

Arni, Patrick

Working Paper

Opening the Blackbox: How Does Labor Market Policy Affect the Job Seekers' Behavior? A Field Experiment

IZA Discussion Papers, No. 9617

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Arni, Patrick (2015) : Opening the Blackbox: How Does Labor Market Policy Affect the Job Seekers' Behavior? A Field Experiment, IZA Discussion Papers, No. 9617, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/130294>

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.

IZA DP No. 9617

**Opening the Blackbox:
How Does Labor Market Policy Affect the
Job Seekers' Behavior? A Field Experiment**

Patrick Arni

December 2015

Opening the Blackbox: How Does Labor Market Policy Affect the Job Seekers' Behavior? A Field Experiment

Patrick Arni

*IZA, DEEP/University of Lausanne
and CAFÉ/Aarhus University*

Discussion Paper No. 9617
December 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Opening the Blackbox: How Does Labor Market Policy Affect the Job Seekers' Behavior? A Field Experiment*

Empirically, not much is known about the mechanisms *how* labor market programs like job search assistance and training operate to support finding a job. This paper provides novel evidence to open the “blackbox”: it causally links the program interventions to the dynamics of search behavior, beliefs and non-cognitive skills. The study is based on a unique combination of a randomized field experiment with detailed register data and a panel of repeated surveys. The tested coaching program, focused on older job seekers, turns out to increase the job finding of participants by 9 percentage points (72% vs. 63% in the control group). The treatment effect is driven by a reduction of reservation wages and an increase in search efficiency. Moreover, I find short-run effects on motivation, self-confidence and beliefs. The job seekers overestimate their chances slightly less with respect to job interviews and salaries. Overall, the focus on realistic expectations and on search strategy appears to be important for program success. The experiment shows that evaluation designs which directly assess behavior can provide a fruitful base for targeted policy design.

JEL Classification: J64, J65, J68, J14

Keywords: field experiment, labor market policy, older workers, reservation wages, job search behavior, non-cognitive skills, biased beliefs, dynamic treatment effects, unemployment insurance

Corresponding author:

Patrick Arni
IZA
P.O. Box 7240
53072 Bonn
Germany
E-mail: arni@iza.org

* I would like to thank Rafael Lalive, Jan van Ours, Lorenz Goette, Michael Rosholm, Gerard van den Berg, Peter Kuhn, Josef Zweimüller, Jaap Abbring, Bas van der Klaauw, Geert Ridder, Arthur van Soest, Marco Caliendo, Michael Kosfeld, Guido Friebel, Brian Krogh Graversen, Konstantinos Tatsiramos, Matteo Picchio and Olivier Deschênes for their valuable comments. Moreover, I thank the seminar participants at UC Santa Barbara, Cornell University, Tilburg University, Tinbergen Institute Amsterdam, Universities of Lausanne, Zurich and Duisburg-Essen, Goethe-University Frankfurt, Aarhus University, and conference participants at SOLE, EALE, EEA, ESPE and the IZA/IFAU Policy Evaluation Conference for their valuable comments. I am grateful to Raphael Weisz and the Office of Economic Affairs and Labour (AWA) of Canton of Aargau, Switzerland, for providing the data and further support. I would like to thank as well the project teams in the regional PES. A part of the work for this paper I carried out during my research stay at Tilburg University. Financial support from the Swiss National Science Foundation (project no. PBLAP1-127652/1) is gratefully acknowledged.

1 Introduction

Most OECD member states spend substantial amounts on active labor market policy (ALMP) with the aim of supporting the unemployed in finding a job. Total spending on ALMP may go beyond 1% of GDP in certain countries, reaching maximum burdens and political focus during economic downswings, most recently during the Great Recession. Despite these costs, institutions like OECD recommended to invest additional resources in supportive ALMP, notably in training and job search assistance programs (OECD 2012 for the U.S.). These could be helpful in mitigating the severe harm of long-term unemployment, especially with respect to human capital. But how should these programs be designed to fulfill this condition given that many negative evaluation results have been found in the past (Card et al. 2015, 2010)? Addressing this question requires a new level of detailedness in policy evaluation: more knowledge needs to be generated on *how* individuals react to supportive interventions. Using a novel combination of data, this study provides the first contribution to open this "blackbox".

Going more detailed into policy assessment means essentially extending the classical policy evaluation setups towards the analysis of *behavioral mechanisms behind policy incentives*. This requires, however, an additional type of data, which is usually not collected within register data sources. The latter can generate some indirect evidence on threat and dislike effects of ALMP programs¹, but they are unable to directly address the question on *how* such programs affect the job seeker's behavior. Particularly in the case of supportive ALMPs –which are supposed to operate via transmitting *content* and not via ex-ante deterrence– knowing how job seekers react to the program is key.

In this study I am able to directly observe a variety of behavioral variables, on top of the unemployment insurance register data. Specifically, I consider three types of behavioral channels: first, measures on *search behavior* –search effort, channels and strategy on the one hand, with reservation wages on the other. Second, measures that track individual *beliefs* (and their bias) –on job chances and on expected wages– are also included. Third are aspects of *non-cognitive skills* –performance-motivation, self-confidence and a component of self-control (reliability). These measures allow assessing how supportive ALMP affects different behaviors in a causal setup.

The study's main contribution is a novel combination of data and design: On the one hand, the design is based on a *randomized controlled trial (RCT)*. The job seekers in the experiment are randomly assigned to an intense coaching treatment. Crucially, the RCT features a rigorous *ex-ante setting of the timing of different treatment periods* in order to identify their specific impact on outcomes. On the other hand, extensive register data are combined with a panel of surveys which were repeated over the unemployment spell. This combination allows tracking the *dynamics of behavior* and analyzing it in parallel with the labor market outcomes.

This setup provides the base for a detailed framework of analysis. The paper studies an RCT intervention featuring intense coaching and counseling of unemployed older workers in Northern Switzerland in 2008-2010. The predetermined timing of the treatment stages delivers an exogenous setting of time periods, which is key for the identification of dynamic treatment effects. We can distinguish treatment effects for four periods during unemployment and for the

¹For example Black et al. (2003), Rosholm and Svarer (2008), Van den Berg et al. (2009) and Graversen/van Ours (2011) report threat and dislike effects of active labor market policy programs.

time thereafter. In particular, we can separately evaluate the effects in the anticipation period, during coaching and in two stages afterwards; moreover, the individual's employment outcome is observed up to 1.5 years after unemployment exit. Proceeding in multiple steps, I assess not only the effectiveness of the program in terms of job placements using a non-parametric technique, but also the behavioral responses and mechanisms driving how participants respond to the program by employing suitable dynamic analyses. Checks for (unobserved) selectivity and robustness are conducted as well.

As a key result with respect to labor market outcomes, the non-parametric analysis shows that the supportive ALMP intervention resulted in a 9 percentage points higher proportion of job finders among the treated. Over the unemployment spell, it takes a while before the positive job finding effect starts dominating: before coaching, I find a significant "attraction" or "waiting" effect, while during the program the common lock-in effect is visible (albeit relatively weak). Then, the positive effect on job finding gradually appears in the post-coaching period. Unemployment durations are not significantly different between treated and controls. In the 1.5 years after unemployment, the positive coaching and counseling effect further materializes, with significantly higher employment stability.

These results raise the question about which reactions in *search behavior* have been driving the positive effect of the supporting measures. The field experiment provides several core insights in this respect: First, the job seekers did *not search more, but more effectively*, due to the support from coaching. The number of applications sent out remained constant or even decreased, but the job finding rate increased after coaching compared to the control group. This implies an improvement in search efficiency, which is supported by further observations: the treated individuals adapted their search strategy (extension of the scope of search). Moreover, they focused the use of search channels by decreasing channel variety and frequency of use and by putting some more weight on informal channels. Second, the treated job seekers gradually *reduced their reservation wages*, whereas the control group members kept their reservation salary about at the level of their pre-unemployment earnings. This is consistent with a model like Burdett et al. (1988), which suggests that job seekers learn over time to correct their idea of the wage and job offer distribution relevant (achievable) for them. Coaching arguably has supported this learning process about "realistic" wage offers.

Interestingly, the analysis reveals that behavioral channels beyond classical search behavior variables generate significant treatment effects as well. The findings with respect to *beliefs about job chances* and *non-cognitive skills* like motivation, self-confidence and reliability suggest that the supportive intervention had a positive short-run impact on these factors. Moreover, the data document substantial ex-ante *upward-bias in beliefs* on job chances, supporting the scarce existing evidence (Spinnewijn 2015). The experiment demonstrates that targeted supportive ALMP is able to affect the bias in the beliefs of job seekers about their chances of success, although in a moderate and non-permanent extent. The same conclusion applies to the analyzed variable components of non-cognitive skills. Further analyses show that the reactions via these non-standard behavioral channels are not just a correlate of search behavior but explain additional variation.

This study demonstrates the potential of combining empirical analysis of behavior and dynamic program evaluation to provide specific evidence on success factors of ALMP. Evidence based on survey data tracking behavior allows the analyst to assess which components of a pro-

gram work. It also delivers the empirical knowledge to understand how the job seekers react to the incentives implied by the program, i.e. which behavioral mechanisms are at play to make the program work. These insights provide the base for the design of specifically targeted policies and programs. Moreover, the tracking of the job seeker’s beliefs could also serve as an early indicator for upcoming improvements (or decreases) in job search success. Finally, the behavioral data reveal that ALMP not only operates via impacts predicted by standard search theory but also via behavioral channels which are usually ignored in these models.

This paper contributes to the still small part of the UI and job search literature which is based on randomized treatments in the field. Whereas in the case of the U.S. most of the published studies date back to the early nineties (e.g., Ashenfelter et al. 2005, Meyer 1995)², the European literature is more recent but features so far only a small number of published experiments. There is, in particular, some evidence based on RCT studies in Denmark (starting with Graversen et al. 2008) and the Netherlands (starting with Van den Berg et al. 2006), which document positive effects of monitoring and job search assistance interventions. A large field experiment in France on job placement assistance for young job seekers resulted in positive short-run effects for the treated but only small net benefits for the total eligible population, due to displacement effects (Crépon et al. 2013)³. For the case of Switzerland, this study is the first that applies an RCT design for the systematic evaluation of UI treatments.

As a further contribution, the paper provides new evidence on the effects of labor market policy programs which are specifically *targeted on older job seekers*. Interestingly, the empirical literature on this issue is very scarce so far. There is some evidence on age heterogeneity in treatment effects of general types of ALMP programs (Card et al. 2015), but studies assessing how programs work which were explicitly designed for the needs of job seekers beyond age 45 are broadly missing in the published literature.

Furthermore, the paper is related to the recently growing literature on behavioral labor economics (e.g., DellaVigna et al. 2015 on UI). The empirical evidence on several behavioral channels allows assessing whether the job seeker’s reaction to supportive ALMP can be fully explained by reasonings of classical job search theory or whether some behavioral reactions go beyond this framework. It turns out that the second is the case; but, nevertheless, many of the identified effects can be best explained without departing from standard rational theory.

The remainder of the paper is structured as follows: Section 2 presents the field experiment and the treatment; then, it discusses the considered outcomes and behavioral channels in the context of existing literature and theory. Section 3 presents the data and descriptive evidence, while Section 4 reports the results of the non-parametric analysis of the RCT. Section 5 introduces the estimation frameworks for the dynamic analyzes. Section 6 discusses the results by treatment period and the employment quality outcomes, while Section 7 assesses cost-benefit and discusses the findings on behavioral channels in a conceptual context. Section 8 concludes.

²We abstract here from the literature strand based on randomized correspondence studies (e.g. Kroft et al. 2013). These field experiments address other research questions like employer behavior and discrimination.

³Note that this discussion focuses on a series of directly related key examples and doesn’t claim comprehensiveness. More RCTs in UI have recently been conducted – for an overview see Card et al. (2015) – but most of them are not yet published, and often they do not focus on randomizing directly an ALMP program but rather on information treatments or randomized assignments to different operators, contract types etc.

2 The Experiment

This section presents the new field experiment, describes the treatment and provides information about implementation and institutional background of the randomized trial. In the second part, we outline all the considered outcomes and behavioral channels and briefly discuss the related concepts and hypotheses. This allows to systematically think about expected results.

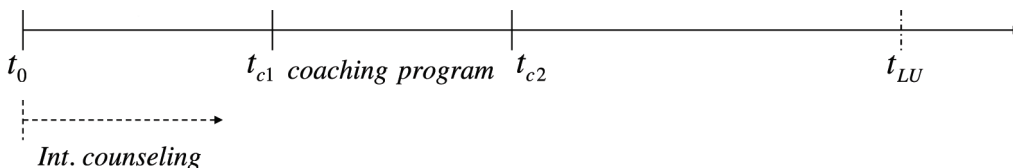
2.1 Setup

2.1.1 Treatment Plan

The treatment plan consists of two combined measures and a specific timing of the interventions. The two measures are high-frequency counseling by the caseworker at the PES and – most importantly – an intense external coaching program performed in small groups.

Timing The timing of the interventions is highly relevant. From a policy point of view, early intervention is crucial in order to fight long-term unemployment. If the interventions start too late, the risk is high that the worker is already on a vicious circle of being disconnected from the labor market too long and therefore facing a decrease in employability – especially in the case of older job seekers who are often confronted with decreasing labor market attractiveness anyway. Beyond this, it is key that the experiment imposes a *strictly structured treatment order* for which the *timing is fixed ex-ante*. The fact that order and timing of the treatments are known from start is crucial for the identification of dynamic treatment effects – it makes this part of the treatment plan exogenous. This feature will be further discussed in section ?? on econometric modeling and identification. The timing of the treatment plan can be visualized in the following way:

Figure 1: The timing of the treatment plan



Treatment *High-frequency counseling* starts right from the beginning of the unemployment insurance spell, from the first meeting onwards. Job seekers meet with their caseworkers every second week – thus in a double frequency compared to the normal monthly rhythm of meetings. High-frequency counseling is applied to the treated during the first four months of the unemployment spell. Then, the frequency goes back to normal (monthly rhythm).

The basic idea behind increasing counseling frequency is that the caseworkers have *more time* available for the respective job seeker⁴. They can spend this additional time on providing more job

⁴It is important to note that the higher counseling frequency for the treated did not result in lower frequency for the controls: The caseworkers in the participating PES agencies got a workload relief of about 20%. The fact that there was no visible "under-treatment" of the control group is confirmed by a comparison to a second external

search assistance to the worker: thus, they can use it to coach the job seeker in job search strategies and to further target their counseling effort. Note that the administrative and monitoring tasks remain the same as for the control group. Therefore the additional time goes predominantly into support; however, this intensified support could imply as well a certain tightening of monitoring (more implicit control).

The *coaching program* – the core treatment measure – starts in median after 50 days (48.5 days for those who really participate, 52 days until potential coaching entry for the others⁵) Thus, the principle of early intervention is taken literally. The coaching was performed in small groups of 10-15 persons. An external, private-sector coaching firm was mandated to perform the coaching program. One coach ran all the coaching programs which took place during the year of sample inflow (December 2007 to December 2008; last program started in January 2009). The coaching program spanned 54 calendar days (in one exception 70 days, due to Christmas/New Year break). Job seekers spent about 3 full days per week in the program; this results in a net participation duration of 20 working days. In addition, homework had to be done. Thus, the coaching program is of high intensity and features a big work load.

The *content of the coaching* is crucial for understanding the treatment effects of this type of supportive labor market policy. It features a combination of intense job search assistance plus elements of human capital development (in terms of core competencies and employability). To achieve the intended skills update, the coaching program addressed five core tasks⁶:

1. *Self-profiling and its consequence for optimizing search strategy*: Detailed collection and analysis of personal strengths and weaknesses; how to communicate them positively; putting the right ones on the CV. Based on the clarified profile, work on optimizing search strategy, i.e. learning where to search; with particular focus on assessing the potential of reorientations (opening scope of search towards other industries, regions, different types of contracts, working times, etc.)
2. *Realistic self-assessment*: Contrast of self-perception and external perception; what is realistic to require and expect from potential jobs; realistic wage demands (at age 45+); what is still feasible in terms of educational updating; risk of long-term unemployment and benefit exhaustion.
3. *Improvement of job application skills*: Interview training and feedback; role plays; use of electronic applications, spontaneous applications by telephone (incl. practical training).
4. *Job search efficiency*: Directed search; hints and lists where to search (focus on internet); use of alternative search channels like spontaneous applications; general search coaching.

control group, see section 7. The controls do not perform worse in job finding than the external controls.

⁵Due to the fact that the timing of the measures was fixed ex-ante, the *potential coaching entry date* can be identified for every person in the study, notably as well for coaching non-participants and the control group. The dates for coaching program starts were fixed with the program provider before project start. Approximately every 1.5th month a new coaching program started; there were 9 in total over the year of inflow. The algorithm for identifying the potential coaching entry date is: (i) take next program start date; (ii) it must be at least 5 days beyond the availability date (by which the individual is available to take up a new job or to join ALMP programs; mostly equal to unemployment registration date), otherwise take the next program start.

⁶This description of the core content is based on an interview with the coach plus written curricula of the coaching program (which were available on the internet during the time of the treatment).

5. *Self-marketing*: Learn how to sell oneself (incl. practical training); motivation to do more self-marketing.

The *control group* followed the 'status quo', i.e. the normal procedures and standard programs. This means in particular that the control individuals met up with the caseworkers only monthly. They took up ALMP programs substantially later since the status quo does not feature an early intervention strategy. A typical ALMP trajectory in the control group starts with participating in a short job search assistance sequence of 3 to 7 working days, roughly after 3 to 4 months of unemployment. Thus, this short program is normally the only ALMP activity in the control group that takes place during the period of intense intervention in the treatment group (first 4 months). After the four months (end of treatment) both groups follow status quo procedures (monthly meetings and further ALMPs, dependent on individual needs). It is important to note that the individuals of the control group had no possibility to enter the coaching program. This newly designed program was exclusively assigned to the treatment group.

2.1.2 Implementation & Background

This field experiment was performed in two PES agencies in north-western Switzerland. The sample population of the experiment consists of job seekers aged 45 to 61.5⁷. The participants entered registered unemployment in the two PES between December 2007 and December 2008 and met the necessary eligibility conditions.

Random assignment The participants were *randomly assigned to treatment and control group at time t_0* , i.e. at registration before the first meeting with the caseworker. The assignment procedure, run separately for each of the two PES, consisted in three steps: First, the complete inflow of the respective PES was filtered with respect to the *eligibility conditions*: Age 45+, employability level medium or low, either full-time or part-time unemployed above 50%, enough language skills to follow the coaching, no top management and no job seekers who have found a longer-term temporary subsidized job (longer than a couple of days). Second, the eligible individuals were *randomly assigned* to the treatment group (60%) and the control group (40%)⁸, by means of a randomized list (by PES)⁹. As a third step, the remaining individuals were assigned to the caseworker pool. 16 caseworkers were involved in the project, whereby 10 bore the main load of cases. The *assignment mechanism* follows a fixed rule: assignment by occupation. It is therefore *exogenous* to the treatment. It is important to mention that, due this mechanism, caseworkers were automatically assigned cases in the treatment *and* the control group. Thus, randomization holds also *within* caseworker cells. Note, moreover, that caseworker fixed effects will be taken into account in the estimations.

⁷The upper age limit was chosen at age 61.5 in order to avoid the special case of direct passage into statutory retirement: individuals above this age have the right to collect 640 benefit days which corresponds to a calendar duration of about 2.5 years; this is enough to switch over to the pension system at (early) retirement age 64.

⁸In the first quarter of 2007, the random assignment ratio was 50%–50%. As a consequence of good economic conditions, inflow was lower than expected. We therefore decided to switch to a 60%–40% assignment rule. This explains why the treatment-control ratio reported in the descriptive analysis in section 3.3 is in-between the two rules. Note that this switch has no impact on the quality of randomization.

⁹The list, generated by the researcher, consists of a simple randomized sequence of "treated" and "control" labels. The head of PES added the eligible individuals in order of inflow (registration date) to that list.

Information It is important to know which *information* was available for the treatment and control group at time t_0 . In their first meeting with the caseworker, the job seekers of both groups were informed in written form that they participate in a project of "quality control" (of PES procedures and advisory). This was necessary since both groups had to fill in repeated surveys over the duration of their unemployment spell (see section 3.1). However, the caseworkers were not allowed to use the terms 'long-term unemployment (risk group)' and 'randomization'. The former was to avoid stigmatization biases, the latter to prevent discussions which could potentially increase the risk of non-compliance.

Economic conditions The participating PES belong to a quite urbanized region in the agglomeration of Zurich (about 45 minutes of commuting distance to the center of the city). Thus, they are part of the biggest and economically most productive local labor market in Switzerland (population: 3.7 million). Therefore, given the relative size of the experiment compared to the size of the labor market, general equilibrium effects of the experimental intervention can be excluded. The unemployment rate in this area for the sample age group and inflow window amounted to 2.2%.

Institutional background The Swiss UI system is a typical representative of an OECD unemployment insurance system. Its structural features and labor market policy design are therefore highly comparable to most other OECD systems (see, e.g., OECD Employment Outlooks). The standard maximum duration of unemployment benefits in the Swiss UI system is 1.5 years (400 benefit days) for individuals who meet the eligibility requirements. It is increased by another half a year (120 benefit days) at the age threshold of 55. The replacement ratio varies between 70 and 80%, depending on income.

It is important to note that all the assignments to ALMP programs and the meeting appointments are compulsory for job seekers¹⁰. Thus, participation in the treatments of the experiment is compulsory as well. If job seekers do not comply to these rules, they risk to be sanctioned (as well if they refuse suitable job offers or do not provide the amount of applications demanded by the caseworker). Sanctioning is comparably frequent in Switzerland (about every sixth job seeker is sanctioned) and implies benefit reductions of 100% during 1-60 days (mostly 5-10 days), for details see Arni et al. (2013). This strict sanctioning regime results in high compliance with the rules. This is the case for the experiment too, see section 3.3 for details. Still, non-compliance by the treated job seeker in terms of intentionally avoiding the coaching program can not be fully excluded. However, as the non-compliance analysis in section 3.3 shows, intentional non-compliance could only be observed in a negligibly small number of cases.

More details on institutional and economic conditions are provided in Appendix A.

2.2 Outcomes & Channels

This section provides an overview on the outcomes and behavioral channels that are observed by the available data and puts them in a conceptual context. We take a basic standard job search model (see, e.g., Mortensen 1986, Van den Berg 1990) as a starting point to consider hypotheses of potential behavioral mechanisms of supportive labor market policy. To support the discussion,

¹⁰During ALMPs all the standard duties (job search effort, meetings at PES) and rights (benefits) remain.

the core equation of (one version of) such a stylized model with endogenous search is reproduced here. The value of staying unemployed, V_t^u , is represented as follows:

$$V_t^u = \rho V_{t+1}^u = \max_{s_t} \left[u(b_t) - c(s_t) + \lambda(s_t) \int_{r_t}^{\infty} \left(\frac{w}{\rho} - V_{t+1}^u \right) dF(w) \right]$$

where the reservation wage r_t results from $\int_{r_t}^{\infty} \frac{w}{\rho} dF(w) = V_{t+1}^u$ in the optimum. The latter represents the value of finding employment in $t+1$, b_t the unemployment benefits, and $\lambda(\cdot)$ is the job offer arrival rate as a function of search behavior.

In this experiment, we consider one set of labor market outcomes and three sets of behavioral variables that will be used to measure the causal effects of coaching and counseling on outcomes and channels. In the following, the variables are first defined, then put into the context of existing literature and the framework above, and finally possible hypotheses are mentioned. Due to the novelty of the presented behavioral data, the brief discussion has explorative character.

Outcomes The key outcome variable is the *probability to find a job* within two years. The effects on *unemployment duration* (as a function of exits from UI in general) are considered as well. Thirdly, the *persistence of found employment* over 1.5 years after UI exit is analyzed, by means recurrence rates (probability to fall back into a UI spell). The considered coaching intervention can be seen as an intense form of job search assistance (JSA), combined with training of labor market skills (see last section). The comprehensive current meta-analysis of ALMP evaluations by Card et al. (2015)¹¹ documents "relatively large" positive short term impacts of JSA, on average. Training programs show smaller impacts in the short run but larger effects in the medium and longer runs. Thus, this evidence suggests, as a hypothesis, that a successful coaching program positively affects unemployment exit and job finding. However, the effect size could be small, if the result by Card et al. (2015) that older workers tend to react less on ALMP applies here too.

When it comes to dynamics, we can distinguish anticipation and during-program effects, as well as different stages post-program. In terms of anticipation, the (small) literature finds evidence of threat effects of JSA with activation in some cases, the opposite has not been documented. Next, one could expect to see a "lock-in effect" (reduced UI exit rates) during the program – as commonly found in the literature. In which direction expected effects post-program – in particular on employment persistence after UI – may go is (theoretically) ambiguous: This depends on whether the coaching effect predominantly results in an increase of job match quality, or whether a possible reductive impact on reservation wage is more important.

In addition, we observe the individual *job interview rate* as a "pre-outcome". This is the number of job interviews that have been obtained based on the applications of the last month, divided by this number of applications. The job interview rate can be seen as an intermediary outcome in the sense that it represents success in the first two steps of the job finding process: arrival of a job offer plus passing the first stage of the matching procedure. There is, to my knowledge, no empirical literature yet on this outcome, as it is not observed in usual datasets.

Search Behavior Search activity is measured in two dimensions: as quantitative search effort and use of search channels. *Search effort* is measured as the number of applications that a

¹¹The analysis considers 207 current studies from around the world, and thus extends the work of Card et al. (2010). For earlier U.S. evidence see as well Meyer (1995).

job seekers sent out in the last four weeks. For each *search channel*, job seekers report their frequency of use within the same time span (see section 3.1 for details). To my knowledge, there is no direct empirical evidence so far on how JSA or coaching causally affects search activity. On one hand, coaching and counseling could motivate the individual to search more and to reduce moral hazard¹². On the other hand, this type of advisory could support directed search; searching more specifically could reflect in lower quantitative effort¹³. By the same argument, the coaching program effect could result in either higher or lower frequency of search channel use. To accommodate several dimensions of search activity, the model above could be complemented by replacing search effort s_t by a search function $s(e_t, \sigma_t^1, \sigma_t^2, \dots)$ which features quantitative effort (e_t) as well as the frequencies of different channels (σ_t). Holzer (1988) presents such a model and argues that channel choices should be related to their costs and expected productivities (as well as to nonwage income and wage offer distribution). Applied to a coaching intervention which is supposed to improve the effectiveness of search, this argument would suggest that one could expect an optimized use of the search channel mix (i.e. stronger weighting of the more productive and less costly ones) as a treatment effect.

Reservation wage is measured by the common survey question which asks for the minimum amount of wage that the job seeker requires to accept a job offer. Coaching could affect the reservation wage in at least three ways. First, it could be seen as an additional cost of unemployment which reduces V_{t+1}^u and therefore r_t . Second, it could impact positively on the individual's human capital (by improving labor market skills) and thus increase reservation wage. A third channel how coaching could affect reservation wage is via its potential correlation with the job seeker's wage expectation. In a recent study Krueger/Mueller (2014) find that reported reservations wages start out "too high" as compared to a calibrated job search model; they presume this to be related to a misjudge of wage prospects. Thus, if the coaching "shock" affects wage beliefs, this could translate in a change of r_t . In total, no unambiguous hypothesis on the direction of the reservation wage effect of the coaching treatment can be derived. We will be able, however, to distinguish to some degree the arguments above by data and dynamic analysis.

Beliefs The available data allows for the distinction of two types of beliefs: expectations on job chances and on wages (salaries). *Beliefs on job chances* are measured by asking the job seekers to indicate how many job interviews they expect based on the applications they sent out in the last month. To assess the bias of this belief, the expected interviews are compared to the realized job interviews. *Wage beliefs* are captured by a simple scale question whether the expected wage in the next job is higher, lower or similar. Some of the few existing contributions on beliefs on job chances (Spinnewijn 2015, Falk et al. 2006) introduce the bias in beliefs by replacing the true job finding probability by a perceived one.

Given the available data and the nature of the job search process, it is useful to decompose the channels of beliefs: On one hand, job seekers may be biased in what they expect concerning

¹²Krueger, Mueller (2012) conjecture – based on their finding that the job seeker's subjective well-being is lower in times of job search – that "job search assistance programs may be successful because they help people overcome the inhibition to search for a new job associated with the unpleasantness of the endeavor."

¹³Note that the assumption of a positive monotonic relation between search effort and job offer arrival, as common in theory, is not systematically supported by the scarce evidence. Krueger/Mueller (2011) find that the amount of search time does "not bear a straightforward relationship with the likelihood of receiving a job offer (...)".

job offer arrival, i.e. act on $\tilde{\lambda}(\cdot)$ rather than $\lambda(\cdot)$. This bias could be proxied by our beliefs on job chances measure. On the other hand, there may be missing information or misperception on the wages offered to the job seeker, i.e. $F(w)$ is replaced by $\tilde{F}(w)$ (following e.g. Burdett et al. 1988). The relatively crude available measure of wage beliefs does not allow a quantitative assessment of the potential bias in wage expectations. But it will be possible to estimate the effect of the coaching treatment on the probability that individuals expect a wage increase and to compare this with the pre-vs.-post-unemployment wage difference. Consistent with evidence in Spinnewijn (2015), we find that job seekers are, on average, too optimistic when it comes to beliefs on job chances, see section 3.2. Based on this observation, the key question of interest will be whether coaching affects the over-estimation in beliefs. There is no causal evidence on this question so far.

Non-Cognitive Skills A third group of variables which is usually ignored in the job search literature are relevant aspects of personality¹⁴, often subsumed by the notion of non-cognitive skills. In this experiment, three aspects of such personality factors are addressed: performance-motivation, self-confidence and self-control (which is a facet of conscientiousness, see Almlund et al. 2011). They are proxied by straightforward scale variables. First, self-declared *performance-motivation* is measured by the question how motivated the job seeker is to search for jobs. Some lab and survey evidence documents a positive relation between motivation and the performance in cognitive tests (e.g. Borghans et al. 2008) or the later labor market outcomes (Duncan/Duniphon 1998). Second, the job seeker’s *self-confidence* is assessed by the case-worker. Different economic models show that self-confidence can improve performance (Compte/Postlewaite 2004) and therefore strategies of self-esteem maintenance – i.e. maintaining high self-confidence by self-deception or positive thinking – may be beneficial (Benabou/Tirole 2002). Third, the case-worker evaluates as well the job-seeker’s *reliability*, an expression of self-control. In particular in education research broad evidence confirms the positive relation between self-discipline and academic achievements (e.g. Heckman/Rubinstein 2001, Duckworth et al., 2005).

Overall, it is thus plausible to hypothesize that the coaching intervention could yield positive causal effects on the presented three aspects of non-cognitive skills. Through which mechanisms could these operate? Whereas motivation (μ_t) probably acts via enhancing search productivity, it is more plausible for the cases of self-confidence (γ_t) and reliability (ϕ_t) to think of them as direct success factors in the acquisition of job offers and job interviews. This means that the former would enter the framework by extending the search function, i.e. $s(\cdot, \mu_t)$, whereas the latter two would complement the specification of the job offer arrival rate, i.e. $\lambda(\cdot, \gamma_t, \phi_t)$.

3 Data & Descriptive Evidence

In the following, we first present the data sources. The second part discusses descriptive evidence on behavioral channels and outcomes. Finally, we assess whether random assignment worked.

¹⁴One of the few empirical exceptions is Caliendo et al. (2015) who analyze the relation between locus of control and search behavior. Their main finding is that individuals with an internal locus of control search more. The paper is silent, however, on how this translates into job finding.

3.1 Data Sources

The analysis of this field experiment is based on a unique combination of detailed administrative records of the unemployment insurance (UI) and a series of repeated surveys which cover the behavioral dynamics unobserved by common UI registers.

Register Data Detailed data from the Swiss UI records provide information on the timing of relevant events within the individual unemployment spell in daily precision: dates of entry, exit, participation in labor market programs etc. The register data include a rich set of observable characteristics like usual socio-demographics, education and occupation as well as past unemployment histories up to three years before UI entry. The tables on covariate balancing in section 4 (Table 1) and on the aggregate effect on unemployment duration (Table 11) in the Appendix document the collection of used observables. The same register data source allows also the observation of potential unemployment recurrence up to 1.5 years after unemployment exit.

Survey Data Linked to the register information, *repeated survey data* have been collected for this field experiment – the LZAR data base (see Arni 2011 for details). After the counseling meetings, the caseworkers had to fill in an online tool which complemented the information of the register data base. Before the meetings, job seekers filled in repeated (paper) surveys as well. The surveys were explicitly designed to track neatly the behavioral reactions of the job seekers in different periods of the treatment.¹⁵

The timing of the repeated surveys is dynamically adapted to the treatment plan. Thus, surveying is more frequent in the period of intense treatment, i.e. the first four months. Specifically, the surveying rhythm is designed as follows: It starts with an entry survey at t_0 before the first meeting with the caseworker, then subsequent surveys for job seekers and caseworkers follow after approximately 1/2/3/4/9/12 months of unemployment and at exit. The job seeker exit survey is administered about three months after unemployment exit and includes as well questions on the salary and hours in the new job. The repeated surveys have been assigned to a timing structure relative to the treatment plan, using the exact date of each survey response. Following Figure 1, we distinguish five (potential¹⁶) treatment periods (plus exit information) to which the surveys are assigned: (i) The *start* survey contains the t_0 information, collected before the first caseworker meeting; it is thus unaffected by treatment or dynamics. (ii) In the *anticipation* period, t_0 to t_1 , the survey closest to coaching entry (t_1) is used. (iii) In the *during coaching* period, t_1 to t_2 , the survey closest to coaching end (t_2) is used. (iv) In the *early post-coaching* period, i.e. the 90 days after coaching end (t_2 to t_3), the survey closest to the middle of the period (coaching end +45 days) is used. Finally, (v) in the *later post-coaching* period, i.e. beyond 90 days after coaching end (after t_3), the survey closest to 100 days after t_3 is chosen.¹⁷

¹⁵Beyond the measures that we will explore in the following, the surveys covered as well items like health state, appearance, quality of applications, language skills, willingness to mobility (assessed by the caseworker), life satisfaction and evaluation of the PES service (job seekers).

¹⁶For the control group, the same timing structure analog to treatment periods is constructed, see footnote 5.

¹⁷Empirically, the timing of the assigned surveys by treatment periods turns out to be as follows (median reported): (ii) anticipation period: 14 days before start of (potential) coaching; (iii) during coaching: 23 days before end of coaching; (iv) early post-coaching (1 to 90 days): 7 days before middle of period; (v) later post-coaching (after 90 days): 5 days beyond the targeted 100 days after t_3 .

The dynamic analyses by treatment period will always follow this schedule. The observed sample in the later treatment periods is naturally subject to dynamic selection – individuals gradually leave unemployment for a job (or non-employment). The issue of the balancing between treatment and control groups in later survey periods will be discussed in the context of the dynamic analysis in section 5.3.

3.2 Descriptive Evidence on Behavior & Outcomes

3.2.1 Outcomes: Unemployment Duration, Recurrence, Job Interview Rate

The sample of the experiment consists of 327 job seekers, 186 in the treatment group, 141 in the control group (sampling see section 2.1.2). The typical *unemployment exit rate path* of this population shows a similar shape as in most European countries. In an early stage, up to 4 or 5 months, the (monthly) exit rate rises quite sharply; here, it tops at 18%, see Figure 2, left panel A. Thus, a relevant subgroup of the population finds a new job quite quickly, within a few months. Thereafter, the exit hazard goes down remarkably and remains on a level of 6 to 12%. In the last months before benefit exhaustion (usually after about 18 months, beyond the x-scale of Figure 2) it typically rises again to levels comparable to the first peak. The median duration in the sample is 140 days, the mean duration 250 days.

[Figures 2 about here]

The *unemployment recurrence* or *employment exit* hazard is plotted in panel B of Figure 2. Post-unemployment durations for individuals who found employment are observed until the end of March 2010 (exogenous censoring date). Thus, the sample which is at risk of recurrence (234 individuals) is observed through 490 days in median; 84% are observed for at least one year, 42% for the full 1.5 years (540 days)¹⁸. The unemployment recurrence rate ranges between 5 and 6 percent within the first eight months of post-unemployment. Then, it tapers off and reaches less than 2 percent after 1.5 years.

Finally, we consider the *job interview rate* as an early outcome indicator. It is defined as the number of realized job interviews per sent application (within the past month; based on the survey data, see section 3.1). At time t_0 , the first meeting with the caseworker, we observe an average job interview rate of 11%. In the subsequent periods, the rate in the control group decreases to 8.5 to 10% with a standard deviation of 11 to 15 percentage points.

3.2.2 Behavioral Channels

We now turn to discussing descriptive evidence on the measures for the behavioral channels that have been conceptually introduced in section 2.2. Step by step, we discuss now the survey measures on search behavior, beliefs and aspects of non-cognitive skills.

¹⁸Note that the sample which is at risk of recurrence is defined as all individuals for whom at least 20 days of post-unemployment duration is observed and who took up employment after unemployment. These are, according to the UI register, all the direct job finders plus 6.4% of individuals who leave UI voluntarily with undeclared exit destination. Past analyses of the latter by means of social security data have shown that this small group usually takes up employment within a few months. Therefore it is sensible to include them into the at-risk population. To avoid the overweight of some long durations, the post-unemployment durations are (exogenously) censored at 540 days (1.5 years).

Search Behavior A first core indicator of search behavior is *job search effort*. The repeated caseworker surveys ask for the number of applications the job seeker has sent out in the last four weeks. Note that the job seekers are required by law to report their application activity to the caseworker. Therefore, this information which is routinely protocolled by the caseworker should be of high reliability. On average over all treatment periods, job seekers in the control group send out 6.9 applications. Figure 3 shows the dynamics (and standard deviations) of the quantitative search behavior: after a lower start, the mean number of sent applications per month stabilizes around 8.5¹⁹. The standard deviation amounts to about 5.7.

The second dimension of search is the *search channel variety*. The job seekers were asked in every survey which specific job search channel they used and how often. The following 11 channels were proposed: newspapers; internet; private recruiters; PES-operated job offer database; job postings found in public spaces; network: strong ties (family, good friends); network: weak ties (colleagues at work, in sports and other associations, from hobbies, neighbors etc.); network: former colleagues from school and other education programs; spontaneous applications by mail; spontaneous applications by telephone; other. To create the measure of channel variety, all channels of the mentioned list which have been used at least "monthly or less often" were counted. On average over all treatment periods, 6.8 channels have been used by the control group. Figure 3 documents that channel variety increases over the unemployment spell.

[Figures 3 about here]

A third considered element of search behavior is *search channel choice*. Based on the same block of items as above²⁰, I analyze for each of the channels its *frequency of use*. The frequency is measured on a 6-point-scale: 3 = "daily", 2.5 = "several times per week", 2 = "weekly", 1.5 = "several times per month", 1 = "monthly or less often", 0 = "never". The third panel of Figure 3 shows that the two formal channels (blue) newspapers and internet are clearly used most often – on average several times per week. Whereas newspaper use is flat over the spell, we observe some increase in the use of the internet. Among the informal channels (green), personal networks are used several times per month on average, with a decreasing tendency. The use of spontaneous applications, written and by telephone, is comparably volatile over the spell – this may indicate a "trail-and-error" strategy.

The fourth aspect of search behavior we observe is *search strategy changes*. The caseworker and the coach have been asked whether they agreed with the job seeker on changes in the scope of search²¹. Detailed analysis revealed that the vast majority, more than 80%, of these changes were *extensions* of the search scope, i.e. the new scope was used supplementary to the existing. On average, the probability of extending the scope of search amounts to 0.18 per period (no graph).

The second fundamental dimension of search behavior is *reservation wages*. They are surveyed by the classical question about the minimum (monthly gross) salary the job seekers still would

¹⁹Note that the Swiss UI features a strict system of monitoring (and sanctions) which includes individual job search requirements that ask for a minimum number of applications per month (usually 6 to 10).

²⁰In the frequency of use analysis, the channels 'PES-operated job offer database', 'job postings found in public spaces' and 'other' are omitted, due to low occurrence and few variations. Descriptives are available on request. Moreover, the three network options are merged into one channel.

²¹Specifically, they could indicate whether there was a change in: industry; occupation; place of work; kind of employer searched for; workload per week; permanent vs temporary jobs; working hours & shifts. For each of the changes, they declared whether it was an extension or change or reduction of the scope of search.

accept. The caseworker asked the job seeker this reservation wage question and reported his/her answer²². We build two measures which are documented in Figure 3. First, the ratio of the reservation salary to the previous salary before unemployment. On average, this ratio amounts to 0.99 for the control group and is remarkably stable over the spell. The increasing standard deviation of about 0.3 is quite substantial. Note that a ratio above one is not forcefully irrational, it can be observed for a series of rational reasons: individuals are willing to work more hours to avoid salary loss; they expect salary growth compared to pre-unemployment; they declare to be indeed unwilling to work for less than their last salary; poor information about salary prospects on the market; or measurement error. The second measure which is analyzed is the level of the reservation salary. To give less weight to outliers, we consider the median level. It turns out to be 5500 CHF (p25: 4215, p75: 6900) and as well quite stable over the spell²³.

Beliefs The repeated surveys allow to construct two measures of beliefs related to labor market chances. On the one hand, we consider the *beliefs about the chances to get job interviews*, which are measured as follows: the difference between the interviews expected and the interviews realized, each based on the applications of the last month (and standardized by the number of sent applications). Thus, this yields the overestimation bias in expected interviews per application. Figure 3 declares that the control group individuals expect on average 0.23 job interviews more than they realize per application. Overestimation first increases and then decreases again over the spell. The second type of beliefs that is tracked by the surveys is *wage expectations*. The specific survey question asks about the expectation to earn more, the same or less (5-point-scale) than before unemployment. For a longer period over the unemployment spell, the job seekers expect to find a job that pays equally well (average scale value 2.9) than the last job. Beliefs turn out only to be remarkably lower in the last considered period. Overall, the descriptive analysis of these measures of beliefs confirm findings of the scarce empirical literature (Spinnewijn 2015) that, on average, individuals tend to be *overconfident* with respect to labor market prospects.

Non-cognitive Skills Three aspects related to personality are considered here: performance-motivation, self-confidence and self-control. As a first measure, we can analyze the *motivation for job search* which is surveyed on a 5-point-scale (very high/high/medium/low/very low), by a direct question to the job seeker. Figure 3 shows that more than 40% of the job seekers are initially "very highly" motivated. This proportion decreases substantially over the spell and reaches less than 26% in the last period. The average scale level decreases from 4.3 to 3.9. To avoid the assumption of a metric scale, we will use the first measure in the econometric analyses. The other two aspects, *self-confidence* and *reliability* (as an expression of self-control), are assessed by the caseworker, to avoid larger self-perception biases. Both are measured on a 4-point-scale (high/rather high/rather low/low degree). According to Figure 3, the proportion of job seekers with a high degree of observed self-confidence varies between 12 and 36%. The high volatility of self-confidence over the spell could express the job seeker's oscillation between positive and

²²The intention behind this setup of reporting is to reduce the risk of unreliable answers. Given that the job seekers must communicate their reservation wage to the caseworker they cannot report any unrealistic number as the caseworker will question the plausibility and ask again.

²³The corresponding figures for the last salary are: median 5500 CHF, p25: 4200, p75: 7500 CHF. 1 CHF = 1.007 USD = 0.93 EUR.

negative signals from the labor market and the caseworker. The job seeker’s reliability, on the other hand, shows upward tendency over the spell. The neat monitoring and monthly caseworker meetings may have a positive impact on the job seeker’s self-discipline.

Note that the figures discussed here are descriptive, i.e. not conditioning on the fact that the composition of the remaining job seekers changes over the spell. This will be taken into account in the dynamic analysis later on. The descriptive analysis, however, reveals interesting insights in the baseline behavior of job seekers. Moreover, it demonstrates that these new measures seem to produce plausible empirical proxies for a uniquely broad collection of behavioral channels.

3.3 Observables: Did the Randomization Work?

In this section, I compare observable characteristics of the treatment and control groups in order to assess if initial randomization worked and to characterize the experimental population. The sub-samples of job finders and of those who reported salaries are discussed as well.

The comparison of observable characteristics between treatment and control group, see Table 1, shows that *randomization worked well*. No significant group differences can be detected for this sample of 327 job seekers. The project eligibility criteria (see section 2.1.2) shape in particular three characteristics of the sample. First, the minimum language and skill requirements for coaching result in higher proportions of (semi-)skilled (95%²⁴) and of Swiss job seekers (85%), compared to a random draw of job seekers. Second, the field experiment is focussed on individuals of middle (3) and low (4) employability, i.e. individuals categorized as relatively easy-to-place were excluded ex-ante from the experiment. Third, the age of the eligible job seekers is restricted to 45 to 62. The average age of the participants in the RCT is 52 years. 40% of the individuals in the sample are of age 45-49, 27.5% of age 50-54, 21.7% of age 55-59 and 10.7% of age 60-62²⁵.

[Table 1 about here]

The median duration of unemployment history in the past three years is zero for both groups. 27.5% of the participants have a positive duration (median 113 days). ‘Duration to availability’ indicates the number of days until a job seekers becomes available for a new job or for ALMP. Usually, individuals are immediately available, the core reason for a delay is that the respective job seeker already registered at the UI during the cancelation period, a practice which is promoted by the UI authorities. A majority of 57% is available within 20 days.

Finally, to which degree did the randomization carry over to the sub-sample of job finders? The full sample of job finders (234, register data) will be used for the analysis of the treatment effect on post-unemployment job stability (recurrence rate). Further, the analysis of effects on salaries will be executed among those who responded to the exit survey (163 individuals). 69.7% of the job finders responded to the salary questions in the exit survey. This response rate is highly balanced between treatment and control group (70.0 vs. 69.1%). The two selective sub-samples are also well balanced in their characteristics. No significant differences in observables between treatment and control group are found, except from the proportion of married people in the salary

²⁴Job seekers are categorized as semi-skilled or skilled as soon as they have any type of professional education.

²⁵Note that none of the last group had the possibility to pass to retirement by means of UI, since their maximum potential benefit duration was not sufficient to reach statutory retirement age.

sample (more among the treated). In total, there is thus no observable indication of a significant response bias. More robustness checks will be done at the end of the dynamic analysis.

4 Non-Parametric Results

What can be learned on the impacts of the field experiment without imposing any econometric structure? Given the successful randomization at t_0 , *causal statements on the net effect of the treatment plan as a whole* can be inferred in a non-parametric manner – by comparing means and Kaplan-Meier survivors. Four main results arise from this analysis.

[Table 2 here]

The most important key finding is that the supportive ALMP treatment caused a *significantly higher proportion of individuals who found a job – an increase by 9 percentage points*. Whereas 63% of the control group (CG) left unemployment to a job, the proportion of treatment group (TG) individuals leaving for a job amounts to 72%, as Table 2 documents.

A more detailed look at the exit destinations in the register reveals interesting supplementary insights. The TG individuals left less often unemployment for non-employment (8.6% vs 13.5% in CG). Moreover, leaving for "unknown status after unemployment exit" is a bit more frequent in the TG (5.4% vs 3.5%). For several reasons it is most probable that a clear majority of these individuals, who left UI voluntarily, found a job too in the near future.²⁶ Therefore, it is useful to consider as well the joint figures on direct and "indirect" job finders: these amount to 77.4% in the TG versus 66.7% in the CG. For this measure, importance and significance of the TG-CG-difference is even higher.

How does the net dynamics of the job finding treatment effect look like? Figure 4 provides a non-parametric survivor analysis of the duration to job finding. It reports the proportion of individuals in the TG and CG who are still without a job; thus, the survivor analysis here defines only those cases as a positive transitions out of the initial status which end up in job finding. The dotted vertical lines indicate the median starting and ending of the (potential) coaching program²⁷. The two curves of the survivor overlap over the course of the first 250 days of unemployment. Beyond, the survivors of treatment and control groups remarkably diverge, leading to a higher job finding proportion in the treatment group in the later stages of unemployment. Thus, this analysis shows that the *tested supportive ALMP takes some time until it develops beneficial effects on job finding*. That the positive treatment effect gradually develops over time, is also visible when considering the margin of long-term unemployment (after 1 year): we find a positive but not (yet) significant effect on the LTU rate of 4 percentage points, see Table 2.

[Figure 4 about here]

²⁶More than two thirds of these individuals renounced to services of the UI voluntarily, in order to avoid the hassle of the control requirements on job search, appointments etc.; the rest left the country to search for a job elsewhere. For a further empirical assessment, see also footnote 18.

²⁷In the upcoming analysis by treatment period I will use, of course, the exact timing by individual.

How did the coaching treatment affect the individual unemployment durations? Have they been prolonged or not? – given the evidence on "lock-in" effects which are commonly observed for training-oriented ALMPs (Card et al. 2015). The non-parametric comparison of unemployment durations, as presented in Table 2, provides a clear answer as second key result: Individual *unemployment duration has not been prolonged* by the coaching and counseling treatment, there is no significant difference between the TG and CG durations, neither for the medians (139.5 vs 138 days) nor for the means (245.1 vs 255.9 days)²⁸.

In the light of the existing ALMP evaluation evidence, the result of no prolongation of unemployment duration due to the new coaching ALMP can be rated positive. It seems that the lock-in effect (less search during the program) is not as dominant here – in spite of the fact that the new treatment implied high workload and time consumption in the first four months of unemployment. This finding could, however, also represent the net result of a reduced lock-in effect and an improvement of effectiveness in job finding after coaching. This can be tested by use of a duration model that distinguishes treatment periods, which will be done in section 6.1.

Next, consider *job quality* as a third key outcome. Does the additional job finding caused by the treatment possibly come at the cost of taking up lower quality employment? This question can only be causally answered without imposing structure if we assume that the sub-sample of job finders is still well balanced between TG and CG. The analysis of observables in the last section 3.3 suggests that this is the case. However, additional robustness checks, in particular as well on the potential importance of unobservables, will be instructive to assess this claim. This will be done within the dynamic analysis. The non-parametric results in Table 2 reveal that no significant difference between gross salaries can be detected among TG and CG. The same applies to declared working hours per week (pensum). In older working age, reestablishment on the labor market after unemployment often implies a wage loss. This is found here as well. On average, a significant pre-to-post-unemployment gross salary loss of 6% (341 CHF) is incurred. However, the loss is not significantly different between TG and CG.

How is job quality affected by the supportive ALMP in the longer run? In the context of this RCT, we can assess *employment stability* within a post-unemployment period of 1.5 years. Table 2 reports the comparison of *recurrence behavior*: Within the given period, 22.1% of the treated job finders recurred to unemployment, whereas the recurrence propensity in the control group amounts to 27.7%. Thus, we find a pretty substantial albeit statistically insignificant *reduction in unemployment recurrence* by the treatment.

The same evidence, but in a time-dynamic perspective, is reported Figure 5. The post-unemployment survivor curve of the treatment group is located clearly above the one of the control group – treated individuals remain on average longer outside unemployment. 300 days after unemployment exit, about 83% of the job finders in the TG remain in employment, whereas the same rate in the CG amounts to about 74%. Note, again, that this result could be biased by the potential selectivity of the sub-sample of job finders. However, taking this into account by a semi-parametric dynamic model will indeed show, in section 6.4, that the treatment effect

²⁸Note that we censor very long unemployment durations to restrict the impact of outlier values, as is common practice in the literature. Durations are exogenously censored after 2 years (730 calendar days, approx. the maximal benefit duration for job seekers aged 55+), which affects 2.4% of the spells. A sensitivity check using the latest possible censoring date (march 31, 2010) shows that the treatment effect is not sensitive to this choice.

on employment stability gets more distinct: the difference in the recurrence (hazard) rates in the post-unemployment period becomes significant. So, in conclusion we can state that the supportive ALMP exerted a *positive effect on job quality: improved employment stability and no significant harm of the post-unemployment salary.*

[Figure 5 about here]

As a fourth key outcome, we shed light on the effect of the coaching treatment on the *job interview rate*. For an aggregate assessment, this early outcome indicator has been averaged over the four treatment periods, as far as observed. 70% (229 individuals) have reported at least once their acquired job interviews – the potential selectivity of this sub-sample will be addressed in the robustness checks as well. The direct non-parametric comparison results in a *significant increase in the job interview rate* by .04 to .14. Thus, this analysis suggests that the counseling and coaching successfully increased the job offer arrival rate of the treated.

To wrap up, the four key results on the main outcomes of the new supportive ALMP can be summarized as follows: *The RCT shows that the counseling and coaching program caused more treatment group individuals to find a job than in the control group.* It took some time after the intense treatment period until the job finding effect materialized. As an early indicator, we find that *the job interview rate has reacted positively on the treatment.* Adopting the UI perspective, we find that individual *unemployment durations have not been prolonged on average*; this hints to a comparably less important lock-in effect in the counseling and coaching program. *Job quality has been positively affected in terms of employment stability and, arguably, has not been harmed in terms of earnings.*

5 Empirical Framework for Dynamic Analysis

These results above document the main final outcomes. However, we want to go a step further and open the behavioral blackbox: *How* did these causal effects of the supportive ALMP come about? To address this question, we need to *directly relate acts of the treatment plan with reactions in behavior and with dynamic outcomes of the job search process.* To gain such insights, we need to put more structure on the evaluation of labor market outcomes and behavioral channels. Therefore, I will develop and apply a series of dynamic treatment effects analyses in the following.

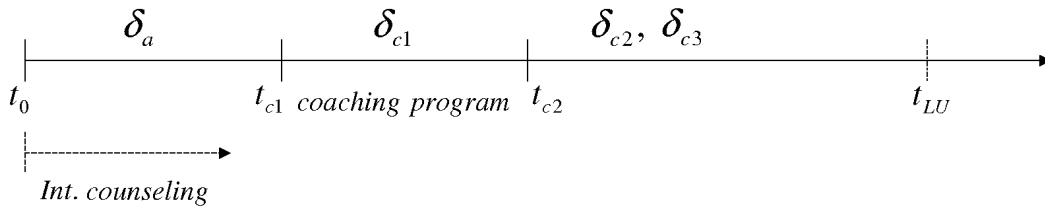
5.1 Dynamic Treatment Effects on Unemployment Exit

To estimate the dynamic treatment effects on the key outcome – transitions from unemployment to jobs – I will apply the *timing-of-events approach* (following Abbring/van den Berg 2003) to the treatment plan. This approach provides at least three key advantages for learning more on the dynamics of the behavioral effects of the supportive ALMP: First, the *identification of sub-treatment-effects* by use of the exact timing of the different treatment periods allows to further explain what really happened during the program. Which part of the treatment plan did contribute in which way to the observed net effect? Second, this duration model approach allows to take *duration dependence* and *dynamic selection* into account. The latter is particularly of importance when analyzing post-unemployment recurrence outcomes as they base on a sub-sample of

job finders, which implies potential additional selectivity. Finally, this modeling approach allows to *estimate and quantify the employment stability effect* (in days of avoided future unemployment) within the same framework, which will be done in section 6.4.

5.1.1 Duration Model with Subsequent Treatment Periods

The subsequent steps of the treatment plan are modeled using a duration framework. For convenience, Figure 1 is reproduced below, complemented by the period-specific treatment effects which we will model. As described earlier (section 2.1.1), intensified counseling starts at t_0 , at t_1 treated job seekers join the coaching which lasts until t_2 .



Following the timing-of-events approach, with extension to an experimental setup with anticipation effect (Abbring et al. 2005), the proportional hazard (PH) model – later extended to a mixed PH model – is constructed as follows:

$$\theta_u(t_u|x, M_j, C_k, D_i, v_u) = \lambda_u(t_u) \exp(x' \beta_u + \sum_{j=1}^6 \tau_j M_j + \sum_{k=1}^{11} \gamma C_k + \sum_i \delta_i D_i(t_u) + v_u) \quad (1)$$

where θ_u is the exit rate from unemployment to a job and t_u is the unemployment duration. x is a vector featuring individual characteristics, including the control for unemployment history in the past 3 years. M_j represents a series of time dummies which control, in 2-months-steps, for the specific time and business cycle conditions at inflow into the sample. C_k are caseworker fixed effects and v_u represents the unobserved heterogeneity component which will be discussed in section 5.3. The component $\sum_i \delta_i D_i(t_u)$ contains the treatment effects of interest.

The duration dependence function $\lambda_u(t_u)$ is designed as a *piecewise-constant function*

$$\lambda_u(t_u) = \exp\left(\sum_k (\lambda_{u,k} \cdot I_k(t_u))\right) \quad (2)$$

where $k = 0, \dots, 5$ time intervals are distinguished and $I_k(t_u)$ represent time-varying dummy variables that are one in the respective intervals. Based on the descriptive hazard for the unemployment exit process (see Figure 2) I define the six time intervals as follows: 0-50/51-100/101-150/151-250/251-350/351+ days. Unemployment durations are exogenously censored at March 31, 2010 (end of observation window), if necessary. Note that the analysis in this paper focuses on *exits to job* rather than on general unemployment exits. This is done in the light of the results found in section 4 that the new policy significantly increased job findings. Thus, θ_u represents the individual job finding hazard²⁹. Accordingly, the non-censoring indicator in this model is 1 for individuals who found a job.

²⁹Estimation results for the case of general unemployment exit are available on request. Section 4 reports exits to other destinations.

Based on this setup, let's consider the *main model with specific treatment effects for every treatment period*. We specify a function of treatment period indicators $\sum_i \delta_i D_i(t_u)$ whereby $i \in \{a; c_1; c_2; c_3\}$ are the treatment effects by subsequent treatment period. Following the figure above, the treatment indicators in the hazard can be defined as follows: $D_a \equiv I(t_u \leq t_{c1})$, $D_{c1} \equiv I(t_{c1} < t_u \leq t_{c2})$, $D_{c2} \equiv I(t_{c2} < t_u \leq t_{c3})$, $D_{c3} \equiv I(t_{c3} < t_u)$, whereby all are conditioned on being in the treatment group.

Let us have a closer look at the content of the sequence of treatment effects: In the early stage of unemployment, from t_0 onwards, the *anticipation effect* δ_a is identified, due to the randomized treatment assignment at time t_0 . δ_a measures potentially two effects: first and foremost the pre-intervention effect, coming from the fact that the individuals in the treatment group are informed about the upcoming coaching program during their first meeting at the PES. Second, a presumably small additional effect may come from the early-stage intense counseling³⁰. δ_{c1} measures the *effect of being in the coaching program*, identified by allowing for a shift in the hazard at the time of entry into the program, t_1 . δ_{c2} and δ_{c3} measure the *post-program effects* of the coaching allowing for a further shift at time of program end, t_2 . Note that t_1 and t_2 are defined as the start or the end of the coaching program plus 14 days each. The reason to do so is that there is a certain delay between having found a job and finally exiting. The 14 days' delay allows to take this into account, such that successful job findings shortly before start or end of coaching are assigned to the right stage of treatment. Allowing for more flexibility, I split the post-coaching effect into an earlier one, δ_{c2} , and a later one, δ_{c3} . The latter starts 180 days after the end of coaching ($t_2 + 166$)³¹ and ends at unemployment exit (or censoring).

Note that, due to the dynamic nature of treatment effects we estimate, there is a potential difference between the effects of having really participated in the program and the intention-to-treat (ITT) effects. Since this paper aims at assessing the behavioral reactions on the supportive treatment, we will focus on the first type of effects. The difference between those and ITT can have two sources. On one hand, there could be intentional non-compliance, i.e. individuals who intentionally avoid the – compulsory – coaching participation. A detail analysis, see Appendix, finds however only a very marginal amount of 3.2% of possible cases of intentional non-compliance. However, the large majority of the non-participants are individuals who announced before t_1 to have found a job (which implies unemployment exit within some weeks) – thus, their non-participation is due to normal reasons of dynamic selection which apply as well to the control group. Therefore, the estimated dynamic treatment effects should not be subject to endogenous non-compliance bias. (Moreover, we will test for unobserved heterogeneity later on.) The ITT treatment effects of the main model are reported in the Appendix. The difference between these "gross program effects" and the reported participation effects can be read as an additional announcement effect (beyond t_1): accelerated job finding as a reaction on the information about the upcoming coaching and intensified counseling at early stage.

³⁰Remind that, in median, job seekers join the coaching after 50 days. Thus, the doubling of counseling frequency only implies one or two additional meetings up to that stage.

³¹The second post-coaching interval is chosen to start later than in the case of the survey-based analyses (90 days after coaching), because register data cover a longer time-span of unemployment (up to 2 years).

5.1.2 Advantages of Randomization in Timing-of-Events Models

The design of this experiment as an RCT brings at least three advantages for the clean identification (and interpretation) of the dynamic treatment effects. Randomization (i) identifies the treatment effect starting right at t_0 , (ii) relieves the no-anticipation assumption, and (iii) allows avoiding a separate modeling of the inflow into later treatment (coaching).

First, randomization at t_0 allows identifying a treatment effect that starts right at t_0 . This is not possible for non-randomized studies since they cannot distinguish between endogenous selection and the real treatment effect in the first period from t_0 onwards (Abbring et al. 2005). In contrast, randomized treatment assignment leads to a balanced distribution of unobserved characteristics at t_0 . This solves the selection issue at t_0 and provides therefore identification.

Second, randomization combined with exogenous timing of treatments and information simplifies the identification of later treatment effects. In the standard case of the timing-of-events approach without randomization, Abbring and Van den Berg (2003) show that the identification of the effect of a treatment starting at $t_1 > t_0$, i.e. a hazard shift at t_1 , requires the *no (probabilistic) anticipation assumption* which basically implies that the counterfactual hazards (for TG and CG) must be equal up to t_1 ³². In the case here, however, we encounter a different situation. Since the sample is fully balanced at t_0 and, in particular, the TG members have full information about the upcoming treatment periods, they can immediately and transparently act on this information – which is captured, without bias, by the anticipation effect δ_a . Thus, the no anticipation assumption is replaced by measurable *“perfect” anticipation*³³. Finally, this full-information-argument carries over to the later treatment periods: Conditional on observables, unobservables, the previous treatment history and full (ex-ante) information about the treatment plan, the anticipation about the treatment in the next period is captured by the treatment effect in the ongoing period.

A third advantage of the randomization and information properties is that they make the additional (joint) modeling of the inflow process into later stage treatment unnecessary. Thus, a test and control of unobserved heterogeneity is enough to cope with the issue of imbalanced dynamic selection. The issue is that inflow into later treatment stages is not necessarily random any more: the relative proportions of unobserved characteristics may change in a potentially different way in treatment and control group. In the standard case of timing-of-events (Abbring/van den Berg 2003), this potential imbalance is also driven by endogenous treatment assignment. This is not the case here. Due to randomization and exogenous ex-ante timing, the ongoing selection is *uncorrelated* to the propensity to enter the later treatment (coaching), *conditional* on the anticipation effect. In other words, the anticipation effect captures changes (related to early treatment) in the propensity to enter later treatment³⁴. Again, this argument carries over to all the later

³²This could be expressed (in simplified notation) as $\theta^T(\tau_0|x, v_u) = \theta^C(\tau_0|x, v_u)$ where θ^T and θ^C are the counterfactual hazard rates at time $\tau_0 \in]t_0, t_1[$. Note, moreover, that the no anticipation assumption refers in fact to no probabilistic anticipation. Deterministic anticipation, i.e. acting on information which is available to everybody at t_0 (like general monitoring behavior of the PES or generally distributed information on a program etc.), does not break the assumption since this information is equally available for treatment and control group. See Arni et al. (2013) for a further discussion and example.

³³So, more formally, the equality $\theta^T(\tau_0|x, v_u, D_a) \exp(\delta_a) = \theta^C(\tau_0|x, v_u, D_a)$ holds here and describes perfect anticipation – as compared to the no anticipation assumption in footnote 32 (using the same notation as there).

³⁴This means that for our main model (1) here the following orthogonality applies: $v_u \perp\!\!\!\perp D_{c1}|x, D_a$. If this

stage treatment parts (D_{c1}, D_{c2}, D_{c3}). By the same line of argumentation, one can conclude that as well issues of potential non-compliance (which are a particular case of imbalanced dynamic selection) can be handled in the same, simplified way.

5.1.3 Extension to Post-Unemployment Job Stability

It is now straightforward to extend the outlined framework to the post-unemployment period. An analog (M)PH model is set up to estimate the causal impact of the supportive ALMP on employment stability. This crucial dimension of post-unemployment jobs is assessed by modeling the recurrence propensity, i.e. the transition rate back into unemployment:

$$\theta_p(t_p|x, M_j, C_k, D_i, v_p) = \lambda_p(t_p)\exp(x'\beta_p) + \sum_{j=1}^6 \tau_j M_j + \sum_{k=1}^{11} \gamma C_k + \delta_p D_p + v_p \quad (3)$$

whereby t_p is defined as the duration from the time of transition from unemployment to a job to the time of reentry into unemployment. The transition (or non-censoring) indicator is therefore 1 if a reentry to unemployment is observed up to 1.5 years (540 days) after unemployment exit (exogenous censoring). As in model (1), the baseline hazard rate $\lambda_p(t_p)$ adopts the form of a piecewise-constant function³⁵. D_p is a dummy variable indicating the treatment group. This means that one constant treatment effect³⁶ is estimated for the post-unemployment period.

It is important to note that equation (3) is estimated on the sub-sample of individuals who found a job after unemployment. The implications will be discussed in section 5.3.

5.2 Dynamic Treatment Effects on Behavior

To set up the econometric analysis of the dynamic treatment effects on different behavioral channels, we again follow the different stages of the treatment plan: anticipation ($\tau = 1$), during coaching ($\tau = 2$), up to 90 days after coaching ($\tau = 3$), beyond 90 days after coaching ($\tau = 4$). The fact that the timing was fixed ex ante and communicated at t_0 provides the base to identify separate treatment effects by treatment period (Abbring et al. 2005). The sequential strategy is refined by use of a dynamic *difference-in-differences* (DiD) of the following type:

$$y_i = \alpha + \gamma^{TG} D_i^{TG} + \gamma_\tau T_\tau + \delta_\tau D_i^{TG} T_\tau + x_i' \beta + \varepsilon_i \quad \text{for } \tau = 1, \dots, 4 \quad (4)$$

whereby D^{TG} is a dummy variable for the treated, T_τ a time period indicator and x the set of the control variables. The coefficient of key interest is the DiD parameter δ_τ which measures the treatment effect in period τ of the intervention on a certain behavioral outcome y .

The sequential DiD provides several advantages to minimize potential biases and to improve the precision of the estimated effects. First, it allows to capture the time trend of behavior in a flexible way, without imposing a functional form. The time effect, γ_τ , captures changes in levels of the behavioral variables over time which are common to the treatment and control groups.

independence is given, no further equation is necessary to model the relation between later treatment inflow and unobserved heterogeneity.

³⁵Following the shape of the descriptive hazard, I estimate four intervals with splits at 210/390/480 days. Note, moreover, that I define a recurrence event as being at least 20 days out of initial unemployment before reentry.

³⁶As a sensitivity analysis, I implemented a more flexible specification which allows for a shift of the treatment effect after 270 days. The two estimated treatment effects were not significantly different in size.

Second, sequential DiD corrects for ex-ante differences in the behavioral outcomes. Even though groups are randomized at t_0 (and randomization worked well, see section 3.3), it can happen by chance that the initial levels of some behavioral variables are not fully balanced; γ^{TG} takes this into account. Finally, the fact that the t_0 data are part of the estimation sample substantially helps to properly identify the effects of x and T_τ and to improve the estimation precision. In case there is an issue of imbalanced dynamic selection in later t_τ , the potential bias could also affect the estimates of β (if the imbalance is correlated to individual characteristics). Thus, using the unbiased t_0 data helps obtaining correct estimates for β . This, in turn, helps reducing potential bias in δ_τ , since the selection on observables is properly taken into account.³⁷

As mentioned earlier, potentially there could be an issue of unbalanced dynamic selection between TG and CG in later stages of the treatment plan: if unemployment exits (and survey non-response) differently affect the composition of *unobservables* relevant for the corresponding outcome. The outlined design cannot fully account for this issue. Therefore, we will perform a series of tests to assess whether there is indication of imbalanced unobservables. The first assessment is to analyze the balancing of observables by treatment period: this is done in Appendix B.1. It shows that also in the later treatment periods almost no observables are imbalanced (exception is, occasionally, nationality). This suggests that the initial randomization (plus the homogeneity of the initial sample) translated to a considerable degree to the later periods. Thus, if the composition of unobservables and observables are correlated, this result bolsters confidence that imbalance is not significantly biasing the later-stage treatment effects on behavior. Further robustness tests are described in the following section.

5.3 Robustness Analyses

5.3.1 Unobserved Heterogeneity

In the context of the duration models for the job- and employment stability outcomes we can directly check whether the composition of unobservables affects estimated treatment effects. The timing-of-events models can be extended to allow for *unobserved heterogeneity*.

We apply the standard non-parametric way of introducing unobserved heterogeneity which consists in modeling a discrete mixture distribution (Heckman/Singer 1984) for v_u and v_p in equations (1) and (3). We implement a version of the non-parametric maximum likelihood estimator (NPMLE) as proposed by Baker and Melino (2000). To start, I choose a design that allows v_u and v_p to have two points of support. For the cases one process (unemployment) and of two correlated processes (incl. post-unemployment), this implies the estimation of the following probabilities of mass points:

$$p_n = P(v_u = v_u^n) \quad \text{with} \quad n = 1, 2 \quad \text{if only process } u \quad (5)$$

$$p_j = P(v_u = v_u^n, v_p = v_p^n) \quad \text{with} \quad j = 1, \dots, 4 \quad \text{if adding process } p \quad (6)$$

These probabilities are designed in a logistic form, i.e. $p_n = \frac{\exp(a_n)}{1+\exp(a_1)}$ for the case (5) and $p_j = \frac{\exp(a_j)}{1+\exp(a_1)+\exp(a_2)+\exp(a_3)}$ for the case (6) (normalizing one parameter to being 0). Thus, a

³⁷Note that the sequential estimation procedure as outlined in equation 4 is more suitable to fulfill this last argument than estimation of all the five periods jointly in one model. In the former case the certainly unbiased t_0 period has more relative weight than in the latter.

maximum of two/four probability parameters a_n/a_j and of two/four baseline hazard intercepts λ_0^n/λ_0^j is estimated in the 1/2 process/es model. We model the correlation of unobservables across the two processes by allowing for all possible combinations of mass points. Like that, we account for selectivity by unobservables, which is generated by the potentially selective sub-sample of job finders. To find the optimal locations, proportions and numbers of mass points, I use an iterative grid search procedure – see Appendix C for details.

5.3.2 IPWRA Estimation

Besides testing for unobservables, we apply an inverse probability weighting estimator with regression adjustment (IPWRA) on the dynamic behavioral variables as a further robustness check. IPWRA can be used to address a variety of selection problems, including that inherent in estimating average treatment effects; nonrandom selectivity and attrition are leading examples (see e.g. Wooldridge 2002 for an overview). Based on the setup represented by equation 4, we weight the data in the periods $\tau > 0$ by inverse probabilities and then perform the same regressions. The corresponding weights ω_i are computed as

$$\omega_i = \frac{D_i^{TG}}{\hat{p}_i} + \frac{1 - D_i^{TG}}{1 - \hat{p}_i}$$

whereby the propensity score $p(x_i)$ is estimated by probit. The results of the robustness checks are discussed in section 6.3.

6 Dynamic Results on Behavior, Exit Rates, Mid-Run Outcomes

This section aims at providing insights about *how* the randomized supportive ALMP affected the *dynamics of outcomes and behavior over time*. These findings let us track directly how treatment "shocks" gradually provoke reactions in behavior and in subsequent outcomes. In the following, we proceed in four steps: first, we discuss the effects on the transition from unemployment to jobs as well as on the job interview rate as early outcome. Second, we assess the treatment effects on the whole series of behavioral channels we consider. Third, the results on mid-run outcomes are discussed, followed by robustness analyses.

6.1 Outcome: Causal Effects on Unemployment Exit Behavior

6.1.1 Transition from unemployment to jobs

We start by discussing the dynamic causal effects of different stages of the coaching and counseling treatment on transitions to jobs.

How do the specific treatment effects by treatment period look like? Table 3 reports these results which are based on model (1).³⁸ The dynamics of the treatment effects reveals indeed a pattern which was not yet visible in the nonparametric analysis (due to overlaps of treatment

³⁸The results for the control variables are reported in Table 11 in the Appendix. Note that the estimation very appropriately fits the shape of the empirical hazard (see Figure 2). When computing the piecewise-constant baseline hazard rates for an "average" individual (see Notes of the Table 11 for the specific calculation) over the different duration pieces, we find that the monthly unemployment exit rate goes from 6.4% to about 15% and then down to 8% and less from 151 days on – as in the empirical hazard.

periods): The found zero net effect on unemployment duration was, in fact, generated by the interplay of a period of lower exit rates, followed by one of higher exit rates. The *anticipation effect* (δ_a) is highly significantly negative. Treated individuals have an on average 37.6% ($= \exp(\delta_a - 1)$) lower transition rate to jobs in the period between unemployment inflow and (potential) coaching entry. Thus, the prospect of being coached obviously results in a smaller propensity to exit early to a job. This suggests that the treated individuals seem to expect a positive outcome or at least some helpful support of the coaching program. Therefore, one may call this negative anticipation effect an *"attraction effect"* – as an opposite to the commonly found *"threat effect"* in the analysis of other kinds of labor market programs (see e.g. Rosholm/Svarer 2008). The analysis of search behavior, as introduced in the next section, can provide some empirical insights if this *"attraction effect"* is rather driven by a smaller job search effort or by being more picky in accepting jobs.

[Table 3 about here]

In the next treatment period – during coaching – a slightly significantly (p-value .12) negative impact on the exit rates is found. Thus, the commonly found *lock-in effect* is present here as well. Individuals participating in the coaching program seem either to exert less job search effort than without coaching, presumably due to the high work load of the program. Alternatively or complementarily, the during-coaching-effect could operate via affecting other channels of behavior that make people find and consider less job offers. Again, the subsequent behavioral analysis will deliver more insights on this open question. In any case, it turns out that the *"lock-in"* effect is restricted to the short time span of the duration of the coaching (60 days in median). Right after, the *early post-coaching effect* (δ_{c2}) is already back to zero. Thus, the coaching design principle *'intense but short'* yields to be beneficial in restricting the lock-in effect.

Six months after coaching end, the treatment effect (δ_{c3}) reveals to be clearly positive but insignificant. The higher exit rate to a job of the coached reflects the insight of the nonparametric analysis: that in later stages of the unemployment the positive impact of the supportive ALMP kicks in. However, given the small sample, the exits to jobs are quite dispersed over time beyond 181+ days after coaching, and thus the standard error is comparably high. As a consequence, the estimated δ_{c3} marginally fails to get significant. However, the dynamic analysis demonstrates how the cross-sectionally significant effect on job finding proportions came about.

Beyond, it is interesting to briefly consider the ITT analysis of the dynamic effects. The dynamic ITT effects are driven as well by those treated individuals who already announced to have found a job and therefore did not enter or not terminate the coaching program. Taking these into account, we estimate the *"gross"* program effect. Consistently, the results reported in Table 12 in the Appendix show more positive treatment effects from δ_{c1} onwards. The later post-coaching effect (δ_{c3}) turns out to become significant at the .15 level. There could be two interpretations of these additional impacts on top of the participant-driven effects: On one hand, they could reflect the *"longer-run anticipation effect"* of the information about the upcoming coaching; i.e. some people reacted on it by early job finding (but then it took a while until they started the new job). On the other hand, the supplementary effect could be caused by the early intensified counseling which may have led to additional job findings.

So, wrapping up, one can state that the nonparametric result of more job finding can be decomposed as follows: on one hand into an attraction effect and a slight lock-in of coaching

which tend to prolong unemployment duration – on the other hand, into a positive post-coaching effect which boosts transitions to jobs due to the supportive treatment.

6.1.2 Effects on the job interview rate

A potentially suitable early indicator of job finding is the propensity to acquire job interviews. Based on the repeated surveys we consider how supportive treatment affects the job interview rate, which is defined as the number of acquired interviews per sent application.

[Figure 6 about here; Table 4]

The job interview rate indeed seems to be a suitable early indicator of successful job finding. The intensified counseling and coaching affects the amount of acquired interviews in early stages already, as Figure 6 shows. The figure plots predicted conditional means of the treatment and control groups and documents the significant treatment effects. Starting at a level of 0.11 interviews per sent application at time t_0 , the treatment effect gradually increases until it becomes significant during coaching, reaching about 0.17. The conditional mean interview rates among the control group appear to have a slight downward tendency, ranging from 0.11 to 0.08.

A direct comparison of the treatment effects on the job interview rate and on job finding by treatment period (see Figure 6) demonstrates well the role of the former as an early indicator. The positive impacts on interviews seem to translate gradually into job finding.

6.2 Channels: Causal Effects on Behavior

This section documents the results representing the dynamic treatment effects of the coaching & counseling intervention on the different behavioral channels. Subsequently, we will discuss effects on search behavior, on beliefs and finally on aspects of non-cognitive skills.

6.2.1 Effects on search behavior

Before we start analyzing the key dimensions of search behavior, let us address a basic question: Was the content of coaching & counseling, as described in section 2.1.1, successfully transmitted to the job seekers? An important part of the content was dedicated to discussing search strategy and search efficiency optimizations. The measure of *search strategy extensions* offers, thus, a direct opportunity to check whether the treated individuals agreed with the coach and/or caseworker to extend the scope of search (for details see section 3.2.2). The Appendix Table 13 shows a distinct picture: Whereas the propensity to extend the scope of search is close to 0.2 for the treatment group (TG) and the control group (CG) during anticipation and after coaching, it is increased as much as 0.4 for the treated during coaching. Thus, this component of the coaching strategy has been substantially transmitted to the participants.

The first fundamental dimension of search we consider is the quantitative *search effort*. The most striking result is that the treated individuals *never searched more than the control group*. On the opposite: The treated searched significantly less in anticipation and during coaching (and also in the 90 days thereafter, however marginally below significance), as Table 5 documents. Figure 7, which reports the conditional predicted means, shows that the CG searches beyond the start t_0 at the level of about 9 applications per month, whereas the TG submits 7 to 8

applications. Now, let us overlap these treatment effect results with the findings on transitions to jobs in section 6.1.1; from this we can conclude the following: The "attraction" effect – i.e. reduced job transitions in the anticipation period – is underpinned by a significant reduction of search effort, as well as the lock-in effect during coaching. Also beyond coaching end the treated never exert more effort than the CG. Thus, the additional job finding success of the treated is not due to quantity but due to quality or *efficiency* of search.

[Figure 7 and Table 5 about here]

A second indication that the treatment positively affected search efficiency is found in the results on the *variety of used search channels*. We find here an analog picture, i.e. that the treated never increased channel variety, but some time they reduced it. In the early post-coaching period, the reduction becomes significant, the effect amounts to -1.2 channels. Whereas the CG individuals use a broad variety of search channels – about 8 out of 11 choices (see Figure 7 and section 3.2) – the treated seem to search in a more directed way.

How did the treatment affect *channel choice* and the *frequency of channel use*? The available data allow the analysis of these questions by looking at the results for each channel of search separately. This is done in Table 5 and Figure 7, where I report the six most important search channels. A first observation is that the negative signs on the DiD coefficients clearly prevail. Thus, as observed for the effort and channel variety dimensions of search, *frequencies of use are in tendency reduced* and not increased. We can distinguish three formal channels – newspapers, internet and private recruiters – and three informal channels – network (weak ties) and spontaneous applications by telephone or by mail. The most prominent result is that the treatment caused several *significant reductions of frequencies of use of formal channels*. In particular, we can observe two patterns. First, whereas the control group tends to increase frequency of newspaper and internet use after t_0 , the treated report significantly lower frequency levels in the anticipation period. The "attraction" effect is reflected here in the sense that, due to the prospect of being coached soon, the treated tend to wait with adapting formal search upwards. Second, the findings on post-coaching effects show that the coaching made the treated refrain from increasing the frequency of formal search. Consistent with the above-discussed results on search behavior, these findings suggest a treatment effect towards more directed or efficient search.

On the side of the *informal channels*, different patterns of reaction are visible. The impact of the supportive ALMP on the use of personal *networks* is zero over all periods. A highly significant and quantitatively important (plus 44 percentage points) upward move is found for *spontaneous applications by telephone* during the coaching period. This has to be linked to the fact that the coach explicitly promoted this type of spontaneous acquisitions. On the opposite, a significantly lower use of *spontaneous written applications* can be observed among the TG individuals after the end of coaching. Looking at the time dynamics, we observe that the CG individuals significantly increased the frequency of use of this channel relative to t_0 . This increase is not visible in the TG – in the anticipation period presumably due to the documented "attraction" or "waiting" behavior, post coaching probably due to the mentioned coach suggestion in favor of phone applications.

[Figure 8 about here]

As a second fundamental dimension of job search behavior, we analyze the evolution of *reservation wages*. In fact, the empirical measures report reservation salaries (i.e. minimal monthly gross earnings that still would be accepted by the job seeker). Let us first consider the control group behavior: Figure 8, supported by the estimates in Table 5, reveals that the control individuals are only willing to reduce the acceptable wage level in the early stage of unemployment. This pattern is consistent with the typical unemployment exit rate profile over time: the exit rate peaks in the first months and then goes down (and only peaks again at exhaustion much later). Thus, the fact that the CG reservation wage profile is not downward-sloping reflects a potential problem: that job seekers are not aware or not willing to adapt their wage requests to the depreciation of their human capital and employability. Note that dynamic job search theory would suggest a downward-sloping reservation wage profile, if the model allows for the adaptation of human capital value (and of the pool of job and wage offers) over the spell time.

Comparing the reservation wage behavior of the treated to the one of the controls, we find two striking differences. First, the TG does *not* reduce the reservation salary level in the anticipation period. I.e., it remains substantially higher (difference significant for median salary levels, marginally insignificant for ratio to last salary). This suggests that as well the reservation wage behavior seems to react in a way consistent with the found "attraction effect": people are not willing to accelerate their job finding in the early period because of the prospect of being coached. As a second treatment effect we find a *significant reduction of reservation salaries in the 'during' and 'early post-coaching' period* (which seems to continue later on, but results are insignificant due to higher standard errors). It is instructive to compare the predicted reservation salaries, see Figure 8, with the realized median salaries in the new jobs: the latter amount to 5470/5280 CHF in the TG/CG. It turns out that the CG reservation salaries in later unemployment stages are above the realized ones, whereas the TG reservation salaries are below. Thus, this suggests that a key treatment effect of the coaching program consisted in *reducing (perceived) reservation wages to a more realistic level*. I.e., the coaching may have helped the treated to take into account that unemployment is often linked with human capital and wage loss, in particular for ages 45+.³⁹

6.2.2 Effects on beliefs

Next, we analyze the dynamic treatment effects of the supportive ALMP on individual beliefs. First, consider *beliefs about job chances*, measured as the deviation between expected and realized job interviews (per application). A look at the CG reveals that the overestimation of interview chances gradually increases in the first periods, up to 0.4 (see Figure 8). Individuals do not seem to take into account that the job interview rate tends to decrease over the spell – on the opposite: they increase expectations. This behavior of increasing overestimation is not visible for the treated. Even though they indeed manage to increase their interview rate due to the treatment, they do not increase overestimation but keep the degree of bias constant. This treatment effect

³⁹We cannot fully exclude that there is a certain degree of overreporting in the surveyed reservation salary figures. However, due to the randomized trial, possible initial overreporting is supposed to be balanced. Note that even if you would assume that a part of the coaching effect consists in a reduction of overreporting behavior, this would not essentially change the interpretation of the treatment effect. At the end, it is highly likely that possible overreporting behavior is correlated with (and driven by) how individuals perceive their own reservation wage level and what their wage expectations are. So, the basic interpretation that some CG individuals may not have accepted certain jobs due to some form of biased wage perception, remains.

is only significant in the short run (during coaching period), however. In later stages, the CG individuals seem to realize their large degree of overestimation as well. Thus, the coaching exerts a *short run effect through stabilizing the degree of bias in beliefs* about job chances.

[Table 6 about here]

As a second type of search-related beliefs, we consider *wage expectations*, measured on a five-point-scale. The main finding is that the TG tends to reduce wage expectations from early on, whereas the CG remains with the same level of belief for a long time. On (conditional) average, the latter report a scale level of 3 – i.e., they expect a pay equal to before unemployment. In the early post-coaching period the difference becomes significant. Like in the case of beliefs about job chances, the coaching treatment causes a *reduction of the wage belief in the short run* (until 3 months after coaching).⁴⁰

Thus, in all, the findings here suggest that a coaching treatment can reduce (biased) beliefs in the short run, but we cannot observe evidence for a persistent reduction of bias.

6.2.3 Effects on non-cognitive skills

As a further type of behavioral reactions we assess now some aspects of personality-related skills. First, let us consider *performance-motivation*. The results provide evidence that the “*attraction*” effect in the anticipation period affects the lower tail of the motivation distribution: we find a significant increase in the probability of low or very low motivation to search (results on request). Moreover, *coaching affects the motivation* of the treated as well *positively*, as Figure 8 demonstrates. The probability for job seekers to be very highly motivated is significantly higher in the 3 months after coaching. The difference is substantial: whereas about 26% of the CG reports the highest motivation level, the (predicted) proportion among the TG amounts to almost 47%⁴¹. This motivation boost tapers off later on.

The *self-confidence* of job seekers (as perceived by caseworkers) is marked by an up and down, as a look on the control group behavior reveals (Figure 8). This volatility in self-confidence seems to be *stabilized* by the coaching and counseling treatment. The probability of high self-confidence among the TG individuals turns out to be higher, whereby the difference to the CG becomes significant in the early post-coaching period.⁴²

Finally, we consider a dimension of personal self-control: *reliability*. In the CG, a proportion of about 60% is assessed (by the caseworker) as being highly reliable. Interestingly, this proportion increases in later stages of unemployment. One way to explain this phenomenon is that being regularly subject to monitoring and caseworker meetings may positively affect the job seeker’s self-control. The treatment effect, which is relatively small, appears through an *earlier increase* in the individual’s reliability levels, the latter becoming significant during coaching.⁴³

⁴⁰If we break up the scale and do not consider means, we find the following: In the anticipation period, the treatment significantly reduces the probability of high wage expectations (scale levels 4 and 5). In the early post-coaching period, the probability of low wage expectations (levels 1 and 2) is significantly increased.

⁴¹The average difference in scale levels is significant too and amounts to .32 in that period.

⁴²The same significance appears as well if we consider the probability of low self-confidence or the scale level. The CG volatility is also (significantly) visible in all the three outcome specifications.

⁴³We don’t find significant differences in the probability of low reliability of in the average scale level; but the same patterns are, in tendency, visible.

To wrap up, it can be stated that we find some meaningful patterns of significant treatment effects on the considered aspects of non-cognitive skills. The strongest impacts are on performance-motivation. However, the measurable effects focus on the short run, mostly during and in the first three months after coaching. Given that we subsequently found significantly increased job finding, this pattern suggests that the short-run boost in becoming more motivated and more disciplined may have been beneficial for the success of search activities later on. However, this cannot be directly shown within this study.

6.3 Robustness Analyses

As discussed in section 5, the dynamic analysis could potentially be affected by a bias in later treatment periods and post-unemployment, due to possible imbalance of unobservables. We run three different robustness tests to assess this issue.

Using the estimation models for the dynamic treatment effects on durations, we can directly check for the (non-)importance of unobservables. We implement the extension for unobservables as outlined in section 5.3. The grid search as part of the non-parametric MLE procedure (following Gaure et al. 2007, see Appendix C) resulted in suggesting a 2-mass-points model as the best choice for estimating equation (1)⁴⁴. This estimation delivers a log likelihood of -1536.16 – whereas the model *without* unobserved heterogeneity yielded a log likelihood of -1455.45 (see Table 3). Therefore, the conclusion is that for our 1-process model there is *no gain in explanatory value by adding unobserved heterogeneity*.

The same procedure is applied to the 2-processes model addressing employment stability, which combines equations (1) and (3) with the unobserved heterogeneity specification (6). The resulting best-choice-specification is reported as estimation 2 in Table 7. Two of the four possible mass point combinations turn out to be non-zero. But again, the log likelihood of -1987.05 is lower than the one resorting from estimation of the 2-processes model without unobserved heterogeneity (log likelihood of two independent processes, i.e. $-1455.45 + (-459.05) = -1914.5$, see Tables 3 and 7). Therefore, the conclusion for the 2-processes model is as well that no gain in explanatory value can be achieved by adding unobserved heterogeneity.

Thus, the analysis of unobserved heterogeneity models reveals that the *size of imbalance in unobservables due to dynamic selection is statistically not relevant here*. We therefore conclude that the models without unobserved heterogeneity can be interpreted causally. There are different possible reasons for the non-importance of unobserved heterogeneity in the context of this study. First, the tight sampling criteria applied in the pre-selection into the sample may have avoided the generation of large heterogeneity. Second, the selection caused by the found treatment effects could be of a balanced nature: i.e., the individuals who found a job due to the program are not fundamentally different from the job finders in the control group.⁴⁵

Does this statistical irrelevance of potentially imbalanced unobservables for outcomes as well apply to the behavioral channels, which are measured by repeated surveys? This is well plausible, if we assume that the unobservables which potentially affect the balancing of behavioral variables

⁴⁴Grid search for a third mass point did not provide any specification yielding a higher log likelihood (on request)

⁴⁵Finally, one could argue that the non-identification of further mass points may be due to the small sample size. However, this is not very probable since Monte Carlo simulations in Baker and Melino (2000) have shown that it is well possible to identify several mass points with 500 observations.

are similar (correlated) to the ones of the outcomes. However, we cannot fully test for this since the framework used for the dynamic behavioral analysis doesn't allow for unobservables. Still, we can run two robustness checks on the balancing of the behavioral variables. The first is to check the balancing of observables in later treatment periods. This analysis has been discussed in section 5.2 and documented in Appendix Table B.1. The observables didn't show any hint of significant imbalance in later stages.

A final robustness check is to re-run all the behavioral regressions by use of inverse probability weighting (IPWRA), see section 5.3. The results are reported in Appendix Table 14. We do not find substantial deviations of the resulting point estimates.⁴⁶

Thus, in total we can conclude from the robustness analyses that we do *not* find empirical indication for (unobservable or observable) imbalance of relevant size and correlation to outcomes or behavioral variables.

6.4 Mid-Run Outcomes: Employment Stability & Earnings

Next, we want to analyze the effects of the supportive ALMP treatment on mid-run labor market outcomes. What does the result that more treated individuals found a job mean for the quality of the found jobs? Nonparametric evidence that compares employment stability and gross salaries across TG and CG suggests a positive but non-significant effect on stability and a zero difference in monthly earnings (see Table 2). However, these results are based on the sub-sample of job finders which potentially is subject to endogenous selectivity. Therefore it is important to analyze these two dimensions of job quality under control for observables and unobservables. The results of last section showed that including unobserved heterogeneity does not increase the explanatory value of the estimation. Due to this insignificant importance of heterogeneity, the best choice is to use the specification without unobserved heterogeneity for the conclusive analysis.

[Table 7 about here]

Table 7 (estimation 1) reports results for employment stability when explicitly modeling post-unemployment job duration and controlling for observables. We find a *significantly positive treatment effect on employment stability* over 1.5 years after unemployment⁴⁷. This result will be further quantified below. For the sake of comparison, Table 7 also reports the results for the model with unobserved heterogeneity (but less explanatory value): it shows that the treatment effect on employment stability does qualitatively not change, it gets slightly stronger.

The second dimension of post-unemployment job quality that we assess is earnings. To check whether the conditioning on observables changes the results of the comparison of median salaries, we run the corresponding regression (available on request). It turns out that the result of *no significant differences in monthly salaries between TG and CG* remains.

⁴⁶The standard errors of the IPWRA results are larger than in the main estimations. They are too high due to the fact that they are not corrected for the use of a prediction (estimated propensity score) within the estimation procedure (Wooldridge 2002). Therefore, they are not used to assess significance; however, the focus of this robustness check is anyway on the point estimates.

⁴⁷I estimated as well a model which splits the treatment effect at 270 days after unemployment. This didn't yield statistically significant differences in the effect size.

7 Discussion: Cost-Benefit, Comparison, Channels

7.1 Cost-Benefit Analysis

Using the estimation results in Table 7 it is possible to quantify the positive impact of the supportive ALMP on employment stability in terms of avoided future unemployment duration. This amounts to calculating the expected values of the post-unemployment duration t_p for the two counterfactuals. The difference between the two yields the average treatment effect on the treated (ATET); it quantifies the *not* realized future unemployment (in days) due to the treatment. Using equation (3), I simulate the following density of post-unemployment employment durations:

$$f_p^D(t_p|x, v_p) = \theta_p^D(t_p|x, v_p)S_p^D(t_p|x, v_p)$$

whereby $D \in \{T, C\}$ indicates the treatment status, i.e. the two counterfactuals. θ_p represents the hazard derived in equation (3), S_p is the corresponding survivor function. Based on this density, the expected value of the employment duration can be calculated as

$$E(t_p|x, v, D_p) = \int_{20}^{\eta} t_p f_p^D(t_p|x, v_p) dt_p + \left[1 - \int_{20}^{\eta} f_p^D(t_p|x, v_p) dt_p \right] \cdot \eta \quad (7)$$

The last part of this equation takes into account that the employment durations are exogenously censored after 540 days or the end of the observation window (after η days). Running the simulation for each of the counterfactuals, we obtain the $ATE_T = E(t_p|x, v, T) - E(t_p|x, v, C)$. We find that, on average, *treated individuals avoid future unemployment of 23 days*. Based on this quantification of the direct benefit (savings) for UI and on additional cost information by the PES administration, we can perform a cost-benefit accounting in order to assess whether the investment in the – quite costly – supportive ALMP pays off or not.

[Table 8 about here]

Table 8 provides the details on this cost-benefit analysis for the UI accounts. It yields a clearly positive result: The avoided future unemployment *pays the additional cost of the new program more than fully*, specifically it covers 1.7 times the additional cost. Thus, in the medium run, one can state that this type of supportive ALMP resulted in (more) jobs of similar salary and increased stability, whereby the latter can pay off the additional program costs.

7.2 Comparison to External Control Group

Besides the issue of cost-effectiveness, another general question that arises for policy design of supportive ALMP is how treatment effects adapt when the program gets scaled up. A comprehensive assessment of external validity and of possible general equilibrium effects is beyond the scope of this paper⁴⁸. However, we can take up one aspect of the former by comparing the baseline results on (dynamic) job finding to external control groups. This comparison, based on register data, allows for at least two useful insights: How comparable is the region, in which the

⁴⁸Given the small size of the RCT, as compared to the dimension of the whole urbanized area (see section 2.1.2), it is rather implausible that displacement effects or externalities on other regions have been realized.

RCT has been implemented, to neighboring regions when it comes to labor market outcomes? Are there some spill-over effects on the control group? It offers as well some indication if the found treatment effects were rather at the lower or upper bound.

The comparison exercise consists in analyzing double-differences in nonparametric survivor curves on the propensity to remain unemployed: we consider, on one hand, the difference between the external control group of neighboring PES regions with the RCT groups and, on the other hand, between the age group of the experiment and a younger one, aged 30 to 44. To select the external control group, the same eligibility criteria as in the RCT (see section 2.1.2) are applied to the other 4 PES within the same state (canton of Aargau).

The results are represented in Figure 10. For the younger age group, the unemployment survivors of the RCT- and the external control regions almost exactly overlap. Thus, the regions are highly comparable in terms of unemployment durations. Looking at the comparison within the age window of the experiment, however, we find a visible difference between the survivors: the control group of the RCT tends to leave slightly faster than the external controls.

This suggests two conclusions. The labor market conditions in the RCT sample are not particular as compared to other regions in the same larger metropolitan area. More interestingly, the fact that the RCT control group performs slightly better than the external controls may hint to two phenomena: a Hawthorne effect – people may have felt under additional monitoring due to the repeated surveys – or positive spill-overs on the control group (e.g. by caseworker learning, exchange among job seekers). In both cases the comparison exercise suggests that the found treatment effects are rather at the lower bound. They could have been higher when abstracting from these external effects (which may be less important in a scale-up without surveys⁴⁹).

7.3 Channels of Effects: Explorative Assessment

The unique data on repeatedly observed behavioral channels allowed us to collect evidence on how supportive ALMP can shape different behaviors. How can these findings be related to theoretical thoughts about modeling job search behavior? Comprehensive assessments of this question go beyond the scope of this paper. But it is still instructive to discuss some hypotheses on how to integrate the (additional) behavioral channels in a search framework. We want to focus on two questions, related to main findings: To which degree can the considered non-cognitive skills be seen as separate, additional channels to explain search outcomes? Second, with which theoretical ideas are the findings on reservation wage (and wage expectation) dynamics consistent?

7.3.1 Job offer arrival rate: the role of non-cognitive skills

Let us consider the job search model as sketched in section 2.2. Our discussion suggests that the considered aspects of non-cognitive skills predominantly operate via affecting the job offer arrival rate: $\lambda[s(e_t, \sigma_t^1, \dots, \mu_t), \gamma_t, \phi_t]$. Understanding non-cognitive skills as a separate channel to affect λ would mean empirically that μ_t , γ_t and ϕ_t have explanatory value *in addition* to the variation which is explained by e_t and σ_t^1, \dots . This can be roughly checked by a series of regressions, which first include only the non-cognitive component as explanatory variable and then add the search variables as further explanatory channel. As the dependent variable we will use the job interview

⁴⁹On the other hand, there is a potential for displacement effects when implementing coaching on a broad scale.

rate. It can quite plausibly proxy the job offer arrival rate: a job interview is the sign of a job offer that successfully passed the first matching between employer and job seeker.⁵⁰

Table 9 presents the results of this explorative assessment⁵¹. The first two columns demonstrate that motivation is significantly correlated to the job interview rate after controlling for search effort and search channel variety. The point estimate gets a bit larger than without controlling for the latter. The R^2 increases visibly. Thus, the motivation measure provides additional explanatory value for the job offer arrival process, beyond of just being another representation of search effort. The same can be observed for the aspect of self-confidence. For the case of reliability we do not find a linear correlation to the job interview rate across time periods.

In conclusion we can state that there is some significant empirical indication that aspects of non-cognitive skills can be seen as *separate channel to "explain" job offer arrival*, beyond the "classical" search variables. Motivation and self-confidence may directly affect the decision which job openings are seen as being suitable and thus followed up. Note that we only considered a simple reduced-form framework; the functional forms on how μ_t , γ_t and ϕ_t affect search outcomes could of course be different. However, the results suggest that it may be useful to further consider the question how to integrate such aspects in job search models.

7.3.2 Explaining the coaching effect on wage behavior

One key result of the analysis is that the coaching treatment had a reductive impact on reservation wages. In addition we find a short-lasting negative effect on wage expectations. With which types of theoretical explanations could these empirical results be consistent?

Following up on the discussion in section 2.2, we can conceptually distinguish three types of mechanisms by which a supportive ALMP could affect wage behavior: First, it is plausible to assume that the treatment by coaching and counseling produces disutility in the short run – through cost of time and work in the program, pressure, etc. This reduces the value of remaining in unemployment and, consequently, the reservation wage r_t . Second, the distribution of wage offers which are (perceived) relevant to the job seeker may change from $F(w)$ to $\tilde{F}(w)$. The fact that coaching may be human-capital-improving calls for an upward adaptation of $F(w)$. However, another function of coaching is information update. Coaching participants may have corrected their misperception or missing information on which wages they realistically will get offered. Thus, $\tilde{F}(w)$ would be lower in mean than $F(w)$. As shown by Burdett et al. (1988), such a learning process would result in a downward-sloping reservation wage path. A third type of mechanism could operate via reference points. Following the model of reference-dependent job search by DellaVigna et al. (2015), the utility of being in UI, $u(b_t|\psi_t)$, is complemented by a gain-loss component⁵². The reference point ψ_t is for example a function of past income and expectations of upcoming income. If coaching operates as a downward shock on ψ_t , comparably more gain utility would be obtained. Thus, due to the lower reference point, finding a job would lose attractiveness compared to staying unemployed. This implies a higher reservation wage or

⁵⁰So, this variable abstracts from arriving job offers which are not taken up by the job seeker.

⁵¹The regressions pool all the available time periods and control for observables.

⁵²Specifically, $u(b_t|\psi_t) = v(b_t) + \eta[v(b_t) - v(\psi_t)]$ if $b_t \geq \psi_t$ and $= v(b_t) + \eta\nu[v(b_t) - v(\psi_t)]$ if $b_t < \psi_t$, whereby η is a weight and $\nu \geq 1$ captures loss aversion. The model implies, thus, additional (gain) utility of achieving an income above the current reference point and additional (higher weighted) negative utility from being below.

reduced job search effort.

The found negative effect of coaching on wage expectations would support the idea of a downward shock on the reference point. However, the prediction on higher reservation wage is not realized (but search effort indeed tends to be reduced). Thus, the other two mechanisms seem to be more dominant in driving the reservation wage. In the short run, the *disutility* explanation is consistent with the evidence: recall the "lock-in" effect in combination with lower r_t . Finally, the finding that the treatment caused lower reservation salaries post coaching is mostly consistent with the Burdett et al. (1988) model of learning about the wage offer distribution relevant to the job seeker. Thus, this type of supportive ALMP probably has caused a *learning process about "realistic" wage offer expectations* which has not taken place in the control group.

8 Conclusion

This paper provides unique evidence on the question of *how* labor market policy affects the *behavior* of job seekers in a real-world context. The key novelty of the paper is the empirical design, which combines four elements: randomized treatment assignment, ex-ante determined timing of treatment periods, detailed register data and a panel of repeated surveys on behavioral variables. This allows for the simultaneous identification of behavioral reactions and effects on labor market outcomes per treatment period —and thus a direct empirical assessment of how ALMP causally impacts the job seeker's behavioral dynamics. Such an additional layer of knowledge on how programs affect behavior offers policy makers the opportunity to become more specific and targeted when doing evidence-based program design.

This study addresses the behavioral "blackbox" by exploiting a unique panel of behavioral survey measures by means of dynamic analysis. The analyzed measures cover three types of behavioral variables: first, *search behavior*, notably, search effort, channel use, search strategy and reservation wages; second, two types of *beliefs*, namely, the expected success in job search and wage expectations; third, aspects of *non-cognitive skills*, notably, effort motivation, self-confidence and self-control (reliability). The surveys also track the individual job interview rate. The randomized field experiment assesses a supportive ALMP program focusing on the coaching of older job seekers. It features a period of intense coaching and high-frequency counseling.

With respect to labor market outcomes I find that the supportive ALMP program significantly increased job finding (by 9 percentage points) and employment stability thereafter. A cost-benefit analysis combined with a counterfactual simulation shows that the program avoids future unemployment of about 20 days p.p., which more than entirely pays the program cost. A further useful insight is that the individual job interview rate can serve as an early indicator for future job finding: positive effects on the rate of acquired interviews precede increased job finding.

The analysis of behavioral variables reveals that the treatment did not operate via increased search effort, but via improving search efficiency and lowering reservation wages. The data document substantial ex-ante upward-bias in beliefs on job chances. The program reduces the bias in beliefs on job interviews and wage, temporarily. I also find short-run effects on performance-motivation and self-confidence. Thus, overall, improving success of search of older job seekers seems less an issue of reducing moral hazard behavior, but more one of supporting realistic wage expectations, search efficiency and motivation.

Adopting a comprehensive view on the findings of this study, several policy conclusions can be drawn. In general, the experiment shows that it is possible to design effective supportive ALMP programs which cause positive impacts that outweigh the common lock-in effects. *Targeted* coaching and counseling is able to improve the employability of older job seekers. Based on the analysis of behavior, we can identify several success factors of counseling-focused ALMP, which should be taken into account when designing policy. First, the findings suggest that improving search *efficiency* and directedness is key, rather than purely focusing on quantitative application effort. Second, supporting job seekers to adopt *realistic wage expectations* is another key factor; learning about what is achievable in terms of wage offers helps participants to take realistic reservation wage decisions and to target on the "right" job vacancies. Third, coaching can serve as an instrument to reduce the bias in beliefs about job chances. Fourth, taking aspects of non-cognitive skills into account can be useful for policy (and for complementing job search theory models); notably, supporting performance-motivation and self-confidence can provide an additional short-run boost in successful job search activities. Fifth, extending the scope of search seems not only to be helpful for attracting more job offers, but also for increasing success of directed search through yielding more stable jobs.

This study demonstrates the potential of combining dynamic program evaluation with empirical analysis of behavior to provide specific evidence on success factors of ALMP. To exploit this potential, it is necessary to invest in policy evaluation designs and data initiatives which incorporate direct empirical measures of the job seeker's behavior. Such evaluation setups can generate specific knowledge for targeted policy design—in ALMP and beyond. Moreover, they may provide real-world tests to assess whether different job seeker behaviors can be best explained by standard rational job search theory or rather by alternative (behavioral) approaches. It would be instructive for future research to run similar data collections and experiments on a larger scale. This could provide broad support for designing more specific labor market programs, aiming at optimizing the individual's incentives to adopt successful job search strategies.

References

- Abbring, Jaap H., and Gerard van den Berg (2003). "The Non-Parametric Identification of Treatment Effects in Duration Models." *Econometrica*, 71, 1491-1517.
- Abbring, Jaap H., and Gerard van den Berg (2005). "Social Experiments and Instrumental Variables with Duration Outcomes," *Tinbergen Institute Discussion Papers 05-047/3*, Tinbergen Institute.
- Almlund M., Duckworth A.L., Heckman J.J., and T. Kautz. (2011). "Personality psychology and economics", in: E. A. Hanushek, S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, Volume 4, 1-181.
- AMOSA (2007). "Langzeitarbeitslosigkeit - Situation und Massnahmen", Arbeitsmarktbeobachtung Ostschweiz, Aargau und Zug (AMOSA), Zurich.
- Arni, Patrick (2011). "Langzeitarbeitslosigkeit verhindern: Intensivberatung für ältere Stellensuchende. Schlussbericht", University of Lausanne and Kanton Aargau. (online: www.patrick-arni.ch)
- Arni, Patrick, Lalive, Rafael, and Jan C. van Ours (2013). "How Effective Are Unemployment Benefit Sanctions? Looking Beyond Unemployment Exit", *Journal of Applied Econometrics*, early view (DOI: 10.1002/jae.2289).
- Ashenfelter, Orley, Ashmore, David, and Olivier Deschênes (2005). "Do Unemployment Insurance Recipients Actively Seek Work? Randomized Trials in Four U.S. States." *Journal of Econometrics*, 125, 53-75.
- Baker, Michael and Angelo Melino (2000). "Duration dependence and nonparametric heterogeneity: A Monte Carlo Study." *Journal of Econometrics*, 96, 357-393.
- Benabou, Roland and Jean Tirole (2002). "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117 (3), 871-915.
- Black, Dan A., Smith, Jeffrey A., Berger, Mark C. and Brett J. Noel (2003). "Is the Threat of Training More Effective than Training Itself? Evidence from Random Assignments in the UI system." *American Economic Review*, 93, 1313-1327.
- Borghans, Lex, Meijers, Huub and ter Weel, Bas (2008). "The Role of Noncognitive Skills in Explaining Cognitive Test Scores", *Economic Inquiry*, 46(1), 2-12.
- Burdett, Kenneth, Vishwanath, Tara (1988). "Declining Reservation Wages and Learning," *Review of Economic Studies*, vol. 55(4), 655-65.
- Caliendo, M., Cobb-Clark, D. A., and Uhlendorff, A. (2015). "Locus of Control and Job Search Strategies", *Review of Economics and Statistics*, 97(1), 88-103.
- Card, David, Kluve, Jochen, and Andrea Weber (2010). "Active Labour Market Policy Evaluations: A Meta-Analysis," *Economic Journal*, 120(548), F452-F477.
- Card, D., J. Kluve and A. Weber (2015): What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations, IZA Discussion Paper 9236.
- Compte, Olivier and Andrew Postlewaite (2004). "Confidence-Enhanced Performance," *American Economic Review*, 94(5), 1536-1557.
- Crépon, Bruno, Duflo, Esther, Gurgand, Marc, Rathelot, Roland, and Philippe Zamora (2013): "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment", *Quarterly Journal of Economics*, 128(2), 531-580.
- Duckworth, Angela Lee and Seligman, M. E. P. (2005). "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents", *Psychological Science*, 16(12), 939-944.
- Duncan, Greg J. and Dunifon, Rachel (1998). "Soft-Skills and Long-Run Labor Market Success." *Research in Labor Economics*, 17, 123-149.
- Falk, A., D.B. Huffman, U. Sunde (2006): Do I Have What It Takes? Equilibrium Search with Type Uncertainty and Non-Participation. IZA Discussion Paper 2531.

- Gaure, Simen, Røed, Knut and Tao Zhang (2007). "Time and Causality: A Monte Carlo Assessment of the Timing-of-events Approach", *Journal of Econometrics*, 141(2), 1159-1195.
- Graversen, Brian K., and Jan C. van Ours (2008). "How to Help Unemployed Find Jobs Quickly; Experimental Evidence from a Mandatory Activation Program", *Journal of Public Economics*, 92, 2020-2035.
- Graversen, Brian Krogh, and Jan C. van Ours (2011). "An Activation Program as a Stick to Job Finding", *Labour*, 25(2), 167-181.
- Heckman, James and Burton Singer (1984). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", *Econometrica*, 52(2), 271-320.
- Heckman, J. J. and Rubinstein, Y. (2001). "The Importance of Noncognitive Skills: Lessons from the Ged Testing Program", *American Economic Review*, 91(2), 145-149.
- Holzer, H. (1988): "Search Method Use by Unemployed Youth". *Journal of Labor Economics*, 1, 1-20.
- Kroft, Kory, Fabian Lange and Matthew J. Notowidigdo (2013): "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment", *Quarterly Journal of Economics*, 128 (3), 1123-1167.
- Krueger, Alan B. and Andreas Mueller (2011). "Job Search and Job Finding in a Period of Mass Unemployment: Evidence from High-Frequency Longitudinal Data," IZA Discussion Papers 5450.
- Krueger, Alan B. and Andreas Mueller (2012). "The Lot Of The Unemployed: A Time Use Perspective," *Journal of the European Economic Association*, 10(4), 765-794.
- Krueger, Alan B. and Andreas Mueller (2014). "A Contribution to the Empirics of Reservation Wages," NBER Working Papers 19870.
- Meyer, Bruce D. (1995). "Lessons from the U.S. Unemployment Insurance Experiments," *Journal of Economic Literature*, 33(1), 91-131.
- Mortensen, D.T. (1986). "Job search and labor market analysis", in: O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, 849-920.
- OECD (2012): "OECD Economic Surveys: United States 2012", June 2012, OECD, Paris.
- Rosholm, Michael and Michael Svarer (2008). "The Threat Effect of Active Labour Market Programmes," *Scandinavian Journal of Economics*, 110(2), 385-401.
- Spinnewijn, J. (2015): "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs", *Journal of the European Economic Association*, 13(1), 130-167.
- Van den Berg, G.J. (1990). "Nonstationarity in job search theory", *Review of Economic Studies*, 57, 255-277.
- Van den Berg, Gerard J., and Bas van der Klaauw (2006). "Counseling And Monitoring Of Unemployed Workers: Theory And Evidence From A Controlled Social Experiment," *International Economic Review*, 47(3), 895-936.
- Van den Berg, Gerard J., Bergemann, Annette, and Marco Caliendo (2009). "The Effect of Active Labor Market Programs on Not-Yet Treated Unemployed Individuals", *Journal of the European Economic Association*, 7(2-3), 606-616.
- Wooldridge, Jeffrey M. (2002). "Inverse probability weighted M-estimators for sample selection, attrition, and stratification," *Portuguese Economic Journal*, 1, 117-139.

Tables & Figures

Figure 2: Unemployment exit hazard (A) and employment exit hazard (recurrence, B): population of the experiment (aged 45 to 61)

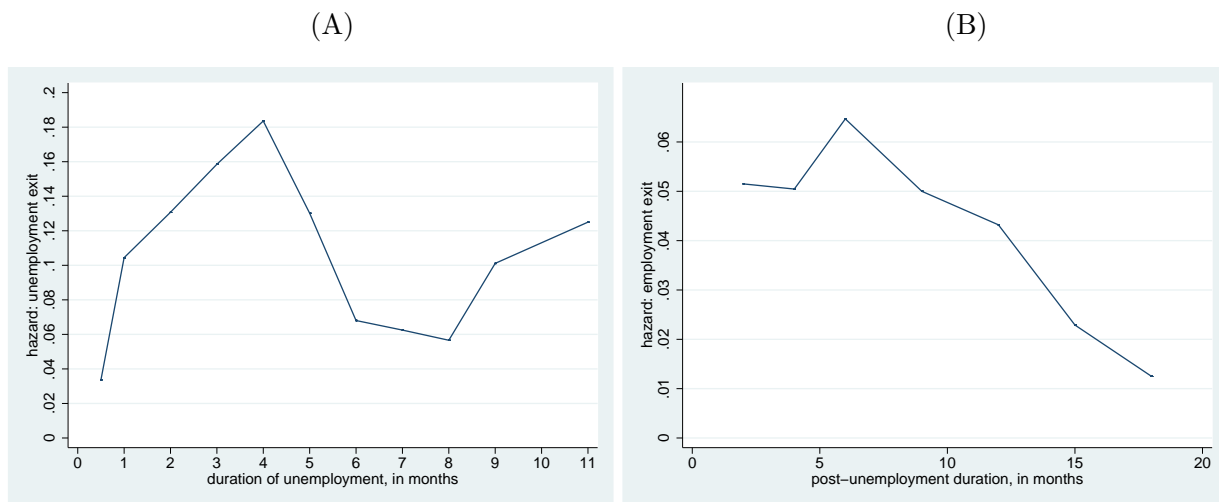


Table 1: Comparison of characteristics of treatment vs control group

	<i>Treatment Group</i>	<i>Control Group</i>	<i>t-values</i>
Gender: Woman	44.1%	43.3%	0.15
Married (incl. separated)	56.4%	49.7%	1.22
Age	52.5	51.9	1.04
Nationality: CH	84.4%	85.1%	-0.17
Qualification: (semi-)skilled	96.2%	95.7%	0.22
Employability: 3/4	77.4% / 21.5%	78.0% / 21.3%	(-)0.05
At least 1 foreign language	55.4%	53.2%	0.39
Job < 100%	17.7%	17.7%	0.00
PES 2	14.5%	10.6%	1.04
Duration to availability (median, days)	11	13	-0.49
Past UE duration (median, days)	0	0	0.00
Observations	186	141	<i>Total: 327</i>

Notes: Frequency percentages for different observable characteristics by treatment and control group are reported. t-values are based on unpaired t-tests with equal variances.

Source: Own calculations based on merged UIR-LZAR database.

Table 2: Non-parametric comparison of main outcomes: unemployment duration, long-term unemployment, job finders, gross salary, recurrence to unemployment, job interview rate

	TG	CG	difference =TE	t-value
Proportion of job finders	0.720	0.631	0.089	1.718
Prop. in job or leaving UI voluntarily	0.774	0.667	0.108	2.173
...job finders, ages 45-54	0.771	0.722	0.049	0.821
...job finders, ages 55+	0.600	0.471	0.129	1.334
Unemployment duration: means	245.1	255.9	-10.73	-0.431
...for ages 45-54	208.9	224.9	-15.98	-0.585
...for ages 55+	331.3	310.5	20.87	0.425
Unemployment duration: medians	139.5	138	1.5	0.060
...for ages 45-54	131	131.5	-0.5	-0.018
...for ages 55+	283	191	92	2.034
Proportion in long-term unemployment	0.280	0.319	-0.040	-0.774
...for ages 45-54	0.206	0.256	-0.049	-0.861
...for ages 55+	0.455	0.431	0.023	0.238
Gross salary: means	5357.6	5392.4	-34.78	-0.105
... difference to pre-UE salary	-402.7	-242.3	-160.37	-0.737
pensum: working hours per week	38.72	37.62	1.10	0.850
Recurrence to unemployment	0.221	0.277	-0.055	-0.962
...for ages 45-54	0.208	0.269	-0.061	-0.926
...for ages 55+	0.265	0.296	-0.032	-0.269
Average job interview rate	0.143	0.102	0.041	1.813
...for ages 45-54	0.136	0.124	0.011	0.395
...for ages 55+	0.157	0.065	0.092	2.703

Note: Means are reported, in the case of the unemployment durations as well medians. Unemployment durations are censored at 2 years (730 days; 2.4% censored). Recurrence to unemployment: within 1.5 years. Job interview rate is averaged over t_0 and the four treatment periods (if observed). Observations: 327, 186 in treatment group (TG) and 141 in control group (CG); subsamples for ages 45-54/55+ amount to 221/106. Observations on salary data: 163; on recurrence: 234; on job interviews: 229. Long-term unemployment = unemployment duration of more than one year.

Source: Merged UIR-LZAR database

Table 3: Dynamic treatment effects of the treatment plan on the exit to job rate. (PH duration model)

	<i>Destination: exit to job</i>		
	coeff.	s.e.	transf.
<i>Treatment effects</i>			
Anticipation effect (δ_a /in %)	-0.499**	0.236	-0.393
During coaching (δ_{c1} /in %)	-0.477 ^o	0.309	-0.379
Post-coaching, 14-180 days (δ_{c2} /in %)	-0.023	0.250	-0.023
Post-Coaching, 181+ days (δ_{c3} /in %)	0.401	0.374	0.494
Control variables	Yes		
Unobserved heterogeneity	No		
-Log-Likelihood	1455.45		
AIC	1508.45		
N	327		

Notes: Coefficients and their transformations are reported: Transformed treatment effects are changes in %. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, ^o $p < 0.15$.

Source: Own estimations based on merged UIR-LZAR database.

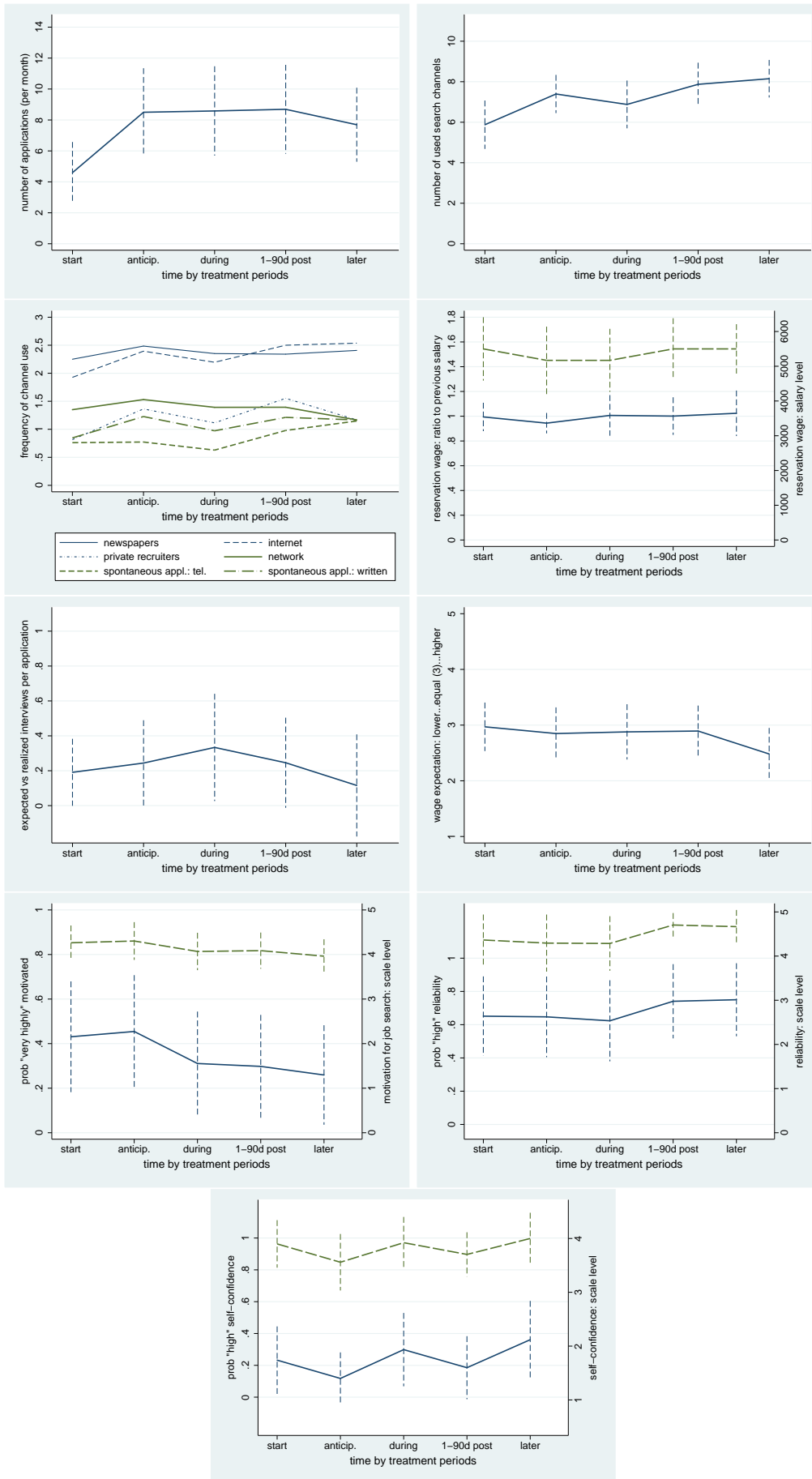
Table 4: Treatment effects (DiD) on job interview rate (job interviews per application)

	(1)		(2)		(3)		(4)	
	<i>anticipation</i>		<i>during coaching</i>		<i>1-90d post-coa.</i>		<i>91+d post-coa.</i>	
	coef	se	coef	se	coef	se	coef	se
DiD	0.056	(0.040)	0.084**	(0.037)	0.026	(0.040)	0.011	(0.058)
treatment (D^{TG})	-0.017	(0.030)	-0.027	(0.027)	-0.012	(0.027)	-0.031	(0.050)
time (T_t)	-0.019	(0.026)	-0.033	(0.027)	-0.023	(0.027)	-0.024	(0.027)
X variables	yes		yes		yes		yes	
Caseworker FE	yes		yes		yes		yes	
Obs. mean	0.121		0.131		0.117		0.080	
Observations	394		473		399		379	
R ²	0.159		0.159		0.148		0.141	

Note: Cluster robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; available observations at t_0 : 301. OLS regression

Source: UIR-LZAR database

Figure 3: Behavioral channels: means and s.d. of control group, by period of treatment plan



Note for Figure 2: Unconditional means are reported in the line graphs, s.d. (centered around means) in the dashed vertical spikes. Frequency of search channel use (s.d. omitted) is measured on a 6 point scale: 3 = daily, 2.5 = several times per week, 2 = weekly, 1.5 = several times per month, 1 = monthly or less often, 0 = never. Reservation wage: blue solid line represents ratio (to past salary), green long dash line the median reservation salary. Wage expectation: blue solid line represents the mean level of a scale ranging from 1 to 5: expected salary clearly higher/a bit higher/equal/a bit lower/clearly lower. Last three graphs: blue solid line represents the proportion of the high(est) scale level, green long dash line the mean scale level; scale is from 1 to 5 (very high/high/medium/low/very low).

Figure 4: Total treatment effect on duration to job finding, survivors treatment vs control group

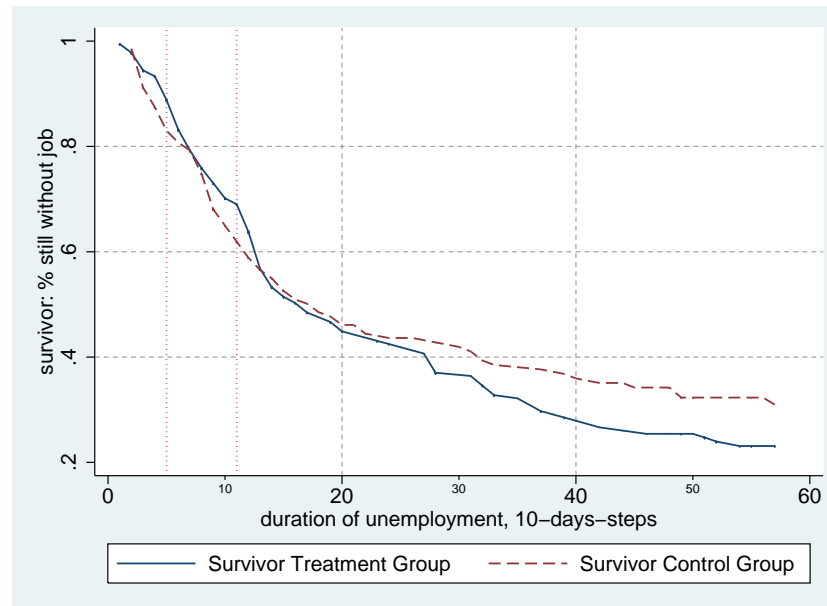


Figure 5: Post-unemployment job stability: Survivor of the reentry rate into unemployment

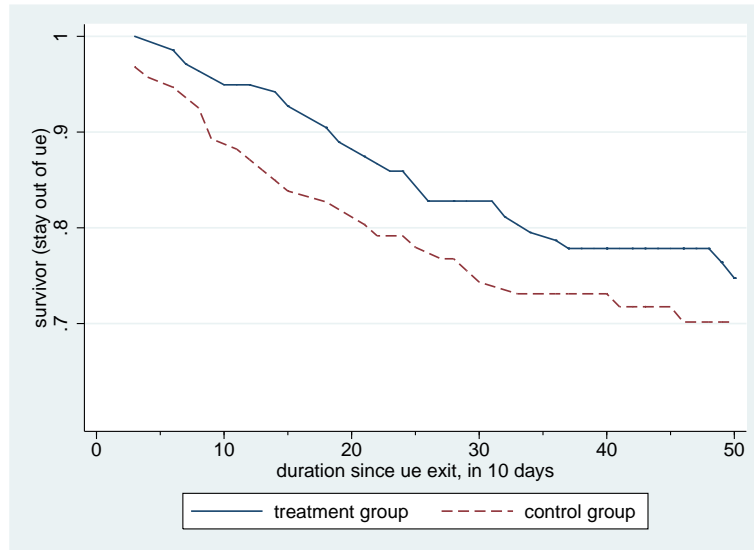
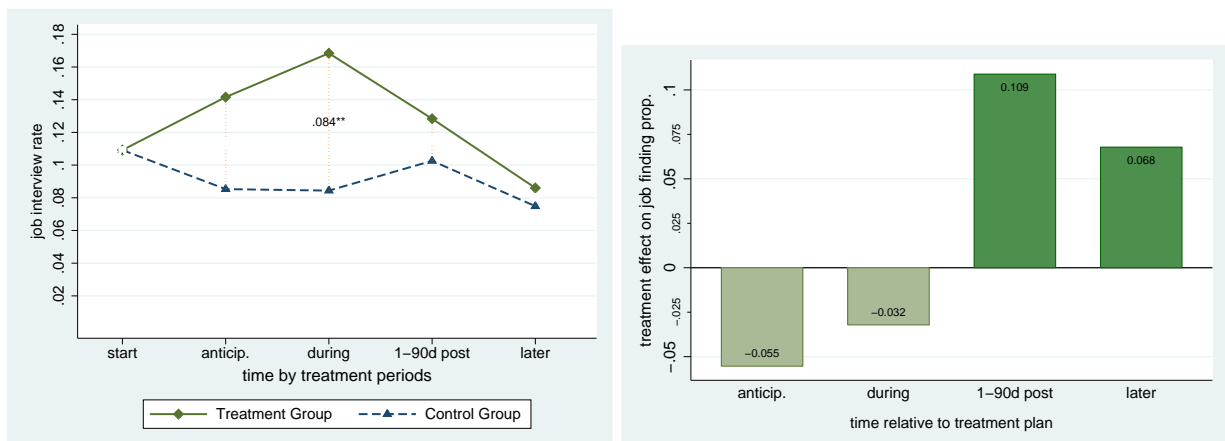


Figure 6: Treatment effects on job interview rate (Diff-in-Diff by treatment period) – compared to the treatment effect of the job finding rate per period



Note: The dots represent predicted conditional means by treatment period for the counterfactuals ($D^{TG} = 1/0$ for treatment/control group).

Table 5: Treatment effects on behavior: search effort, search channels, reservation wage

	(1) <i>anticipation</i>		(2) <i>during coaching</i>		(3) <i>1-90d post-coa.</i>		(4) <i>91+d post-coa.</i>	
	coef	se	coef	se	coef	se	coef	se
<i>Search effort (applications per month)</i>								
DiD	-1.719*	(1.042)	-1.653**	(0.829)	-1.527	(1.093)	-0.479	(1.036)
time	3.885***	(0.839)	4.132***	(0.673)	4.076***	(0.830)	3.337***	(0.703)
TG	0.672	(0.486)	0.797	(0.494)	0.760	(0.515)	0.665	(0.510)
<i>Search channels: number</i>								
DiD	-0.645	(0.454)	-0.049	(0.359)	-1.207***	(0.429)	-0.433	(0.555)
time	1.475***	(0.370)	1.019***	(0.278)	2.150***	(0.353)	2.224***	(0.401)
TG	0.337	(0.282)	0.396	(0.280)	0.408	(0.280)	0.404	(0.286)
<i>Search channels: frequency of use - Formal channels</i>								
<i>newspapers</i>								
DiD	-0.266*	(0.161)	-0.015	(0.127)	-0.141	(0.152)	-0.370**	(0.181)
time	0.198	(0.125)	0.082	(0.088)	0.131	(0.123)	0.160	(0.133)
TG	0.148	(0.102)	0.147	(0.102)	0.121	(0.103)	0.129	(0.102)
<i>internet</i>								
DiD	-0.408*	(0.216)	-0.158	(0.164)	-0.361*	(0.214)	-0.455*	(0.248)
time	0.505***	(0.160)	0.272**	(0.130)	0.588***	(0.162)	0.596***	(0.194)
TG	0.287**	(0.137)	0.310**	(0.136)	0.309**	(0.138)	0.313**	(0.138)
<i>private recruiters</i>								
DiD	-0.346	(0.239)	-0.190	(0.176)	-0.511**	(0.198)	-0.083	(0.254)
time	0.542***	(0.198)	0.342***	(0.131)	0.751***	(0.152)	0.352*	(0.189)
TG	0.142	(0.125)	0.161	(0.126)	0.185	(0.124)	0.188	(0.125)
<i>Search channels: frequency of use - Informal channels</i>								
<i>network</i>								
DiD	-0.120	(0.191)	0.131	(0.145)	0.000	(0.165)	0.078	(0.206)
time	0.267*	(0.149)	0.030	(0.102)	0.060	(0.115)	-0.167	(0.162)
TG	-0.046	(0.118)	-0.028	(0.117)	-0.036	(0.117)	-0.026	(0.117)
<i>spontaneous appl.: tel.</i>								
DiD	0.039	(0.202)	0.436***	(0.159)	0.032	(0.181)	-0.074	(0.227)
time	0.054	(0.160)	-0.091	(0.127)	0.241*	(0.141)	0.445**	(0.178)
TG	-0.095	(0.114)	-0.070	(0.111)	-0.092	(0.112)	-0.078	(0.113)
<i>spontaneous appl.: written</i>								
DiD	-0.328	(0.210)	-0.036	(0.152)	-0.342*	(0.185)	-0.194	(0.208)
time	0.358**	(0.165)	0.129	(0.112)	0.368**	(0.145)	0.296*	(0.172)
TG	0.027	(0.112)	0.028	(0.109)	0.034	(0.109)	0.041	(0.111)
<i>Reservation wage (ratio to last salary)</i>								
DiD	0.039	(0.029)	-0.055*	(0.031)	-0.074*	(0.045)	-0.073	(0.062)
time	-0.036*	(0.021)	0.032	(0.027)	0.040	(0.039)	0.040	(0.048)
TG	-0.026	(0.023)	-0.022	(0.024)	-0.026	(0.024)	-0.032	(0.024)
<i>Reservation salary</i>								
DiD	451.93*	(254.87)	-359.18*	(215.77)	-505.76*	(297.87)	-506.82	(422.08)
time	-390.68*	(212.99)	137.01	(156.65)	133.27	(219.30)	244.65	(332.07)
TG	-65.06	(186.35)	-151.19	(172.64)	-109.77	(193.98)	-77.21	(205.95)
X variables	yes		yes		yes		yes	
Caseworker FE	yes		yes		yes		yes	
Obs. effort	394		473		399		379	
Obs. channels	388		466		409		365	
Obs. res'wages	358		435		363		342	

Note: Cluster robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; available observations at t_0 : 301/298/265. DiD=treatment effect, diff-in-diff coefficient ($D^{TG}T_t$); time=indicator T_t , TG=indicator D^{TG}

Source: UIR-LZAR database

Figure 7: Treatment effects on behavior, illustration: search effort, search channels, frequencies of channel use

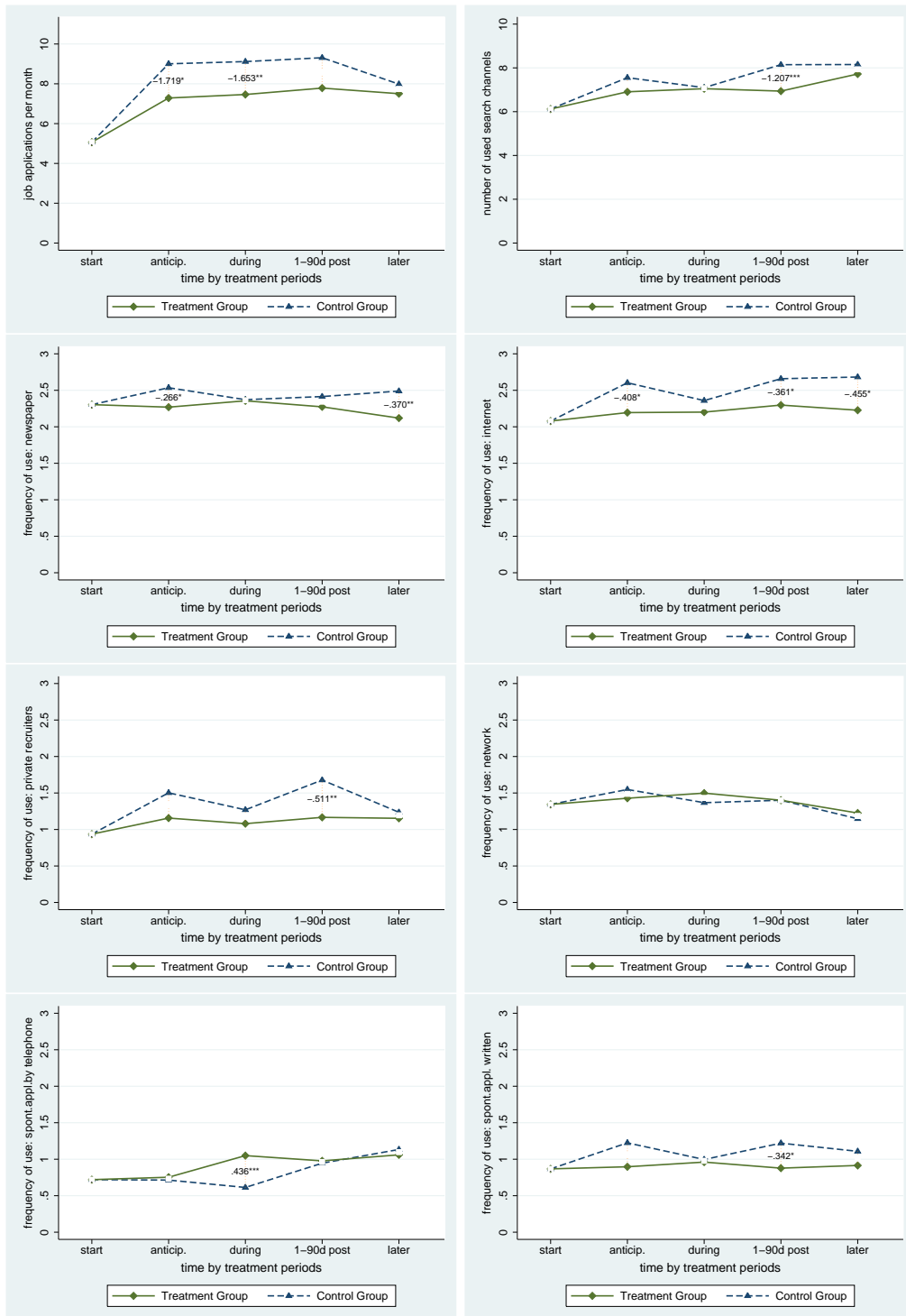


Table 6: Treatment effects on behavior: beliefs, motivation, self-confidence, reliability

	(1)		(2)		(3)		(4)	
	<i>anticipation</i>		<i>during coaching</i>		<i>1-90d post-coa.</i>		<i>91+d post-coa.</i>	
	coef	se	coef	se	coef	se	coef	se
<i>Beliefs: difference expected vs. realized interviews</i>								
DiD	-0.086	(0.107)	-0.173*	(0.092)	-0.060	(0.110)	0.028	(0.126)
time	0.039	(0.088)	0.171**	(0.078)	0.082	(0.081)	-0.063	(0.105)
TG	0.128***	(0.047)	0.126***	(0.049)	0.114**	(0.048)	0.117**	(0.048)
<i>Beliefs: wage expectation (scale 1 to 5, 3=equal)</i>								
DiD	-0.166	(0.198)	-0.102	(0.142)	-0.298*	(0.181)	0.018	(0.268)
time	-0.110	(0.171)	-0.086	(0.107)	0.014	(0.129)	-0.432**	(0.208)
TG	0.166	(0.104)	0.184*	(0.104)	0.182*	(0.105)	0.199*	(0.105)
<i>Motivation for job search (prop. very high)</i>								
DiD	-0.098	(0.108)	0.055	(0.074)	0.209**	(0.095)	0.109	(0.122)
time	0.011	(0.086)	-0.101*	(0.054)	-0.133*	(0.070)	-0.159*	(0.093)
TG	-0.093	(0.057)	-0.098*	(0.057)	-0.106*	(0.057)	-0.098*	(0.058)
<i>Self-confidence (prop. high)</i>								
DiD	0.115	(0.073)	0.063	(0.045)	0.240**	(0.111)	0.009	(0.116)
time	-0.099*	(0.056)	-0.044	(0.030)	-0.148*	(0.086)	-0.054	(0.074)
TG	-0.090*	(0.052)	-0.085*	(0.051)	-0.086*	(0.052)	-0.079	(0.051)
<i>Reliability (prop. high)</i>								
DiD	-0.059	(0.084)	0.082*	(0.046)	0.090	(0.122)	-0.047	(0.123)
time	-0.006	(0.061)	-0.018	(0.033)	0.056	(0.093)	0.116	(0.087)
TG	0.003	(0.055)	-0.002	(0.055)	-0.009	(0.056)	0.001	(0.055)
X variables	yes		yes		yes		yes	
Caseworker FE	yes		yes		yes		yes	
Obs. beliefs, motiv.	388		466		409		365	
Obs. self-c., reliab.	394		473		399		379	

Note: Cluster robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; available observations at t_0 : 298/301. DiD=treatment effect, diff-in-diff coefficient ($D^{TG}T_t$); time=indicator T_t , TG=indicator D^{TG} . Note that for beliefs on interviews, 7.6% of the observations are imputed (by regression). The results are robust to omitting these few observations, as a sensitivity check shows (available on request).

Source: UIR-LZAR database

Figure 8: Treatment effects on behavior, illustration: reservation wage (ratio to last salary; median salary level), beliefs (on job interviews; on wage), motivation, self-confidence, reliability

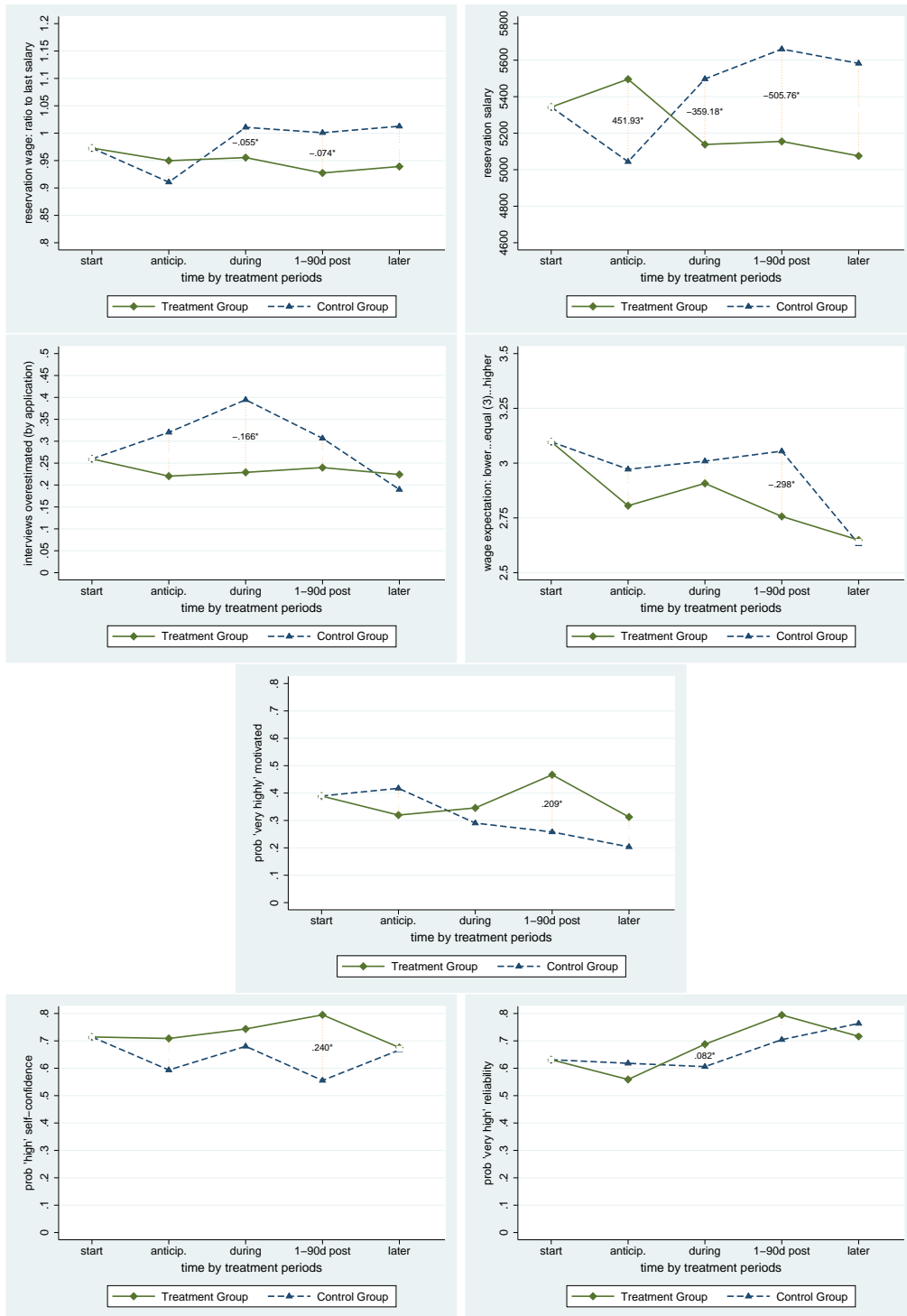


Table 7: Employment stability: Effect of new policy on reentry rate into unemployment (20–540 days after UE); sensitivity analysis: model with unobserved heterogeneity

	1: Employment stability			2: Both processes: UE & post-UE		
	coeff.	s.e.	transf.	coeff.	s.e.	transf.
<i>Treatment effects</i>						
Unemployment reentry (δ_p/in %)	-0.590*	0.341	-0.446	-0.629°	0.408	-0.467
Anticipation effect (δ_a /in %)				-0.865**	0.353	-0.579
During coaching (δ_{c1} /in %)				-0.696°	0.431	-0.502
Post-coaching, 14-180 days (δ_{c2} /in %)				-0.222	0.325	-0.199
Post-Coaching, 181+ days (δ_{c3} /in %)				0.247	0.393	0.280
<i>Reentry rate into unemployment</i>						
$\lambda_{b,a}/exp(u_{b,a})$, 20-210 days	-6.112***	0.834	2.58	-7.344***	1.107	0.973
$\lambda_{b,b}/exp(u_{b,b})$				-5.859***	1.001	4.298
$\lambda_1/exp(u_{1,a})$, 211-390 days	-0.152	0.406	2.22	-0.094	0.417	0.886
$exp(u_{1,b})$						3.912
$\lambda_2/exp(u_{2,a})$, 391-480 days	-1.257°	0.798	0.73	-1.234	0.981	0.283
$exp(u_{2,b})$						1.252
$\lambda_3/exp(u_{3,a})$, 481+ days	-0.404	0.818	1.72	-0.438	1.020	0.628
$exp(u_{3,b})$						2.773
<i>Probabilities:</i>						
p_1 (type aa)				0.644	0.036	
p_4 (type bb)				0.356	–	
Unobserved heterogeneity		No			Yes	
All control variables UE process		–			Yes	
-Log-Likelihood		459.05			1987.05	
AIC		496.05			2080.05	
N UE/N post-UE		–/234			327/234	

Notes: Coefficients and their transformations are reported: Transformed coefficients are changes in %. Transition rates are in % per month (for the respective piece of the hazard). Note that λ_b is the intercept of the baseline hazards, the further steps are incremental; the transformations represent the monthly transition rate for an "average" individual: $u_{j,g} = \lambda_{b,g} + \lambda_j + \bar{x}'\beta_j + \sum_i \tau_i \bar{M}_i + \sum_k \gamma_k \bar{C}_k$ where $j = 1, \dots, 6$ and $g \in \{a, b\}$ ($\lambda_j = 0$ for first segment) and the bars are means, except for the past unemployment and the duration until availability where medians are used. (post-)UE=(post-)unemployment. Probabilities: Model with 4 mass points whereby $p_2 = p_3 = 0$ is optimal; type aa=baseline hazards a in UE and post-UE, type bb=baseline hazards b in UE and post-UE. Note that in the post-UE process the occupation variables and the ones for non-German speaking and for support function are omitted (due to high collinearity to comparable variables) in order to avoid overparametrisation. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, ° $p < 0.15$.

Source: Own estimations based on merged UIR-LZAR database.

Table 8: Analysis of costs vs benefits of new policy for the UI accounts: avoided future unemployment vs. additional program cost; in CHF per job seeker in treatment group (TG)

Benefits		Cost	
		Additional cost of new program (compared to status quo):	
Average increase of duration until reentry into unemployment (up to 540 days after UE)	23.16 days	Coaching seminar instead of short job search assistance sequence	4500 CHF
... times average daily benefit rights	189.43 CHF	... times proportion of coaching participants in TG	53.80%
		Cost for additional counseling	115.38 CHF
Total savings for UI	4387.01 CHF	Total additional cost for UI	2534.74 CHF
Savings per job seeker due to avoided future unemployment: 1852.27 CHF			

Notes: Average duration of avoided future unemployment is calculated by means of the simulation described in the respective section of the text. Average daily benefit rights are calculated according to the legal rules, based on the salary information in the survey. The calculation of the cost for additional counseling is based on the following data: Assume 100 cases per caseworker; median unemployment duration is 140 days; caseworkers in the new program got a reduction of the caseload by 20%; this results in a caseload reduction to 208 instead of 260 job seekers per year; this caseload reduction is multiplied by the average employment cost of a caseworker per year. 1 CHF=0.766 EUR. UI=unemployment insurance, UE=unemployment. *Source:* Own calculations based on merged UIR-LZAR database.

Table 9: Analysis of the correlation between non-cognitive skills and search activity to explain job offer arrival (proxied by job interview rate)

	<i>job interview arrival per application</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
motivation	0.013 (0.010)	0.023** (0.011)				
self-confidence			0.085*** (0.021)	0.095*** (0.021)		
reliability					-0.012 (0.017)	0.000 (0.018)
# applications p.mt.		-0.008*** (0.002)		-0.008*** (0.003)		-0.007*** (0.003)
# search channels		-0.002 (0.004)		-0.000 (0.004)		0.000 (0.004)
Constant	0.083 (0.051)	0.146** (0.058)	0.017 (0.044)	0.031 (0.048)	0.052 (0.048)	0.066 (0.051)
<i>socio-demogr.</i>	yes	yes	yes	yes	yes	yes
<i>spell time FE</i>	yes	yes	yes	yes	yes	yes
<i>Observations</i>	676	676	646	646	646	646
<i>R</i> ²	0.118	0.146	0.154	0.176	0.127	0.144

Notes: Cluster robust standard errors; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; max. 4 observations (spell intervals) p.p.

Appendices

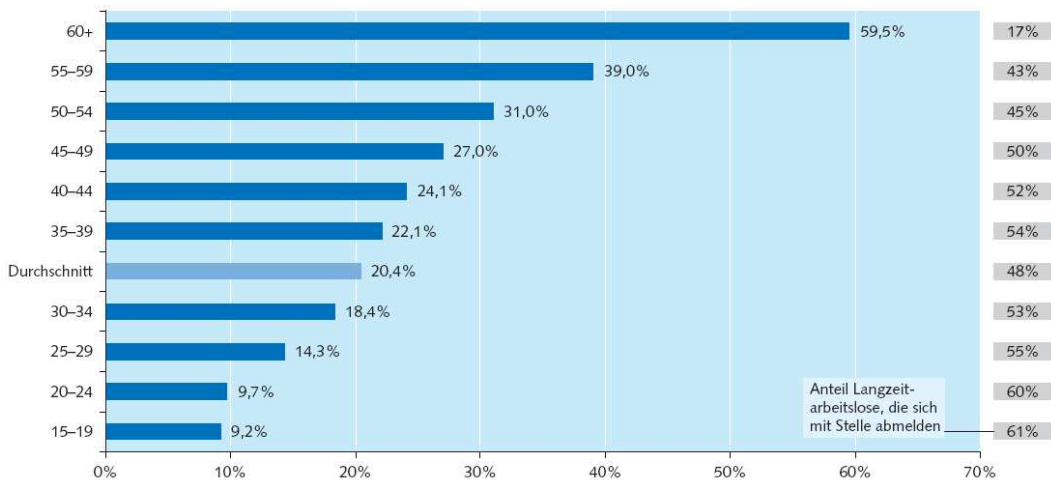
A Further Facts on Institutional & Economic Background

Benefits & eligibility This field experiment for individuals aged 45+ was performed in the frame of the rules of the Swiss unemployment insurance (UI). The maximum duration of unemployment benefits in the Swiss UI system is 1.5 years (400 days) for individuals who meet the eligibility requirements.

The two requirements are (i) that they must have paid unemployment insurance taxes for at least 12 months in the two years prior to entering registered unemployment, and (ii) that they must be 'employable' (i.e. fulfill the requirements of a regular job). After this period of two years or in the case of non-employability the unemployed have to rely on social assistance. From the 55th birthday on, job seekers profit of a benefit duration which is prolonged by about half a year (120 working days). Beyond the age of 61, benefit rights get extended by another 120 days.

The marginal replacement ratio is 80% for job seekers with previous monthly income up to CHF 3797 (about 2550 €)⁵³. For income between 3797 CHF and 4340 CHF (2900) the replacement ratio linearly falls to 70%. For individuals with income beyond 4340 CHF the ratio is 70%, whereby the insured income is capped at 10500 CHF (7000 €). For job seekers with dependent children, the marginal replacement ratio is always 80% (up to the same maximal insured income cap). Job seekers have to pay all income and social insurance taxes except for the unemployment insurance contribution.

Figure 9: Incidence of long-term unemployment by age groups



Daten: Anteil Langzeitarbeitslose, AMOSA-Kantone, Jahresdurchschnitt 2006 (Quelle: AVAM, SECO)

Note: The bars represent the proportion of long-term unemployed (1 year or more) individuals among the registered unemployed of the respective age category. The figure to the right reports the age-related proportions of the long-term unemployed who deregister from unemployment insurance due to having found a job.

Source: AMOSA 2007.

⁵³The figures on potential benefit durations and income thresholds apply to the situation before the last major reform of UI in 2011, thus to the period of execution of the field experiment.

Long-term unemployment & age LTU incidence is highly age-dependent. For the region under consideration, Figure 9 shows this strong pattern in terms of proportion of LU in the unemployed population of a certain age category. Figure 9 (AMOSA 2007) reveals that this proportion amounts to 18.4% for individuals aged 30-34 – and increases up to 39.0% for individuals aged 55-59. Note that the last figure may be affected by the above-mentioned fact that job seekers of age 55+ and 61+ receive a benefit duration extension. The percentage numbers to the right of Figure 9 represent the age-related proportions of the long-term unemployed who deregister from unemployment insurance due to having found a job. This percentage remarkably decreases from age 45 on, from around 50% to less than 30% beyond age of 60. Figure 9 clearly shows that individuals of age 45+ face a markedly increased risk of long-term unemployment.

B Dynamic Analysis: Further Robustness Analyses

B.1 Balancing of observables by treatment period

Table 10: Repeated surveys: Balancing of observables, by treatment (TG) and control group (CG) and periods of the treatment plan

<i>Job seeker surveys</i>	Start		Anticipation		During Coaching		1-90d post Coaching		Later		After Exit	
	TG	CG	TG	CG	TG	CG	TG	CG	TG	CG	TG	CG
Gender: woman	44.05%	43.08%	45.61%	42.42%	46.81%	44.59%	43.75%	36.17%	47.50%	37.04%	52.81%	44.62%
Married (incl. Separated)	55.95%	47.69%	66.67%	51.52%	57.45%	48.65%	64.06%	51.06%	50.00%	37.04%	57.30%	55.38%
Age	52.45	52.38	53.23	52.85	52.51	52.64	52.98	53.19	54.65	53.48	52.04	51.66
Nationality: CH	85.12%	86.15%	87.72%	78.79%	80.85%**	93.24%**	84.38%	85.11%	87.50%	88.89%	89.89%	84.62%
Qualification: (semi-)skilled	97.02%	96.15%	98.25%	100.00%	96.81%	95.95%	96.88%	95.74%	100%*	92.59%*	95.51%	98.46%
Employability: 4	20.83%	21.54%	15.79%	15.15%	21.28%	17.57%	18.75%	21.28%	22.50%	14.81%	16.85%	23.08%
At least 1 foreign language	58.33%	55.38%	66.67%	63.64%	57.45%	59.46%	60.94%	59.57%	55.00%	66.67%	57.30%	47.69%
Job < 100%	17.26%	17.69%	21.05%	18.18%	15.96%	21.62%	15.63%	19.15%	22.50%	11.11%	20.22%	20.00%
PES 2	14.29%	10.77%	10.53%	6.06%	14.89%	12.16%	20.31%**	6.38%**	12.50%	7.41%	17.98%	16.92%
Observations	168	130	57	33	94	74	64	47	40	27	89	65
... in %	56.38%	43.62%	63.33%	36.67%	55.95%	44.05%	57.66%	42.34%	59.70%	40.30%	57.79%	42.21%
<i>Caseworker surveys</i>												
Gender: woman	43.60%	43.85%	45.00%	44.12%	45.45%	45.68%	41.38%	34.88%	43.90%	35.00%	45.24%	47.92%
Married (incl. Separated)	56.40%*	46.15%*	63.33%	55.88%	55.56%	49.38%	63.79%	48.84%	48.78%	42.50%	53.17%	43.75%
Age	52.53	52.28	52.97	52.71	52.58	52.53	53.12	53.58	54.63	53.98	52.20	51.92
Nationality: CH	84.88%	85.38%	91.67%*	79.41%*	81.82%**	93.83%**	82.76%	86.05%	85.37%	85.00%	100.00%	100.00%
Qualification: (semi-)skilled	97.09%	96.15%	96.67%	97.06%	97.98%	96.30%	96.55%	95.35%	100%	97.50%	96.03%	94.79%
Employability: 4	21.51%	22.31%	18.33%	17.65%	23.23%	17.28%	17.24%	27.91%	21.95%	17.50%	22.22%	23.96%
At least 1 foreign language	58.72%	55.38%	60.00%	61.76%	54.55%	60.49%	58.62%	53.49%	51.22%	57.50%	60.32%	52.08%
Job < 100%	17.44%	18.46%	23.33%	14.71%	15.15%	20.99%	17.24%	18.60%	21.95%	15.00%	20.63%	17.71%
PES 2	13.95%	10.77%	8.33%	5.88%	15.15%	11.11%	22.41%**	4.65%**	9.76%	2.50%	16.67%	12.50%
Observations	172	130	60	34	99	81	58	43	41	40	126	96
... in %	56.95%	43.05%	63.83%	36.17%	55.00%	45.00%	57.43%	42.57%	50.62%	49.38%	56.76%	43.24%

Notes: All TG-CG differences are not significantly different from zero, except from those marked: *** 1%, ** 5%, * 10%

Source: LZAR database.

C Estimation of Unobserved Heterogeneity Mass Points by Grid Search

In this section of the Appendix I describe the systematic procedure I applied to search for unobserved heterogeneity in the context of the models developed in the sections 5.1.1, ?? and ??. Such a procedure amounts to searching for additional mass points in order to establish a discrete mixture distribution for v_u and v_p (described in section ??). Thus, the benchmark and starting point is the model with 1 mass point, i.e. with no unobserved heterogeneity in the baseline hazard profile. In the following, I demonstrate step-by-step the iterative procedure – an interplay between grid search and estimation – I use to establish a second mass point and then to search for further ones.

1. Use the results of the separate estimations of the two processes (unemployment and post-unemployment) without unobserved heterogeneity as starting values.
2. Start with an initial set of 2 mass points (per process), i.e. the aim is to estimate their probabilities and locations (=intercept of the transition rate/baseline hazard): p_1 and λ^a as well as p_2 and λ^{b54} .
3. *Grid search (over the probabilities' space)*: Run systematically through all possible combinations of probabilities, using a loop. I.e., pick a probability combination, fix it and estimate the corresponding location of the mass points. More specifically, I use a double loop:
 - (a) Loop over the sign (i.e. 2 runs) of the difference between the two locations. Note that this loop is used to set the starting values for the location estimation: I.e., set $\lambda^b = \lambda^a \pm 3$, whereby λ^a is the location (intercept) of the baseline hazard of the model without unobserved heterogeneity⁵⁵.
 - (b) Loop over the i increments (here of 0.01) of the probabilities which are to be grid-searched: $p_1 = 1 - i \cdot 0.01$, whereby $p_2 = 1 - p_1$. Choice criterion: Take the set (p_1^*, p_2^*) with the corresponding estimated $(\lambda^{a*}, \lambda^{b*})$ which yields the highest likelihood⁵⁶.
4. *Estimation of the probabilities*: Fix the location of the mass points at λ^{a*} and λ^{b*} . Use p_1^* and p_2^* to calculate the starting values for the parameters a_1 and a_2 (the probabilities are designed in a logistic form, see section ??). Estimate these parameters (i.e. the probabilities) in the model.
5. *Fully free estimation*: Un-fix the location of the mass points, and use them and the estimated probabilities as starting values for the fully free estimation. If this estimation yields a higher likelihood, continue with the next step; otherwise stop and choose the model without unobserved heterogeneity as the best one.

⁵⁴Extension to two processes u and p implies four probabilities and location combinations: p_1 for type aa (i.e. λ_u^a and λ_p^a), p_2 for type ab (i.e. λ_u^a and λ_p^b), p_3 for type ba (i.e. λ_u^b and λ_p^a), p_4 for type ab (i.e. λ_u^b and λ_p^b).

⁵⁵Note that the difference, 3, can be chosen arbitrarily. It should be sufficiently big in order to allow the estimation to distinguish the two locations.

⁵⁶In the grid search performed for this paper, this criterion always corresponded to choosing the lowest AIC. See Gaure et al. (2007) for a discussion of choice criteria. They opt for the use of the likelihood.

6. *Increase the set of mass points:* Add a third mass point to every process (this can be done gradually, following Gaure et al. 2007). Redo steps 3 to 5.
7. *Stopping rule:* After having performed step 6, check whether the chosen model with 3 mass points yields a higher likelihood. If no, stop and take the previous model as the best. If yes, continue by adding a fourth mass point... and so on.

D Additional Tables & Figures

Table 11: Effects on exit to job rate (PH model): coefficient estimates for control variables

	<i>Destination: exit to job</i>		
	coeff.	s.e.	transf.
<i>Exit rate from unemployment</i>			
$\lambda_b/exp(u_b)$, 1-50 days	-6.532***	0.442	6.44
$\lambda_1/exp(u_1)$, 51-100 days	0.823***	0.236	14.67
$\lambda_2/exp(u_2)$, 101-150 days	0.802***	0.250	14.37
$\lambda_3/exp(u_3)$, 151-250 days	0.214	0.260	7.98
$\lambda_4/exp(u_4)$, 251-400 days	0.162	0.283	7.57
$\lambda_5/exp(u_5)$, 401-550 days	-0.413	0.381	4.26
$\lambda_6/exp(u_6)$, 551+ days	-1.010 ^o	0.633	2.35
<i>Control variables</i>			
UE duration in past 3 years	0.000	0.001	0.000
duration until availability	-0.001	0.003	-0.001
age: 50-54 (base: 45-49)	-0.336*	0.202	-0.286
age: 55-59	-0.657***	0.207	-0.482
age: 60+	-1.481***	0.354	-0.772
married (base: unmarried)	0.136	0.199	0.146
divorced	0.062	0.242	0.064
female	0.361 ^o	0.243	0.434
non-Swiss	0.308	0.260	0.360
low employability (base: medium)	0.419 ^o	0.289	0.521
semi-skilled (base: skilled)	-0.041	0.393	-0.040
unskilled	0.112	0.547	0.118
non-German-speaking	-0.012	0.340	-0.011
1 foreign language (base: 0)	-0.126	0.254	-0.118
2+ foreign languages	0.177	0.285	0.194
PES 2 (base: PES 1)	0.194	0.516	0.214
management (base: professionals)	-0.293	0.408	-0.254
support function	-0.076	0.546	-0.073
part-time (but above 50%)	0.246	0.232	0.279
Controls occupation groups		Yes	
Controls month of entry in UE (bi-monthly)		Yes	
Caseworker fixed effects		Yes	
Unobserved heterogeneity		No	
-Log-Likelihood		1468.01	
AIC		1517.01	
N		327	

Notes: Coefficients and their transformations are reported: Transformed treatment effects are changes in %. Transition rates are in % per month (for the respective piece of the hazard); note that λ_b is the intercept of the baseline hazard, the further steps are incremental; the transformations are calculated for an "average" individual: $u_j = \lambda_b + \lambda_j + \bar{x}'\beta_j + \sum_i \tau_i \bar{M}_i + \sum_k \gamma_k C_k$ where $j = 1, \dots, 6$ ($\lambda_j = 0$ for first segment) and the bars are means, except for the past unemployment and the duration until availability where medians are used. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, ^o $p < 0.15$.

Source: Own estimations based on merged UIR-LZAR database.

Table 12: Effects of the treatment plan on the exit to job rate: ITT (intention-to-treat) model

	ITT		
	coeff.	s.e.	transf.
<i>Treatment effects by period</i>			
Anticipation effect (δ_a /in %)	-0.472**	0.240	-0.376
During coaching (δ_{c1} /in %)	0.174	0.229	0.190
Post-coaching, 14-180 days (δ_{c2} /in %)	0.079	0.254	0.082
Post-Coaching, 181+ days (δ_{c3} /in %)	0.510 ^o	0.360	0.666
Control variables		Yes	
Unobserved heterogeneity		No	
-Log-Likelihood		1463.48	
AIC		1515.48	
N		327	

Notes: Coefficients and their transformations are reported: Transformed treatment effects are changes in %. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, ^o $p < 0.15$.
Source: Own estimations based on merged UIR-LZAR database.

Table 13: Effect of treatment on probability to extend scope of search

	(1) <i>anticipation</i>		(2) <i>during coaching</i>		(3) <i>1-90d post-coa.</i>		(4) <i>91+d post-coa.</i>	
	coef	se	coef	se	coef	se	coef	se
treatment (D^{TG})	0.057	(0.092)	0.413***	(0.070)	0.043	(0.087)	0.012	(0.139)
X variables	yes		yes		yes		yes	
Caseworker FE	yes		yes		yes		yes	
Mean CG	0.176		0.129		0.190		0.282	
Observations	93		186		98		78	

Note: Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. CG=control group. No observations at t_0 since the outcome variable refers to a change.
Source: UIR-LZAR database

Note: Due to the fact that at t_0 no strategy changes are possible yet, this regression is not DiD as modeled in section 5.2. But given the zero level of the outcome at t_0 , the direct regression per period is equivalent here to DiD.

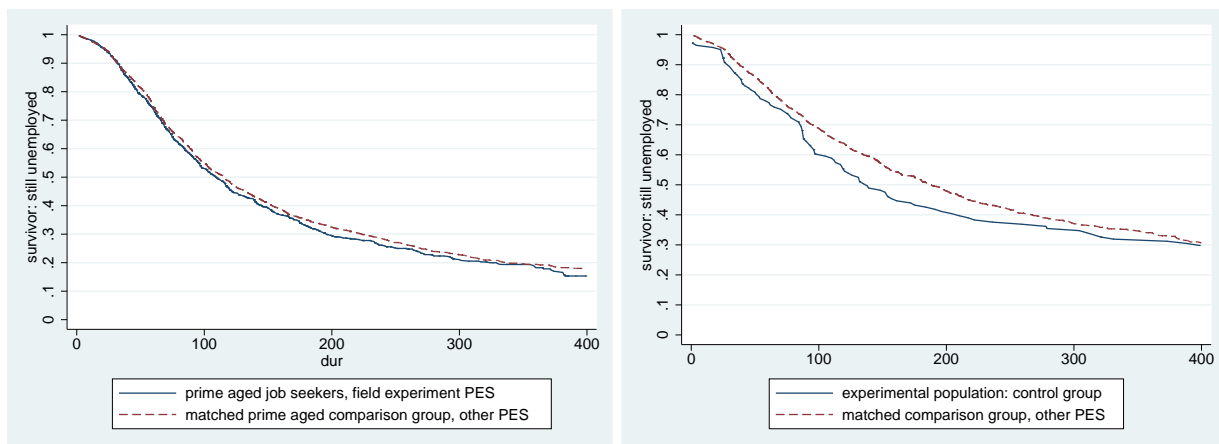
Table 14: Robustness check: effect of treatment on behavioral variables, IPWRA regressions

	(1) <i>anticipation</i>		(2) <i>during coaching</i>		(3) <i>1-90d post-coa.</i>		(4) <i>91+d post-coa.</i>	
	coef	se	coef	se	coef	se	coef	se
<i>Job interview rate</i>								
DiD	0.048	(0.045)	0.076	(0.037)	0.002	(0.037)	0.016	(0.050)
<i>Search effort (applications per month)</i>								
DiD	-1.359	(1.054)	-1.463	(0.822)	-1.277	(0.944)	-0.062	(1.030)
<i>Search channels: number</i>								
DiD	-0.745	(0.425)	0.113	(0.391)	-1.071	(0.492)	-0.211	(0.612)
<i>Search channels: frequency of use – formal channels: newspapers</i>								
DiD	-0.384	(0.150)	-0.050	(0.139)	-0.247	(0.185)	-0.519	(0.187)
<i>Search channels: frequency of use – formal channels: internet</i>								
DiD	-0.462	(0.212)	-0.160	(0.188)	-0.244	(0.289)	-0.462	(0.247)
<i>Search channels: frequency of use – formal channels: private recruiters</i>								
DiD	-0.352	(0.256)	-0.129	(0.188)	-0.415	(0.242)	0.046	(0.268)
<i>Search channels: frequency of use – informal channels: network</i>								
DiD	-0.126	(0.188)	0.202	(0.148)	-0.023	(0.170)	0.045	(0.249)
<i>Search channels: frequency of use – informal channels: spontaneous appl.: telephone</i>								
DiD	-0.065	(0.203)	0.467	(0.185)	-0.024	(0.237)	-0.117	(0.316)
<i>Search channels: frequency of use – informal channels: spontaneous appl.: written</i>								
DiD	-0.382	(0.223)	-0.003	(0.171)	-0.211	(0.216)	-0.088	(0.201)
<i>Reservation wage (ratio to last salary)</i>								
DiD	0.052	(0.030)	-0.054	(0.030)	-0.066	(0.047)	-0.090	(0.059)
<i>Beliefs: difference expected vs. realized interviews</i>								
DiD	-0.140	(0.125)	-0.144	(0.087)	-0.113	(0.114)	-0.073	(0.132)
<i>Beliefs: wage expectation (scale 1 to 5, 3=equal)</i>								
DiD	-0.301	(0.195)	-0.107	(0.146)	-0.304	(0.198)	0.083	(0.231)
<i>Motivation for job search (prop. high)</i>								
DiD	-0.055	(0.111)	0.066	(0.080)	0.206	(0.102)	0.148	(0.131)
<i>Self-confidence (prop. high)</i>								
DiD	0.087	(0.071)	0.064	(0.054)	0.185	(0.111)	0.073	(0.118)
<i>Reliability (prop. high)</i>								
DiD	-0.045	(0.089)	0.089	(0.053)	-0.013	(0.114)	0.044	(0.124)
X variables	yes		yes		yes		yes	
Caseworker FE	yes		yes		yes		yes	

Note: The reported standard errors are too high (since not corrected for the use of predictions in the first stage, see Wooldridge (2002) and footnote 46). Therefore, no significance levels are reported. However, the focus of the robustness check is on the point estimates anyway. Per regression, only the treatment effect (DiD, i.e. $D^{TG}T_i$) is shown in the table. The median regressions for reservation salaries could not be estimated since they do not allow for the necessary type of weights. Numbers of observations are reported in the main estimation tables.

Source: UIR-LZAR database

Figure 10: Comparisons between experimental PES/experimental control group and external control group (other PES): Kaplan-Meier survivors on being unemployed



Note: The left panel compares the survivor curves of prime-aged unemployed (30-44) in the PES agencies where the experiment was conducted (N=830) and in other PES of the state (N=1499). The right panel compares the survivor curves of the unemployed belonging to the age group of the experiment (45+) in the experimental control group (N=141) and in other PES of the state (N=1471). The comparison groups are matched by selecting on the same eligibility criteria which apply to the experimental population, specifically: skilled (professional certificate), employability middle or low, German mother tongue, last job full-time or part-time $\geq 50\%$, same inflow window. The curves are censored at 400 days of unemployment because of restricted observation window in the register data for the external control groups. PES = public employment service unit.