and, therefore, often referred to as (un)intended consequence.
* **Displacement effects:** Potential positive employment effects within the target group are offset by negative employment effects for individuals/firms outside the target population.

Box 8.3.1: Definition of the three different employment effects, Kluve and Stöterau (2014).

8.3.2 The calculation of employment effects of human capital interventions

In the context of employment effects, the terms gross and net effects are commonly used. However, the understanding of gross and net effects in the multiplier literature differs from how the terms are used in impact assessments or monitoring and evaluation. In the multiplier literature, gross effects are often those that do not take into account indirect and induced employment effects, whereas net effects do (see for instance European Commission, 2007). In impact assessment, however, net effects are referred to direct employment effects that can be causally attributed to an intervention, i.e., taking into account the counterfactual outcomes of beneficiaries in the absence of the intervention, whereas gross effects typically refer to employment figures measured through before-after comparisons (Kluve and Stöterau, 2014). In the following sections, the impact assessment definition of net effects is used. This definition of a net effect is also equivalent to the direct employment effect and, thus, implies that net effects are a necessary ingredient for the calculation of economy-wide employment effects.

The formula to convert net- or direct effects into (economy-wide) employment effects provided by Kluve and Stöterau (2014) is:

$$Employment\ Impact = Net\ Effect\ x\ (1 \pm indirect\ Effects\ \pm induced\ Effects)$$

$$= Net\ Effect\ x\ \ (1 - Substitution)\ x\ (1 - Displacement)\ x\ (1 + Multiplier)$$

Based on this formula, the net effect is multiplied with the relevant employment effect parameters. In the aggregated form this includes the multiplication of the net (direct employment) effect with the parameters for indirect and induced employment effects. More specifically, the indirect and induced multipliers are disassembled into their respective positive components, captured by the general multiplier, and negative components, captured by the substitution (indirect) and displacement (induced) effects.

8.4 E4D Application

Because the estimation of E4D-specific multipliers is not feasible as this would require data from individuals outside of the target population as well as a sophisticated model of the Ugandan economy and effect mechanisms of E4D interventions within this economy, multipliers from the literature on employment effects of human capital interventions were gathered and applied to the estimated E4D net effects from chapters 2 and 5. Specifically, multipliers from RWI (2013) were adopted and are presented in Table 8.4.1.

The application of these multipliers to the estimated E4D net effects serves two purposes. First, to illustrate how multipliers could be applied, if available, and, second, to highlight how much multipliers can vary and, therefore, provide very uncertain and, to some extent, even arbitrary estimates of employment impacts.

In Table 8.4.1, most parameters are based on papers with employment effects calculated for European countries and interventions that contain either only partially a human capital component or different types of human capital interventions, e.g., wage subsidy programs. Hence, the employment effects and, thus, multipliers of these interventions may differ from typical E4D-interventions in African economies. The magnitudes provided by these multipliers are very broad and present only rough reference values for initial orientation, they do not present ultimately valid parameter ranges.

It is important to emphasize that the multipliers presented in Table 8.4.1 are highly context-specific and an appropriation of multipliers which are calculated for other countries, regions, interventions, and years is not advisable as common practice for GIZ monitoring or reporting.

Table 8.4.1 presents a range of parameter values from RWI (2013) that are used as examples of multipliers for human capital interventions. A parameter of 0.3 implies that one direct job supports 0.3 additional jobs. Values are presented as ratios of the net effect. Thus, in the case of negative effects, such as the displacement and substitution effects, the parameter X is used as a factor of (1-X) of the direct employment effect. For positive indirect or induced effects, the parameter X is used as a factor of (1+X) of the direct employment effect (see above formular).

Table 8.4.1
Employment multiplier estimates from the literature

Human Capital Development	
(positive) Indirect Effect	0.05 - 0.1
(positive) Induced Effect	0.3 - 0.5
Substitution	0.2 - 0.4
Displacement	0.1 - 0.3

Source: Based on RWI (2013). The search included Bondonio and Martini (2012), Centre for Strategy and Evaluation Services (2006), Criscuolo et al. (2012), English Partnerships (2008), European Union (2013), Greenberg et al. (2011), Maré (2005), Mouqué (2012), National Audit Office (2003).

Using the parameters of Table 8.4.1 and the formula provided by Kluve and Stöterau (2014), one KAM and one S4C net effect of the impact evaluations were identified to be used in the calculation of (economy-wide) employment effects. In order to compare the employment effects of the two programs, the net effects would ideally be available for the same output, the same intervention - e.g., skills training and internship placement versus control or skills training and placement versus skills training only -, the same subpopulation, as well as the same follow-up timing of the data collection. Unfortunately, this cannot be implemented as no significant effects were identified for exactly the same characteristics in both projects. As a second-best solution, as many characteristics as possible were held constant and the effects that seemed of most importance in each project were selected. For the KAM project, this is the main effect of the training and placement in comparison to the no-intervention control group on decent employment as measured 10 to 15 months after the baseline survey for the whole study population, which is 28.1 percentage points or a parameter of 0.281. For S4C, the net effect of the training and placement in comparison to only the training on decent placement as measured in November 2020 for participants aged 25 years and older was selected, which is 17.0 percentage points or a parameter of 0.170. For these two effects the interventions and sub-populations differ, whereas the outcome as well as the timing of the outcome measurement were held constant to the extent possible.

Based on the identified KAM and S4C net effects and the formula provided by Kluve and Stöterau (2014) as discussed above, the calculation of lower bound employment effects, using the lower end of the range of the respective multipliers in Table 8.4.1, and upper bound employment effects, using the upper end of the range of the respective multipliers in Table 8.4.1, are illustrated. The calculation is shown in depth for the KAM project, whereas the calculation results for the S4C project are presented in Table 8.4.2.

The lower bound employment effect calculation is:

$$\begin{aligned} Employment\ Effect &= Net\ Effect\ x\ (1 - Substitution)\ x\ (1 - Displacement)\ x\ (1 + Multiplier) \\ &= Net\ Effect\ x\ (1 - Substitution)\ x\ (1 - Displacement) \\ &\qquad x\ (1 + (positive\ indirect + induced\ Effect)) \\ &= 0.281\ x\ (1 - 0.4)\ x\ (1 - 0.3)\ x\ (1 + (0.05 + 0.3)) \\ &= \underline{0.159} \end{aligned}$$

The upper bound employment effect calculation is:

$$\begin{aligned} Employment\ Effect &= 0.281\ x\ (1 - 0.2)\ x\ (1 - 0.1)\ x\ (1 + (0.1 + 0.5)) \\ &= \underline{0.324} \end{aligned}$$

The calculation yields an estimated overall KAM employment effect range of 0.159 (lower bound) to 0.324 (upper bound), which means that the KAM program increases the probability of decent employment by 16 to 32 percentage points taking into account direct, indirect and induced employment effects. Conducting analogous calculations for the S4C program yields an estimated employment effect range of 0.096 (lower bound) to 0.196 (upper bound), i.e. the S4C program increases the probability of decent employment by 10 to 20 percentage points taking into account direct, indirect and induced employment effects. As the multipliers are constant parameters, the KAM effect is just a multiple of the S4C effect by a factor of how much the net KAM and S4C effects differ. Table 8.4.2 summarizes the net effects and calculated (economy-wide) employment effects for the KAM and the S4C project.

Table 8.4.1
Calculation results of (economy-wide) employment effects

Project	Net Effect	Employment Effect	
		Lower bound	Upper bound
KAM	0.281	0.159	0.324
S4C	0.17	0.096	0.196

Source: Own calculations.

The lower bound effect estimates are lower than the net effects due to the small positive indirect and induced parameters and large negative substitution and displacement parameters applied in the lower bound calculation. The reverse is true for the upper bound employment effects, i.e., large positive indirect and induced parameters and small substitution and displacement parameters were applied in the upper bound calculation. The upper bound employment effects are more than twice as large as the lower bound employment effects. The differences between lower and upper bound effects illustrate the substantive variation in employment effects that can arise from the choice of multipliers and highlights how arbitrary the calculation of employment effects can be.

8.5 Conclusion and Recommendation

The literature on economy-wide employment effects does not offer a practical approach for the calculation of employment multipliers for human capital-focused interventions similar to those in the E4D intervention portfolio. The top-down approaches based on IOTs or SAMs commonly used in the employment effect literature, are not suitable for applications in the context of human capital interventions. For E4D interventions targeted at firms, the IOT/SAM method would require a measurement of output increases induced by the E4D intervention and knowledge of beneficiary firms' supply chains. For E4D interventions targeted at individuals, the logic of IOTs is not applicable at all as there are no supply chains that can be analyzed. RWI (2013) and Kluve and Stöterau (2014) noted the inadequacy of top-down approaches in the measurement of employment effects of human capital interventions and put forward a conceptualization of employment effects for such types of interventions. Following Kluve and Stöterau (2014), employment multipliers were applied to net direct employment effects estimated in the KAM and S4C evaluations using multipliers previously estimated for human capital interventions implemented in Europe or other high-income contexts. However, multipliers are highly context-specific and it is not recommended to integrate this approach into regular E4D monitoring and reporting. The adoption of those multipliers would raise severe doubts about the validity of the calculated employment effects. Here, the calculation exercise served the purpose to illustrate how multipliers could be applied, if available, and to highlight how much multipliers can vary and, therefore, provide very uncertain and, to some extent, even arbitrary estimates of employment impacts.

9 Lessons Learned for Future Program Designs

Sections 2 to 5 studied the impact of the KAM program in Kenya and the RtW and S4C programs in Uganda. All three programs follow a similar structure, comprising two components – a skills training and, for a subset of participants, an internship placement. The KAM and RtW programs were particularly similar because they both targeted TVET graduates and the training consisted of a two- to three-day work readiness training. In contrast, the S4C program targeted a broader beneficiary group and the skills training lasted for 6 to 10 weeks and focused on technical construction skills.

The main take-aways of the quantitative evaluations of the KAM (section 2), RtW (section 4), and S4C (section 5) programs are:

1. The KAM program and, specifically, the KAM skills training rather than the internship placements significantly improved TVET graduates' decent and formal employment probability and incomes.
2. The RtW program did not affect the labor market outcomes of recent TVET graduates in the short time considered (later data collections were not conducted due to the coronavirus pandemic).
3. The S4C program improved incomes among employed and, specifically, the internship placement component rather than the technical skills training improved the probability of decent employment, formal employment, and employment with a contract of participants who were at least 25 years, TVET educated, and previously had worked in a job for at least six months.

A more detailed summary of the results of the three quantitative evaluations is presented in Table 9.1.

The results of the three quantitative evaluations engender mixed implications for the effectiveness of specific program components. The results of the KAM program suggest that the work readiness training was effective in improving employment outcomes, whereas the internship placements were less effective. In contrast, the results of the S4C program suggest that the internship placements were effective in improving employment outcomes, at least in some subgroups, whereas the comprehensive technical skills training was less effective. The results of the RtW program evaluation do not support either of the KAM or S4C findings.

Despite the heterogeneity in the effectiveness of the three programs to improve labor market outcomes, the results suggest that both skills training and internship placement can be effective to improve employment outcomes in specific settings. Further, the results suggest that the effectiveness of the respective program components depends on the local context, program design aspects, and the target group.

Table 9.1
Summary of the results of the quantitative evaluations of the KAM, RtW, and S4C programs

Country	Program/ Intervention	Main results
Kenya	KAM program	– Vocational training graduates who participated in the KAM training and internship placement experienced a significant improvement in their labor market outcomes. Particularly striking are the large and persistent effects on decent and formal employment as well as on income. Moreover, results show an improved job search performance, a reduced financial dependency, and an increased probability of having a bank account. – The positive effects on labor market outcomes seem to be mainly driven by the work readiness training component rather than the labor market attachment component. – The impact of receiving an internship placement in addition to participating in the KAM training alone is small and insignificant with regards to most employment and labor outcomes or the effects do not sustain in the longer term.
Uganda	RtW program	– The RtW training and placement program did not affect TVET graduates' employment probability or earnings in the short time considered. – Secondary outcomes, such as employment aspirations or migration intentions, were also not affected.
	S4C program	– The S4C program did not affect participants' employment probabilities, but it significantly improved incomes of employed participants. The income increase appears to be driven by a mixture of higher wages and longer working hours. – The S4C internship placements improved the probability of decent employment, formal employment, and employment with a contract of participants who were at least 25 years old, TVET educated, and previously had worked in a job for at least six months. – The S4C program increased participants' employment aspirations and internal migration intentions. – The S4C placement seems to be the most effective program component which drives the program's overall impact.

Source: Own illustration.

The case-specificity of the programs' impacts mirrors well existing evidence. Overall, the evidence for stand-alone training programs is mixed (Betcherman et al., 2004; Blattman and Ralston, 2015; Kluve et al., 2017, 2019), although they showed to be effective in certain populations and contexts (Adoho et al., 2004; Alcid et al., 2014; De Mel et al., 2014). More comprehensive multi-component programs, such as those that combine training and employment services, tend to be more effective (Betcherman et al., 2004; Attanasio et al., 2011; Chakravarty, 2016; J-Pal, 2017; Kluve et al., 2017, 2019; Datta et al., 2018); although the evidence suggests that the inclusion of soft-skills trainings does not improve outcomes further (Kluve et al., 2017, 2019). Programs of higher implementation quality and longer duration tend to perform better (Kluve et al., 2017, 2019). Due to large informal sectors in low-income contexts, positive effects hinge on the demand for skilled labor and a growing economy (Betcherman et al., 2004; Vivarelli, 2014; J-Pal, 2017; Ibarraran et al., 2019; Escudero et al., 20019). Private sector involvement can be effective, for example, through support in curriculum development, as mentors or networking facilitators, or by absorbing trainees in their own companies as apprentices or employees (J-PAL, 2017; Kluve et al., 2017; Datta et al., 2018).

Lending from the knowledge base of the literature, the following subsections discuss specific aspects of the program designs and target groups that potentially contributed to the effectiveness or ineffectiveness of specific program components. The local context touches upon both topics, i.e., design features and the target group, and, therefore, is not discussed separately but as part of the two topics.

9.1 Program design

9.1.1 Trainings

Surprisingly, the short soft skills training of the KAM program was effective, whereas the much more comprehensive technical skills-oriented S4C training was not, or at least to a lesser extent. The S4C training was much longer, lasting for 6 weeks if only the level 1 training was completed and 10 weeks if also the level 2 training was completed. The level 1 training included a 9-day general work readiness training, a 10-day training in basic construction skills, and a two-week international certification program in health, safety, and environmental standards. The level 2 training consisted of training and certification in either rigging or pipe fitting. Thus, the S4C training was longer and more intensive than the KAM training. Further, the S4C program also included work readiness training and dedicated even more training days to the topic of work readiness than the KAM training, which focused mostly on work readiness, comprised overall. Therefore, the S4C skills training's lack of effectiveness cannot be explained by a lack of work readiness skills training.

A context specific factor that is key for the effectiveness of a skills training program is that there is demand for and a supply shortage of the skills that are trained to beneficiaries (Betcherman et al., 2004; Vivarelli, 2014; J-Pal, 2017; Ibarraran et al., 2019; Escudero et al., 20019). One reason why the S4C training was less effective might have been that - although construction skills were effectively transferred to beneficiaries and signaled to employers - construction skills were not in net demand at the time the program's impact was measured or at least not in the places where beneficiaries were searching for jobs. The theories of change of many E4D programs in Uganda assumed large scale investments in the resource sector which would in consequence increase the demand for skilled construction workers. If those investments were delayed or reduced, beneficiaries might have been better off seeking jobs outside of the construction sector. Alternatively, participants may not have been searching for jobs in places where infrastructure investments had been made. This could also explain why the level 2 rigging or pipe fitting training had no additional effect on participants' labor market outcomes in comparison to those who were only trained in basic construction skills (level 1). Potentially, rigging and pipe fitting skills were difficult to market at the time labor market outcomes were surveyed. If the demand for construction skills or, specifically, rigging and pipe fitting skills increases when the anticipated resource sector investments are made, positive S4C training effects may show in the long-term.

A feature of the KAM skills training that might have contributed to its success in comparison to the RtW and S4C trainings could be the organized opportunities for trainees and employers to meet, such as at job bazaars (J-Pal, 2017; Kluve et al., 2017, 2019; Datta et al., 2018). Job bazaars took place after the last day of the KAM training and offered opportunities for trainees and company representatives to meet and mingle. The job bazaars potentially served as a platform to individually seek an internship or job placement. Indeed, the results of the KAM evaluation showed that internship placements did not improve the probability or number of job interviews in comparison to beneficiaries who only participated in the KAM training.

Although the KAM and RtW programs were very similar in structure and length a key difference – in addition to the networking opportunities – existed in who conducted the training. As part of the RtW program, TVET teachers were trained to implement the work readiness training and schools were free to conduct work readiness trainings at times of their convenience. While the training of trainers can have advantages with respect to the sustainability of the implementation of the RtW program, imposing training responsibilities on schools can have implications for the quality of trainings; for example, schools may be less restrictive with respect to the class size of trainings. In contrast, KAM's work readiness training was implemented by a private service provider who had extensive experience in conducting such trainings and who potentially was required to follow stricter implementation protocols than project partners whose business activities did not depend on E4D assignments. Overall, this difference in the training implementation potentially resulted in different training qualities and, hence, varying program effectiveness (Kluve et al., 2017, 2019).

9.1.2 Placements

The qualitative evaluation of KAM's "sustainability program" (see section 3) suggested that internship-offering companies took up interns regularly already before the KAM program started and also employed other interns at the same time as they employed interns facilitated through E4D's KAM program. Similarly, the Kenyan Association of Manufacturers, which was the project partner that facilitated the internship matching of the KAM program (therefore, the program name), also regularly facilitated internships irrespective of the support by E4D. Thus, participants who were not placed into an internship through E4D's KAM program, were potentially successfully seeking internships through job bazaars themselves or through internship matching services offered by KAM and other providers outside of E4D's KAM program. Indeed, over 80% of the pure control group completed an internship. Among beneficiaries who only received the work readiness training but no internship placement through the KAM program, about 20% completed an internship. These insights suggest that the KAM program might not have added significantly to increase internships beyond internships that would be ongoing irrespective of E4D's KAM program, i.e., that would take place also in the absence of the KAM program. Note, this does not suggest that internship placements are not effective in general, but it would suggest that the infrastructure for internship facilitation may be in place irrespective of the placements facilitated through E4D's KAM program.

The internship matching of the RtW program in Uganda was, similar to the KAM program in Kenya, implemented through the local association of manufacturers, i.e., the Ugandan Manufacturers Association. Although qualitative evidence similar to the one available for the Kenyan context is missing, it may well be that internships are also very common among target group members of the RtW program in Uganda, so that also the RtW program would not add significantly to internship placements beyond what is in place already. The RtW study results show that a large but yet smaller share of control group participants completed an internship and control group participants were invited for interviews at similar rates.

In contrast to the RtW and KAM programs, the internship placements of the S4C programs were not facilitated through a manufacturers' association but through a private service provider, who potentially can manage internship placements more efficiently and monitor its quality better. The direct payments for the provider's services might have resulted in more resources for the implementation of placements and its monitoring and, potentially, also generated higher accountability. Anecdotal evidence suggests that the monitoring of KAM internship placements was deficient as the records of the Kenyan Manufacturers Association did not allow to quantify

how many interns were placed in total or at which companies. The monitoring records as received from the service provider of S4C were also often unclear, although, to a lesser extent than KAM records.

9.2 Target group

Participation in the RtW program was conditional on being a final year student at selected vocational training institutes. Similarly, participation in the KAM program was conditional on being a final year student or recent graduate, i.e., in the past five years, of selected vocational training institutes. The target group of the S4C program was broader. The S4C outreach consisted of an open call and interested individuals had to fulfill some eligibility criteria with respect to age, having a national ID, willingness to stay at the training sites in Packwach or Kampala for the duration of the training, and a test on skills required during the training such as, for example, on basic math skills.

These differences in target groups could also have contributed towards the effectiveness of selected program components. The S4C program's impact is most pronounced among individuals who were at least 25 years old, TVET educated, or had previously worked in one job for at least six months. Further, in contrast to the RtW program, the KAM program targeted individuals who were older and had gained first work experiences already, given that beneficiaries were not required to be final year students, but instead their graduation may have been a few years back already. However, the heterogeneity analyses of the program impacts of KAM and RtW by age and work experience do not suggest that older and experienced participants benefited more from the program. In fact, the results of the KAM evaluation suggest that participants without prior work experience benefited relatively more from the KAM program.

9.3 Recommendations:

Based on the discussion of design features and target groups in the respective program contexts, it is difficult to pin down the precise aspects of the programs that drive their effectiveness or ineffectiveness. Yet, some broader recommendations can be deduced:

1. **A careful assessment of the context and the needs of the target groups should be conducted prior to the planning and implementation of interventions:**
   It is important to first define a target group and target outcomes to be able to precisely study and assess the needs of the target population. For example, if the target group are final year TVET students and the target outcome is an increase in the probability of decent employment, then a detailed understanding is required about what the barriers of TVET graduates in finding decent employment are in the specific local context; i.e., what do their job search patterns and methods entail; what kind of jobs are they looking for in terms of salary, position, tasks, and location; are they willing to relocate for a job; are they willing to switch sectors; what were the experiences of former TVET graduates who successfully found and secured a job and what were the experiences of former TVET graduates who did not successfully secure a job.
   Similarly, the requirements on the company side need to be carefully studied with respect to labor demands and skills shortages in the specific context (Betcherman et al., 2004; Vivarelli, 2014; J-Pal, 2017). To ensure that a program meets the needs of the demand side, private sector companies could be involved in the curriculum design of an intervention. Finally, an assessment of whether the needs of the supply and demand sides match should be conducted and how each side's needs can be reconciled, which is where an effective intervention can be targeted at.

Such a needs assessment can be based on previous empirical and quantitative studies in similar contexts, own data analyses of the local or national labor force or household surveys, and qualitative research through interviews and focus group discussions with target groups and companies.
In addition, in the planning of an intervention, it is important to critically review what beneficiaries' outside options are if they do not participate in the E4D intervention. An important aspect of this is to reflect on whether the activity would be implemented without E4D support or whether the program supports existing structures. If the E4D program supports existing structures, then assessing how it improves the existing structures to enhance participants' employment probabilities is required.

2. **Project partners must have sufficient resources for program implementation and monitoring:**
   Most E4D interventions are structured as public-private partnerships and, therefore, E4D programs rely on the respective private or public partner to effectively implement the intervention. In contrast, only a few programs are implemented through service providers, such as the S4C program. The concept of public-private partnerships is promising with respect to the potential sustainability of the program when E4D ends its activities and exits the program. However, the implementation of interventions through private or public partners also involves risks. As opposed to a service provider, the daily business of partners does not rely on funding. Therefore, it is more difficult to hold the project partner accountable if project activities are not implemented as expected or to the desired standard of quality. The consequence of this is not, of course, to stop working with private and public project partners, but to ensure that the project partners have sufficient resources for the effective implementation of project activities as well as for the monitoring of these activities. In particular, if these activities cause additional workload as, for example, the monitoring and reporting standards in accordance with the E4D M&E logframes or result matrices require.

3. **Incorporate quality assurance as project outputs:**
   Project activities and outputs are monitored throughout each E4D project based on the previously determined partner contracts and project-specific result matrices. In addition, project activities and outputs can be designed to enable quality assurance. For example, when an intervention includes training of trainers, the program could design and conduct refresher trainings for the trainers as a program output. Such refresher trainings could, at the same time, improve the sustainability of the program and improve the quality of trainings of end-users through improved training qualities of trainers.

4. **Private sector involvement in the design of the program:**
   The private sector could be involved more intensively in the planning and implementation of an intervention to improve its context suitability and quality (J-PAL, 2017; Kluve et al., 2017; Datta et al., 2018). As discussed in the first recommendation, companies can be involved in the assessments of the market's labor demand. The results of the qualitative evaluation of the KAM program highlighted that the involvement of companies in the planning of an intervention is key for companies that are involved in the intervention implementation to take ownership of the intervention. Further, companies can directly form part of an intervention, for example, through networking forums or job bazaars, through on-the-job training, and by absorbing trainees as apprentices and employees.

10 Lessons Learned for Future Impact Evaluations

Rigorous evaluations of development projects are demanding. They require a close collaboration between researchers, the E4D team at GIZ's headquarter, the local E4D team in the respective GIZ country office, and the local partners or service providers. Despite the overarching aim of learning about effective programs for employment promotion and poverty reduction, impact evaluations often serve or are geared towards different interests, such as reporting, impact communication, or academic rigor. Accommodating all interests is challenging and may collide with implementation realities. For example, the researchers' primary interest may be academic rigor, which often requires very strict and rule-based intervention enrollment and implementation structures that, ideally, should not be deviated from. However, in practice these rules are sometimes difficult to follow or cause additional workload, which is difficult to integrate with daily project management activities and already restrictive time capacities.

Example 1: Targeting and enrollment of study participants into an intervention.

In rigorous impact evaluations, the group of beneficiaries is compared to a group of individuals who are similar to the group of beneficiaries but who do not participate in the intervention. In the researchers' ideal world, the allocation of individuals who participate in the intervention and who do not participate would be based on a random selection mechanism, e.g., either by rolling a dice or using a random computer algorithm, and the assignment to participate in the program or not may not deviate from the result of the random selection mechanism.[64] In contrast, local GIZ program management often do not have fully clear-cut decision rules for intervention enrollment, but are concerned with enrolling the most disadvantaged individuals into the intervention, and have non-random replacement procedures if participants decide not to participate or drop out. In addition, random enrollment is sometimes not perceived as fair by project staff.

Although these two perspectives seem contradicting, they can be reconciled. Random allocation is per se not unfair and can be targeted at the most disadvantaged individuals.[65] For example, the S4C program preceded an outreach and from a pool of interested candidates individuals were systematically selected based on certain eligibility criteria. In order to enable randomization, the outreach could be increased in order to increase the pool of eligible candidates and within this pool of eligible candidates a random lottery can decide who of these equally eligible people is enrolled into the program. Such randomization may also work in strata, i.e., the group is divided into certain groups that should be represented among beneficiaries, such as gender, location of residence, age etc., and within these strata random selections can be made. In most cases randomization can address and concerns of project management. However, it typically requires joint efforts of and additional work for the people involved in the planning and implementation of the intervention and its impact evaluation.

---

[64] *Individuals who were randomly assigned to the beneficiary group may decide not to participate in the intervention. This is not a problem as rigorous impact evaluation methods have ways to deal with non-compliance with the treatment assignment as long as the person is not arbitrarily replaced by a control group individual.*

[65] *Randomization is not always an appropriate evaluation method and can be fair in some settings but not in others. For example, providing a random half of a final year VTI class with wage subsidy vouchers may not be fair as hierarchies arise within classes and schools, whereas randomization may be less of a concern if schools were randomized into the wage subsidy voucher treatment.*

Example 2: Study duration and timing of data collection.

GIZ program phases are typically structured in three-year cycles and study results are often desired to be completed early in order to directly inform program activities. However, the lifecycle of an impact evaluation often takes several years. The impact evaluation design phase ideally starts in the program design phase and follow-up data collections take place only after the intervention had been implemented. Depending on the intervention and the outcome of interest, this could be immediately after an intervention - for example, when the outcomes are soft skills that were transferred in a soft skill training – or a year – for example, when the outcomes are employment status or incomes after a soft skills training. Thus, in some cases it will be difficult to deliver study results in the same program phase that the intervention was implemented. However, if researchers are onboarded early on, the evaluation could focus on the first few months of intervention implementation and not include the last one or two intervention years. Further, multiple follow-up data collections could be conducted, one in which short term outcomes, e.g., soft skills or immediate labor market outcomes, are measured and one in which long term outcomes, e.g., persistent soft skills and medium-term labor market outcomes, are measured. This way, immediate outcomes can already partially inform program activities or the program design for the phase and longer-term outcomes contribute to the systematic learning of GIZ's program impacts and, of course, can also inform program activities of subsequent phases.

These examples illustrate the importance of bringing together the perspectives of practitioners' and researchers' early on to facilitate a productive collaboration for all. **Empathy for the respective other perspective and a strong collective effort have made the collaboration for the quantitative evaluations of sections 2 to 6 a successful example which generated tailor made solutions for the project and evaluation needs at hand, demonstrating that quantitative evaluations of policy relevant projects can be brought into practice.**

The experiences during the evaluation design, planning and implementation in the field led to a number of lessons learned for future impact assessments. Many of these lessons learned are no news but have been stated in previous RWI evaluation reports and elsewhere (Bachmann et al., 2019). **The repeated efforts to bring rigorous evaluation into practice are commendable. They highlight that the way to mainstream evaluation in German development work is a long and complex one. But they also highlight that the process of mainstreaming evaluation is underway and is improving through continual learning of researchers, practitioners, and institutional parameters.**

The lessons learned are grouped in the following four themes:

1- **IIntegration of project management and impact evaluation**

It is key for the planning and implementation of an intervention to go hand-in-hand with the impact evaluation from the start. The beginning of an impact evaluation is already in the conception phase of an intervention, i.e., when the intervention type might be known but its implementation has not been planned. The early inclusion of researchers is essential to design a tailor-made evaluation method that can address the research questions previously identified in a rigorous manner in order to generate the largest possible learning outcomes. It also allows to plan the intervention implementation in accordance with specific methodological requirements of the rigorous impact evaluation, such as sample size, targeting of beneficiaries, set-up of treatment and control groups, timing of data collections and alike.

To integrate project management and impact evaluation well, the practitioners' and researchers' perspectives should be explained and motivated to the respective other party and jointly discussed. For researchers that entails to onboard local practitioners about why rigorous impact evaluations are important, what the potential methods are and what the project implementation and data collection demands as well as advantages and disadvantages of the potential methods are. On the practitioners side it entails to explain the intervention's objectives and target group, the theory of change of the intervention detailing the assumptions, inputs and processes, and providing details about what possible outreach and implementation structures and procedures could be. This allows each side to understand the needs of the others as well as to react to and accommodate the other ones needs to the extent possible in order to enable a productive collaboration.

The most important level of communication is between the local project manager and the researcher. While other practitioners, e.g., in GIZ's head quarter or local M&E officers, play important roles, the project manager is best informed about the project as he or she is involved in the project planning and implementation on a daily basis and is in contact with project partners. Intensive and continual exchange throughout the planning and implementation phases between project managers and researchers is highly important so that the researcher stays thoroughly informed and can react to any changes that occur in the implementation of the program. In order for the project manager to acknowledge and realize how small aspects of or changes in the project implementation can have important implications for the impact evaluation, the previous onboarding is key. Other important actors are local M&E officers who can facilitate the formation of a knowledge base of existing and potential data sources, e.g., data collected as part of the project monitoring or national survey data, and with whom the researcher can closely coordinate data collections that serve the purpose of the regular M&E and the evaluation.

The integration of impact evaluation with project implementation creates workload for E4D staff that is on top of their daily activities and available or additional resources are typically scarce. Thus, when impact evaluations are initiated it is important to discuss how the responsible E4D staff can manage the additional tasks and how other workload can be reduced. In the reduction of the workload the researchers themselves can build an essential part by supporting or assuming responsibilities for some of the M&E activities, such as integrating M&E data collection with data collection of the impact evaluations and conducting data analyses along the E4D reporting guidelines to be used in E4D's regular reports.

2. Specification of research questions

Rigorous impact evaluation can tackle a plethora of research questions. It can evaluate the overall effectiveness of a program as a "black box" or it can speak to the effectiveness of a specific intervention design aspect. If specific aspects of an intervention should be evaluated, then the impact evaluation needs to be set up in a very different way than when the overall impact were to be evaluated. The important point is that the effect of the specific program aspect or component needs to be isolated from other program components. This is ideally done by randomly allocating beneficiaries in different beneficiary groups, one for each aspect of an intervention practitioners are interested in, while holding all other project aspects constant.

For example, one could test how the delivery mode of an intervention influences its effectiveness. Consider a work readiness training which is delivered through a training of trainers at schools (similar to the RtW program) for some beneficiaries and through a professional service provider (similar to the KAM program) for other beneficiaries. In this case, all eligible individuals would be randomly allocated to one of three groups, one training of trainers-treatment group, one service provider-treatment group and one no intervention-control group. This set-up would allow to evaluate the overall effectiveness of each delivery mode in comparison to the no intervention-control group and would also allow to compare the two delivery modes.

The example highlights that the exact research interests need to be discussed ideally before an intervention is planned and implemented. It further highlights that rigorous impact evaluations are limited to speak to the driving factors of an intervention's effectiveness or ineffectiveness if these factors had not been specifically addressed through the research design previously. To gain deeper insights into the how's and why's of an effect or null-effect, qualitative research methods may be useful and can be used complementary to quantitative methods.

The early and clear commitment to research interests and specific research questions is further important to understand ex-ante what can be learned from rigorous impact evaluations. One single impact evaluation may allow to evaluate whether a certain intervention or intervention aspect *was* effective or not, but it is no guarantee that the intervention *will be* effective again at different times or in different settings. The results of one study never provide the answers to the questions raised. Instead, one study can only add to the policy relevant scientific literature and dialogue, which is continually evolving as more rigorous studies are being conducted and which as a whole can provide answers to certain questions.

3. Timing of intervention implementation and data collection

The quality and rigor of an impact evaluation is highly sensitive to the intervention implementation. In order to understand the extent to which treatment and control groups are comparable and where potential treatment effects are stemming from it is important to hold as many aspects of the intervention implementation constant. An intervention type where this is particularly difficult are demand-driven program components such as internship placements where the timing and the match in general depend on companies and beneficiaries. An alternative example is the implementation of trainings at schools through teachers who have previously received a trainers' training and can deliver a training at any time to as many participants as desired. In such cases it is very difficult to judge what the opportunity costs and outside options of beneficiaries are in order to model an adequate counterfactual scenario. If the evaluation method consisted of a randomized controlled trial which measures an intention to treat-effect, these concerns can be reduced. In an ideal scenario, the intervention delivery would take place simultaneously for all participants, but there are many customized implementation designs possible that come close to this ideal, in some cases this could be implementation in a cohort structure. In order to integrate rigorous impact evaluation with program implementation, it is essential to onboard researchers right from the start of the planning phase of a program. The integration of rigorous impact evaluation with intervention implementation is of course challenging in reality as it requires, among many other things, coordination with project partners and training institutes about training dates and sites in different regions and for different cohorts. However, researchers typically have experience

with the integration of intervention implementation and rigorous impact evaluation of large scale programs and can, therefore, be a valuable resource to enable and fundamentally support this process and provide required customized solutions.

Another important aspect poses the timing of data collections. As indicated in example 2 of this section, sufficient time is required between the intervention completion and the follow-up data collection for treatment effects to unfold. It is important that the relevant reference time is not the beginning of intervention implementation but the completion of the final intervention component. For example, if a beneficiary completes a technical training as part of an E4D program and subsequently waits three months for an internship placement of six months, the intervention is only fully completed 9 months after the training was completed. If data were to be collected 9 to 12 months after the training or shortly after the internship completion, then potential positive treatment effects may not be visible as the participant had no time to search for a job. Treatment effects may even be negative as no intervention-control group members had 9 months to search for a job. As impact evaluation timelines may extend the period of a project phase, it is important to allow the impact evaluation to continue beyond the project phase in order to adequately measure intervention impacts.

Beyond the overall long timeframe required, the timing of follow-up data collection is difficult to determine in demand-driven interventions such the internship placements. If all participants are surveyed around the same time, and participants had different placement waiting times and also different lengths of placements, then the time participants had to search for a job varies considerably across beneficiaries, making it difficult to draw conclusion about whether reduced job search times were a disadvantage for beneficiaries. If the time of follow-up data collection would be held constant at 9 months after the placement completion, then beneficiaries would be surveyed at different times which can be problematic if outcomes such as employment status and incomes, are prone to seasonality. Further, it is unclear what would be a consistent and appropriate time to survey control-group members. The simultaneous intervention delivery to beneficiaries or a cohort delivery structure would alleviate these concerns.

Some research designs also require the measurement of participants' characteristics prior to the implementation of the intervention. In addition, sample sizes can be reduced if data that explain variation in the outcomes measured in follow-up surveys was collected prior to the intervention implementation. In order to have sufficient time to collect pre-intervention data, researchers have to be involved in the project already during the intervention planning phase.

4. Sample size

Reliable results hinge on large sample sizes, which allow to attach a degree of certainty to the estimated intervention impacts. The relevant sample size is the one that is included in the final estimation sample and is, thus, net of intervention attrition and survey non-response. Further, the required sample size depends on how large the effect sizes will potentially be and which and how many outcome indicators are to be considered. If in addition to the overall program impact the impact of specific population groups or program components should be evaluated, then the estimation sample size in the relevant group of beneficiaries is the relevant sample size.

Although, the required estimation sample size depends on the respective setting and factors mentioned, it is rare that sample sizes of less than a few hundred people is

sufficient, typically rather several thousand participants are required. Such sample sizes may preclude the evaluation of smaller projects. For example, in the E4D portfolio many of the interventions targeted at business or firms included less than 50 or even 30 companies. In most cases, these numbers would not allow a rigorous quantitative evaluation and other methods, such as qualitative analyses, would be needed. Alternatively, if at program conception it is known that a quantitative evaluation of a specific firm level intervention is of interest, the program could be set up at larger scale with a sample size that suffices for quantitative evaluation purposes.

These lessons learned are not a comprehensive list of considerations that need to be made when planning and preparing a rigorous impact evaluation but are aspects that seemed particularly relevant in the context of the quantitative evaluations of sections 2 to 6. Many aspects that went well in these evaluations were not mentioned here but could be relevant “risk factors” in other set-ups. All of the above points essentially boil down to the fact that the early onboarding of researchers and the integration of the impact evaluation with the planning and implementation of interventions is vital.

11 References

Adoho, F., Chakravarty, S., Korkoyah, D.T., Lundberg, M., & Tasneem, A. (2014). *The Impact of an Adolescent Girls Employment Program: The EPAG Project in Liberia.* The World Bank.

Alcid, A. (2014). A Randomized Controlled Trial of Akazi Kanoze Youth in Rural Rwanda. *Education Development Center Report.*

Attanasio, O., Kugler, A., & Meghir, C. (2011). Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics*, 3 (3), 188–220.

Asian Development Bank (2010). *Post-Completion Sustainability of Asian Development Bank-Assisted Projects.* Asian Development Bank. https://www.adb.org/documents/post-completion-sustainability-asian-development-bank-assisted-projects

Aurini, J. D., Heath, M., & Howells, S. (2016). How to Design a Qualitative Project and Create a Research Question: Selecting the Right Tools for the Job. In J. D. Aurini, M. Heath, S. Howells (Eds.), *The How To of Qualitative Research.* Sage. (pp. 34-76).

Ballantyne, P. (2003). Ownership and Partnership – Keys to Sustaining ICT-enabled Development Activities, *IICD Research Brief*, 8, 1-8.

Bachmann, R., Kluve, J., Martinez Flores, F., & Stöterau, J. (2019). *Employment impacts of German development cooperation interventions: A collaborative study in three pilot countries*. RWI Projektberichte.

Bacon, R. and Kojima, M. (2011). Issues in estimating the employment generated by energy sector activities. World Bank.

Barajas, A., Beck, T., Belhaj, M., Naceur, S. B., Cerra, V., & Qureshi, M. S. (2020). Financial Inclusion: What Have We Learned So Far? What Do We Have to Learn? IMF, 157.

Barnett, C., Barr, J., Christie, A., Duff, B., & Hext, S. (2010). Measuring the Impact and Value for Money of Governance & Conflict Programmes. Final Report, ITAD.

Betcherman, G., Olivas, K., & Dar, A. (2004). *Impacts of active labor market programs: New evidence from evaluations with particular attention to developing and transition countries*. The World Bank, No. 29142.

Besley, M. (2016). *Value for Money Assessment Oxfam South Africa Australia Africa Community Engagement Scheme (AACES) Capacity Building component*. Oxfam.

Blattman, C. and Ralston, L. (2015). Generating Employment in Poor and Fragile States: Evidence from Labor Market and Entrepreneurship Programs. *SSRN Electronic Journal*.

BMZ – Federal Ministry for Economic Cooperation and Development (2017). Africa and Europe – A new partnership for development, peace and a better future. BMZ.

Bondonio, D. and A. Martini (2012). Counterfactual Impact Evaluation of Cohesion Policy: Impact, cost-effectiveness and additionality, Final Report to DG Regional Policy "Counterfactual Impact Evaluation of Cohesion Policy", Associazione per lo Sviluppo della Valutazione e l'Analisi delle Politiche Pubbliche: Turin.

Bossert, T. J. (1990). Can they get along without us? Sustainability of donor-supported health projects in Central America and Africa. *Social Science and Medicine,* 30 (9), 1015-1023.

Brinkmann, V., Diembeck, K., Feldmann, L., Messinger, C., & Raabe, T. (2014). *Sustainability Assessment of Improved Household Cookstove Dissemination (HERA).* Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ). http://digicollection.org/eebea/documents/gh029e/gh029e.pdf

Brolin, T. (2017). Ownership and Results in Swedish General Budget Support to Mozambique. *Forum for Development Studies,* 44 (3), 377-399.

Bryman, A. (2012). *Social Research Methods* (4th ed). Oxford University Press.

Bryman, A. (2016). *Social Research Methods* (5th ed). Oxford University Press.

Card, D., Kluve, J., & Weber, A. (2017). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Associations*, 16 (3), 894-931.

Chakravarty, S., Lundberg, M. K. A., Nikolov, P. V., & Zenker, J. (2016). *The role of training programs for youth employment in Nepal: Impact evaluation report on the employment fund*. The World Bank, No. WPS7656.

Cekan, J. (2020, April 20). *Sustaining "Sustainable Development"?* Valuing Voices Blog. http://valuingvoices.com/sustaining-sustainable-development/

Cekan, J. (2016a, February 12). *What happens after the project ends? Lessons from post-project sustained impacts evaluations (Part 1).* Valuing Voices Blog. http://valuingvoices.com/what-happens-after-the-project-ends-lessons-from-post-project-sustained-impact-evaluations-part-1/

Cekan, J. (2016b, February 19). *What happens after the project ends? Country-national ownership lessons from post-project sustained impacts evaluations (Part 2).* Valuing Voices Blog. http://valuingvoices.com/what-happens-after-the-project-ends-country-national-ownership-lessons-from-post-project-sustained-impact-evaluations-part-2/

Cekan, J. (2016c, February 29). *What happens after the project ends? Lessons about Funding, Assumptions and Fears (Part 3).* Valuing Voices Blog. http://valuingvoices.com/what-happens-after-the-project-ends-lessons-about-funding-assumptions-and-fears-part-3/

Cekan, J. (2015, March 13). *When Funders Move On.* Stanford Social Innovation Review (SSIR). https://ssir.org/articles/entry/when_funders_move_on

Centre for Strategy & Evaluation Services (2006). Study on Measuring Employment Effects – Final Report. Centre for Strategy & Evaluation Services (CSES): Sevenoaks.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., & Zamora, P. (2013). Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomizes Experiment, *The Quarterly Journal of Economics* (2013), 531–580.

Criscuolo C., Martin, R., Overman, H., & van Reenen, J. (2012). The Causal Effects of an Industrial Policy. NBER Working Papers 17842, National Bureau of Economic Research, Inc.

Datta, N., Assy, A. E., Buba, J., & Watson, S. M., 2018. *Integration: A New Approach to Youth Employment Programs - General Guidelines for Project Teams*. The World Bank, No. 135323.

de Mel, S., McKenzie, D., & Woodruff, C. (2014). Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka. *Journal of Development Economics*, 106, 199–210.

Demirguc-Kunt, A., Klapper, L., & Singer, D. (2017). Financial Inclusion and Inclusive Growth. A Review of Recent Empirical Evidence. Policy Research Working Paper 8040.

DFID (2011). *DFID's Approach to Value for Money (VfM)*. Department for International Development, United Kingdom.

Dhaliwal et al. (2013). *Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education*. Education Policy in Developing Countries, edited by Paul Glewwe, University of Chicago Press.

Dupas, P., Karlan, D., Robinson, J., & Ubfal, D. (2018). Banking the Unbanked? Evidence from Three Countries. *American Economic Journal: Applied Economics*, 10 (2), 257-297.

Dunbar, M. (2013). *Engaging the private sector in skills development*. Oxford Policy Management, 1-48.

Edgren, G. (2003). Donorship, Ownership and Partnership. Issues Arising from Four Sida Studies of Donor-Recipient Relations. Sida Studies in Evaluation (03/03). https://publikationer.sida.se/contentassets/9eb50ac50e7d4494bdccd4ef52d09fa7/13364.pdf

English Partnerships (2008). Additionality Guide: A standard approach to assessing the additional impact of interventions, third edition, English Partnerships: London.

Escudero, V., Kluve, J., López Mourelo, E., & Pignatti, C. (2019). Active Labour Market Programmes in Latin America and the Caribbean: Evidence from a Meta-Analysis. *The Journal of Development Studies*, 55 (12), 2644–2661.

European Commission (2007). Measuring structural funds employment effects, Working Document Nr. 6, January 2007.

European Union (2013). Expert evaluation network deliver- ing policy analysis on the performance of Cohesion policy 2007-2013 – Job creation as an indicator of outcomes in ERDF programmes – Synthesis report, Report to the European Commission Directorate-General for Regional and Urban Policy, European Union: Brüssel.

Fereday, J. and Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods,* 5 (1), 80-92.

Fleming, F. (2013). *Evaluation methods for assessing Value for Money*. BetterEvaluation, Australian Evaluation Society.

GIZ (2018). *Operational Plan of project promoting youth employment through technical human capital development.* Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ). Unpublished Document, 1-6.

GIZ (2017). *Promoting youth employment through technical human capital development – Kenya, Factsheet.* Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ). Unpublished Document, 1-2.

Greenberg, D. H., Knight, G., Speckesser, S., & Hevenstone, D. (2011). Improving DWP Assessment of the Relative Costs and Benefits of Employment Programmes, U.K. Department for Work and Pensions Working Paper No. 100.

Grohman, A., Klühs, T., & Menkhoff, L. (2018). Does financial literacy improve financial inclusion? Cross country evidence. *World Development*, 111, 84-96.

Hamukwala, P., Muuka, G. N., Wheelock, G., & Bukenya, J. (2008). NGO Project Managers' Perception of the Sustainability of Development Projects: Empirical Evidence from Sub-Saharan Africa. *Journal of African Business,* 9 (1), 191-217.

Hasselkong, M. and Schierenbeck, I. (2017). The Ownership Paradox: Continuity and Change. *Forum for Development Studies,* 44 (3), 323-333.

Hofisi, C. and Chizimba, M. (2013). The Sustainability of Donor-funded Projects in Malawi. *Mediterranean Journal of Social Sciences,* 4 (6), 705-714.

Ibarrarán, P., Kluve, J., Ripani, L., & Rosas Shady, D. (2019). Experimental Evidence on the Long-Term Effects of a Youth Training Program. *ILR Review*, 72 (1), 185–222.

Independent Commission for Aid Impact (ICAI) (2011). *ICAI's Approach to Effectiveness and Value for Money (1).*

International Finance Corporation, IFC (2013). IFC Jobs Study. Assessing private sector contributions to job creation and poverty reduction, January 2013.

International Labour Organization, ILO (2009). ILO School-to-Work Transition Survey: A methodological guide. *Youth Employment Programme, International Labour Office.*

International Labour Organization, ILO (2013). *Measuring informality: A statistical manual on the informal sector and informal employment*. Geneva: International Labour Office.

International Labour Organization, ILO (2020a). Reference guide for Employment Impact Assessments (EmpIA), Employment Policy Department, Geneva: ILO 2020.

International Labour Organization, ILO (2020b). Guide for Monitoring Employment and Conducting Employment Impact Assessments (EmpIA) of Infrastructure Investments, Employment Intensive Investment Programme, Development and Investment Branch, Employment Policy Department, Version 1.0.

IPA (2020). Comprehensive Impact Evaluation and Monitoring. Progress Report No. 5.

Jackson, P. (2012). *Value for money and international development: Deconstructing myths to promote a more constructive discussion.* OECD Development Co-operation Diretorate. https://www.oecd.org/dac/effectiveness/49652541.pdf

J-PAL (2017). *J-PAL Skills for Youth Program Review Paper*. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.

J-PAL (n.d.). J-PAL Costing Guidelines. https://www.povertyactionlab.org/sites/default/files/research-resources/costing-guidelines.pdf

Kagendo, R. (2015, November 19). *Whose responsibility is it to sustain project activities?* Valuing Voices Blog. http://valuingvoices.com/who-is-responsible-for-sustaining-development/

KAM (Kenyan Association of Manufacturers) (2020a). *KAM Membership.* https://kam.co.ke/membership/

KAM (Kenyan Association of Manufacturers) (2020b). *About KAM.* https://kam.co.ke/about-kam/

KAM (2017). *E4D/SOGA- Employment and Skills for Eastern Africa Annual Report (Jan - Dec 2017).* Unpublished Document, 1-11.

Kaplan, M. (2013, August 15). *Local ownership: Who, what and how.* https://www.devex.com/news/local-ownership-who-what-and-how-81623

Karanja, G. M. (2014). Influence of management practices on sustainability of youth income generating projects in Kangema District, Murang'a County, Kenya. *International Journal of Education and Research,* 2 (2), 1-12.

King, J. (2018). *OPM's approach to assessing Value for Money*. Oxford Policy Management.

Kinoti, E. K. (2011). *Factors influencing ownership of development projects by the communities: A case of Kiambaa Constituency, Kenya.* [Dissertation, University of Nairobi]. http://erepository.uonbi.ac.ke:8080/handle/123456789/4583

Kluve, J. (2011). Measuring employment effects of technical cooperation interventions – some methodological guidelines, 2nd revised edition, Eschborn: GIZ/BMZ.

Kluve, J., Puerto, S., Robalino, D., Romero, J.M., Rother, F., Stöterau, J., Weidenkaff, F., & Witte, M. (2017). Interventions to improve the labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions. *Campbell Systematic Reviews*, 13 (1), 1–288.

Kluve, J., Puerto, S., Robalio, D., Romero, J.M., Rother, F., Stöterau, J., Weidenkaff, F., & Witte, M. (2019). Do youth employment programs improve labor market outcomes? A quantitative review. *World Development*, 114, 237-253.

Kluve, J. and J. Stöterau (2014). A systematic framework for measuring employment impacts of development cooperation interventions, Berlin: GIZ / Federal Ministry for Economic Cooperation and Development.

Kuckartz, U. and Rädiker, S. (2019). *Analyzing Qualitative Data with MAXQDA. Text, Audio, and Video.* Springer.

Kuria, E. and Wanyoike, D. M. (2016). Assessment of factors influencing sustainability of donor-funded projects in Nakuru county, Kenya. *International Journal of Economics, Commerce and Management,* VI (10), 472-501.

Kwangware, J., Mayo, A., & Hoko, Z. (2014). Sustainability of donor-funded rural water supply and sanitation projects in Mbire district, Zimbabwe. *Physics and Chemistry of the Earth,* 76–78, 134–139.

Lopes, C. and Theisohn, T. (2003). *Ownership, leadership, and transformation: Can we do better for capacity development?* Earthscan.

MacGillivray, A., R. Kim, T.v. Moorsel, & Kehoe, A. (2017). Measuring Total Employment Effects: a lean data methodology for a portfolio of investments in developing countries. CDC Group research paper.

Maier, S. D., Beck, T., Vallejo, J. F., Horn, R., Söhlemann, J., & Nguyen, T. T. (2016). Methodological Approach for the Sustainability Assessment of Development Cooperation Projects for Built Innovations Based on the SDGs and Life Cycle Thinking. *Sustainability,* 8 (1006), 1-26.

Maré, D.C. (2005). Indirect Effects of Active Labour Market Policies, Motu Working Paper 05–01. Motu Economic and Public Policy Research Trust: Wellington.

Matuella, J. (2010, December 09). *Well intended, yet harmful: Unless local people assume ownership, development projects are meaningless.* D + C Development and Cooperation. https://www.dandc.eu/en/article/unless-local-people-assume-ownership-development-projects-are-meaningless

McEwan (2012). Cost-effectiveness analysis of education and health interventions in developing countries. *Journal of Development Effectiveness*, 4 (2), 189-213. https://doi.org/10.1080/19439342.2011.649044

MercyCrops (2013). *Elevating the Community's Agenda: Honing USAID Forward's Approach to Local Ownership.* Mercy Crops, 1-20.

Miguel, Edward and Michael Kremer (2004). Worms: Identifying Impacts on Education and Health in the Presence of Externalities. *Econometrica*.

Mouqué, D. (2012). What are counterfactual impact evaluations teaching us about enterprise and innovation support? Regional Focus Paper 2/2012, European Commission, Regional and Urban Policy: Brussels.

Morán, M. D. and Ferguson, M. (2013). Sustainability: the dream of international development projects. *Sustainable Development and Planning,* VI (173), 431-442.

National Audit Office (2003). Regional Grants In England, The Department For Trade And Industry: Report By The Comptroller And Auditor General, Hc 702 Session 2002- 2003, NAO: London.

National Audit Office. (n.d.). What is a value for money study?. https://www.nao.org.uk/about-us/wp-content/uploads/sites/12/2016/10/What-is-a-value-for-money-study.pdf

Nyanjura, C. (2010). *Is self help the best help? Comparing donor-funded and locally funded (self help) development projects in view of sustainable development. A comparative study from Uganda.* [Dissertation, University of Agder] https://uia.brage.unit.no/uia-xmlui/handle/11250/135153

Ochunga, F. O. and Awiti, L. H. (2017). Influence of Stakeholder Participation on Sustainability of Community Development Projects Implemented by Plan International in Homa Bay Town Sub-County, Kenya. *International Journal of Academic Research in Business and Social Sciences,* 7 (4), 375-400.

OECD (2019). *Better Criteria for Better Evaluation. Revised Evaluation Criteria Definitions and Principles for Use.* OECD/DAC Network on Development Evaluation. OECD Publications, 1-12. http://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf

OECD (2002). *Glossary of Key Terms in Evaluation and Results Based Management.* OECD Publications, 1-37. https://www.oecd.org/dac/evaluation/2754804.pdf

Oino, P. G., Towett, G., Kirui, K. K., & Luvega, C. (2015). The Dilemma in Sustainability of Community-based Projects in Kenya. *Global Journal of Advanced Research,* 2 (4), 757-768.

Okun, M. A. (2009). *Factors Affecting Sustainability of Donor-funded Projects in Arid and Semi-arid Areas in Kenya: A Case of Marsabit Central District.* [Dissertation, Kenyatta University] https://ir-library.ku.ac.ke/handle/123456789/3494

Olsson, N., Johansen, A., Langlo, J. A., & Torp, O. (2008). Project ownership: implications on success measurement. *Measuring Business Excellence,* 12 (1), 39–46.

Ostrom, E., Gibson, C., Shivakumar, S., & Andersson, K. (2002). *Aid, Incentives, and Sustainability, An Institutional Analysis of Development Cooperation.* Sida Studies in Evaluation (02/01). https://publikationer.sida.se/contentassets/34b423c6b93e41798038c901e30cd4d1/12924.pdf

Rey-Moreno, C., Sabiescu, A. G., & Siya, M. J. (2014). Towards self-sustaining community networks in rural areas of developing countries: Understanding local ownership. In J. Steyn, D. Van Greunen, (Eds.), *ICTs for inclusive communities in developing societies. Proceedings of the 8th International Development Informatics Association Conference* (pp. 63-77).

RWI (2013). *Nachweis der Beschäftigungswirkungen von Maßnahmen der deutschen EZ, Pilotstudie Marokko*, Endbericht, 31. Dezember 2013.

RWI (2014). *Integrierte Evaluationsansätze zur Messung von Beschäftigungseffekten – Anschlussstudie Marokko*. Studie im Auftrag von BMZ/GIZ, Berlin/Essen, 2014.

RWI (2019). *Employment impacts of German development cooperation interventions – A collaborative study in three pilot countries*. Project report commissioned by Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Berlin/Essen, 2019.

Sustainable Development Goals Fund (2020). *How we work. Ensuring national ownership.* https://www.sdgfund.org/ensuring-national-ownership

Seppey, M., Ridde, V., Touré, L., & Coulibaly, A. (2017). Donor-funded project's sustainability assessment: a qualitative case study of a results-based financing pilot in Koulikoro region, Mali. *Globalization and Health,* 13 (86), 1-15.

Steiger, J. (2015, August 24). *Three Approaches to Increase Local Ownership of Development Innovations, ICT works.* https://www.ictworks.org/3-approaches-to-increase-local-ownership-of-development-innovations/

Uganda Bureau of Statistics (2018). Uganda National Household Survey 2016/2017. Kampala, Uganda; Uganda Bureau of Statistics.

USAID (2016). Job Creation Assessment, USAID/Jordan Monitoring and Evaluation Support Project (MESP), September 2016.

Weeks, J., Andersson, D., Cramer, C., Geda, A., Hailu, D., Frank, M., Rizz, M., Ronge, E., & Stein, H. (2002). *Supporting Ownership: Swedish Development Cooperation with Kenya, Tanzania and Uganda. Volume II: Country Studies.* Sida Evaluation (02/03). https://publikationer.sida.se/contentassets/6851b9f4d3f945dca429c27109b40738/13115.pdf

Vivarelli, M. (2014). Innovation, Employment and Skills in Advanced and Developing Countries: A Survey of Economic Literature. *Journal of Economic Issues*, 48 (1), 123–154.

World Bank (2012). *World Development Report 2013: Jobs*. The World Bank.

World Bank Institute and Public-Private Infrastructure Advisory Facility (2013). *Value for Money analysis - practices and challenges: How governments choose when to use PPP to deliver public infrastructure and services*. Report from World Bank Global Round-Table. World Bank, Washington, DC.

A Appendix

A1 Appendix of the quantitative KAM evaluation

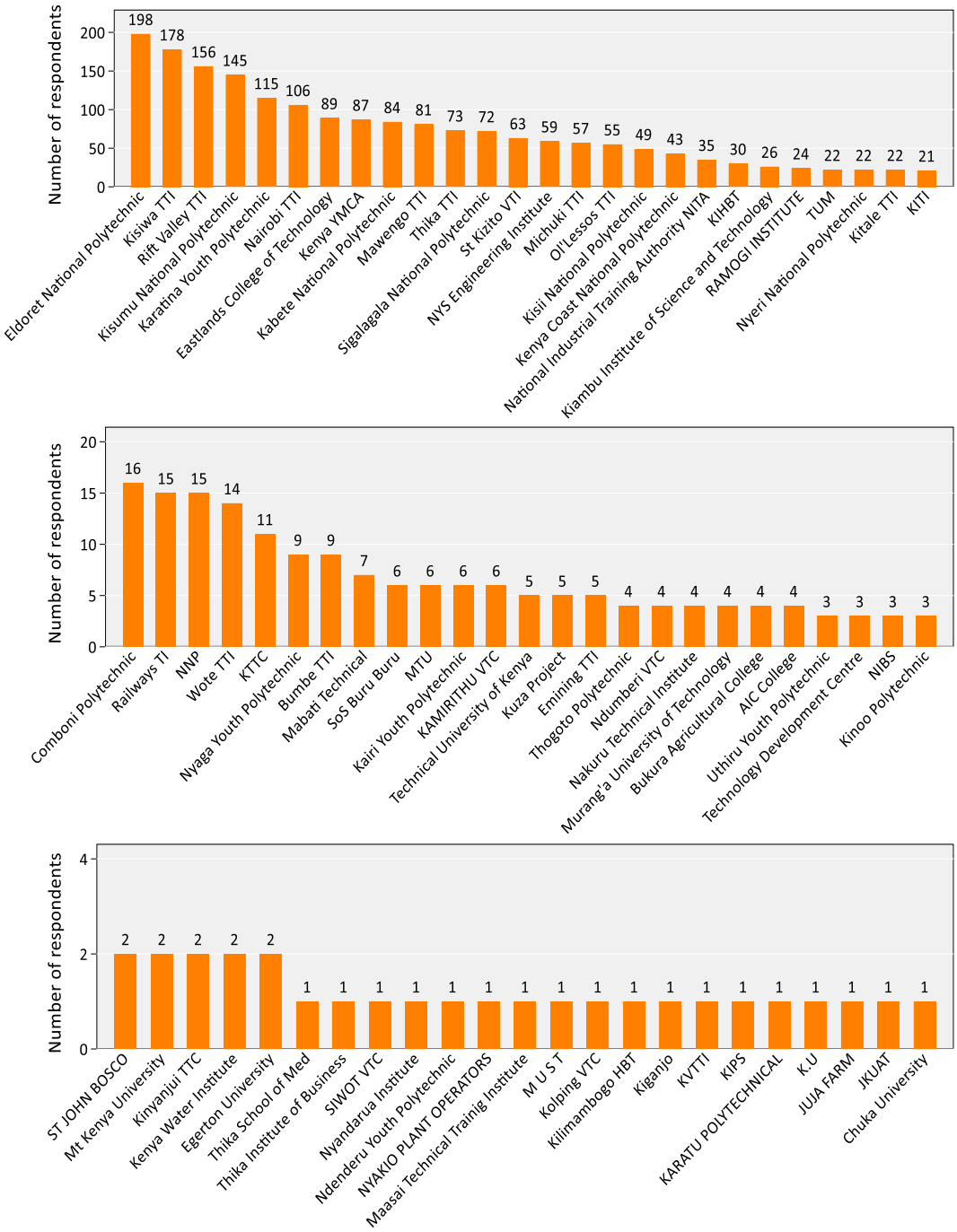
A1.1 Additional descriptive statistics

Table A1.1.1
Assumed thresholds of weekly working hours to define the working days

# of working days	# of working hours
1	1 – 10
2	11 – 20
3	21 – 29
4	30 – 36
5	37 – 47
6	48 – 55
7	> 56

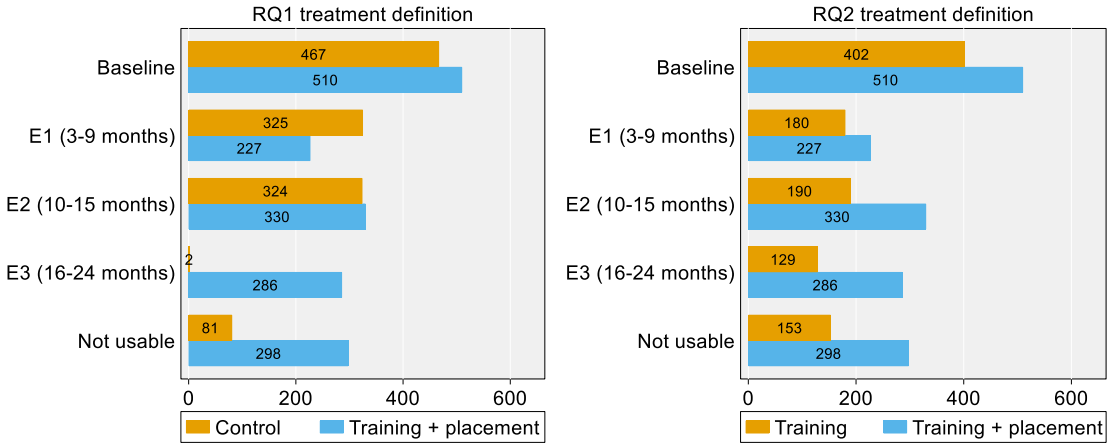
Notes: At 30 hours per week the number of jobs indicated as fulltime increases sharply which is more plausible if they work 4 days per week. - Source: Own calculations based on KAM survey.

Figure A1.1.1
Number of study participants in baseline, by TVET



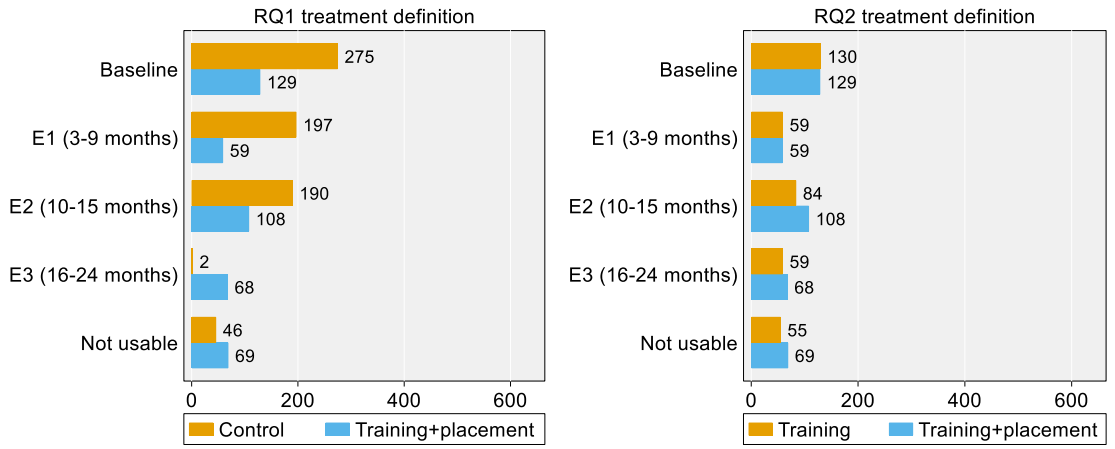
Notes: The graphic displays the number of baseline study participants separated by TVET from which the respondents graduated from. - Source: Own calculations based on KAM survey.

Figure A1.1.2
Number of observations by wave and study group, all TVETs but no CBET



Notes: The graphics show the number of observations by treatment and control group for the baseline survey and each modified follow-up wave when excluding CBET control group participants. The left graphic displays the number of observations by treatment and control group according to the RQ#1 definition. The right graphic displays the number of observations by treatment and control group according to the RQ#2 definition. "Not usable" refers to interviews conducted less than 3 months after the baseline or individuals who were interviewed twice within the same caliper. - Source: Own calculations based on KAM survey.

Figure A1.1.3
Number of observations by wave and study group, only mixed TVETs and no CBET



Notes: The graphics show the number of observations by treatment and control group for the baseline survey and each modified follow-up wave when reducing the sample to mixed TVETs and excluding CBET control group participants. The left graphic displays the number of observations by treatment and control group according to the RQ#1 definition. The right graphic displays the number of observations by treatment and control group according to the RQ#2 definition. "Not usable" refers to interviews conducted less than 3 months after the baseline or individuals who were interviewed twice within the same caliper. - Source: Own calculations based on KAM survey.

Table A1.1.2
Sociodemographic characteristics of study participants, mixed TVETs

	No treat-ment	Training only	Training and place-ment	Std. Diff.	Std. Diff.
	(1) % or Mean	(2) % or Mean	(3) % or Mean	(4) (1)-(3)	(5) (2)-(3)
Respondent characteristics					
Gender	17.0	13.2	18.6	-0.016	-0.054
Age at interview	23.9	26.2	25.8	-1.950***	0.39
Family situation					
Single/never married	89.8	77.7	79.8	0.099***	-0.022
Has children, yes/no	14.9	23.1	24.0	-0.092**	-0.01
Children, number	0.2	0.3	0.4	-0.163***	-0.01
Children aged 6-14 are all enrolled in school	99.0	99.2	97.7	0.013	0.016
Children aged 6-18 are all enrolled in school	97.8	98.5	95.3	0.024	0.031
Other dependants (excl. Children), yes/no	6.9	9.2	7.8	-0.008	0.015
Other dependants, number	0.2	0.1	0.2	-0.05	-0.063
Dependency on household head, yes/no	73.2	47.7	55.8	0.174***	-0.081
Education of female household head					
None or pre school	6.9	5.4	3.9	0.03	0.015
Primary standards 1 to 6	5.0	3.1	8.5	-0.035	-0.055*
Primary standard 7	7.3	5.4	8.5	-0.013	-0.031
Primary standard 8 or secondary forms 1 to 3	21.6	13.8	10.1	0.115***	0.038
Secondary form 4 or higher	35.6	39.2	41.1	-0.054	-0.019
No female household head or refusal	23.5	33.1	27.9	-0.044	0.052
Living situation					
Live alone	30.3	35.4	41.1	-0.108**	-0.057
Live with family	59.5	58.5	51.9	0.076	0.065
Live with room mates	9.9	6.2	7.0	0.029	-0.008
Live with non-relatives	0.3	0.0	0.0	0.003	N/A
Number of observations	578	129	129		
Test of joint orthogonality					
F-test statistics				3.146***	1.392
p-value				0.000***	0.141

*Notes: Information displayed as indicated by study participants during the baseline survey. Observations include only participants from mixed TVETs. The value displayed for t-tests are the differences in the means across the groups (***, **, and * indicate significance at the 1, 5, and 10 percent critical level). - Source: Own calculations based on KAM survey.*

Table A1.1.3

Pre-intervention employment characteristics of study participants, mixed TVETs

	No treat-ment	Training only	Training and placement	Std. Diff.	Std. Diff.
	(1)	(2)	(3)	(4)	(5)
	% or Mean	% or Mean	% or Mean	(1)-(3)	(2)-(3)
Employment					
Employment	0.68	0.85	0.79	-0.109**	0.06
Decent employment	0.12	0.20	0.11	0.01	0.091**
Self-employment	0.18	0.35	0.16	0.02	0.191***
Formal employment	0.07	0.22	0.21	-0.142***	0.01
Fulltime employment	0.52	0.75	0.75	-0.228***	0.00
Total hours worked per week	32.87	42.28	45.26	-12.393***	-2.99
Income					
Total monthly income	10,167.86	12,156.65	9,253.01	914.85	2,903.64
Average hourly earnings	67.85	103.71	152.04	-84.194**	-48.33
N	578	130	129		
Test of joint orthogonality					
F-test statistics				7.264***	3.226***
p-value				0.000***	0.002***

*Notes: Information displayed as indicated by study participants during the baseline survey. Observations include only participants from mixed TVETs. The indicated number of observations refers to the first row of the table. The value displayed for t-tests are the differences in the means across the groups (***, **, and * indicate significance at the 1, 5, and 10 percent critical level). - Source: Own calculations based on KAM survey.*

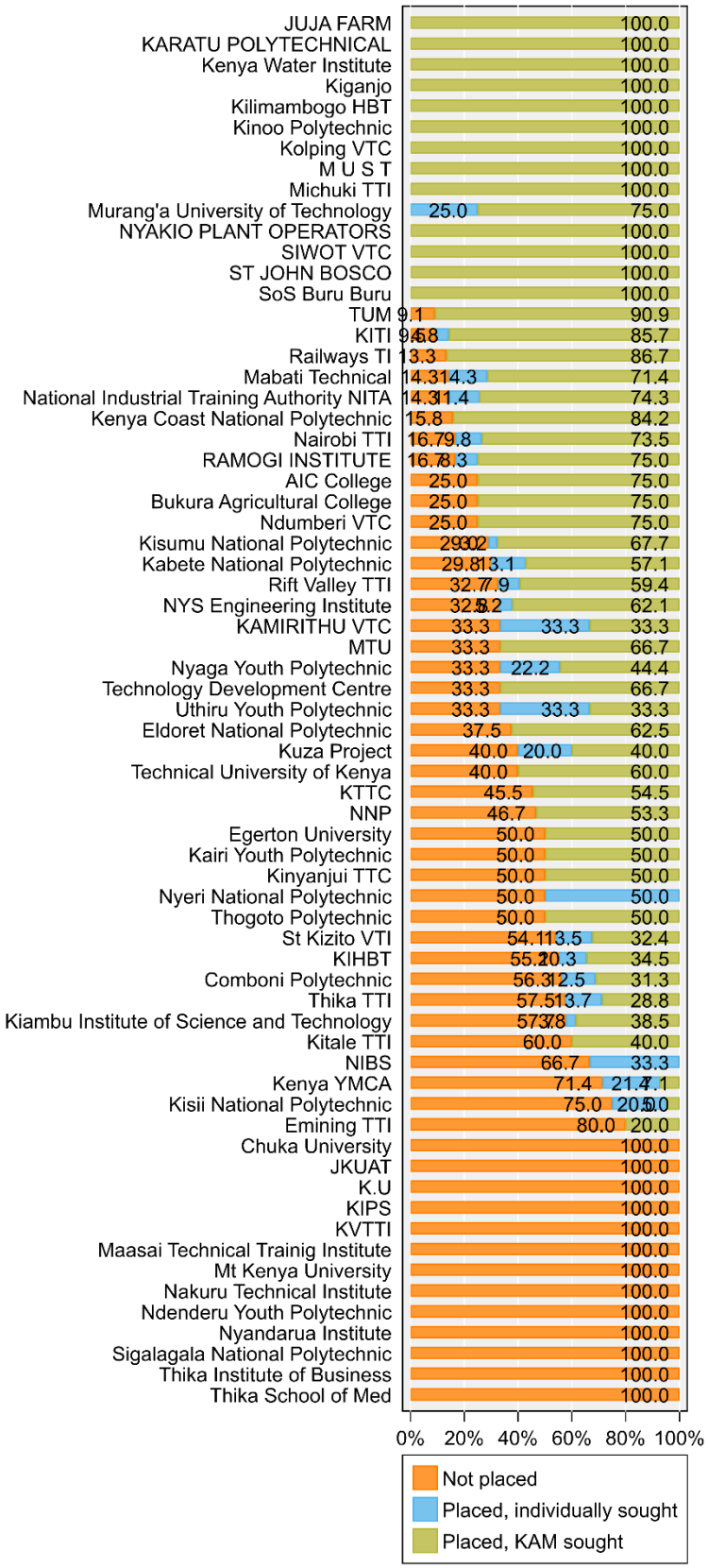
Table A1.1.4

Pre-intervention employment characteristics of employed study participants, mixed TVETs

	No treat-ment	Training only	Training and placement	Std. Diff.	Std. Diff.
	(1)	(2)	(3)	(4)	(5)
	Mean	Mean	Mean	(1)-(3)	(2)-(3)
Employment					
No. of jobs	1.239	1.300	1.333	-0.095*	-0.033
Total hours worked per week	48.221	49.964	57.245	-9.024***	-7.281*
Income					
Total monthly income	14,916.296	14,366.945	11,702.333	3,213.963	2,664.612
Average hourly earnings	99.773	122.569	192.285	-92.512	-69.716
N	394	110	102		
Test of joint orthogonality					
F-test statistics				3.545***	2.321*
p-value				0.007***	0.058*

*Notes: Information displayed as indicated by study participants during the baseline survey. Observations include only participants from mixed TVETs who had a job at the time of the baseline. The indicated number of observations refers to the first row of the table. The value displayed for t-tests are the differences in the means across the groups (***, **, and * indicate significance at the 1, 5, and 10 percent critical level). - Source: Own calculations based on KAM survey.*

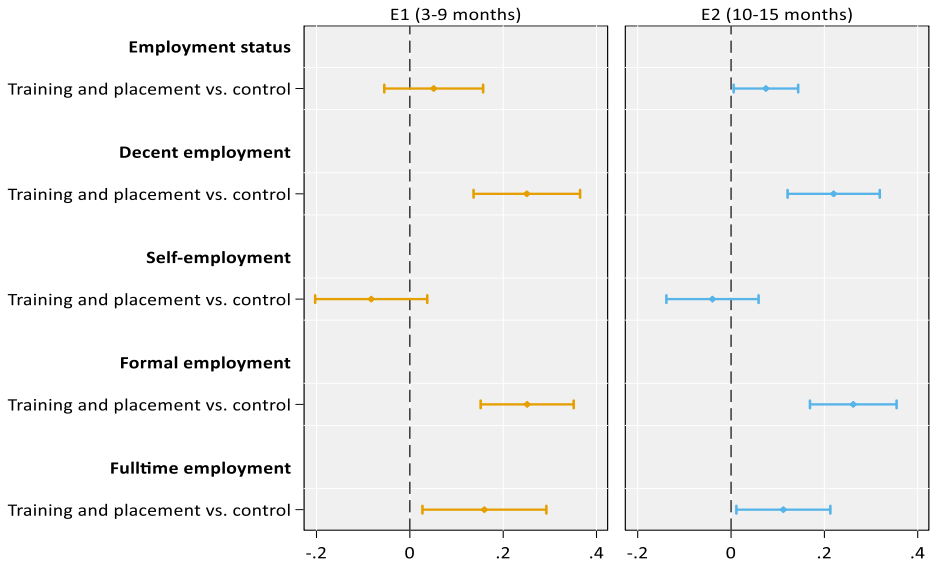
Figure A1.1.4
Placement status by TVET



Source: Own calculations based on KAM survey.

A1.2 Results from alternative sample compositions

Figure A1.2.1
Estimated effects on employment indicators, no CBET participants

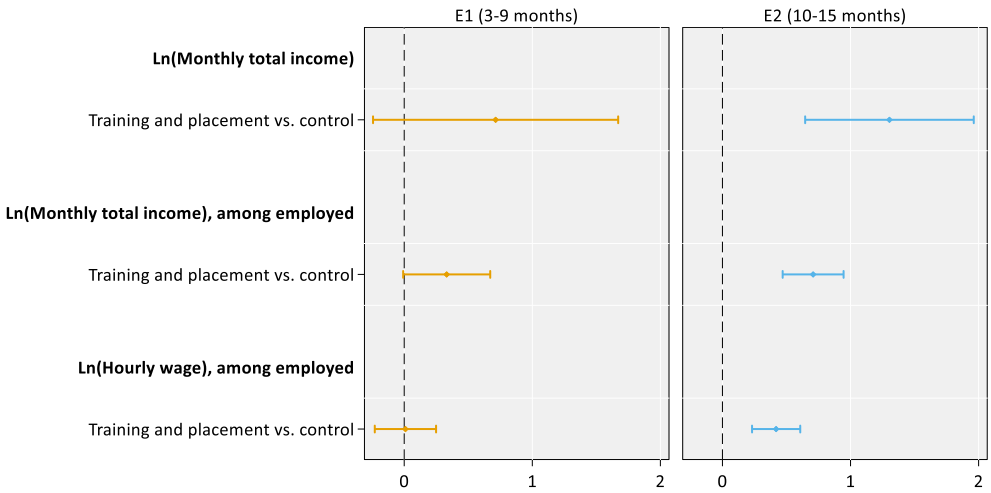


Notes: Regressions are based on observations from study participants of all TVETs and apply course fixed effects. Control group participants who indicated to have been enrolled in a CBET program were excluded. Regressions refer only to E1 and E2 because when excluding CBET participants only two control group observations from participants who did not receive any KAM benefits remain for the period 16 to 24 months after the baseline (E3). - Source: Own calculations based on KAM survey.

Figure A1.2.2
Estimated effects on employment indicators, no CBET participants and mixed TVETs only

Notes: Regressions are based on observations from study participants of mixed TVETs only and apply course fixed effects. Control group participants who indicated to have been enrolled in a CBET program were excluded. Regressions refer only to E1 and E2 because when excluding CBET participants only two control group observations from participants who did not receive any KAM benefits remain for the period 16 to 24 months after the baseline (E3). - Source: Own calculations based on KAM survey.

Figure A1.2.3
Estimated effects on income and wages, no CBET participants

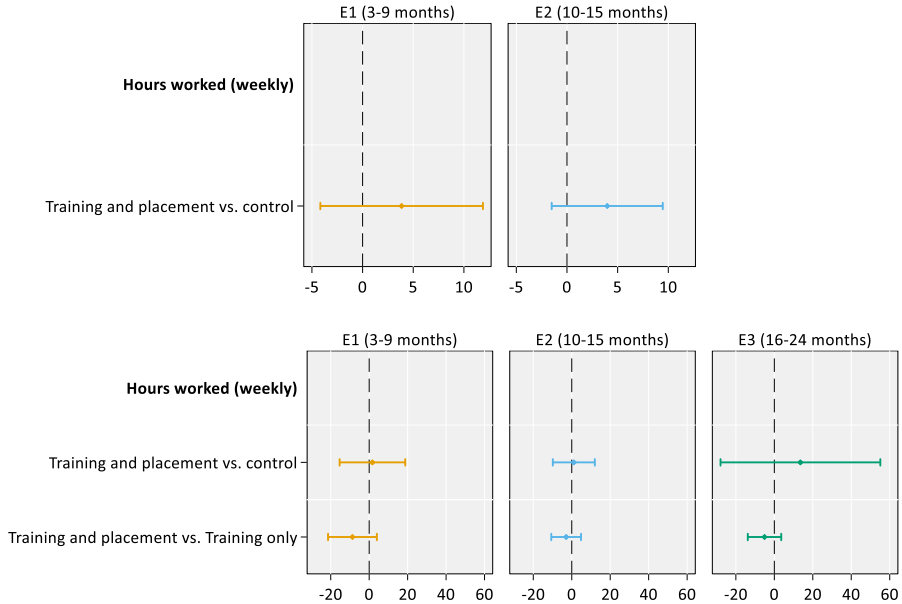


Notes: Regressions are based on observations from study participants of all TVETs and apply course fixed effects. Control group participants who indicated to have been enrolled in a CBET program were excluded. Regressions refer only to E1 and E2 because when excluding CBET participants only two control group observations from participants who did not receive any KAM benefits remain for the period 16 to 24 months after the baseline (E3). - Source: Own calculations based on KAM survey.

Figure A1.2.4
Estimated effects on income and wages, no CBET participants and mixed TVETs only

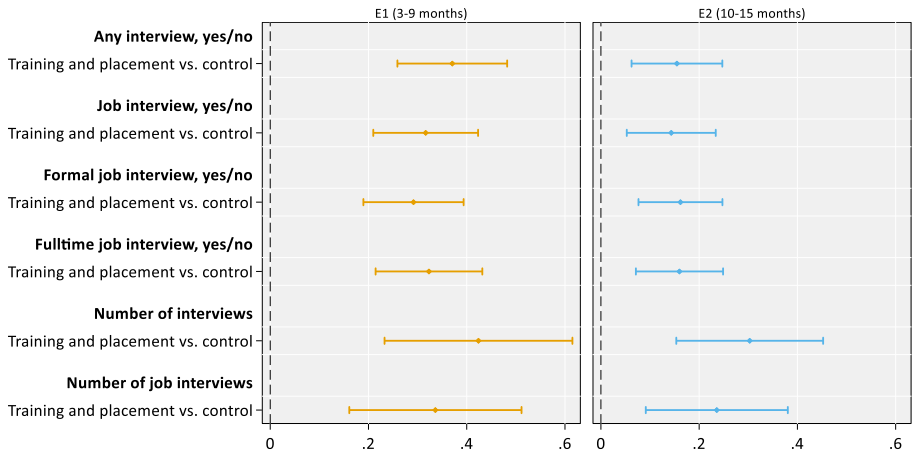
Notes: Regressions are based on observations from study participants of mixed TVETs only and apply course fixed effects. Control group participants who indicated to have been enrolled in a CBET program were excluded. For RQ#1, regressions refer only to E1 and E2 because when excluding CBET participants only two control group observations from participants who did not receive any KAM benefits remain for the period 16 to 24 months after the baseline (E3). - Source: Own calculations based on KAM survey.

Figure A1.2.5
Estimated effects on working hours among employed



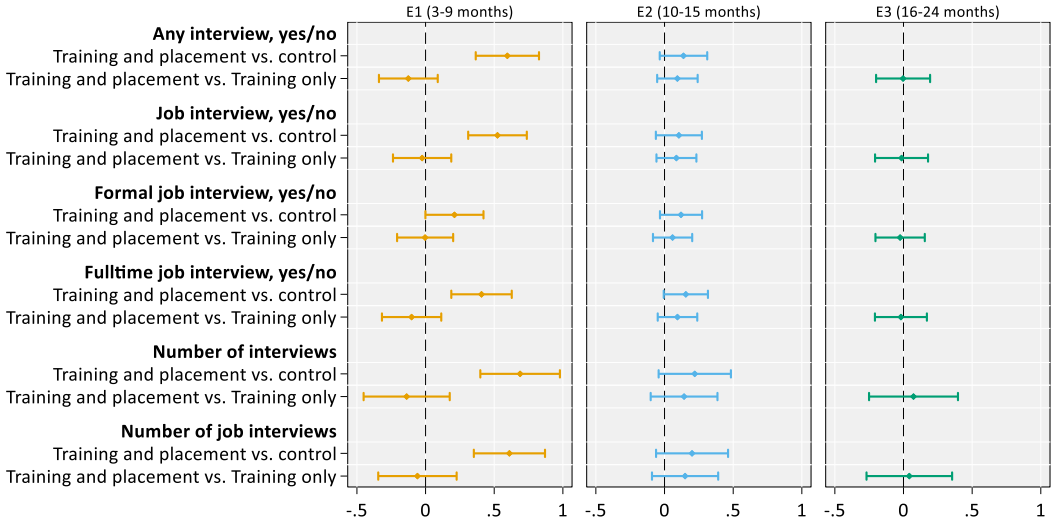
Notes: For the upper graphs, regressions are based on observations from study participants of all TVETs and apply course fixed effects. Control group participants who indicated to have been enrolled in a CBET program were excluded. The results of the lower graphs refer to regressions based on observations from study participants of mixed TVETs only and again excluding control group participants who indicated to have been enrolled in a CBET program were excluded. For RQ#1, regressions refer only to E1 and E2 because when excluding CBET participants only two control group observations from participants who did not receive any KAM benefits remain for the period 16 to 24 months after the baseline (E3). The lower graphs are Regressions. - Source: Own calculations based on KAM survey.

Figure A1.2.6
Estimated effects on job search, no CBET participants



Notes: Regressions are based on observations from study participants of all TVETs and apply course fixed effects. Control group participants who indicated to have been enrolled in a CBET program were excluded. Regressions refer only to E1 and E2 because when excluding CBET participants only two control group observations from participants who did not receive any KAM benefits remain for the period 16 to 24 months after the baseline (E3). - Source: Own calculations based on KAM survey.

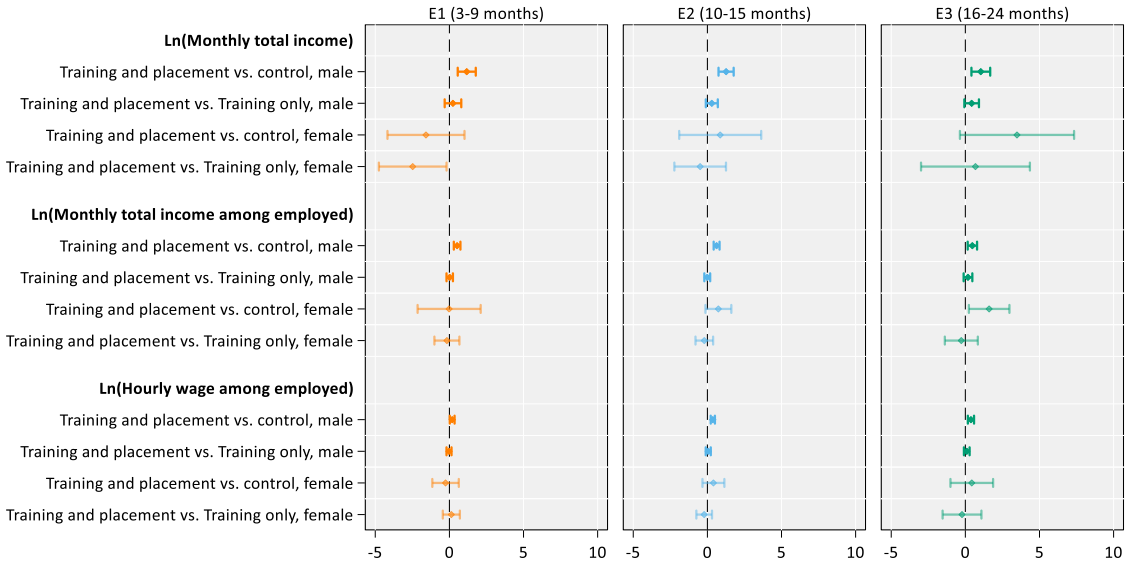
Figure A1.2.7
Estimated effects on job search, no CBET participants and mixed TVETs only



Notes: Regressions are based on observations from study participants of mixed TVETs only and apply course fixed effects. Control group participants who indicated to have been enrolled in a CBET program were excluded. For RQ#1, regressions refer only to E1 and E2 because when excluding CBET participants only two control group observations from participants who did not receive any KAM benefits remain for the period 16 to 24 months after the baseline (E3). Source: Own calculations based on KAM survey.

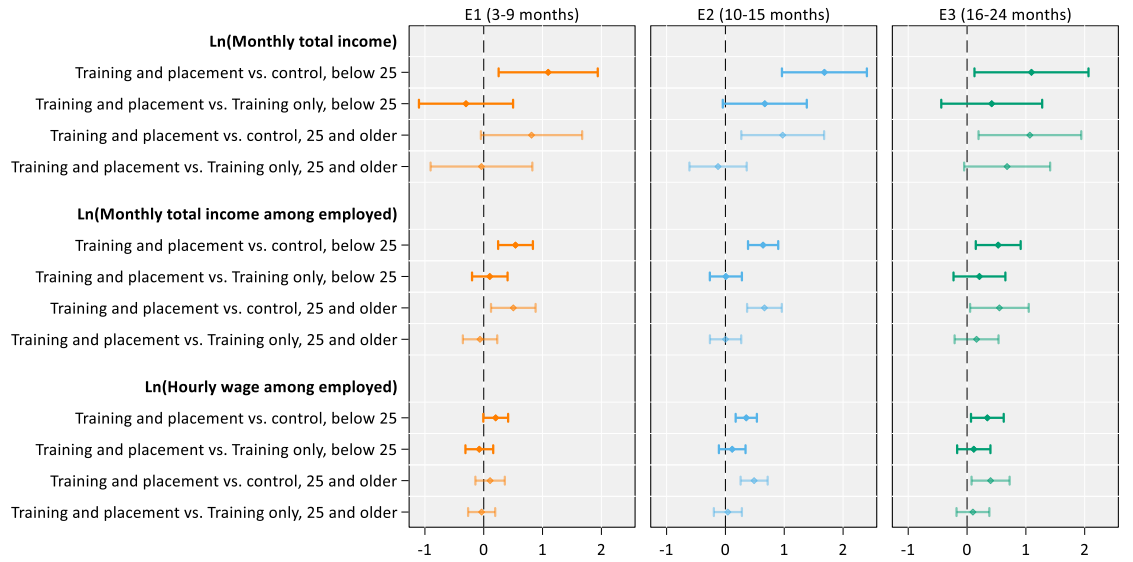
A1.3 Subsample estimations on income, wages, working hours, and job search

Figure A1.3.1
Estimated treatment effects on income, by gender



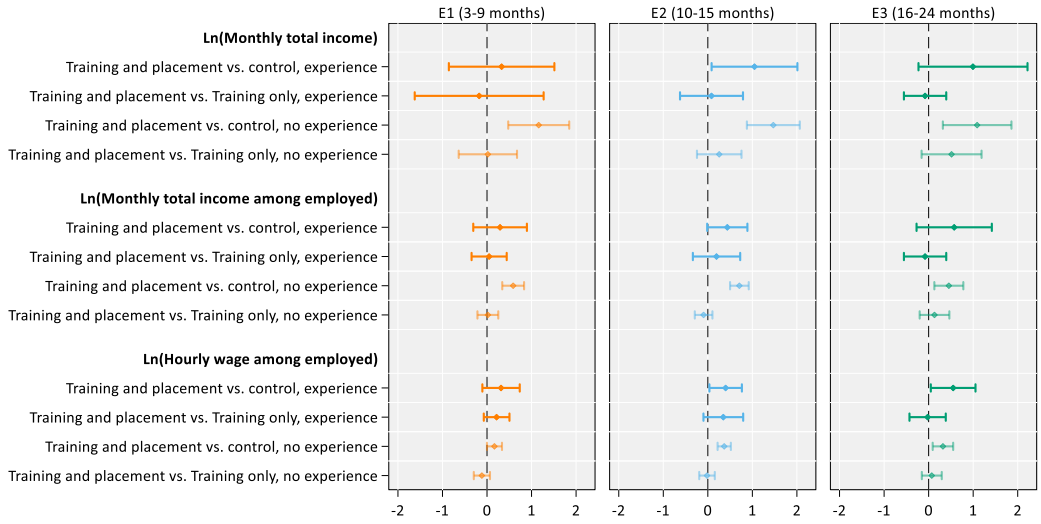
Source: Own calculations based on KAM survey.

Figure A1.3.2
Estimated treatment effects on income, by age group



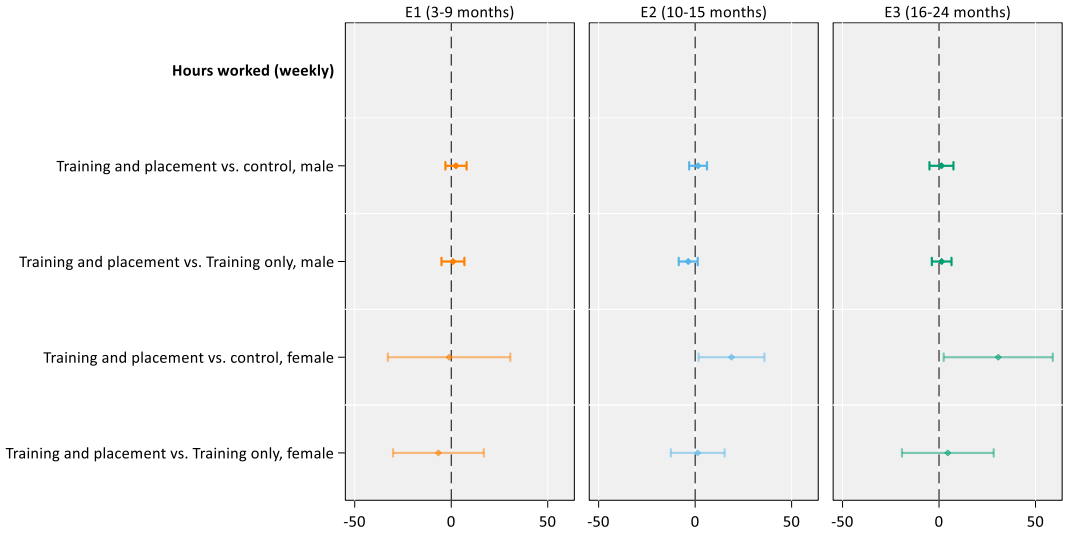
Source: Own calculations based on KAM survey.

Figure A1.3.3
Estimated treatment effects on income, by experience



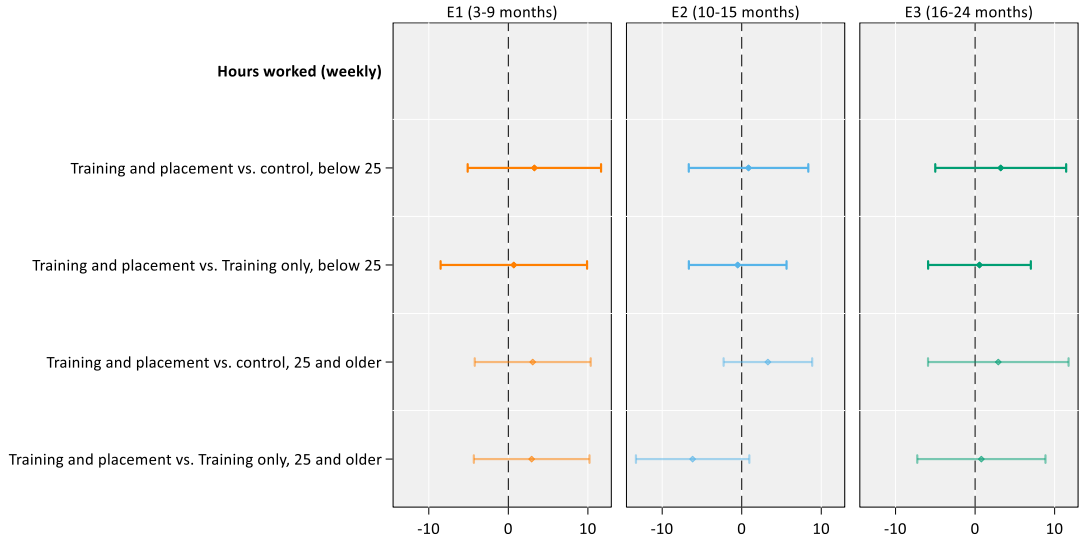
Source: Own calculations based on KAM survey.

Figure A1.3.4
Estimated treatment effects on working hours among employed, by gender



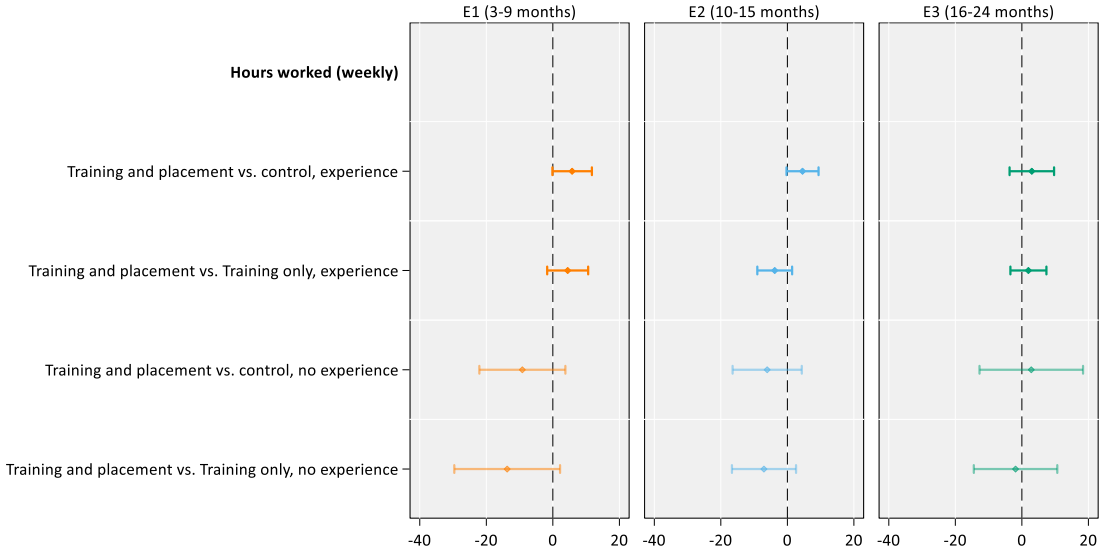
Source: Own calculations based on KAM survey.

Figure A1.3.5
Estimated treatment effects on working hours among employed, by age group



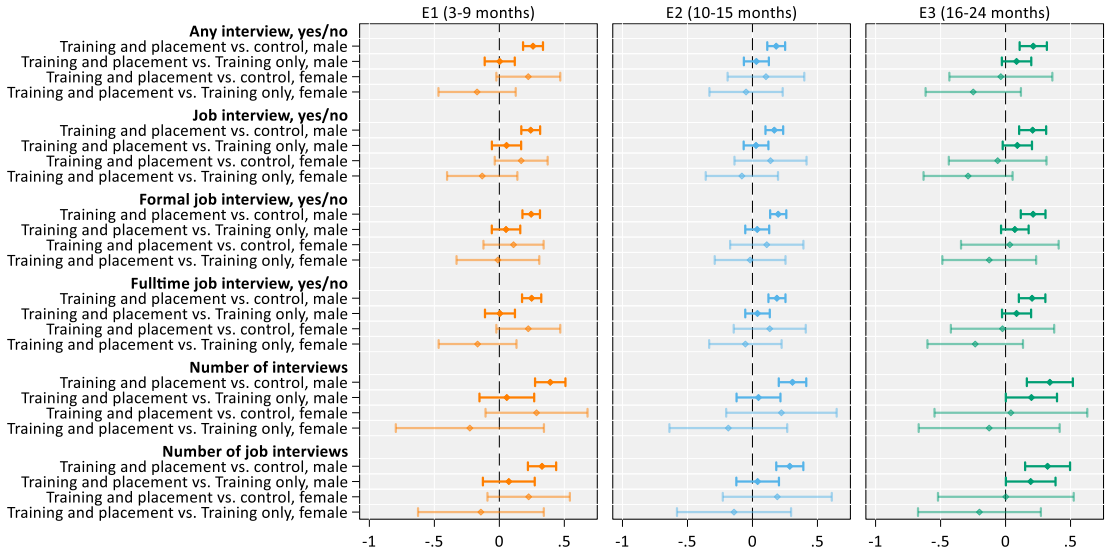
Source: Own calculations based on KAM survey.

Figure A1.3.6
Estimated treatment effects on working hours among employed, by experience



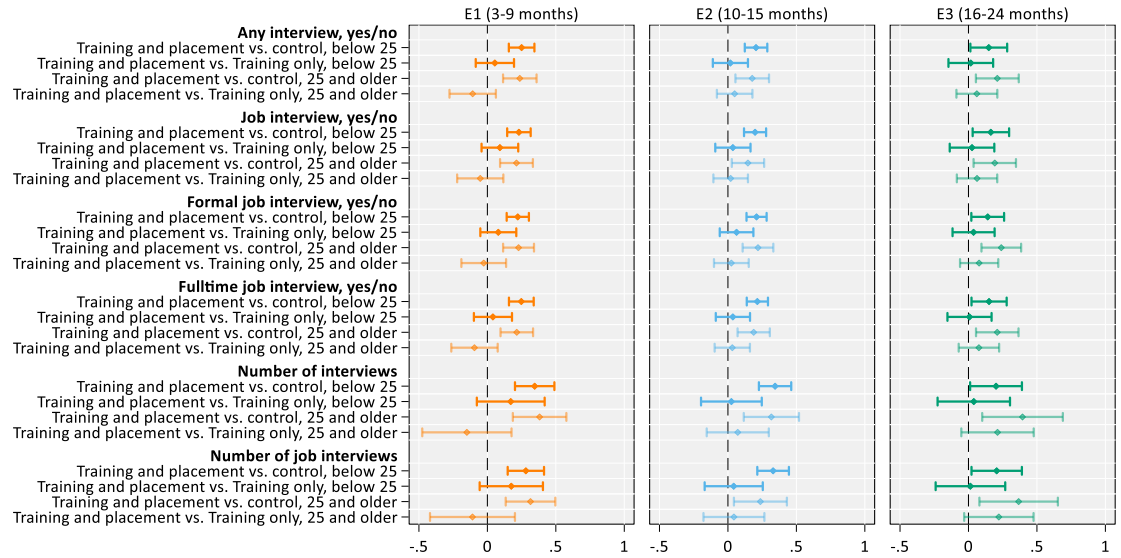
Source: Own calculations based on KAM survey.

Figure A1.3.7
Estimated treatment effects on job search, by gender



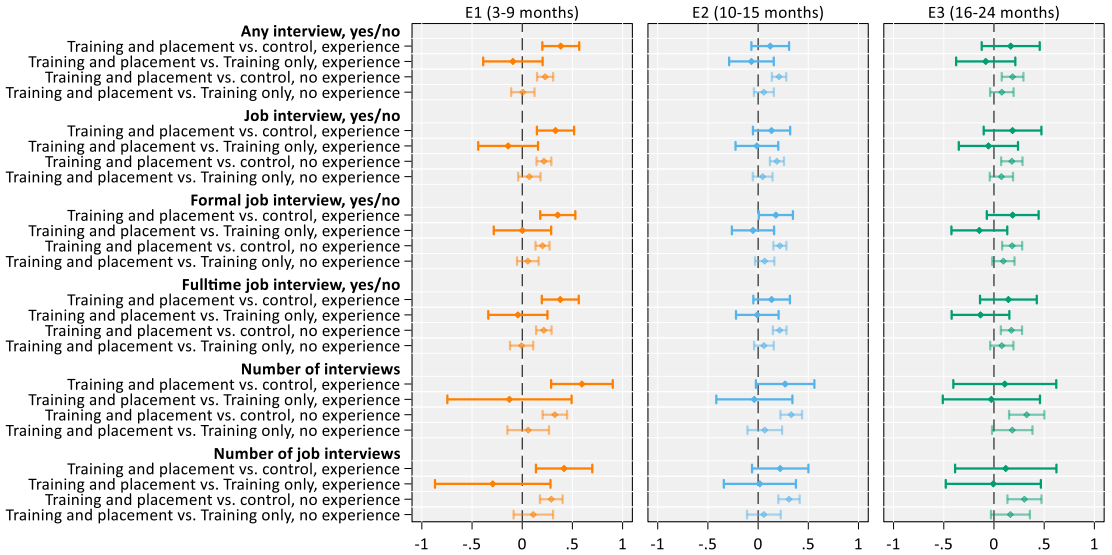
Source: Own calculations based on KAM survey.

Figure A1.3.8
Estimated treatment effects on job search, by age group



Source: Own calculations based on KAM survey.

Figure A1.3.9
Estimated treatment effects on job search, by experience



Source: Own calculations based on KAM survey.

A2 Appendix of the qualitative KAM evaluation

A2.1 Overview of interviews

Table A2.1.1
Referencing and overview of interviews

Ref.	Interview type	Interview partners' position	Location	Interview date
[1]	Background Interview 1	GIZ project staff	Nairobi	February 12, 2020
[2]	Background Interview 2	KAM TVET program coordinator	Nairobi	February 13, 2020
[3]	Background Interview 3	KAM TVET program staff	Nairobi	February 14, 2020
[4]	Background Interview 4	GIZ project staff	Nairobi	March 6, 2020
[5]	Company 1, Interview 1	HR manager & managing director	Thika	February 17, 2020
[6]	Company 2, Interview 1	Training manger	Thika	February 18, 2020
[7]	Company 2, Interview 2	Head of HR	Thika	February 18, 2020
[8]	Company 3, Interview 1	HR manager	Limuru	February 20, 2020
[9]	Company 4, Interview 1	Factory manager	Ruiru	February 21, 2020
[10]	Company 5, Interview 1	HR Manager	Thika	February 21, 2020
[11]	Company 5, Interview 2	Supervisor of interns	Thika	February 21, 2020
[12]	Company 6, Interview 1	HR manager	Thika	February 25, 2020
[13]	Company 7, Interview 1	Former and current HR manager	Thika	February 25, 2020
[14]	Company 8, Interview 1	Managing director & HR manager	Thika	February 26, 2020
[15]	Company 8, Interview 2	Supervisor of interns 1 & Supervisor of interns 2	Thika	February 26, 2020
[16]	Company 9, Interview 1	Head of HR & HR manager	Kikuyu	February 27, 2020
[17]	Company 10, Interview 1	Managing director & HR manager	Thika	March 6, 20202
[18]	Company 11, Interview 1	Head of HR & HR manager	Ruiru	March 12, 2020
[19]	Company 12/CCC 1, Interview 1	CEO; CCC representative	Ruiru	February 26, 2020
[20]	Company 13/CCC 2, Interview 1	CEO; CCC representative	Thika	February 26, 2020
	Observation of CCC meeting		Thika	February 27, 2020

Source: RWI.

A2.2 Blueprints of consent forms for interviews with companies

CONSENT FORM FOR COMPANIES PARTICIPATING IN FUNDING SWITCH

Assessment of the GIZ/KAM Central Chapter TVET internship program implemented by the Kenyan Association of Manufacturers (KAM)

Leibniz Institute for Economic Research

BACKGROUND AND USE OF INFORMATION:

The Kenyan Association of Manufacturers (KAM) is implementing internship placement services, supported by the German Development Cooperation (GIZ). In 2019, a switch in the internship funding scheme has taken place where the GIZ stipend funding for the KAM interns was cancelled and the companies themselves have taken over the payment of the stipends for the interns as part of a sustainability program. RWI, Leibniz Institute for Economic Research from Germany is carrying out an assessment of that switch as part of the GIZ monitoring.

You have been selected to be interviewed today because you are involved with the KAM internship program at your company and we believe that your views and experiences are a very valuable. We would like you to participate in an interview of about 1-2 hours with Stephanie Ihring, an analyst from RWI – Leibniz Institute for Economic Research. The interview will be recorded.

The purpose of the interview is to capture your personal views on the internship program and to understand your perspectives on how the switch from the GIZ stipend scheme to the sustainability program has affected the internship program at your company. Your participation in the interview is voluntary, which means that your refusal will have no negative consequences for you.

The information provided by you in the interview will be anonymized in a manner which will not allow a disclosure of your identity. Any results, reports or publications resulting from the analysis of the interview will be anonymous. Your personal data in combination with your responses will not be shared with third parties and will only be accessible by selected personnel of the RWI team, who will use the data confidentially. If the interview data will be used further by ourselves or others, the same measures as described above will be taken to protect confidentiality of the data and anonymity of the respondents. If you have any further questions regarding the interview or the use of your information, please contact Stephanie Ihring, RWI – Leibniz Institute for Economic Research via email at [EMAIL ADDRESS].

CONSENT

- I, the undersigned, have read and understood the background of this assessment and the use of information explained above. I have been given the opportunity to ask questions, which have been answered satisfactorily.
- I understand that the participation in this interview is voluntary and that my refusal has no negative consequences.
- I understand that I will be interviewed and that the interview will be audio recorded.
- I understand that my personal details such as my name will not be revealed to people outside the RWI team.
- I understand that my words may be used, in an anonymized manner, in publications, reports, web pages and other outputs but my name will not be used.
- I agree to assign the copyright I hold for any agreed upon material related to the internship program to RWI.

SIGNATURES

_____	_____	_____
Full name of participant	Signature	Date
_____	_____	_____
Name of interviewer	Signature	Date

CONSENT FORM FOR COMPANIES NOT PARTICIPATING IN FUNDING SWITCH

Assessment of the GIZ/KAM Central Chapter TVET internship program implemented by the Kenyan Association of Manufacturers (KAM)

Leibniz Institute for Economic Research

BACKGROUND AND USE OF INFORMATION:

The Kenyan Association of Manufacturers (KAM) is implementing internship placement services, supported by the German Development Cooperation (GIZ). In 2019, a switch in the internship funding scheme has taken place where the GIZ stipend funding for the KAM interns was cancelled and the companies themselves were asked to over the payment of the stipends for the interns as part of a sustainability program. RWI, Leibniz Institute for Economic Research from Germany is carrying out an assessment of that switch as part of the GIZ monitoring.

You have been selected to be interviewed today because you are involved with the KAM internship program at your company and we believe that your views and experiences are a very valuable. We would like you to participate in an interview of about 1-2 hours with Stephanie Ihring, an analyst from RWI – Leibniz Institute for Economic Research. The interview will be recorded.

The purpose of the interview is to capture your personal views on the internship program and to understand your perspectives on the switch in funding from the GIZ stipend scheme to the sustainability program. Your participation in the interview is voluntary, which means that your refusal will have no negative consequences for you.

The information provided by you in the interview will be anonymized in a manner which will not allow a disclosure of your identity. Any results, reports or publications resulting from the analysis of the interview will be anonymous. Your personal data in combination with your responses will not be shared with third parties and will only be accessible by selected personnel of the RWI team, who will use the data confidentially. If the interview data will be used further by ourselves or others, the same measures as described above will be taken to protect confidentiality of the data and anonymity of the respondents. If you have any further questions regarding the interview or the use of your information, please contact Stephanie Ihring, RWI – Leibniz Institute for Economic Research via email at [EMAIL ADDRESS].

CONSENT

- I, the undersigned, have read and understood the background of this assessment and the use of information explained above. I have been given the opportunity to ask questions, which have been answered satisfactorily.
- I understand that the participation in this interview is voluntary and that my refusal has no negative consequences.
- I understand that I will be interviewed and that the interview will be audio recorded.
- I understand that my personal details such as my name will not be revealed to people outside the RWI team.
- I understand that my words may be used, in an anonymized manner, in publications, reports, web pages and other outputs but my name will not be used.
- I agree to assign the copyright I hold for any agreed upon material related to the internship program to RWI.

SIGNATURES

_____	_____	_____
Full name of participant	Signature	Date
_____	_____	_____
Name of interviewer	Signature	Date

A2.3 MAXQDA code system used for analysis of interviews

As mentioned in section 3.3 the raw data of the interview transcripts was analyzed with the help of MAXQDA, a 'Computer Assisted Qualitative Data Analysis Software' (CAQDAS (Kuckartz/Rädiker 2019). MAXQDA helps to identify themes and categories in the data by coding. 'Coding' essentially means that data is broken down into named an labeled segments. This helps to organize the data and to detect broader themes and patterns in it (cf. Bryman 2012). To code the data, a primary code system was developed based on the research interests from the proposal and theoretical factors on ownership and sustainability identified from the literature.

In an iterative process that consisted of several cycles of reformation of the existing codes and recoding of text segments the final code system presented in Figure A2.3.1-A2.3.6 emerged. The code system consists of 160 codes (cf. Figure A2.3.2-A2.3.6) summarized under 19 top level codes (cf. Figure A2.3.1). In total 1.646 text segments were coded across all interview transcripts.

Figure A2.3.1
Overview of MAXQDA code system used for analysis

Code System		1646
Broader Cooperation with KAM beyond TVET program		4
KAM TVET program under GIZ stipend		86
Intern numbers		73
Existence of other attachés/interns outside of KAM TVET program		60
Compnies cooperation's with other TVET institutions	L	17
Existance of internship/attachment program before KAM TVET		10
Participation under sustainability program (yes/no)		133
Willingness to take interns under sustainab. prog. in future	L	31
Willingness to hire interns as regular employees in future		8
Payment of interns		137
Reasons to take interns under sustainability program	L	63
Reasons not to take interns under sustainability program	L	73
Possible future participation under sustainability program	L	77
Companies' organization, structure & design of internships		454
Satisfaction with KAM TVET program (pro & contra)	L	44
General satisfaction with the KAM interns	L	149
General suggestions for KAM TVET program improvements	L	76
Boarder topics raised by companies	L	130
Function of interview partner in firm		21

Source: RWI.

Figure A2.3.2
Detailed MAXQDA code system used for analysis (1/5)

Code System		1646
Broader Cooperation with KAM beyond TVET program		4
KAM TVET program under GIZ stipend		0
Introduction of KAM TVET program to CCC		5
Introduction of KAM TVET program to companies	L	12
Reasons by industries to take-up the KAM TVET program		10
Reasons by KAM to offer TVET program		6
General reasons for companies to offer internships		34
Companies perception of TVET program under GIZ stipend		19
Intern numbers		0
# of interns under GIZ stipend		25
# of interns hired as regular employees under GIZ stipend		18
# of interns under the sustainability program		23
# of interns hired as regular employees under sustainability		7
Existence of other attachés/interns outside of KAM TVET program		27
# of attachés/interns outside KAM		18
# of other attachés/interns hired as regular employees		10
Existance of CBET interns in companies		5
Compnies cooperation's with other TVET institutions	L	17
Existance of internship/attachment program before KAM TVET		10
Participation under sustainability program (yes/no)		15
Companies perception of switch/sustainability program	L	29
Process of funding switch		14
Explanation of funding switch by KAM to companies		14
Reasons for funding switch from CCC side		6
Minimum requirement given by KAM?		4
Minimum payment requirement should be set	L	2
Consequences of switch/sustainability program	L	5
Companies take less interns when they have to pay themselves	L	16
Temporary workers are substituted		2
Consequences for low paying companies - interns leave		16
Students have money expectations because of GIZ stipend		10
Willingness to take interns under sustainab. prog. in future	L	31
Willingness to hire interns as regular employees in future		8

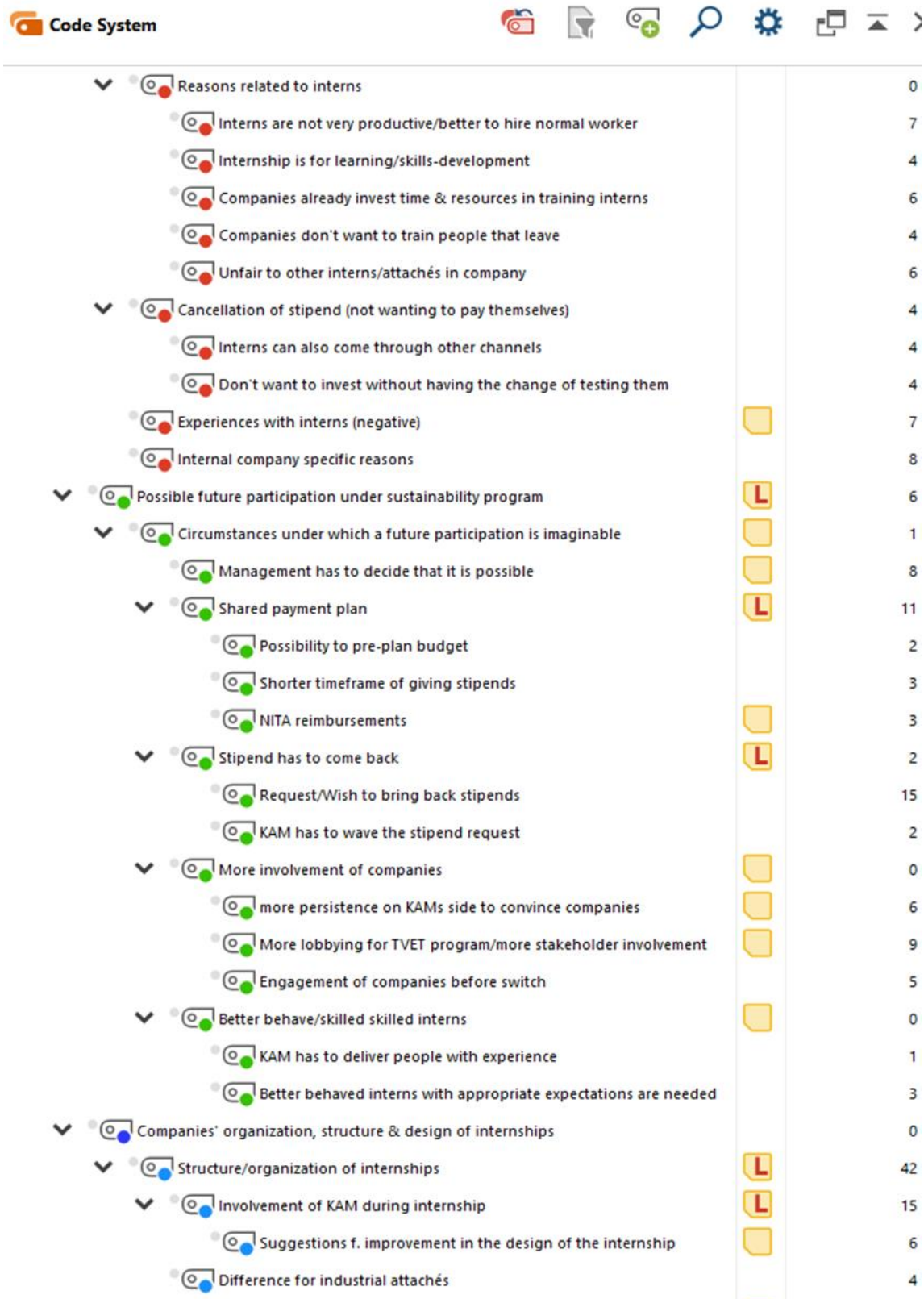
Source: RWI.

Figure A2.3.3
Detailed MAXQDA code system used for analysis (2/5)

Code System		
Payment of interns		0
Amount of stipend paid by companies		17
Justification of (not) paying certain amount		29
Role of reimbursement by NITA		7
Stipends paid to other interns/attachés	L	19
Willingness to pay/increase stipends in the future	L	23
Suggestions for appropriate amount	L	8
Assessment of sufficiency of stipend to live		11
Personal evaluation: payment of interns		8
Usage of stipend by interns		3
Payment level of regular employees		3
GIZ stipend		3
Delayed payments of stipends		6
Reasons to take interns under sustainability program	L	0
Altruistic reasons		0
Stipend needed for survival of interns		14
TVET interns have already graduated		2
"Giving back"/ companies duty		5
Self-serving reasons		0
It motivates interns/better performance		7
Investing in interns is worth it in the long run		9
Interns add value to company		3
Giving a stipend is the companies policy		17
Because of sustainability program/KAM expectation		4
Management supports the decision		2
Experiences with interns (positive)		0
Reasons not to take interns under sustainability program	L	0
Company related reasons		0
Paying stipends is not the company's policy		5
Decision to pay interns is up to management		7
Not calculated in budget/switch too short-term		5
Bad economic year/unpredictable economy		2

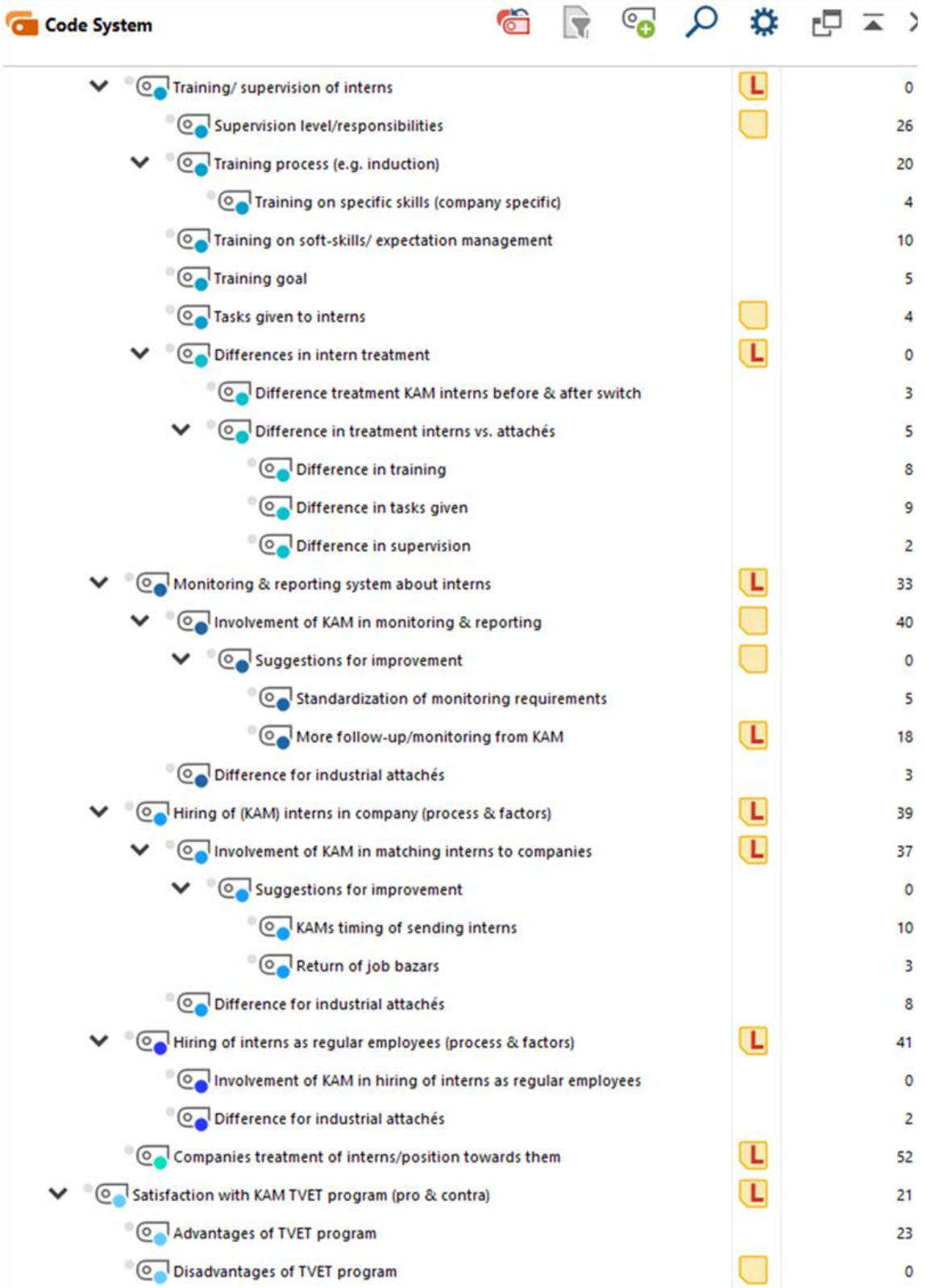
Source: RWI.

Figure A2.3.4
Detailed MAXQDA code system used for analysis (3/5)



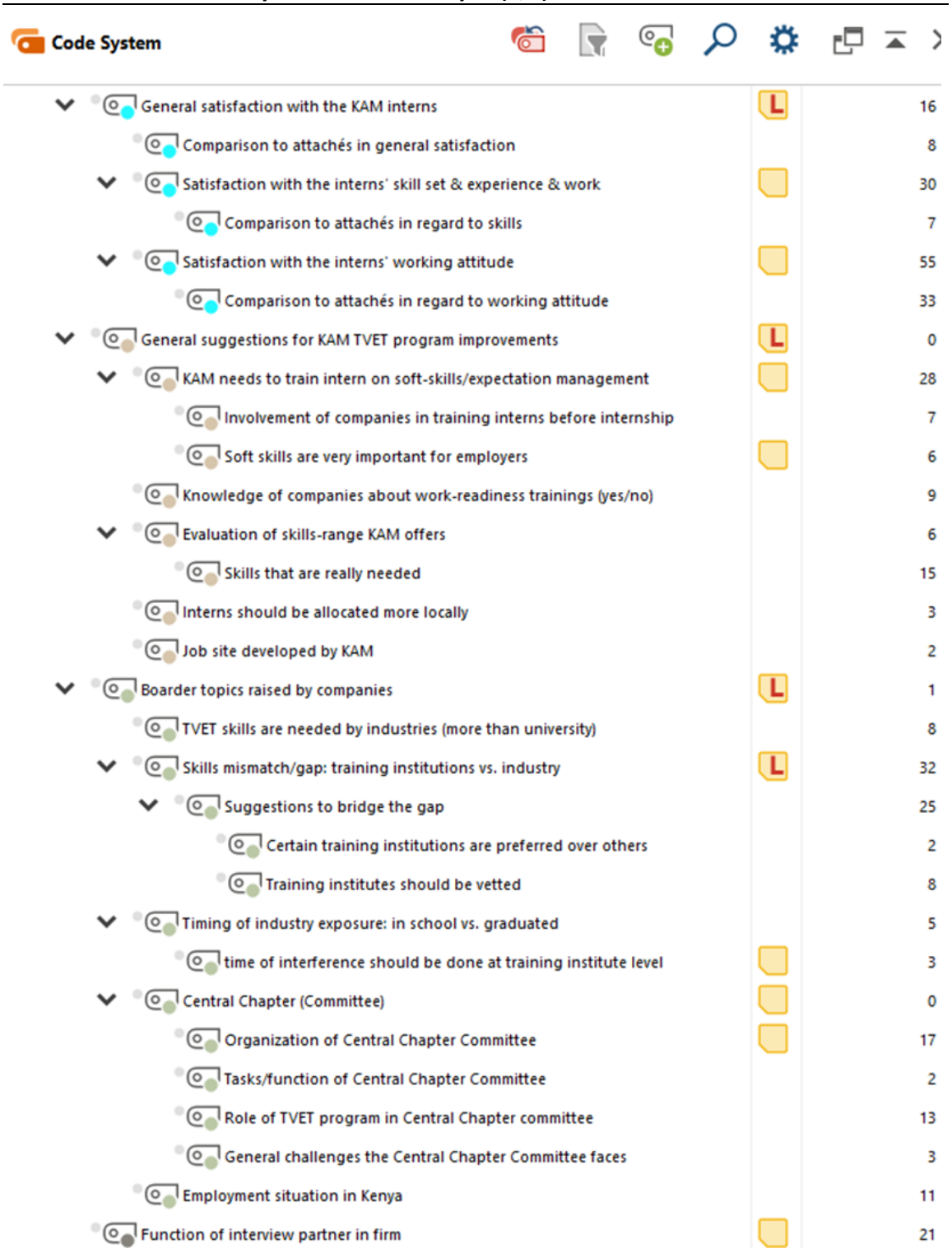
Source: RWI.

Figure A2.3.5
Detailed MAXQDA code system used for analysis (4/5)



Source: RWI.

Figure A2.3.6
Detailed MAXQDA code system used for analysis (5/5)



Source: RWI.

A2.4 Detailed description of factor model on ownership and sustainability

A factor model on ownership and sustainability was developed as a theoretical framework for the analysis. Since the factor model was only very briefly described in section 1.4, more detailed descriptions of the individual success factors on ownership and sustainability are provided in the following. The factors are based on findings from various studies that have researched ownership and/ or sustainability in different development cooperation contexts and project areas.

A2.4.1 Factors promoting ownership

Since ownership is a pre-requisite for sustainability (cf. Brolin 2017), the factors relating to ownership are discussed first. The factors are structured along the same topics and in the same order as presented in Table A (cf. section 1.4).

Factor O.1: Project/ activities must be in the interest of and useful for local stakeholders

One of the most crucial pre-requisites for achieving ownership of development projects or project activities is that stakeholders, who are supposed to take ownership, are interested in the project and perceive it as useful. Hasselkong and Schierenbeck (2017) argue that before a new project or project activity is designed, it should be considered whether local stakeholders and beneficiaries even have any interest in taking ownership of it (Hasselkong and Schierenbeck 2017). In cases where it is obvious from the beginning that local stakeholders or beneficiaries show no or only little interest in taking ownership, such projects should not be implemented or should be redesigned accordingly (cf. Edgren 2003). If not possible otherwise, there should be a focus on those "willing and able to take the lead and fulfil other demands of ownership" (Hasselkong and Schierenbeck 2017, 325).

Rey-Moreno et al. (2014) found in their study, which measures local stakeholders' ownership in a community-based project, that community members were more likely to accept a project when they perceived it to be useful, e.g., because it was relevant to their needs. If the project then actually delivered on its promises, people internalized a sense of ownership and committed to it in the sense that they took over responsibility and control (Rey-Moreno et al. 2014).

Factor O.2: Local stakeholders should already be involved in the design phase of projects/ activities

Since ownership is determined by the interest and the acceptance of the people for whom development projects are created, it is crucial that local stakeholders, i.e., the prospective owners, "drive the [project's] creation process from the beginning" (Ballantyne 2003, 3). Ballantyne (2003) argues further that it must be understood who they are and what demands, incentives and motivations they have to ensure their participation in the delivery and preservation of the project activities. Such end-user involvement is important because designing whole projects or project activities 'for someone else to take charge of', especially when they are not involved in it, makes it unlikely that they will take ownership of it (cf. Ballantyne 2003; cf. also Ostrom et al. 2002). Besides, when local stakeholders do not take ownership in the design stage, it becomes much more difficult to foster ownership among them when the project is already implemented (Weeks et al 2002).

Therefore, external development agents should let local stakeholders identify their own challenges and needs and account for their understanding of the problem. This helps to ensure that the project is useful to them, which goes a long way to evoke a sense of ownership among them later on (cf. Matuella 2010; Kinoti 2011). To ensure that all stakeholders can form an opinion and

formulate their ideas and expectations, the initial identification and design phases of a project or activity should not be rushed (cf. Edgren 2003).

Factor O.3: Possible evolvement of local ownership of a project/ activity over time

While the ideal way to achieve ownership is to involve end-users or local stakeholders as early as in the design phases of a project, in practice this is unfortunately rather the exception. In most cases, projects are developed by donors, which means they are the initial owners (cf. Ballantyne 2003). However, Edgren (2003) found that "it seems to be less important to ownership who originally came up with the idea than who took an interest in it and what forces supported it on either side of the partnership" (Edgren 2003, 12). Thus, owners at the outset of a project are not necessarily the same people as at the end because ownership can also develop over time when local stakeholders grow in capacities, understanding and commitment (cf. Ballantyne 2003).

However, in the academic literature agreed upon that ownership is something that is taken, not given, which means that taking ownership is not an automatic process but rather an active decision to take responsibility and control. Ballantyne (2003) and Edgren (2003), for example, argue both that local ownership develops generally best, when it is not dictated by external timetables or actors. In cases where there is no local ownership, ways must be found to transfer ownership to the local level (cf. Ballantyne 2003; Edgren 2003).

Factor O.4: Local stakeholders should be actively involved in all project stages/ activities

Ownership can develop particularly well when local stakeholders, or essentially all those that are expected to take ownership, are continuously and actively involved throughout *all* stages of the project, not just during the initial design phase (Weeks et al. 2002; Steiger 2015). This implies that "participation of all the stakeholders in project planning, implementation, monitoring and evaluation [is necessary to] foster ownership" (Kuria and Wanyoike 2016, 477).

Some studies even emphasize that a simple involvement of local stakeholders is not enough, but that an active participation in the decision-making processes throughout the life cycle of a project is crucial to ensure successful ownership (Ostrom et al. 2002; Kinoti 2011; Sustainable Development Goals Fund 2020). Kwangware et al. (2014) argue, for example, that when local stakeholders and beneficiaries have the opportunity to make decisions that influence program activities, they develop a greater commitment, which consequently fosters their ownership of these activities (Kwangware et al. 2014). Ballantyne (2003) stresses further that to achieve local ownership not only the direct intermediaries (i.e., the implementing agency) but especially the end-users (which are prospective owners) must be involved in all project decisions (Ballantyne 2003). The reason for that is that end-users "usually see things quite differently and have different incentives and motivation" (ibid., 3).

Factor O.5: Financial involvement of local stakeholders, a strong indicator of local ownership

Many researchers agree that a key indicator for ownership is that local stakeholders contribute with their own resources, either financial or in kind, to a development project or an activity of it (cf. Weeks et al. 2002; Kuria and Wanyoike 2016; Kinoti 2011; Olsson et al. 2008; Ballantyne 2003; Ostrom et al. 2002). It is assumed that, when recipients contribute "to the total costs of the project […], [it] is a measure of [their] commitment to the whole undertaking" (Edgren 2003, 13; cf. also Weeks et al. 2003; Olsson et al. 2008).

In the literature, it is further discussed whether ownership can be achieved when project costs are shared between donors and recipients or whether the costs have to be completely taken

over by local stakeholders or end-users of the project. Olsson et al. (2008), for example, equate the project owner to the main financing party of a project (Olsson et al. 2008). While Kuria and Wanyoike (2016) argue that when local stakeholders share the costs for project implementation, then it already contributes to a sense of ownership among them (Kuria and Wanyoike 2016). Thus, to promote ownership, local stakeholders as well as partner organizations should be encouraged to inject some of their own resources into a project (cf. Kinoti 2011). However, Ballantyne (2003) stresses, "since not all partners have the same resources, different types of contributions by each partner need to be valued" (Ballantyne 2003, 3).

Some authors also describe that it can be very challenging to convince local actors to take over any part of the project funding either because they have limited resources or because they do not see the urgent need to do so as most donors are willing to offer a financial safe when a project threatens to fail (Matuella 2010). The challenge for donors is, therefore, to find ways to encourage local stakeholders to invest their own resources (cf. Ballantyne 2003).

Factor O.6: Donors should limit their direct involvement in projects/activities

Focusing on the involvement of local stakeholders is not the only way to ensure ownership. Donors can also contribute to promote local ownership. For example, to help local stakeholders take ownership, donors need to "limit their direct roles and responsibilities and, as far as possible, encourage local stakeholders to participate and take the lead" (Ballantyne 2003, 3 For this to happen, Hasselkong and Schierenbeck (2017) argue that donors must have "trust in the recipient […] and [the] willingness to cede control" (Hasselkong and Schierenbeck 2017, 328). If this is not the case, then local ownership of a project or project activity is extremely unlikely.

Factor O.7: Donors should be accountable to and transparent with local stakeholders

To ensure local ownership, donors must be "transparent and clear in regard to [their] objectives, interests and expectations" (Ballantyne 2003, 8). If the project's objectives and next steps are transparent and understandable for end-users, then they are more likely to take ownership of the project (cf. Kinoti 2011; Ballantyne 2003; SDGF 2020).

The same level of transparency should also be offered to local implementing agencies. For example, donor agency policies (e.g., in regard to implementation or monitoring and evaluation) should be made more accessible and understandable to local partners if they are expected to take ownership of the implementation and evaluation of project activities (cf. Kaplan 2013; MercyCrops 2013). Lastly, to achieve local ownership, donors must also be accountable towards local stakeholders and project participants (cf. Cekan 2016b). Ballantyne (2003) recommends that implementing agencies should be accountable primarily to the recipients they are working with instead of their foreign donors (Ballantyne 2003).

Factor O.8: Incorporating projects/activities into local structures, values, and knowledge

Turning away from what donors can do, the next factors focus on the circumstances at project level that can be beneficial for promoting ownership. Some authors state that new projects or project activities should be implemented into existing local structures and built upon already existing local knowledge and values (Rey-Moreno et al. 2014; Ballantyne 2003). Rey-Moreno et al. (2014) found that local stakeholders were much more willing to take ownership of a development activity if it was incorporated into existing local values and norms "because [they] resonated with local ways of being and thinking" (Rey-Moreno et al. 2014, 75). Ballantyne (2003) advocates similarly for building new projects or project activities on structures that already exist

and that are ideally already owned by the people who are supposed to take ownership of the new development activity (Ballantyne 2003).

Factor O.9: Local leadership and management of projects/activities should be encouraged

Project management and leadership are also important factors that influence ownership at project level. Scholars agree that in order to promote ownership, local stakeholders should take over leadership and management roles in the project (cf. Edgren 2003; Kinoti 2011). A development project or activity can be said to be truly 'owned' by end-users if they are the ones who manage and lead the design, implementation as well as the monitoring and evaluation. Donors and implementing agencies should only support them in areas where local stakeholders lack skills and knowledge (Edgren 2003). Edgren (2003) argues that a "sense of ownership depends to a high degree on budget authority" (ibid., 13). If local managers are given the authority to deal with funds as they see fit, "they can be expected to take personal responsibility for results in a way they would not, […] if the donor is the real decision-maker" (ibid., 13).

Factor O.10: Capacity building for local stakeholders to promote ownership

Building on the argument above, capacity building is needed whenever local stakeholders lack certain capacities or knowledge. Capacity building helps to ensure that they can become "drivers and owners, instead of mere passengers, in the project […] process" (Ballantyne 2003, 3; cf. also Rey-Moreno et al. 2014). Ballantyne (2003) notes further that only thorough capacity building, e.g., in (financial) management or administration, can empower local stakeholders to actually take ownership (Ballantyne 2003). Moreover, capacity building is central because taking ownership can also be a major burden for local stakeholders (Weeks et al. 2002). Therefore, donors should not only consider whether local stakeholders show an interest in a project, but also whether they have the capacities to take ownership (Hasselkong and Schierenbeck 2017). Edgren (2003) argues that many donors tend to blame local partners after project failures in a 'backward logic' for "not having the capacity to assume ownership of [the] project [when it was actually] the designers […] who should be blamed for building it on unrealistic capacity assumptions" (Edgren 2003, 16).

Factor O.11: Open and transparent communication between all stakeholders involved

The academic literature on ownership also touches upon issues concerning relationships between external and local stakeholders involved in a project.

Ballantyne (2003) stresses that spaces or mechanisms need to be created where different owners can regularly meet and engage in dialogue to discuss decisions and express their demands and expectations (Ballantyne 2003). Edgren (2003) argues that a "critical instrument for establishing ownership and […] produc[ing] successful outcomes is dialogue" (Edgren 2003, 4). Good communication and dialogue help to ensure that objectives, time horizons and planned activities as well as changes within the project are made transparent for the parties involved, which in turn mitigates potential conflicts (Edgren 2003; Ballantyne 2003). Weeks et al. (2002) mention that the ideal project model would ensure that "all stakeholders, beneficiaries, technical staff, donors and policy makers come together to discuss and agree on action or strategy" (Weeks et al. 2002, 204). Of course, the larger the project, the more difficult it will become to achieve this ideal. Nonetheless, at least major new ideas and activities must be coordinated and aligned with all stakeholders to enable all relevant stakeholders to take ownership (ibid.).

Factor O.12: Clear allocation of roles and responsibilities among all stakeholders involved

Finally, if shared ownership between different stakeholders of a project is the objective, a clear allocation of roles and responsibilities is necessary. Roles and responsibilities should be clearly defined in the beginning and must be reviewed, if "the setting change[s] during implementation and new actors and roles emerge" (Edgren 2003, 20). It is helpful to establish internal governance structures for the project, where responsibilities and roles are clearly set out for each stakeholder (Ballantyne 2003). A stakeholder analysis prior to project implementation can ensure that "each stakeholder can take appropriate types and levels of ownership" (Ballantyne 2003, 4; cf. also Edgren 2003).

A2.4.2 Factors promoting project sustainability

Since ownership is a pre-requisite for sustainability (cf. Brolin 2017), many factors relating to sustainability are similar to those discussed in regard to ownership. However, local ownership alone is not sufficient to ensure the sustainability of a development project (Ostrom et al. 2002; Edgren 2003). Therefore, the factors mentioned in the literature on sustainability are in some areas wider-ranging compared to those mentioned in the literature on ownership. To ensure a comparability to the ownership factors, the factors on sustainability include the same topics and are presented in the same order as the factors on ownership.

Factor S.1: Projects/activities must be in the interest of and useful for local stakeholders

The interest of local stakeholders in a development project is not only important for local ownership (cf. Factor O.1), but it is also often referred to as a crucial factor to ensure sustainability of development projects. Maier et al. (2016) argue that local interest in a project or project activity and an acceptance of it "[are] crucial for the success and long-term effectiveness of the development cooperation project" (Maier et al. 2016, 11). Bossert (1990) found that projects which could demonstrate effectiveness in reaching their goals and objectives were much more likely to be perceived as useful by local stakeholders. This led to the finding that those projects were more likely to be sustained by "those expected to fund and manage the activities after the donor funding stops" (Bossert 1990, 1019).

Factor S.2: Local stakeholders should already be involved in the design phase of projects/activities

Similar to the literature on ownership (cf. Factor O.2), it is also often argued in the literature on sustainability that stakeholders, who are supposed to sustain project activities, should be involved as much as possible right from the start "when decisions are being made about what type of project is required" (Oino et al. 2015, 764). Okun (2009) states that "designs which are expected to result in sustainable benefits should build on local demand and initiatives" (Okun 2009, 18). He explains that projects are more likely to be sustained when they actually meet the demands and needs of the people that are supposed to sustain them (ibid.). To achieve this, the beneficiaries as well as the staff implementing the project should "play a core role in the identification and design process" (ibid.). Some authors go even further and state that sustainability cannot be achieved unless local players are truly involved in the design phase and any further project transformations (cf. Lopes and Theisohn 2003; Kuria and Wanyoike 2016).

Factor S.3: Local stakeholders should be actively involved in all project stages activities

Same as in the literature on ownership (cf. Factor O.4), active stakeholder participation is with overwhelming quantity mentioned by scholars as a key factor to achieve project sustainability, not only in the project design phase, but also in all subsequent project stages (cf. Oino et al. 2015; Hamukwala et al. 2008; Okun 2009; Ochunga and Awiti 2017; Kuria and Wanyoike 2016; Hofisi and Chizimba 2013; Karanja 2014; Morán and Ferguson 2013). A number of studies researching sustainability of development projects argue that "donor-funded projects can only be sustainable if they allow for participatory processes from identification to completion" (Hofisi and Chizimba 2013, 713; cf. also Oino et al. 2015; Kuria and Wanyoike 2016; Steiger 2015). Studies that evaluate how different forms of participation influence the sustainability of projects find that an interactive participation, where stakeholders have decision-making and planning power, lead to the highest rates of project sustainability (cf. Ochunga and Awiti 2017; Hamukwala et al. 2008 for further details). Oino et al. (2015) remark that simply informing local stakeholders of decisions that are already made is not sufficient for genuine participation (Oino et al. 2015). Actively involving local stakeholders is therefore an important factor for sustainability because people who have a voice in determining the project's objectives and the possibility to influence the implementation are more likely to take over the project activities after the external assistance has ended (cf. Seppey et al. 2017; Oino et al. 2015; Okun 2009).

Factor S.4: Financial involvement of local stakeholders - one way to ensure project sustainability

Involving local stakeholders or project beneficiaries financially in a project is seen as one possibility to ensure the sustainability of a project or project activity after the donor withdraws funding (cf. Kuria and Wanyoike 2016; Okun 2009, Bossert 1990). However, contrary to what is discussed in the literature on ownership (cf. Factor O.5), continuous and stable funding is seen as more important for sustainability than actual financial involvement of local stakeholders. Essential for sustainability is that a continuous flow of funding from some source is ensured in order to sustain the project activities and, thus, the project benefits (Okun 2009; Kuria and Wanyoike 2016). However, possibilities to ensure such a continuation of funding are discussed separately below (cf. Factor S.12).

Nevertheless, Kuria and Wanyoike (2016) argue that contributions by local stakeholders or beneficiaries to project costs should be encouraged because it ensures project sustainability (cf. Kuria and Wanyoike 2016). However, Okun (2009) remarks that, if financial involvement of local stakeholders is the goal, "a rigorous and realistic assessment of the local capacity to meet these costs" (Okun 2009, 19) is necessary. Local stakeholders often have limited resources which makes them unable to sustain the project financially after donor support is withdrawn. Shen et al. (2016) find that an equal level of funding distribution leads to the best level of project sustainability and hence suggest cost sharing between different parties involved as a promising solution when local stakeholders only have limited resources. Similarly, Okun (2009) argues that local stakeholders should, to the best of their ability, contribute to project costs. While donors should provide the supplementary finances needed to ensure the most sustainable outcomes, at least until a permanent system can be implemented or the financial capacities of local stakeholders can be enhanced (Okun 2009).

However, as explained in Factor O.5, sometimes the problem is not the inability of local stakeholders to contribute financially, but rather the unwillingness to commit their own resources (cf. Matuella 2010). Studies have shown that local stakeholders or beneficiaries only commit their resources (particularly their financial resources) to a project or project activity when they

perceive the project as valuable (Kuria and Wanyoike 2016). Okun (2009) mentions that "if a program or project does *not* [emphasis added] deliver clear and equitable financial or economic benefits, which are apparent to the stakeholders, it is most unlikely to be sustained after donor funding finishes" (Okun 2009, 14). In other words, if local stakeholders do not find a project or its activities financially or economically lucrative, they will most likely not use their own resources to sustain it (cf. Kuria and Wanyoike 2016; Okun 2009). Thus, finding ways to ensure that projects meet the needs of local stakeholders goes a long way in ensuring that it is perceived as economically or financially beneficial by them. This in turn also contributes to the (financial) sustainability of a project (cf. Factor S.1).

Factor S.5: Donors should show flexibility to changing priorities of local stakeholders

Not only local stakeholders need to be involved to ensure project sustainability, but also donors can contribute to its achievement. While for ownership it is crucial that donors limit their direct roles and responsibilities and are accountable and transparent (cf. Factor O.6 and Factor O.7), for sustainability it is important that donors do not just push through donor-led top-down projects (Weeks et al. 2002; Kuria and Wanyoike 2016). If donors implement inflexible top-down projects, the projects will "fail to bring sustainable benefits because they do not lead to stakeholder ownership and commitment" (Kuria and Wanyoike 2016, 481; cf. also Oino et al. 2015). However, if donors recognize that local stakeholders needs change over time and if they subsequently show flexibility and responsiveness to adapt the project activities to the changing priorities of local stakeholders, then projects are likely to be sustained in the long term (Weeks et al. 2002; Kuria and Wanyoike 2016). This means that projects designed by donors must be flexible in the implementation and evolve when lessons are learned (cf. Okun 2009).

Factor S.6: Incorporating projects/activities into existing local structures, values, and knowledge

Turning away from what donors can do to help promote sustainability, the following factor addresses the external circumstances at project level that can be beneficial to achieve sustainability. Mirroring the academic literature on ownership (cf. Factor O.8), scholars focusing on project sustainability also find that development projects should build on existing local structures and take into account values and knowledge of local people (cf. Bossert 1990; Hofisi and Chizimba 2013; Okun 2009; Oino et al. 2015; Hamukwala et al. 2008). Bossert (1990) who conducted a sustainability assessment of development projects in five countries finds that projects were more likely to be sustained if they are "integrated into the existing institutional hierarchies" (Bossert 1990, 1019). He argues that projects which are integrated in existing administrative structures may even create actors that "continue project activities with their own resources or may help lobby for additional national resources to cover the loss of donor funding" (ibid., 1020). Oino et al. (2015) as well as Okun (2009) find that especially projects that are built on and integrated in local management structures have good prospects to achieve project sustainability (Oino et al. 2015; Okun 2009). Moreover, if project designs build on knowledge of local stakeholders, they become more relevant to them because "local people understand their problems better and can therefore use their skills and resources to find flexible solutions that are tailored to suit their unique needs" (Oino et al. 2015, 760). This in turn makes local stakeholders more likely to sustain such projects (ibid.)

Factor S.7: Local leadership and management of projects/activities should be encouraged

Same as for ownership (cf. Factor O.9), encouraging local project management and leadership can also help to promote sustainability (cf. Karanja 2014; Okun 2009). If a project is well managed and led during the implementation, it can be ensured that sufficient capacities and resources are available to continue the project when the donor finally decides to retreat (Oino et al. 2015). Furthermore, when local stakeholders or the staff of a local implementing agency already take over essential leadership and management positions while the donor is still involved in the project, a significant weakening of the key project activities can be avoided when the donor withdraws. To prepare for such a take-over, a good organizational culture and well-funded knowledge-management structures should already be developed during project implementation. This is necessary to ensure that the required knowledge is accessible to the people that are supposed to sustain the project activities (cf. Oino et al. 2015; Okun 2009). Furthermore, local project managers and staff involved in the project should receive adequate training to ensure effective project delivery, which in turn supports project longevity (Karanja 2014).

Factor S.8: Capacity building for local stakeholders to ensure project sustainability

In close connection to the Factor S.7 described above stands the issue of capacity building which is not only central to ensure ownership (cf. Factor O.10) but also for the sustainability of projects (cf. Okun 2009, Karanja 2014, Oino et al. 2015). Okun (2009) argues, for example, that capacity building for local stakeholders is central to ensure project sustainability, because their capacities distinguish whether and how well a project or project activity is handled after the donor withdraws (Okun 2009). Oino et al. (2015) state that "capacity building is an empowering tool that enables [local stakeholders] to manage challenges on their own, rather than depending on the help of the [implementing] organizations or donor" (Oino et al. 2015, 763).

Projects should generally only be handed over when the "target beneficiaries and stakeholder[s] have adequate capacity, knowledge and skills to effectively run the project" (Okun 2009, 61f.). Thus, adequate time and resources should be invested to ensure that capacity building and training for all people involved in the project can be facilitated. Local capacity building is especially important in areas of effective implementation, financial management, monitoring and evaluation as well as in day-to-day management of the project and its activities. Capacity building should be initiated at the start of the project and should continue throughout the project until the donor exits. It is particularly important when new staff members at partner level are onboarded or when new stakeholders join the project (cf. Oino et al. 2015; Karanja 2014; Nyanjura 2010).

Factor S.9: Thorough monitoring and evaluation can be helpful to achieve sustainability

Kuria and Wanyoike (2016) find that rigorous "monitoring and evaluation greatly improves the sustainability of donor-funded projects" (Kuria and Wanyoike 2016, 496). Thorough monitoring and evaluation make it possible to identify potential problems as well as successes and provide the basis for corrective action that is needed to ensure long-term sustainability (ibid.). Karanja (2014) argues along the same line when stating that monitoring and evaluation is "a major aspect that cannot be overlooked because it determines the sustainability of any venture or project" (Karanja 2014, 4). He understands monitoring and evaluation as a means to learn from past experience which can positively influence the delivery of sustainable benefits in the future (ibid.). Kagendo (2015) states that all those involved in program activities should be responsible for monitoring and evaluation (Kagendo 2015).

Therefore, there should be a monitoring and evaluation framework for every project to allow an ongoing review of the project's effectiveness in reaching its benefits (Kuria and Wanyoike 2016). However, Hofisi and Chizimba (2013) warn that when such frameworks are designed, they should not be so complicated that they require continued outside expertise. The framework should be kept simple enough to be sustained even after the donor withdraws (Hofisi and Chizimba 2013; cf. also Okun 2009).

Factor S.10: Regular sustainability assessments are helpful to plan for and enhance future sustainability

Okun (2009) explains that if achieving project sustainability is a major objective, then monitoring and evaluating the targeted outputs and activities is not enough. Instead, thorough monitoring and evaluation should also include assessments that focus on how project sustainability can be achieved (cf. Okun 2009). So-called sustainability assessments "should be used as a management tool to identify any deficiencies and develop an action plan for sustainability" (Kuria and Wanyoike 2016, 478; cf. also Kagendo 2015). Particularly the perspectives of local stakeholders should be included in a sustainability analysis because feedback about ongoing activities from local stakeholders is important for sustainability (cf. Kuria and Wanyoike 2016; Cekan 2015). A study from the OECD finds that "it may be useful to evaluate sustainability even while [donor] funding or activities are ongoing" (OECD 2019, 12). A sustainability analysis can also help to integrate sustainability strategies in existing projects.

Factor S.11: Implementation of sustainability strategies and exit plan to plan for project phase-out

While sustainability assessments can be helpful for projects that do not have (well thought-out) sustainability strategies, such strategies should ideally be in place right from the project start (cf. Oino et al. 2015). For example, Hofisi and Chizimba (2013) found in their analysis that projects were more sustainable if sustainability strategies were integrated right from the start. Evidence from other scholars even suggests that projects cannot achieve long-term sustainability, unless sustainability strategies are included from the beginning (cf. Nyanjura 2010; Seppey et al. 2017). Ochunga and Awiti (2017) observed that while "many projects highlight elements of sustainability in their proposal stage, the actual implementation seems to lack emphasis on sustainability" (Ochunga and Awiti 2017, 382). Therefore, achieving project sustainability should be regarded as an "ongoing process [that] needs to be reviewed and updated as circumstances change and lessons are learned from experience" (Okun 2009, 11).

A central part of such a sustainability strategy is an exit plan or, as it is also called, a 'project phase-out strategy'. Many donors do not plan for a project phase-out which often leads to an abrupt end of, for example, the funding and in turn significantly impairs the continuation of the project (cf. Okun 2009; Ochunga and Awiti 2017). Therefore, many authors argue that having an exit strategy or a succession plan is "necessary to ensure that the target beneficiaries and the stakeholders are well prepared to effectively run the project after withdrawal of donor support" (Okun 2009, 61; cf. also Oino et al. 2015; Dunbar 2013; Cekan 2016a; Hofisi and Chizimba 2013).

Moreover, it is important that exit strategies are transparent for local stakeholders. This includes, for example, that they are informed from the beginning about the donor's time horizons for project phase-out or what the exit plan actually entails (Hofisi and Chizimba 2013). Cekan (2016b) argues that exit strategies should be very explicit and thorough to ensure sustained benefits. Requests of local stakeholders in terms of what they need for a successful continuation of

the project should also be considered by the donor when planning the transition period (cf. Cekan 2016b).

Apart from planning an exit strategy, the implementation of the actual project handover to the local level is also essential for sustainability (Cekan 2016a). To sustain the activities, a successful exit strategy should include, for example, extra training and support for local stakeholders (cf. Cekan 2015). Okun (2009) argues that sometimes it is even necessary that donors provide some 'limited follow-on assistance' after the donor-funded project activities have already ended (Okun 2009). This can include continued technical guidance for a transition period, e.g., frequent advisory visits, or even "supplementary financial support to enhance the prospects for sustainability and to consolidate achievements" (ibid., 11f.). Exit strategies should also include propositions how continued funding of project activities can be ensured without donor involvement (cf. Cekan 2015).

Factor S.12: Ensuring a continuation of funding is essential for project sustainability

As mentioned above, an important part of a successful exit strategy should also include plans or at least propositions for a continuation of project funding after donor withdrawal (Ballantyne 2003). As already mentioned in Factor S.4, ensuring a continuation of funding from reliable sources is absolutely essential for project sustainability, since every project depends on resources for survival. For example, after reviewing 491 completed projects, the Asian Development Bank concludes that continued project funding is the most crucial factor to ensure sustainability (Asian Development Bank 2010). Because only with adequate funds can project activities continue and project benefits be sustained (cf. Kuria and Wanyoike 2016; Cekan 2016b, Okun 2009).

However, a major problem is that most donor-funded projects have short-term funds that are tied to a pre-planned project duration. Very rarely do development project designs factor in how these projects can become sustainable after donor funding has been withdrawn (Okun 2009). Assessing the sustainability of projects funded by Plan International, Ochunga and Awiti (2017) find that 41 percent of the respondents warned that the projects ceased to operate altogether if donor funding would be withdrawn. Another 31 percent of the respondents stated that the projects would be at least significantly negatively affected (cf. Ochunga and Awiti 2017).

Thus, the question is how a continuation of funding after donor withdrawal can be achieved. As mentioned above, the best way to ensure continued funding is to include a financial take-over or transition plan in the project's exit strategy (cf. Karanja 2014; Okun 2009; Oino et al. 2015). Bossert (1990), for example, emphasizes that project sustainability depends to a large extent on whether planning for future funding was developed during the life cycle of a project. Ideally, projects should be designed in a way that they become financially autonomous before the donor withdraws funding (Bossert 1990).

To help projects to become financially autonomous, funds can be raised from national sources in the recipient country. Bossert (1990) explains that "projects which were funded by national sources (private or public) after the [donor] funding ceased were clearly sustained" (ibid. 1017). Thus, establishing partnerships with national government ministries or large private organizations that are willing to invest their resources can be a possibility to ensure continued funding and, thus, sustainability. Kuria and Wanyoike (2016) argue that the funding responsibility should be transferred to the host country organization, i.e., the local implementing agency, which should either take over the funding itself or should acquire new funds from other sources (Kuria and Wanyoike 2016). Furthermore, as mentioned in Factor S.4, local stakeholders or beneficiaries of a project can also be asked to continue the funding of the project activities in order to sustain

them. However, issues such as limited resources and an unwillingness to contribute with their own resources often makes them an unreliable source.

Factor S.13: Close coordination and shared understanding between all stakeholders involved

Mirroring the academic literature on ownership (cf. Factor O.11; Factor O.12), scholars studying sustainability also refer to issues concerning the relationships of different stakeholders involved in a project.

Good communication between different stakeholders as well a clear allocation of roles is also regarded as helpful to ensure project sustainability (cf. Cekan2016b; Kagendo 2015). Okun (2009) stresses that roles and responsibilities should be clearly documented in an agreement between the donor and the implementing agency to ensure sustainability (Okun 2009, 19).

Studies researching sustainability additionally name close coordination between the different stakeholders, especially in regard to the content of the project, as a crucial factor to ensure sustainability (cf. Hofisi and Chizimba 2013; Cekan 2016b). A shared understanding of project objectives and goals between donors and local stakeholders is particularly important because projects will only be sustainable if their objectives are responding to the needs and priorities of the people that are supposed to sustain them (Kuria and Wanyoike 2016; Cekan 2016b). Such a shared understanding can be achieved when projects are "approved in a mutually respectful negotiating process" (Bossert 1990, 1021). Thus, sustainability seems to be determined by an agreement about the content of the project, while ownership concerns more the "organisation [and form] of aid-related activities, rather than their contents" (Hasselkong and Schierenbeck 2017, 330).

Factor S.14: Wider networks and partnerships should be built to ensure project sustainability

An additional factor which is also related to the area of relationships between different stakeholders is building networks and widening partnerships. Building networks and partnerships with a wider group of relevant stakeholders and institutional actors is regarded in the academic literature on sustainability as being helpful for achieving project sustainability. Collaborating with important ministries and embedding the project within relevant institutional contexts creates a supportive environment in which project benefits can be sustained even after the donor withdraws funding (cf. Oino et al 1015, Kuria and Wanyoike 2016; Hofisi and Chizimba 2013).

If projects are linked with organizations and/or with individual stakeholders that have similar objectives, they "are more likely to be sustainable […] [because] they [can] support and learn from each other, and are able to exploit [each] others' agendas, for example, for new funding opportunities" (Oino et al 2015, 765). Okun (2009) states that "generating an understanding of, and support for, a program or project's objectives among a wide group of stakeholders should be a component of any sustainability strategy" (Okun 2009, 22). Raising such support from other institutions and stakeholders on political and administrative levels should be facilitated through workshops, seminars, newsletters as well as through personal contacts and should already start in the beginning of a project (ibid.).

A2.5 Overview of data contradictions at each company by source of information and discrepancies in these numbers across sources

Table A2.5.1

Legend for Table A2.5.2 "Overview of data contradictions at each company by source of information and discrepancies in these numbers across sources"

Sources	
M&E records	≙ Data of the project's M&E records collected by KAM The records used consist of databases where detailed information of the interns under the GIZ stipend were collected (called **'Master Database'**) and a preliminary **'Sustainability program list'** where just intern numbers per company are listed. Both databases were very poorly maintained and were, thus, often incomplete or contradicting.
GIZ/KAM	≙ (verbal or written) information provided by GIZ or KAM It has to be noted that some of the information provided changed in some instances over the course of the field research.
Companies	≙ Information provided by interviewed companies. The company rows are only filled for the interviewed companies.
IPA	≙ Data collected by Innovations for Poverty Action (IPA) IPA collected quantitative data on the interns that participated under the KAM TVET program (e.g., on internship and job placements)
Color coding	
	≙ Company names marked in mud green indicate the interviewed companies
	≙ Cells marked in red indicate when contradictions between the different sources
	≙ Cells marked in green indicate when the different sources march
N.A.	≙ Cells marked with N.A. indicate when information should be available but is not
-	≙ Cells marked in grey indicate when no information is required E.g., a particular source is not required to have that information, or a particular source was not asked to provide it.
'Conclusion'	≙ The last column on the right named "Conclusion", provides an explanation whether it could be verified if a particular company was recruited to participate under the sustainability program and whether the company then *actually* participated in the pilot (i.e., took in interns and paid them a stipend). The assessment is based on the availability of the information in the other columns as well as on the consistency of the four different sources. Certain sources were regarded as more trustworthy than others (e.g., information obtained directly from the companies is regarded as more trustworthy than the often incomplete KAM M&E records).
Color coding 'Conclusion'	
	≙ Conclusion cells marked in light green indicate that it could be verified that a particular company participates under the sustainability program
	≙ Conclusion cells marked in yellow indicate that it could be verified that a particular company was recruited to participate under the sustainability program but has not yet started to actually participate
	≙ Conclusion cells marked in light red indicate that it is assumed that a particular company was never asked to participate in the pilot of the 'sustainability program'

Source: RWI.

Table A2.5.2

Overview of data contradictions at each company by source of information and discrepancies in these numbers across sources

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
1	Company 14	*M&E records:*	yes	2[1]	0	1[2] (1)[3]	3 (3)[4]	0	N.A.	N.A.	Kiamwangi	N.A.	N.A.	*confirmed that interns are taken under new funding scheme	Verifiable that recruited to participate & that company subsequently participated in new funding scheme Reasons: verbal confirmation that company was asked to participate & M&E records of interns placed under new funding scheme exist
		KAM/GIZ:	no, then yes*	3	-	-	3	-	-	N.A.		101-250 mil.	147		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	yes	2	0	2 (1)[3]	2	0	1(1)[3]	-		-	-		
2	Company 15	*M&E records:*	N.A.	N.A.	N.A.	1**	N.A.	N.A.	N.A.	N.A.	Thika	N.A.	N.A.	*not in original list distributed by KAM/GIZ but marked in M&E database as part of pilot program **assumption based on KAM M&E records: listed because one intern found employment	Assumption that not recruited to participate & subsequently not participating in funding switch Reason: reliable database is missing
		KAM/GIZ:	N.A. then yes*	N.A.	-	-	N.A.	-	-	-		N.A.	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	N.A.	-	-	-	-	-	-	-		-	-		

1 *In their M&E records, KAM counts interns as "placed in the company" even when the interns never reported to the company. Here, only actual placements in the companies were counted, i.e., either when the internship was finished or when an intern started an internship but dropped after some time.*

2 *Other than in the KAM M&E records, job placements are only counted if the job was found after intern was attached to the particular company (i.e., either finished internship or dropped after some time) but not when intern was not available for placement i.e., never reported (never started) because he/she found a job somewhere else.*

3. ***Numbers** without parenthesis ≙ total number of interns placed in jobs after attachment to company; number in parenthesis ≙ number of interns placed in jobs in same company (number comparable with company accounts); e.g., 4 (2) ≙ interns placed in company X, 4 placed in jobs, 2 of them in the company they did their internship.*

4. *For the intern numbers placed under the sustainability program, two different M&E databases exist. Number without parenthesis ≙ listed in main M&E database; number in parenthesis ≙ based on numbers in incomplete 'sustainability placement list'. The sustainability list is less reliable because the single interns are not identifiable; only intern number placed per company is listed.*

5. *If answers are put in parenthesis, information is less reliable because it was only available in the (incomplete) sustainability M&E database and not in the main database.*

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
3	Company 3	*M&E records:*		6	0	3 (0)[3]	-2	N.A.	N.A.	N.A.	Limuru	N.A.	N.A.		Verifiable that recruited to participate & that company subsequently participated in new funding scheme
		KAM/GIZ:	no, then yes	3	-	-	3	-	-	N.A.		5-10 billion	954		
		Company:	yes	N.A.	N.A.	N.A.	02. Mrz	N.A.	N.A.	12.000KES		N.A.	N.A.		Reason: Verified personally with company that they are participating
		IPA:	yes	10	1	12(2)[3]	2	1	2 (2)[3]	-		-	-		
4	Company 11	*M&E records:*	(yes)[5]	15	2	3 (1)[3]	(2)[4]*	N.A.	N.A.	N.A.	Ruiru	N.A.	N.A.	*Company hired two interns directly from job bazar, pays them 30.000 KES salary not stipend. Those two were counted by KAM as placed under the new funding scheme.	Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	yes	N.A.	-	-	2 *	-	-	30.000KES		>10 billion	N.A.		Reasons: Verified personally with company; KAM M&E records wrongly reported the company as participating in funding switch
		Company:	no	12	N.A.	≈ 2	0	0	2*	0 KES		N.A.	>3000		
		IPA:	yes	12	1	12(2)[3]	2	0	2 (2)[3]	-		-	-		
5	Company 16	*M&E records:*	no	1	0	1(0)[3]	0	0	0	N.A.	Ruiru	N.A.	N.A.	*confirmed that recruited to participate but never took interns	Verifiable that recruited to participate but did not participate in
		KAM/GIZ:	yes, then no*	N.A.	-	-	N.A.	-	-	-		251-500 million	35		Reasons: M&E records and (verbal) information given by GIZ/KAM match
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	no	1	0	1(0)[3]	0	0	0	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
6	Company 17	*M&E records:*	N.A.	4	1	2(0)[3]	N.A.	N.A.	N.A.	N.A.	Ruiru	N.A.	N.A.	*confirmed that company was recruited to participate and that the company took one intern under the sustainability program in the beginning of March 2020 (during field research)	Verifiable that recruited to participate & that company subsequently participated in new funding scheme Reasons: verbal confirmation that company started to participate during the field research; IPA data confirm placement
		KAM/GIZ:	yes*	N.A.	-	-	1*	-	-	N.A.		251-500 million	120		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	yes	4	1	4 (1)[3]	1	0	0	-		-	-		
7	Company 18	*M&E records:*	(yes)[5]	N.A.	N.A.	1	(1)[4] **	N.A.	N.A.	N.A.	Kikuyu	N.A.	N.A.	*not in original list distributed by KAM/GIZ but listed in M&E records ** assumption: intern him-/herself found job in that company but was counted as placed under sustainability	Assumption that not recruited to participate & subsequently not participating under sustainability program Reasons: reliable database is missing (e.g. not available in IPA data; KAM sustainability placement list is unreliable
		KAM/GIZ:	N.A. then yes*	N.A.	-	-	N.A.	-	-	N.A.		-	-		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	N.A.	-	-	-	-	-	-	-		-	-		
8	Company 2	*M&E records:*	yes	3	1	1(0)[3]	6 (6)[4]	0	N.A.	N.A.	Thika	N.A.	N.A.		Verifiable that recruited to participate & that company subsequently participated in new funding scheme Reason: Verified personally with company & all sources report that company participates
		KAM/GIZ:	yes	6	-	-	6	-	-	N.A.		3-5 billion	130		
		Company:	yes	10	0	3	6	0	3	10.000KES		N.A.	680		
		IPA:	yes	2	0	2 (2)[3]	4	0	3 (1)[3]	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
9	Company 4	*M&E records:*	(yes)[5]	18 **	1 **	4 (1)[3] **	(2)[4] ***	N.A.	N.A.	N.A.	Tatu (Ruiru)	N.A.	N.A.	*Not in original list distributed by KAM/GIZ but later added by KAM ** not possible to differentiate between intern # at company in Nairobi and Tatu City (Kiambu county)	Verifiable that recruited to participate & that company subsequently participated in new funding scheme
		KAM/GIZ:	N.A. then yes	N.A.	-	-	2	-	-	30.000 KES		N.A.	N.A.	***company hired two people directly form job bazar as regular workers, but counted by KAM as placed in internship under sustainability program	Reason: Verified personally wit company. However, special case, since it was not part of original pilot program. The company received a stipend until Jan. 2020 because interns were allocated via the headquarter in Nairobi. Just started sustainability program in Feb. 2020.
		Company:	yes	4	0	2	1	0	2 ***	10.000 KES		N.A.	40		
		IPA:	no	18 **	2 **	16(6)[3] **	0	0	0	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
10	Company 7	*M&E records:*	no	4	0	1(0)[3]	0	0	0	N.A.	Thika	N.A.	N.A.		Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	no	N.A.	-	-	N.A.	-	-	-		251-500 million	N.A.		
		Company:	no	4	0	2	0	0	0	-		N.A.	N.A.		Reasons: Verified personally with company & all four different sources match
		IPA:	no	3	0	3 (2)[3]	0	0	0	-		-	-		
11	Company 19	*M&E records:*	no	4	0	0	0	0	0	N.A.	Thika	N.A.	N.A.	*confirmed that recruited to participate but company had no need for new interns under sustainability program because internships under GIZ stipend were prolonged	Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	yes, then no*	N.A.	-	-	N.A.	-	-	-		251-500 million	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-		Reasons: M&E records & IPA data match
		IPA:	no	5	0	5 (3)[3]	0	0	0	-		-	-		
12	Company 12	*M&E records:*	no	4	3	3 (1)[3]	0	0	0	N.A.	Ruiru	N.A.	N.A.	*confirmed that no interns under sustainability program were taken by the time of the field research	Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	yes, then no*	N.A.	-	-	N.A.	-	-	-		N.A.	2200		
		Company:	no	6	1	2	0	0	0	-		N.A.	N.A.		Reasons: Verified personally with company & all four different sources match
		IPA:	no	4	0	3 (2)[3]	0	0	0	-		-	-		
13	Company 20	*M&E records:*	N.A.	3	0	0	N.A.	N.A.	N.A.	N.A.	Thika	N.A.	N.A.	*confirmed that recruited to participate	Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	yes*	N.A.	-	-	N.A.	-	-	N.A.		N.A.	1400		
		Company:	-	-	-	-	-	-	-	-		-	-		Reasons: Verbal confirmation that company was recruited to participate but no records for actual placements available
		IPA:	no	4	0	3 (0)[3]	0	0	0	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
14	Company 21	*M&E records:*	N.A.	3	0	2 (0)[3]	N.A.	N.A.	N.A.	N.A.	Kikuyu	N.A.	N.A.	*confirmed that company was recruited to participate but change of management → process stopped.	Verifiable that recruited to participate but did not participate in funding switch Reasons: Verbally confirmed that company was recruited to participate but no records for actual placements available
		KAM/GIZ:	yes*	N.A.	-	-	N.A.	-	-	N.A.		20-50 million	12		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	no	3	0	3 (2)[3]	0	0	0	-		-	-		
15	Company 1	*M&E records:*	yes	6	0	1(1)[3]	1(4)[4]	N.A.	N.A.	N.A.	Thika	N.A.	N.A.		Verifiable that recruited to participate & that company subsequently participated in new funding scheme Reasons: Verified personally with company & all sources report that company participates
		KAM/GIZ:	yes	4	-	-	4	-	-	N.A.		101-250 million	305		
		Company:	yes	4	N.A.	1	5	N.A.	1	Lunch		N.A.	≈ 250		
		IPA:	yes	4	0	4 (1)[3]	2	0	1(1)[3]	-		-	-		
16	Company 5	*M&E records:*	N.A.	26 **	6 **	11(9)[3] **	N.A.	N.A.	N.A.	N.A.	Thika	N.A.	N.A.	*at the time of the research the company had not yet taken interns but requested some ** not possible to differentiate between interns placed in Nairobi & interns placed in Thika (Kiambu county)	Verifiable that recruited to participate but did not participate in funding switch Reason: Verified personally with company that no interns under the sustainability program were taken yet
		KAM/GIZ:	yes	N.A.	-	-	N.A.	-	-	N.A.		3-5 billion	77		
		Company:	no *	16	2	13	0	0	0	-		N.A.	N.A.		
		IPA:	yes	18**	1**	18(14)[3]**	1**	0	1(0)[3]**	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount	Location	Approx. revenue/	Employee numbers (total)	Comments	Conclusion
										pilot program	(Pilot region)	turnover			
				Total	Dropped	Job	Total	Dropped	Job			(in KES)			
17	**Company 13**	*M&E records:*	no	-	-	-	-	-	-	-	Thika	N.A.	N.A.	*company was interviewed because it is the company of the current CCC chair	Verifiable that company was not part of any funding scheme (neither GIZ stipend nor sustainability program) → Thus, not part of the pilot scheme Reasons: Verified personally with company & all sources report that company does not participate
		KAM/GIZ:	no	0	-	-	0	-	-	-		N.A.	N.A.		
		Company:	no	0	0	0	0	0	0	-		N.A.	>1000		
		IPA:	no	0	0	0	0	0	0	-		-	-		
18	**Company 22**	*M&E records:*	yes	4	0	0	3 (3)[4]	3	N.A.	N.A.	Thika	N.A.	N.A.	*confirmed that company took three interns under sustainability program but none of them finished because company did not pay them → Thus, did not follow premise of sustainability program	Verifiable that recruited to participate but did not participate in funding switch Reasons: verbal confirmation that company took three interns; However, all interns dropped because company did not adhere to premise of sustainability program. → Thus, not counted as participating
		KAM/GIZ:	yes*	3	-	-	3	3	0	no payment*		51-100 million	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	no	2	0	2(0)[3]	0	0	0	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount	Location	Approx. revenue/	Employee numbers (total)	Comments	Conclusion
										pilot program	(Pilot region)	turnover			
				Total	Dropped	Job	Total	Dropped	Job			(in KES)			
19	Company 23	*M&E records:*	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	Limuru	N.A.	N.A.	*confirmed that company was probably misplaced in original list provided by KAM/GIZ; one intern found employment independently there	Assumption that not recruited to participate & subsequently not participating in funding switch
		KAM/GIZ:	yes, then no*	N.A.	-	-	N.A.	-	-	N.A.		N.A.	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-	**intern found internship independently at that company	Reasons: reliable database is missing & verbal confirmation that company was wrongly reported as participating
		IPA:	no	1**	-	-	-	-	-	-		-	-		
20	Company 10	*M&E records:*	no	4	0	0	0	0	0	N.A.	Thika	N.A.	N.A.	*confirmed that company was recruited to participate but did not take up any interns	Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	no*	N.A.	-	-	N.A.	-	-	-		101-250 million	50		
		Company:	no	4	N.A.	1	0	0	0	-		N.A.	60		Reasons: Verified personally with company & all four different sources match
		IPA:	no	6	0	6 (0)[3]	0	0	0	-		-	-		
21	Company 6	*M&E records:*	no	19	1	9 (8)[3]	0	0	0	N.A.	Thika	N.A.	N.A.		Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	no	N.A.	-	-	N.A.	-	-	-		1-3 billion	1431		
		Company:	no*	≈ 10	≈ 1	02. Mrz	0	0	0	-		N.A.	>600		Reasons: Verified personally with company & all four different sources match
		IPA:	no	13	1	13 (9)[3]	0	0	0	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
22	Company 24	*M&E records:*	no	1	0	1(0)[3]	0	0	0	N.A.	Banana Hill	N.A.	N.A.	*confirmed that company was recruited to participate under the sustainability program	Verifiable that recruited to participate but did not participate in funding switch Reasons: M&E records & IPA data match
		KAM/GIZ:	no	N.A.	-	-	N.A.	-	-	-	Limuru	1-3 billion	205		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	no	1	0	1(0)[3]	0	0	0	-		-	-		
23	Company 25	*M&E records:*	N.A.	N.A.	N.A.	1(0)[3] **	N.A.	N.A.	N.A.	N.A.	Ruiru	N.A.	N.A.	*Not in original list that was provided by KAM/GIZ	Assumption that not recruited to participate & subsequently not participating in funding switch Reasons: reliable database is missing & IPA data only record an intern that found a placement independently there - without the help from KAM
		KAM/GIZ:	N.A.	-	-	-	-	-	-	N.A.		N.A.	N.A.	**assumption: intern found employment independently	
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	no	-	-	-	1 ***	-	-	-		-	-	***intern found internship independently at that company	
24	Company 26	*M&E records:*	N.A.	1	0	0	N.A.	N.A.	N.A.	N.A.	Ruiru	N.A.	N.A.	*confirmation through phone call that company was recruited to participate but that no intern under sustainability program was taken so far	Verifiable that recruited to participate but did not participate in funding switch Reasons: M&E records & IPA data match; verbal confirmation that company has not taken interns under the sustainability program
		KAM/GIZ:	yes, then no *	1	-	-	0 *	-	-	N.A.		N.A.	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	no	1	0	1(0)[3]	0	0	0	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
25	Company 27	*M&E records:*	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	Limuru	N.A.	N.A.	*Confirmed this is a training institute, and not a company. Thus, it is probably misplaced in original list provided by GIZ/KAM; might be that one intern found employment independently there	Assumption that not recruited to participate & subsequently not participating in funding switch
		KAM/GIZ:	yes, then no*	N.A.	-	-	N.A.	-	-	-		N.A.	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-	** intern found internship independently there	Reasons: verbal confirmation that this is not even a company & that it was wrongly reported as recruited to participate
		IPA:	no	1**	-	-	-	-	-	-		-	-		
26	Company 18	*M&E records:*	(yes)[5]	N.A.	N.A.	N.A.	(1)[4] **	N.A.	N.A.	N.A.	Ruiru	N.A.	N.A.	*Not in original list provided by KAM/GIZ but listed in M&E records	Assumption that not recruited to participate & subsequently not participating in funding switch
		KAM/GIZ:	N.A. then yes*	-	-	-	-	-	-	N.A.		N.A.	N.A.	**assumption: intern found employment independently but was counted as placed under sustainability program	Reason: reliable database is missing
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	N.A.	-	-	-	-	-	-	-		-	-		

Table A2.5.2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
27	Company 8	*M&E records:*	yes	2	1	1 (1)[3]	3 (3)[4]	0	N.A.	N.A.	Thika	N.A.	N.A.	*Company was not aware of introduction of sustainability program	Verifiable that recruited to participate & that company subsequently participated in new funding scheme Reasons: Verified personally with company & all sources report that company participates
		KAM/GIZ:	yes	3	-		3	-	-	3.000 KES		501million - 1 billion	550		
		Company:	yes*	5	2	0	5	3	1	1000 KES		N.A.	700		
		IPA:	yes	0	0	0	2	0	1 (1)[3]	-		-	-		
28	Company 29	*M&E records:*	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	Banana Hill	N.A.	N.A.		Assumption that not recruited to participate & subsequently not participating in funding switch Reason: reliable database is missing
		KAM/GIZ:	no	N.A.	-	-	N.A.	-	-	-	Limuru	501million - 1 billion	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-		
		IPA:	no	1	0	1 (1)[3]	0	0	0	-		-	-		
29	Company 9	*M&E records:*	yes	11	4	5 (4)[3]	4 (5)[4]	0	N.A.	N.A.	Kikuyu	N.A.	N.A.		Verifiable that recruited to participate & that company subsequently participated in new funding scheme Reasons: Verified personally with company & all sources report that company participates
		KAM/GIZ:	no, then yes	5	-	-	5	-	-	N.A.		N.A.	N.A.		
		Company:	yes	16	≈ 5	5	4	N.A.	0	3.000-3.600 KES		N.A.	≈ 425		
		IPA:	yes	10	2	10 (6)[3]	3	0	0	-		-	-		

Table A2.5.2 2 continued

	Company name	Sources	Participates in funding switch	Intern numbers under GIZ stipend			Intern numbers under sustainability program			Stipend amount pilot program	Location (Pilot region)	Approx. revenue/ turnover (in KES)	Employee numbers (total)	Comments	Conclusion
				Total	Dropped	Job	Total	Dropped	Job						
30	Company 30	*M&E records:*	N.A.	N.A.	N.A.	N.A.	(0)[4]	N.A.	N.A.	N.A.	Thika	N.A.	N.A.	*Not in original list provided by KAM/GIZ but listed in M&E records	Verifiable that recruited to participate but did not participate in funding switch
		KAM/GIZ:	N.A. then no**	N.A.	-	-	N.A.	-	-	N.A.		N.A.	N.A.		
		Company:	-	-	-	-	-	-	-	-		-	-	**Confirmed that recruited to participate but currently not willing to take interns; company might be willing to take interns the future	Reasons: based on verbal KAM records
		IPA:	N.A.	-	-	-	-	-	-	-		-	-		
31	Company 31	*M&E records:*	N.A.	6	0	3 (0)[3]	N.A.	N.A.	N.A.	N.A.	Thika	N.A:	N.A.	*Argued by KAM that company was recruited to participate and is participating	Verifiable that recruited to participate but did not participate in funding switch Reasons: verbal KAM records insist that company is participating but no M&E records about actual placements exist; Thus, participation wrongly reported by KAM

Source: RWI. - Legend: See Table H for a legend of this table, including descriptions of rows, columns and color coding.

A2.6 Zooming in on companies' main reasons for (non-)participation in pilot of funding switch

As discussed in section 1.8.4., the main reasons for companies' (non-)participation in the sustainability program are mostly related either to their attitude towards having to pay KAM interns or to their (dis-)satisfaction with former KAM interns. Most often it is a combination of both factors, thus, by just looking at the willingness to pay or the satisfaction alone, it cannot be determined whether the companies participate in the sustainability program (i.e., the funding switch).

Therefore, Table A2.6.1 visualizes for each interviewed company whether it participated in the switch to the sustainability program and how this relates to the two categories (willingness to pay and satisfaction with KAM interns) introduced in section 1.8.4..The table is divided into quadrants. The upper half of the table contains the companies that perceive the switch positively, either because they are generally in favor of funding KAM interns or because they are generally satisfied with KAM interns or both. The lower half of the table presents the companies that perceive the switch negatively, either because they are opposed to funding KAM interns or because they are generally dissatisfied with their interns or both. On the left-hand side are those companies that participate in the pilot of the funding switch and on the right-hand side are those that do not. Each side is again divided to display whether the company already had a policy of paying *other* interns/attachés at the company before KAM and GIZ introduced the switch.

Table A2.6.1
Overview of companies' main reasons to (not) participate in the new payment scheme

		Participates in new funding scheme (has taken interns)		Does not participate in new funding scheme	
Reasons for (non-) participation in the funding switch		(A) Pays other interns/attachés outside KAM	(B) Does not pay other interns/attachés outside KAM	(C) Pays other interns/attachés outside KAM	(D) Does not pay other interns/attachés outside KAM
Positive perception of switch	(1) In favor of paying KAM interns	Company 9		Company 10	
	(2) Satisfied with KAM interns				
	(3) In favor of paying KAM interns & satisfied with KAM interns		Company 2 Company 4		
Negative perception of switch	(4) Opposed to paying KAM interns	Company 1 Company 8	Company 3	Company 5	Company 11
	(5) Dis-satisfied with KAM interns			Company 12	
	(6) Opposed to paying KAM interns & dis-satisfied with KAM interns			Company 6	Company 7

Legend: red ≙ companies with a strong negative perception of (KAM) interns; yellow ≙ companies with mixed perceptions of (KAM) interns; green ≙ companies with high satisfaction of (KAM) interns. - Source: RWI.

There are several observations that can be deduced from Table J. For example, eight of the 12 interviewed companies perceive the funding switch to the sustainability program rather negatively, in most cases because they are opposed to investing their own resources to pay KAM interns (cf. [5]; [14]; [8]; [10]; [18]; [12]; [13]).

Only the three companies in the *upper left quadrant* participate in the sustainability program because they support its premise that interns should be paid. Since they are in favor of paying interns, they perceive the funding switch and, thus, the implementation of the sustainability program, where the companies have to pay, generally positively (cf. [6]; [7]; [9]; [16]).

The other three companies in the *lower left quadrant* are also participating under the sustainability program because they have taken KAM interns after March 2019, but they are all generally rather opposed to having to invest their own resources to pay the KAM interns, even though none of them had distinct negative experiences with them (cf. [5]; [14]; [8]).

The company in the upper half of cell A4 gives their interns (and attachés) no monetary stipend (just provides lunch) and is, therefore, logically opposed to paying stipends to the KAM interns (cf. [5]). The other company in the lower half of cell A4 unknowingly participates under the sustainability program because it was never informed about the funding switch but just kept on taking interns from KAM. That company gives any intern or attaché an unsubstantial stipend, which it finances through reimbursements it receives from the National Industrial Training Authority (NITA)[66]. Therefore, it is opposed to having to invest any extra company-internal resources into paying the KAM interns[67] (cf. [14]). The company in cell B4 in the lower left quadrant of Table A2.6.1 differs from the first two companies because it started to give a substantial stipend to KAM interns after the implementation of the funding switch. However, the company substitutes casual workers with KAM interns to give them their salary:

> *"the KAM group […] comes and fill[s] in the gaps that we want to fill, as opposed to hiring temporary workers. We put them into those positions and therefore we pay them those stipends that we'd pay a temporary worker in that position. [...] And then we don't bring on board the temporary staff, that we might have wanted to bring in to bridge the gaps in certain areas."* ([8], pos. 99).

While this ensures that the KAM interns receive a decent stipend, it also means that the company is not willing to invest any additional company-internal resources. Instead, they reallocate the resources planned for the casual workers to KAM interns. This practice exerts negative employment effects for the casual workers since they are substituted by KAM interns.

Two further observations can be made by looking at the six participating companies on *the left-hand side* of Table A2.6.1. First, the participation of the three companies in the *lower left quadrant* goes against the original premise of the sustainability program (i.e., the funding switch). The premise was that only those companies willing to invest their own resources to pay the KAM interns were supposed to receive further interns via the KAM TVET program (cf. section 1.2.). However, these three companies are either not paying the KAM interns a monetary stipend or they do not invest their own resources to finance the stipend. Yet, they still received KAM interns

[66] *NITA reimbursements ≙ the National Industrial Training Authority (NITA) collects a small fee (around 50 KES) for each employee from manufacturing companies for 'training purposes'. If the companies register their interns/attachés with NITA, they get 3.000 KES (approx. 25€) per month per intern reimbursed (cf. [2])*

[67] *Since that particular company gives only a small amount of the reimbursements they receive from NITA to interns, they even make profit with interns, instead of investing company-internal resources into paying them.*

after the introduction of the sustainability program because the original premise of the sustainability program was not consistently implemented.

The second observation can be deducted from *column B:* The introduction of the funding switch has only led in three instances to a 'change of heart', in the sense that companies that had no stipend policy prior to its implementation, started to actually pay a stipend because of KAM's and GIZ's 'switch request' (Table A.2.6.1 column B; cf. [6]; [9]; [8]). The other three participating companies in *column A* already had a stipend policy in place (in one case just lunch provision) for other interns/attachés at their companies (cf. [5]; [14]; [16])[68].

On the contrary, by looking at the six *non-participating* companies on the *right-hand side* of Table A2.6.1, *column C* shows that four of these six companies still have a policy of paying other interns/attachés at their companies[69] (cf. [17]; [10]; [19]; [12]). Thus, generally opposing payment of interns or attachés is not the reason why these four companies are not participating under the sustainability program. However, the three companies in the cells C4, C5 and C6 were particularly opposed to paying the *KAM* interns because they were:

a) either reluctant to invest their own resources when there is a chance that GIZ continues to pay the stipends:

   "We were appreciating what KAM was doing. When they were doing the stipend, we were sure, we are not wasting our resources, you know, every time you feel like you're wasting a resource, it doesn't feel good." ([12], pos. 201; cf. also [10])

b) or had particularly negative experiences with former KAM interns (cf. [19]).

Only the non-participating company in the *upper right quadrant* argued without restrictions in favor of paying KAM interns. Here, company-internal concerns[70] were the reason why they did not take any KAM interns after the implementation of the funding switch in 2019 (cf. [17]).

The remaining two companies in *column D* that do also not participate in the pilot scheme are generally opposed to paying interns or attachés. This is reflected in the fact that they both do not pay any other intern or attaché at their company (cf. Table A2.6.1; cells D4 & D6). The difference between the two is that the company in cell D4 was very satisfied with former KAM interns, while the company in cell D6 was not ([18]; [13]).

Finally, as already mentioned, participation patterns in the funding switch cannot be determined solely by looking at the companies' satisfaction with KAM interns. Among the five companies that are generally very satisfied with KAM interns (marked in green), three are participating in the pilot program while two are not (cf. Table A2.6.1, cells A4, B3 vs. C4, D4). A similar picture emerges for the four companies with mixed experiences or perceptions about KAM interns (marked in yellow) (cf. Table A2.6.1, cells A1, A4, B4 vs. C1). However, neither of the three companies with the strongest negative perceptions about KAM interns (marked in red) participates in the pilot (cf. Table A2.6.1, cells C5, C6, D6). Thus, negative perceptions of KAM interns seemed to generally outweigh positive or mixed ones.

---

[68] *These three companies pay the KAM interns very minimal stipends (ranging between lunch provision to 3.000 KES (approx. 25€), they already had in place prior to the introduction of the funding switch (cf. [5]; [14]; [16]). In contrast, the companies that introduced a stipend at KAM's request pay substantial stipends ranging between 10.000 to 12.000 KES per month (approx. 80€ - 90€) (cf. [6]; [9]; [8]).*

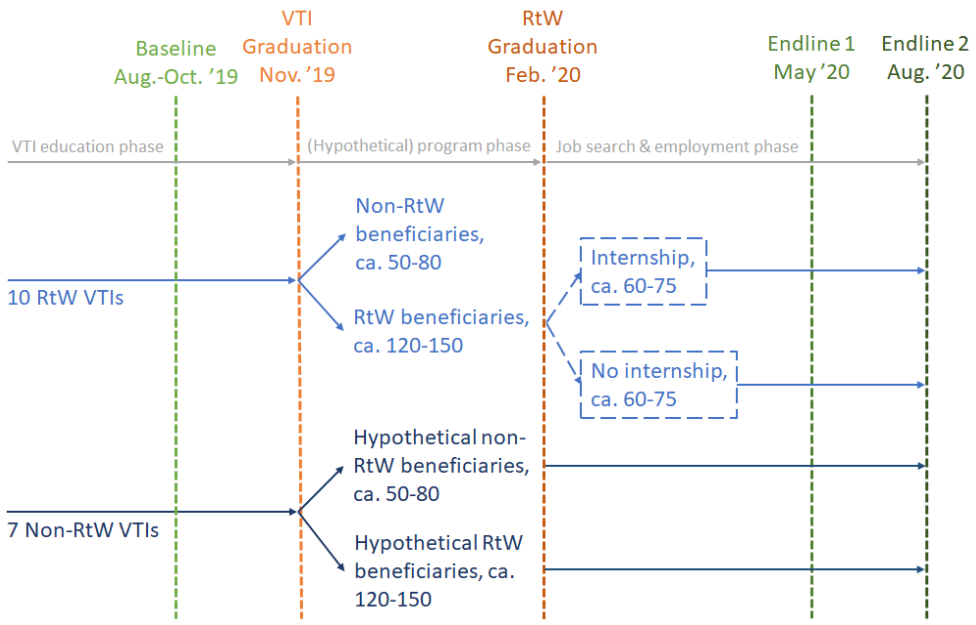
[69] *All four even pay substantial stipends between 8.000 to 15.000 KES (approx. 60€ - 110€) (cf. [17]; [10]; [19]; [12]).*

[70] *The company had negative experiences with industrial attachés in 2018, which led them to hire much less interns or attachés after that incident ([17]).*

A3 Appendix of the quantitative RtW evaluation

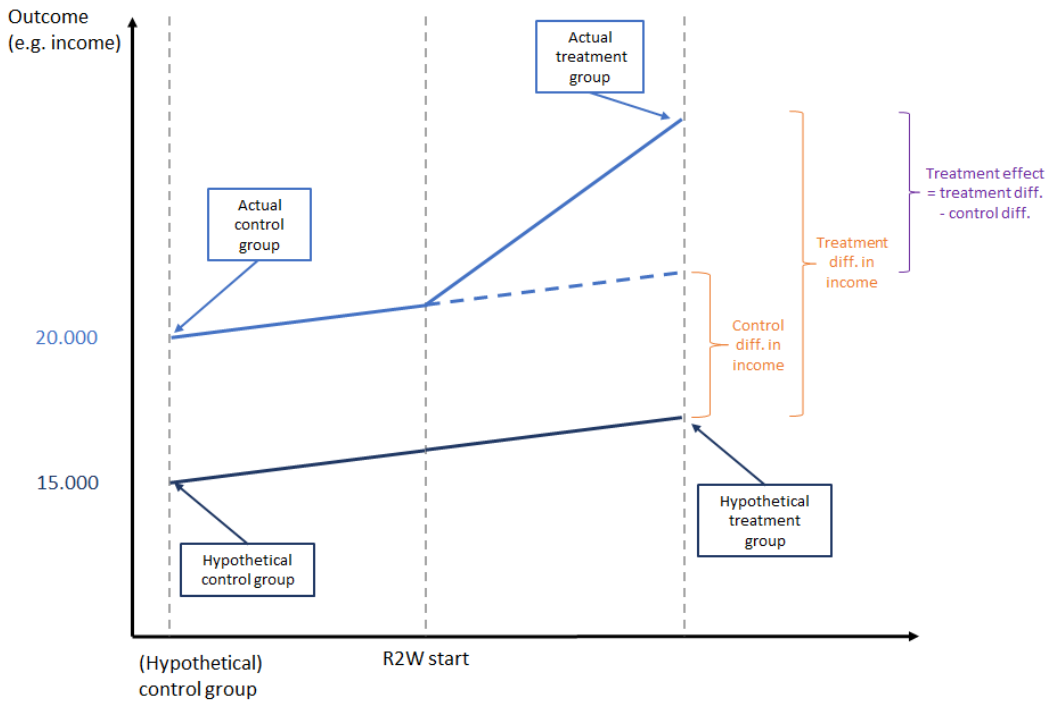
A3.1 Original RtW cohort differences-in-differences evaluation design

Figure A3.1.1
Set-up and timeline of the RtW cohort difference-in-differences design



Source: Own illustration.

Figure A3.1.2
Estimation strategy of the cohort difference-in-differences design



Source: Own illustration.

A3.2 Results from Double Machine Learning estimations

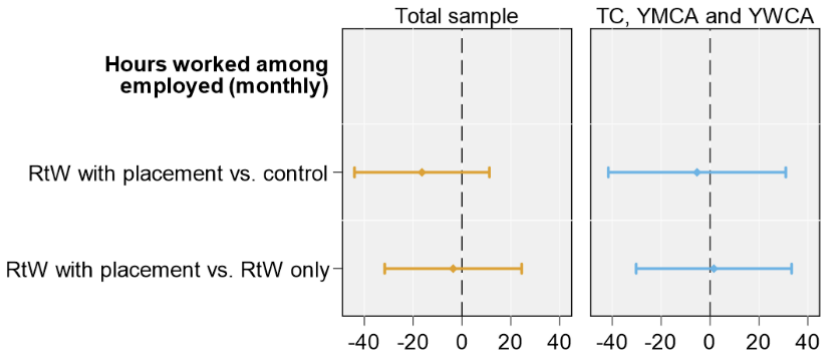
Figure A3.2.1
The effect of the RtW program on binary employment indicators using Double Machine Learning estimations

Source: Own calculations based on RtW Survey.

Figure A3.2.2
The effect of the RtW program on income and wages using Double Machine Learning estimations

Source: Own calculations based on RtW Survey.

Figure A3.2.3
The effect of the RtW program on working hours using Double Machine Learning estimations



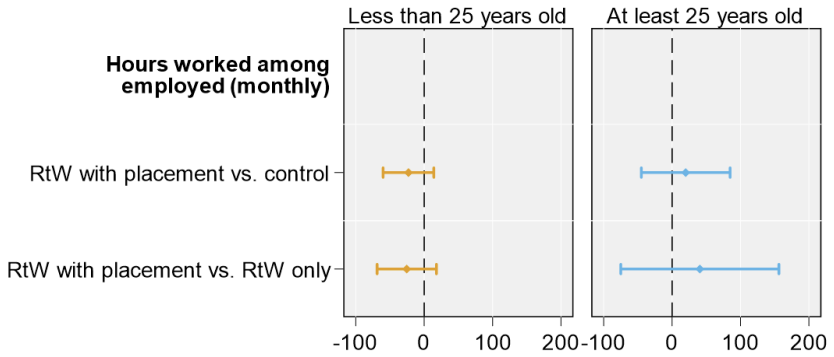
Source: Own calculations based on RtW Survey.

A3.3 Subsample estimations on income, wages, and working hours

Figure A3.3.1
The effect of the RtW program on income and wages by respondents' age

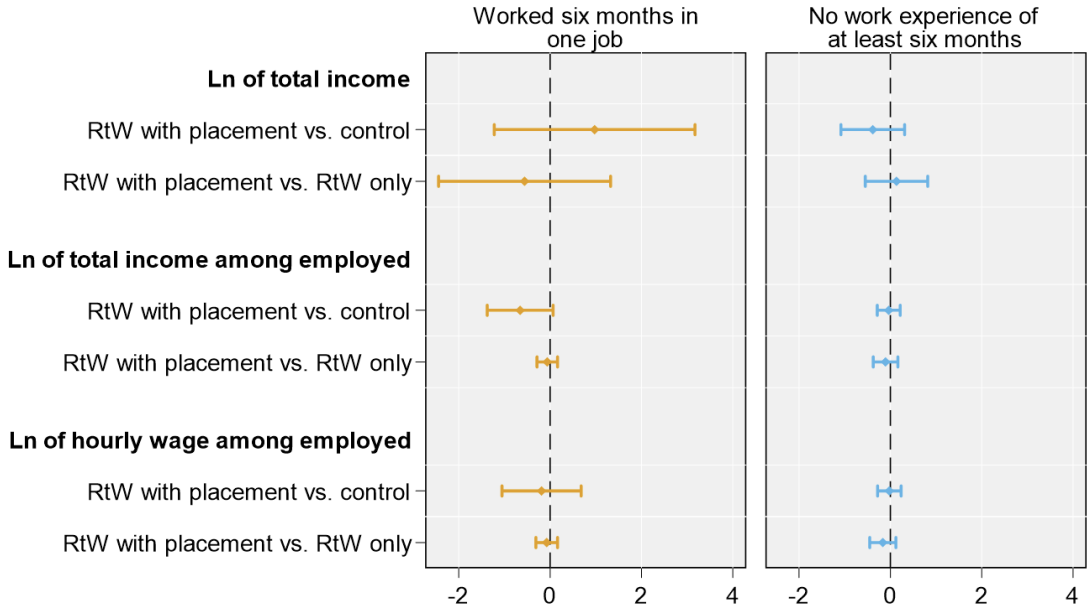
Source: Own calculations based on RtW Survey.

Figure A3.3.2
The effect of the RtW program on hours by respondents' age



Source: Own calculations based on RtW Survey.

Figure A3.3.3
The effect of the RtW program on income and wages by respondents' sustained work experience

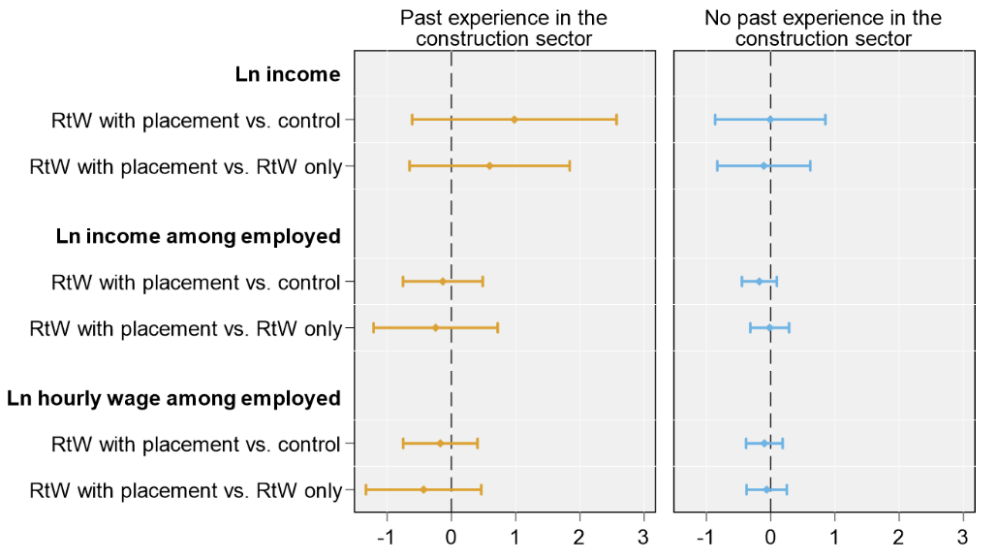


Source: Own calculations based on RtW Survey.

Figure A3.3.4
The effect of the RtW program on hours by respondents' sustained work experience

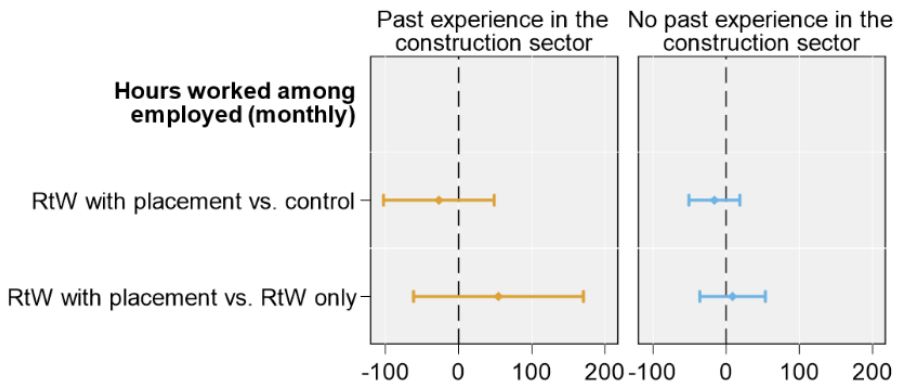
Source: Own calculations based on RtW Survey.

Figure A3.3.5
The effect of the RtW program on income and wages by respondents' construction sector experience



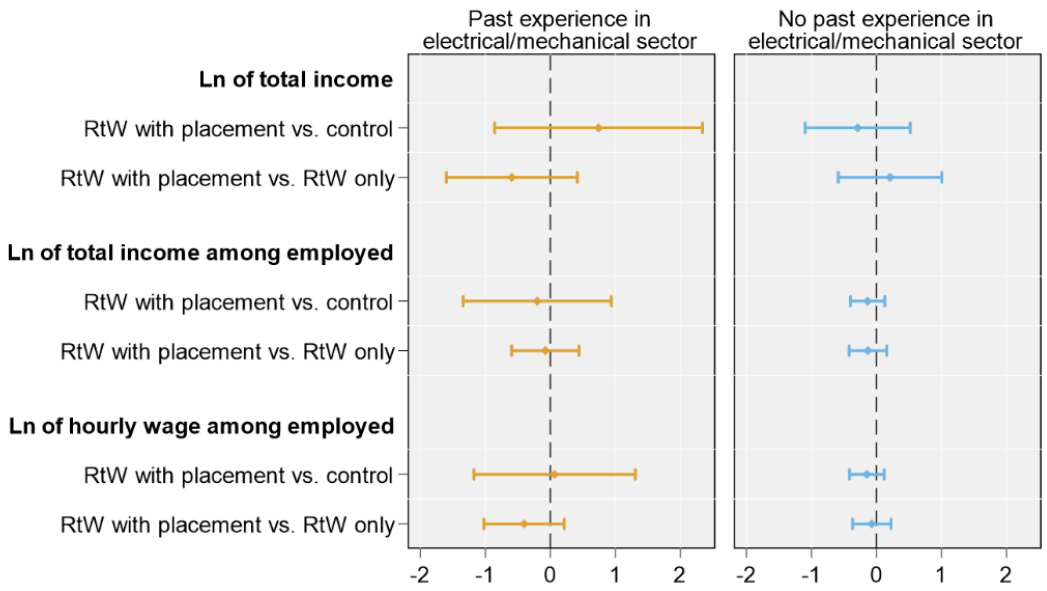
Source: Own calculations based on RtW Survey.

Figure A3.3.6
The effect of the RtW program on hours by respondents' construction sector experience



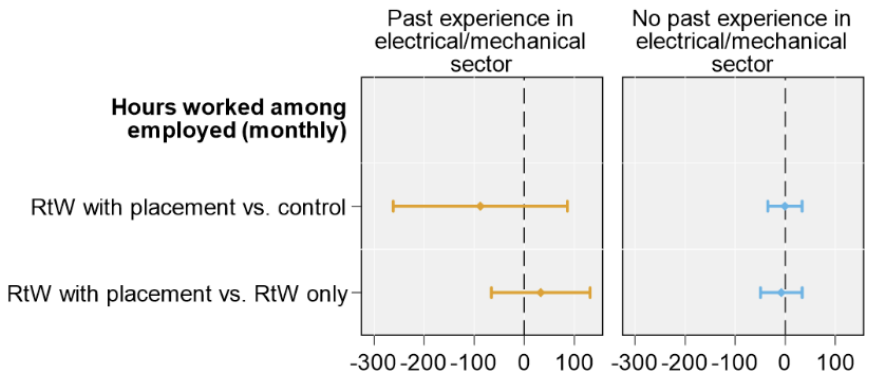
Source: Own calculations based on RtW Survey.

Figure A3.3.7
The effect of the RtW program on income and wages by respondents' experience in electrical or mechanical work



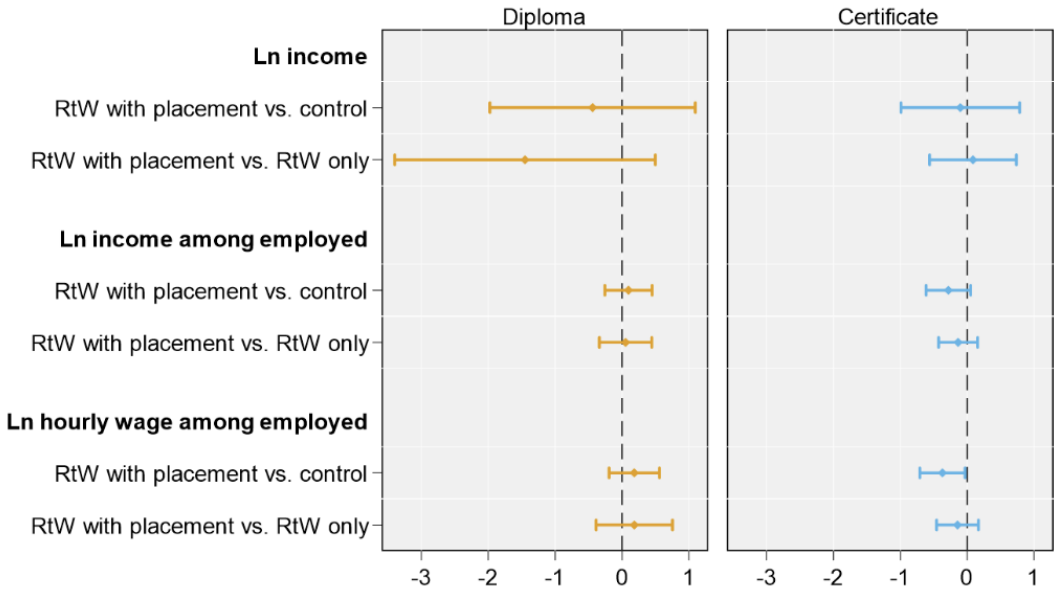
Source: Own calculations based on RtW Survey.

Figure A3.3.8
The effect of the RtW program on hours by respondents' experience in electrical or mechanical work



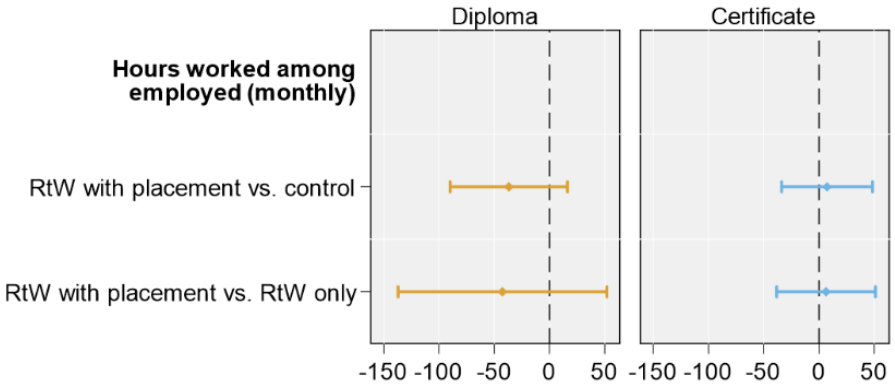
Source: Own calculations based on RtW Survey.

Figure A3.3.9
The effect of the RtW program on income and wages by respondents' vocational training level



Source: Own calculations based on RtW Survey.

Figure A3.3.10
The effect of the RtW program on hours by respondents' s vocational training level

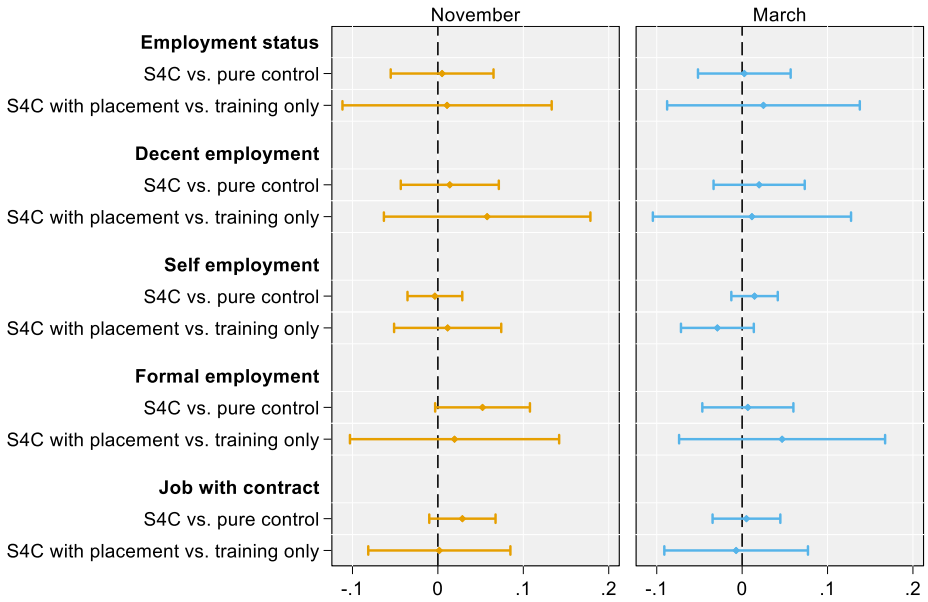


Source: Own calculations based on RtW Survey.

A4 Appendix of the quantitative S4C evaluation

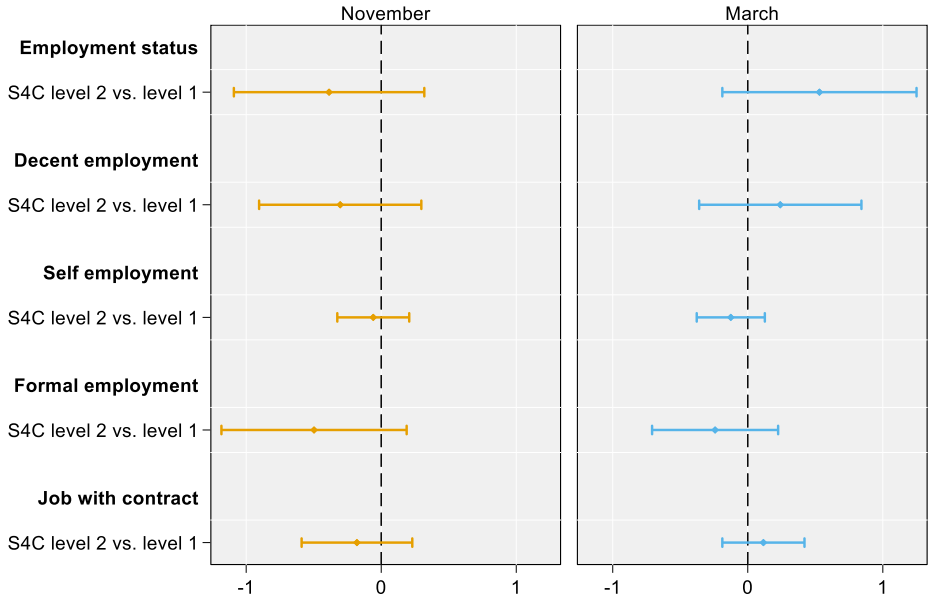
A4.1 Program impact for subsample of respondents who completed training

Figure A4.1.1
Estimated treatment effects of S4C program and placement component on binary outcomes (alternative treatment definition)



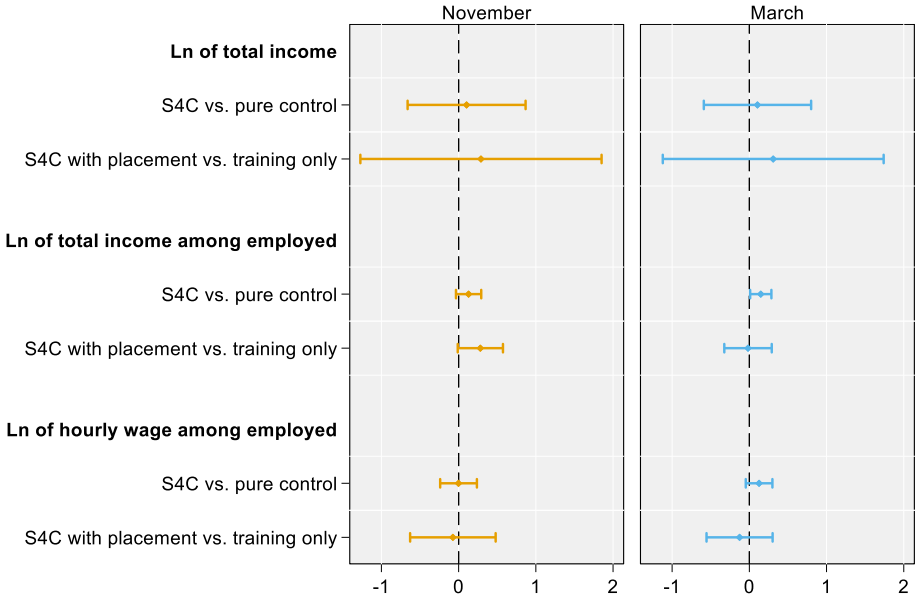
Source: Own calculations based on S4C survey.

Figure A4.1.2
Estimated treatment effects of S4C level 2 training on binary outcomes (alternative treatment definition)



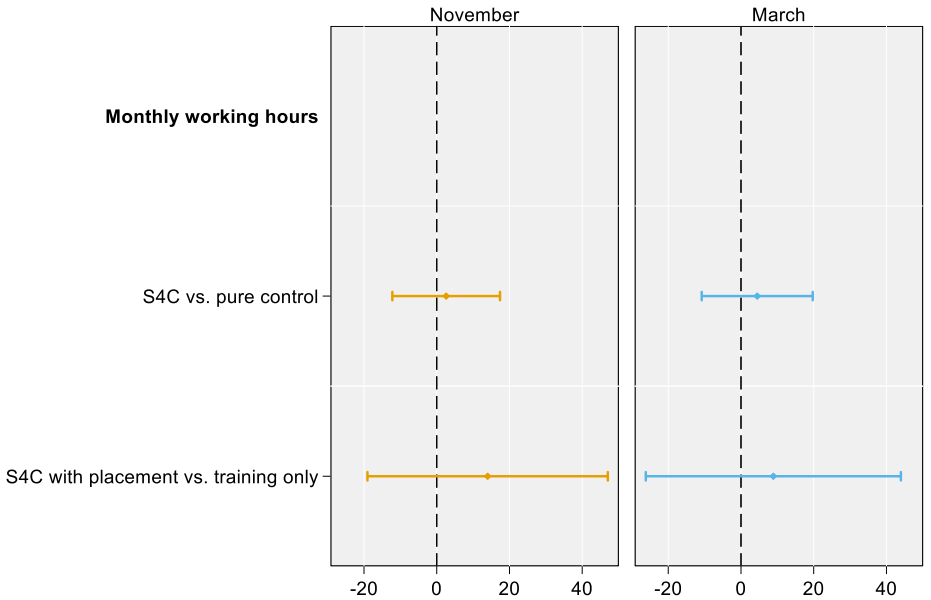
Source: Own calculations based on S4C survey.

Figure A4.1.3
Estimated treatment effects of S4C program and placement component on incomes and wages (alternative treatment definition)



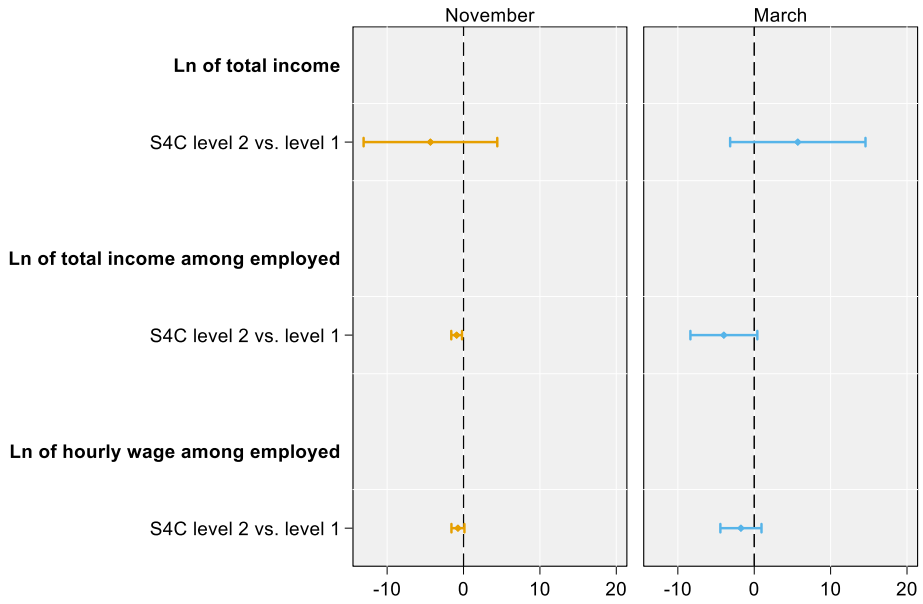
Source: Own calculations based on S4C survey.

Figure A4.1.4
Estimated treatment effects of S4C program and placement component on working hours (alternative treatment definition)



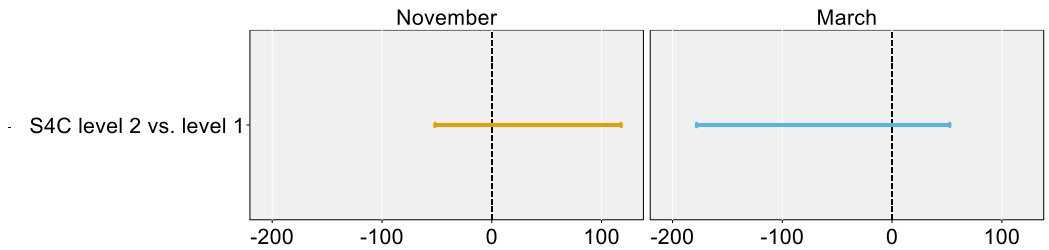
Source: Own calculations based on S4C survey.

Figure A4.1.5
Estimated treatment effects of S4C level 2 training on incomes and wages (alternative treatment definition)



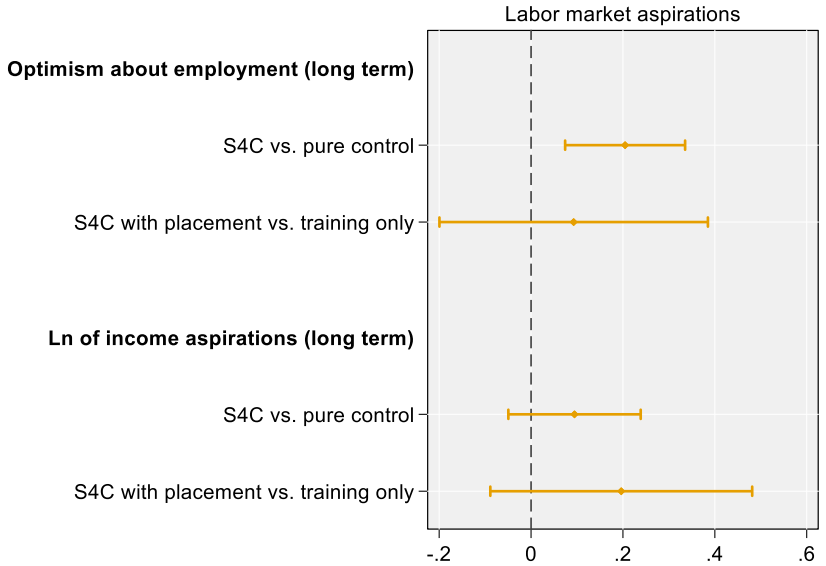
Source: Own calculations based on S4C survey.

Figure A4.1.6
Estimated treatment effects of S4C level 2 training on working hours (alternative treatment definition)



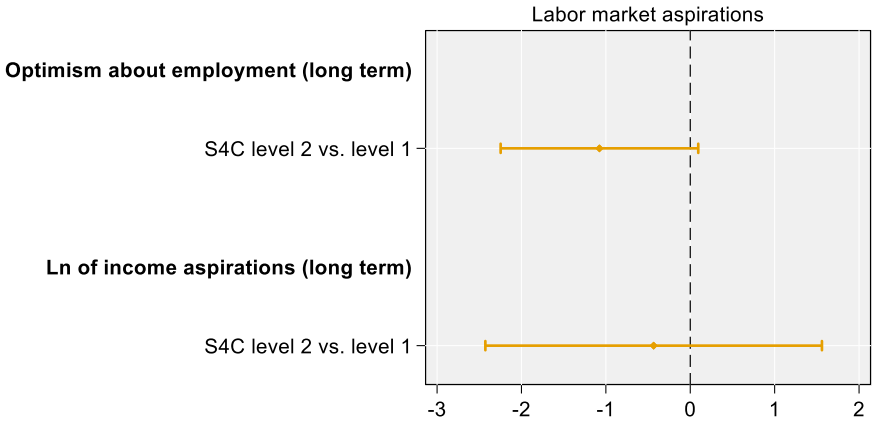
Source: Own calculations based on S4C survey.

Figure A4.1.7
Estimated treatment effects of S4C program and placement component on labor market aspirations (alternative treatment definition)



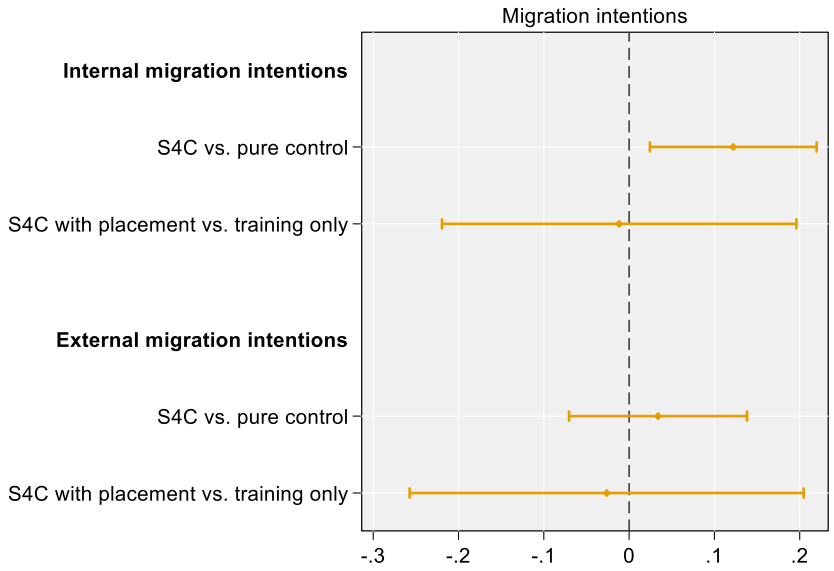
Source: Own calculations based on S4C survey.

Figure A4.1.8
Estimated treatment effects of S4C level 2 training on labor market aspirations (alternative treatment definition)



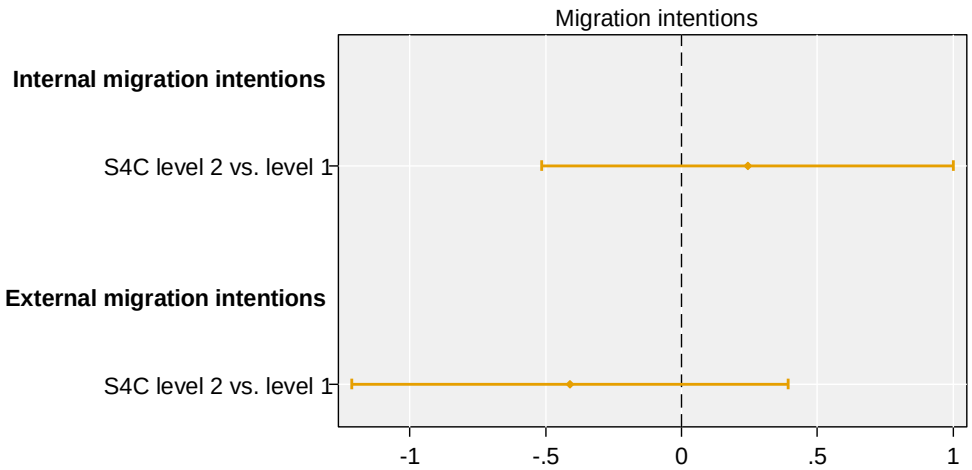
Source: Own calculations based on S4C survey.

Figure A4.1.9
Estimated treatment effects of S4C program and placement component on migration intentions (alternative treatment definition)



Source: Own calculations based on S4C survey.

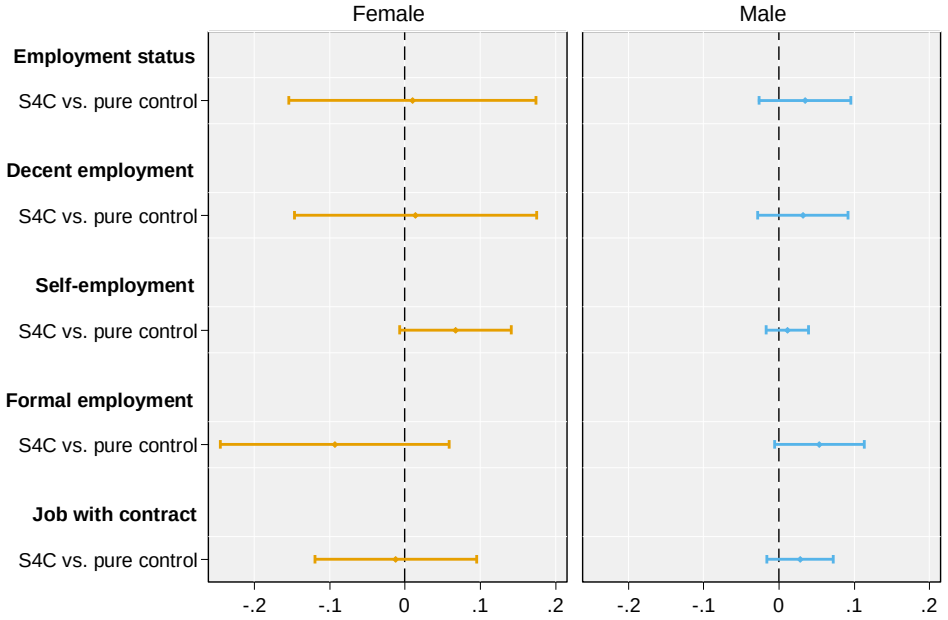
Figure A4.1.10
Estimated treatment effects of S4C level 2 training on migration intentions (alternative treatment definition)



Source: Own calculations based on S4C survey.

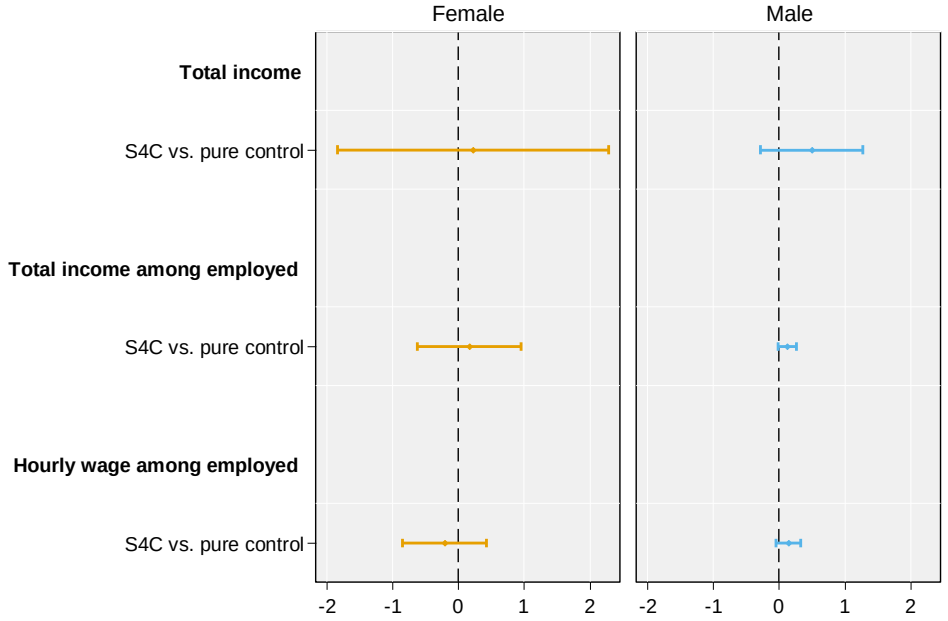
A4.2 Program heterogeneity with respect to endline in March 2020

Figure A4.2.1
Program effect heterogeneity on binary outcomes by respondents' gender in March



Source: Own calculations based on S4C survey.

Figure A4.2.2
Program effect heterogeneity on incomes and wages by respondents' gender in March

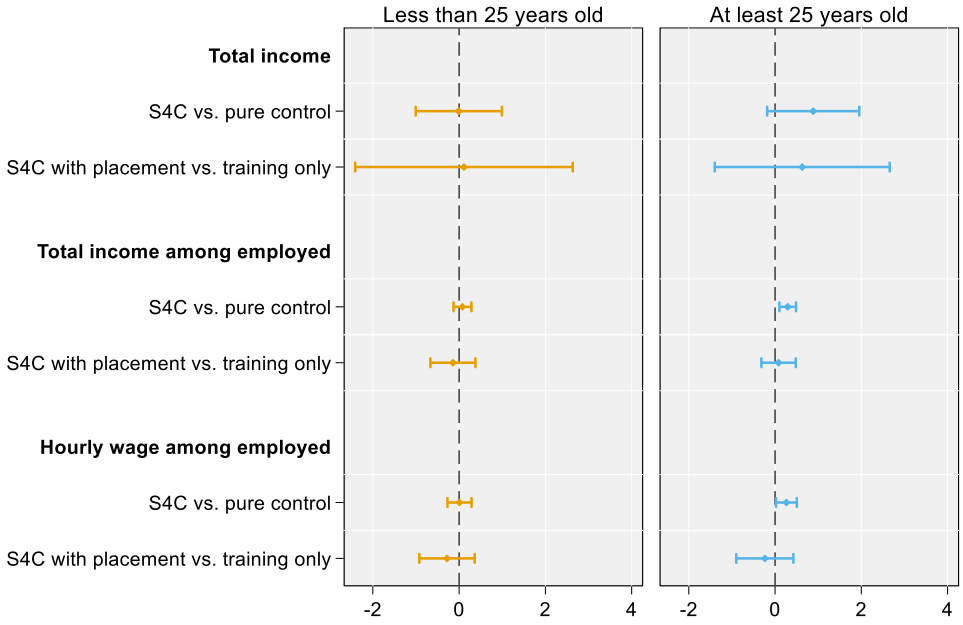


Source: Own calculations based on S4C survey.

Figure A4.2.3
Program effect heterogeneity on binary outcomes by respondents' age in March

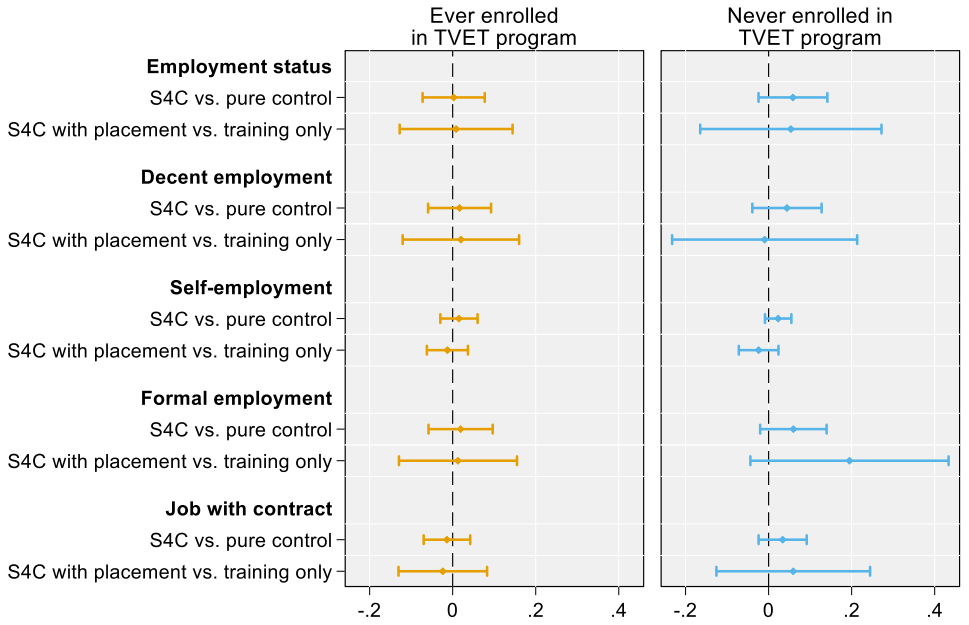
Source: Own calculations based on S4C survey.

Figure A4.2.4
Program effect heterogeneity on incomes and wages by respondents' age in March



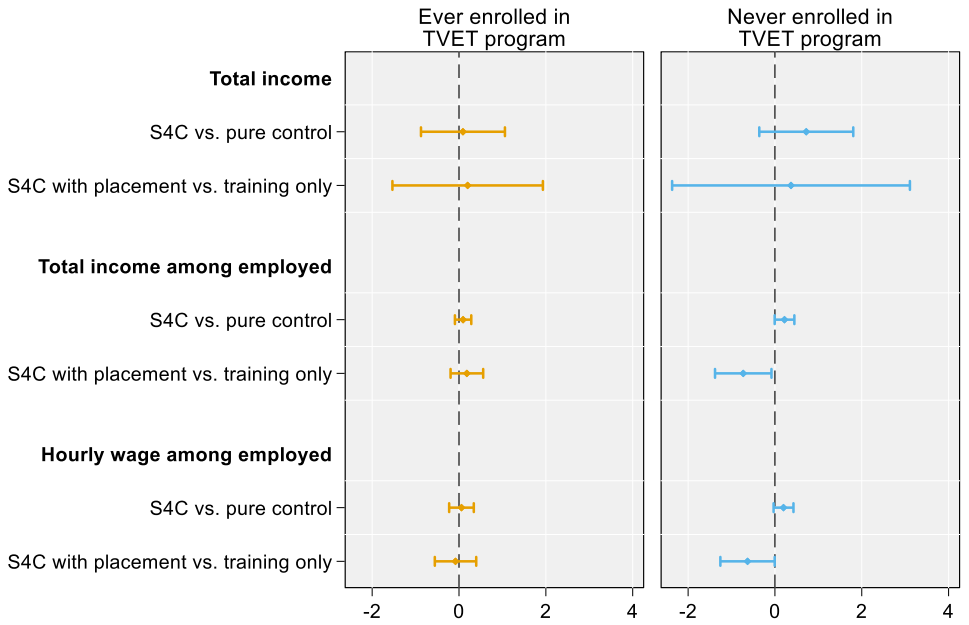
Source: Own calculations based on S4C survey.

Figure A4.2.5
Program effect heterogeneity on binary outcomes by respondents' vocational training experience in March



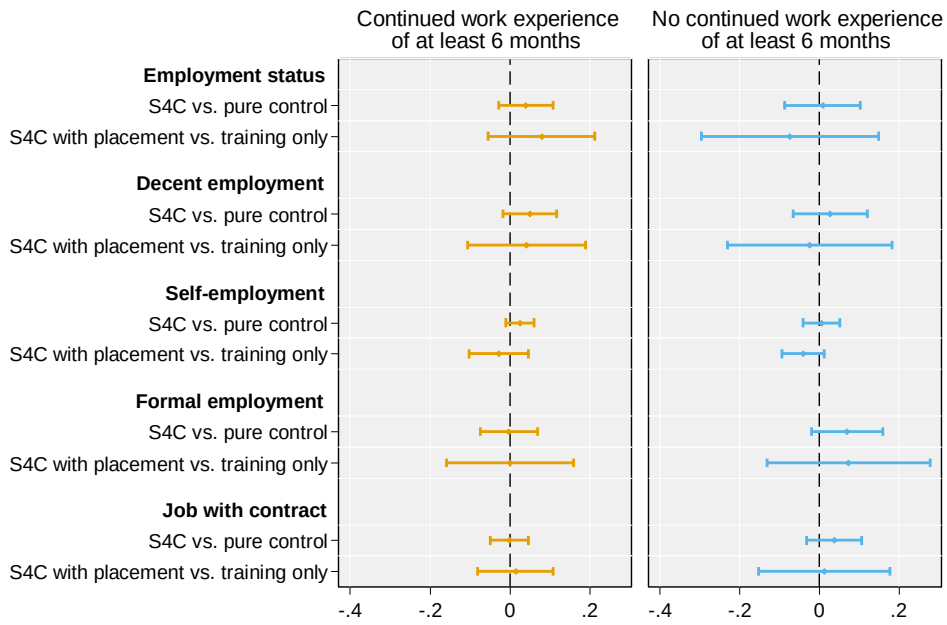
Source: Own calculations based on S4C survey.

Figure A4.2.6
Program effect heterogeneity on incomes and wages by respondents' vocational training experience in March



Source: Own calculations based on S4C survey.

Figure A4.2.7
Program effect heterogeneity on binary outcomes by respondents' work experience in March

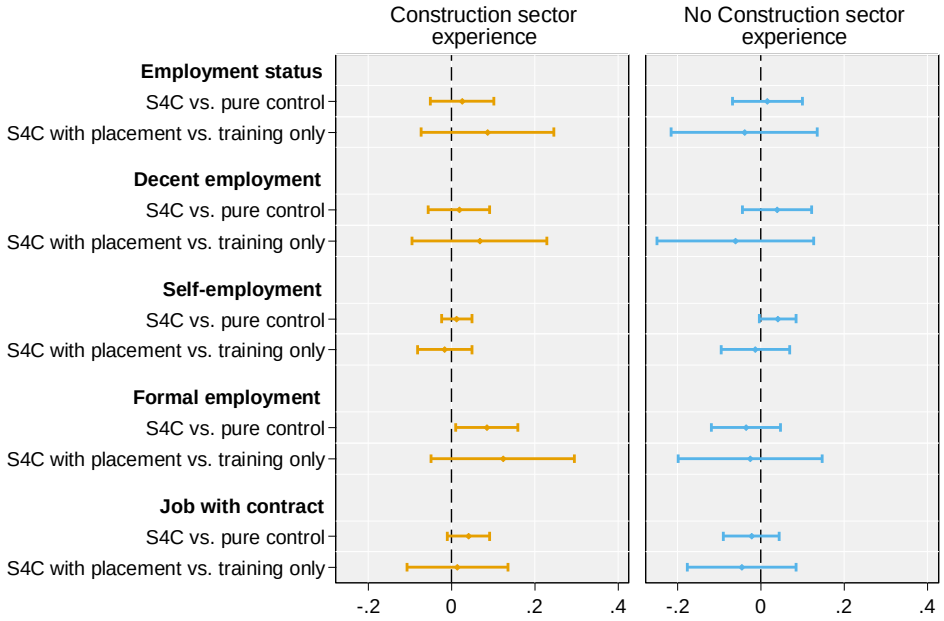


Source: Own calculations based on S4C survey.

Figure A4.2.8
Program effect heterogeneity on incomes and wages by respondents' work experience in March

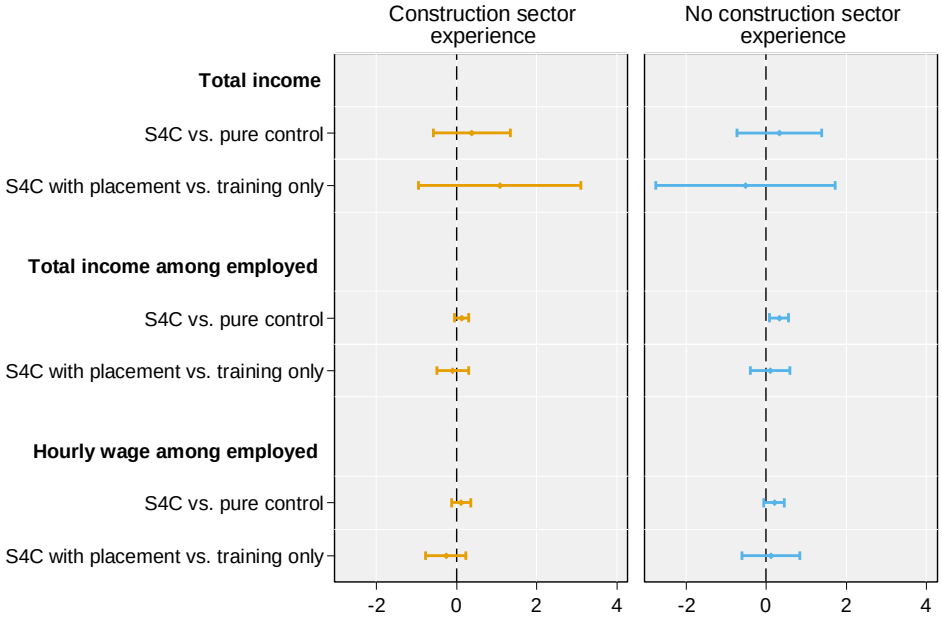
Source: Own calculations based on S4C survey.

Figure A4.2.9
Program effect heterogeneity on binary outcomes by respondents' construction sector experience in March



Source: Own calculations based on S4C survey.

Figure A4.2.10
Program effect heterogeneity on incomes and wages by respondents' construction sector experience in March



Source: Own calculations based on S4C survey.

RWI Leibniz Institute for Economic Research

The RWI is funded by the Federal Government
and the State of North Rhine-Westphalia.