

Du, Shihan

Doctoral Thesis

Essays in Empirical Studies Based on Administrative Labour Market Data

PhD Series, No. 1.2019

Provided in Cooperation with:

Copenhagen Business School (CBS)

Suggested Citation: Du, Shihan (2019) : Essays in Empirical Studies Based on Administrative Labour Market Data, PhD Series, No. 1.2019, ISBN 9788793744455, Copenhagen Business School (CBS), Frederiksberg,
<https://hdl.handle.net/10398/9702>

This Version is available at:

<https://hdl.handle.net/10419/209095>

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.



<https://creativecommons.org/licenses/by-nc-nd/3.0/>

COPENHAGEN BUSINESS SCHOOL
SOLBJERG PLADS 3
DK-2000 FREDERIKSBERG
DANMARK

WWW.CBS.DK

ISSN 0906-6934

Print ISBN: 978-87-93744-44-8
Online ISBN: 978-87-93744-45-5

ESSAYS IN EMPIRICAL STUDIES BASED ON ADMINISTRATIVE LABOUR MARKET DATA

PhD Series 1-2019

Shihan Du

ESSAYS IN EMPIRICAL STUDIES BASED ON ADMINISTRATIVE LABOUR MARKET DATA

PhD School in Economics and Management

PhD Series 1.2019

CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

ESSAYS IN EMPIRICAL STUDIES BASED ON ADMINISTRATIVE LABOUR MARKET DATA

Shihan Du

Supervisors: Ralf Wilke and Birthe Larsen

PhD School in Economics and Management

Copenhagen Business School

Shihan Du
ESSAYS IN EMPIRICAL STUDIES BASED ON
ADMINISTRATIVE LABOUR MARKET DATA

1st edition 2019
PhD Series 1.2019

© Shihan Du

ISSN 0906-6934
Print ISBN: 978-87-93744-44-8
Online ISBN: 978-87-93744-45-5

The PhD School in Economics and Management is an active national and international research environment at CBS for research degree students who deal with economics and management at business, industry and country level in a theoretical and empirical manner.

All rights reserved.

No parts of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage or retrieval system, without permission in writing from the publisher.

Acknowledgement

This thesis would not have been possible without the guidance and help of several individuals who in one way or another contributed to the preparation and completion of my study.

First and foremost I would like to thank my parents. Pure love and missing to them.

Then, great thanks to my fiancé Dang Yibo, who loves me, supports me and encourages me in my difficult times. It is my greatest fortune to be with you.

I would like to express my special thanks of gratitude to my supervisor Ralf Wilke and Birthe Larsen for all their guidance. They teach me so many knowledge and principles of doing research work, which will benefit me throughout the rest of my life. I would also want to thank David, Fane, Dario, Lisbeth, Mauricio and all people at ECON for their great advice. Besides, I want to thank Mrs. Guo Qi who teaches me a lot in life.

Special thanks to Lone, Torben, Agata, Anja and Linea for their great support on administration since I came to CBS.

Finally, many thanks to my pretty buddies: Song Siqi and Vanessa, Han Jing and Sun Yigui, Zhuang Xinru and Feng Kai, He Lan, Zhao Ruijia, Cheng Shu and all the other friends who support me during my PhD life.

Abstract

This PhD thesis, entitled "Essays in Empirical Studies Based on Administrative Labour Market Data", is composed of three independent chapters, a general introduction for all three chapters at the beginning, and a brief conclusion in the end. While all three chapters are independent research papers and can be read as such, each chapter applies and compares different econometric frameworks by using individual-level administrative labour market data, addressing important topics within the field of labour economics.

The first chapter of my thesis, entitled "On Omitted Variables, Proxies and Unobserved Effects in Analysis of Administrative Labour Market Data", is written together with Ralf Wilke and Pia Homrighausen. We present a unified framework that nests various approaches aiming at reducing omitted variable bias in linear regression analysis. Linked administrative labour market data in Germany is used for our two empirical applications—wage regression and labour market transition model. We find empirical evidence for sizeable omitted variable bias in a wage regression, while only a small number of coefficients is systematically affected in the transition analysis. Benefit from the available linked administrative and survey data, it is found that additional survey variables contribute only to the wage model, while the use of work history variables and panel models lead to changes in coefficients in the two models. Overall, panel data models with a restricted regressor set are found to control for more unobserved effects than cross-sectional analysis with an extended variable set.

The second chapter, entitled "Impact of Immigration on the Wages of Native Workers in Denmark", examines the impact of immigrants on the wages of natives by using administrative data from Denmark Statistics on the full population in Denmark for the period from 2004 to 2013. Following Malchow-Møller et al. (2012), I apply OLS, FE and IV(2SLS) models for the empirical analysis. Then I extend their study by investigating into the quantile regression model, as not much previous literature has focused on the impact of immigrants with different skill levels and different wage quantiles on the wages of natives in Denmark. I find that high-skilled immigrants have a positive impact on natives, based on results from all estimation models. I also obtained evidence from the quantile regression indicates that the positive wage effect is mainly on natives who earn higher wages. In addition, according to the estimation results from the FE, FE-IV, and quantile regressions, it is found that low-skilled immigrants also have a positive effect on the wages of natives, and they have a more positive impact on low-wage natives. Through OLS, FE, and quantile estimations, I find that medium-skilled immigrants bring negative wage effects and the negative effects dominants for the medium-wage native group. I show that my hypothesised mechanisms—the wage efficiency theory as well as the demand-supply model—are strongly supported by the empirical evidence I obtained.

The last chapter in my thesis, entitled "Analysis on native-immigrant wage gap in Denmark", empirically investigates the native-immigrant wage gap as well as discrimination against immigrants for male workers in the labour market in Denmark. This is the first study to empirically examine the native-immigrant wage gap in the aspect of different skill levels and countries of origins, for the period of the year 2004 to 2013 in Denmark. I compare and apply Oaxaca-Blinder and Melly (2005) decomposition approaches using Danish register data. I find that the size of the wage gap is largely dependent on the skill level, and whether the wage gap is positive is more associated with an immigrant's country of origin. Wage differentials were generally the smallest within the low-skilled group. After controlling education, I find that a substantial part of this gap can be explained by the coefficient effect which is fully regarded

as potential discrimination in this study. By comparing across different ethnic groups, I find strong empirical evidence showing that the measured potential discrimination is the strongest and most positive for immigrants from less developed countries, most of which are non-EU countries. It is found through the decomposition approach based on quantile regressions that stronger potential discrimination occurs at the upper wage quantiles in each group of origin.

Abstract-Danish

Denne ph.d.-afhandling, med titlen "Essays i empiriske studier baseret på administrativ arbejdsmarkedsdata" er sammensat af tre uafhængige kapitler, en generel introduktion til alle tre kapitler i begyndelsen og en kort konklusion til slut. Mens alle tre kapitler er selvstændige forskningspapirer og kan blive læst som sådan, anvender og sammenligner hvert kapitel forskellige økonometriske strukturer ved at bruge individniveau administrativ arbejdsmarkedsdata, der adresserer vigtige emner inden for feltet arbejdsmarkedsøkonomi.

Det første kapitel af min afhandling, med titlen "Om udeladte variabler, proxier og uobserverede effekter i analyser af administrativ arbejdsmarkedsdata" er skrevet i samarbejde med Ralf Wilke og Pia Homrighausen. Vi præsenterer en forenet struktur, der indlejrer forskellige tilgange fokuseret på at reducere udeladt variabel bias i lineær regressionsanalyse i hinanden. Tysk linked administrativ arbejdsmarkedsdata bliver brugt til vores to empiriske anvendelser - lønregression og arbejdsmarkedsovergangsmodeller. Vi finder empirisk bevis for betydelige udeladte variabler i en lønregression, mens kun et lille antal af koefficienter er systematisk påvirket i overgangsanalysen. Med udgangspunkt i den tilgængelige linked administrativ- og surveydata finder vi at supplerende surveyvariabler kun bidrager til lønmodellen, mens brugen af arbejdshistorievariabler og panelmodeller fører til ændringer i koefficienter i de to modeller. Overordnet set bliver paneldatamodeller med en begrænset uafhængig variabel fundet til at kontrollere for flere uobserverede effekter end tværsnitsanalyse med et udvidet variabelsæt.

Det andet kapitel, med titlen "Påvirkning af lønninger for indfødte arbejdere i Danmark på immigration", undersøger hvilken indflydelse lønninger for indfødte arbejdere i Danmark har på immigration ved at bruge administrative data fra Danmarks Statistik på den fulde befolkning fra perioden 2004 til 2013. Efter Malchow-Møller et al. (2012), bruger jeg OLS, FE og IV(2SLS) modeller til den empiriske analyse. Derefter udvider jeg deres studie ved at undersøge fraktilregressionsmodeller, da der ikke er meget af den forudgående litteratur der har fokuseret på påvirkning af immigranter med forskellige færdighedsniveauer og lønfraktiler på lønningerne for de indfødte i Danmark. Jeg kan konstatere at immigranter med højt færdighedsniveau har en positiv indflydelse på de indfødte baseret på alle estimeringsmodeller. Jeg har også fundet bevis for at fraktilregressionen indikerer at the den positive løneffekt er fokuseret på indfødte med højere lønninger. Dertil skal det tilføjes at ifølge estimeringsresultaterne for FE-, FE-IV- og fraktilregressionerne finder jeg at immigranter med et lavt færdighedsniveau også har en positiv effekt på de indfødtes lønninger og at de har en mindre positiv indflydelse på lavtlønnede indfødte. Igennem OLS, FE og fraktilestimeringer kan jeg konkludere at immigranter med et færdighedsniveau på middel medbringer negative løneffekter og disse negative effekter dominerer for gruppen af indfødte med middelniveau færdigheder. Jeg viser at min hypotetiske mekanisme-løn effektivitetsteori, såvel som udbud- og efterspørgselsmodellen, er stærkt støttet af mit empiriske materiale.

Det sidste kapitel i min afhandling, med titlen "Analyse af indfødt-immigrant lønforskellen i Danmark", undersøger empirisk indfødt-immigrant lønforskellen såvel som diskrimination mod immigranter for mandlige arbejdere på arbejdsmarkedet i Danmark. Dette er det første studie der undersøger indfødt-immigrant lønforskellen med henblik på forskellige færdighedsniveauer og oprindelseslande empirisk for perioden 2004 til 2013 i Danmark. Jeg sammenligner og anvender Oaxaca-Blinder og Mellys (2005) nedbrydningstilgange med brug af dansk registerdata. Jeg finder at størrelsen på lønforskellen er i høj grad afhængig af færdighedsniveau og hvorvidt lønforskellen er positiv er mere forbundet med immigrantens

oprindelsesland. Lønforskelle var generelt set mindst inden for gruppen med lavt færdighedsniveau. Efter der blev kontrolleret for uddannelse, fandt jeg at en substantiel del af forskellen kan forklares med effektkoefficient, hvilket er betragtet som potentiel diskrimination i dette studie. Ved at sammenligne på tværs af forskellige etniske grupper for immigranterne, kunne jeg konkludere at der er stærkt empirisk bevis der viser at den målte potentielle diskrimination er stærkest og mest positiv for immigranter fra mindre udviklede lande, hvoraf de fleste ikke er EU-lande. Jeg fandt bevis gennem nedbrydningstilgangen baseret på fraktilregression for at øget potentiel diskrimination foregår i de øvre lønfraktiler inden for hvert oprindelsesland.

Contents

Acknowledgement	iii
Abstract	v
Abstract-Danish	viii
Introduction	1
Chapter 1 On Omitted Variables, Proxies and Unobserved Effects in Analysis of Administrative Labour Market Data	9
Chapter 2 Impact of Immigration on the Wages of Native Workers in Denmark .	59
Chapter 3 Analysis on Native-Immigrant Wage Gap in Denmark	126
Conclusion	313

Introduction

The recent decades have seen the increasing availability of rich data source, which has sparked a wave of both innovation and research. Within the field of labour economics, the increasing significance of administrative labour market data has undoubtedly extended the range for empirical studies. Comparisons on different econometric methodologies as well as tests on different theoretical hypothesis become possible with the access to large-scale individual-level data. Widely discussed issues, from generally discussed omitted variables bias to specific empirical analysis on labour markets, can be investigated and implemented by the use of more informative datasets.

In my PhD thesis (entitled "Essays in Empirical Studies Based on Administrative Labour Market Data"), different econometric methodologies are compared and then applied using administrative labour market data. The findings through three empirical studies provide insights into the topics within labour economics. More specifically, various econometric frameworks and empirical approaches to mitigate omitted variable bias are investigated in Chapter 1, the impact of immigration on the wages of native workers in Denmark is examined in Chapter 2, then the native-immigrant wage gap in Denmark is studied in Chapter 3, and finally a brief conclusion for all chapters is presented.

Administrative datasets from Germany and Denmark are used in this thesis. Both datasets are highly detailed and contain individual-level labour market data. Omitted variable bias in the analysis of administrative labour market data in the first chapter is studied using specific

linked data from Germany, while the other two empirical studies, on wage and immigration in Chapter 2 and Chapter 3, are based on registered data provided by Statistic Denmark. All three chapters are motivated by and developed based on existing literature, making a unique and novel contribution within their perspective fields.

The linkage of administrative and survey data has generated an abundance of additional information that was not previously accessible, which has induced a surge of extensive economic research into the topics of labour economics. I start the thesis by investigating into on common empirical strategies for reducing omitted variable bias in labour market research, together with Ralf Wilke and Pia Homrighausen, in Chapter 1 (entitled "On Omitted Variables, Proxies and Unobserved Effects in Analysis of Administrative Labour Market Data").

When there are one or more relevant variables missing in a model, the omitted variable bias in the estimation results will occur. In the empirical studies based on labour market data, researchers often use constructed variables from the individual work history, add survey-based variables to the administrative data, or use available panel data to mitigate omitted variable bias. However, little systematic research has been conducted to assess how the additional information contributes to reducing the omitted variable bias. Attempts to investigate their role are restricted to sensitivity analysis (e.g. Lechner and Wunsch, 2013; Arni et al., 2014; and Caliendo et al., 2014). Motivated by this research gap, we conduct a study to test how such additional information from labour market data contribute to the omitted variable bias in estimation. With the access to administrative data which is linked to extensive survey data from Germany, we are enabled to obtain relevant empirical evidence for our study.

In Chapter 1, We contribute to the existing literature by providing a unified framework that nests various approaches aiming at reducing omitted variable bias in linear regression analysis. Our approach exceeds a sensitivity analysis as it tests a number of relationships and restrictions that can be partly derived from a panel model. This is helpful in obtaining a more profound understanding of the viability of the different approaches. Moreover, we apply our framework to wage regression and a linear probability model for labour market transition anal-

ysis. It is found that the availability of longitudinal information for key variables appears to add more to the analysis than an exceedingly but possibly unfocused set of additional (survey) variables at some point. Our results are not only crucial for empirical researchers but also for data providers.

Based on the general guidance provided by Chapter 1, I then conduct two specific empirical applications on wages and immigration in Chapter 2 and Chapter 3. In Chapter 2 (entitled "Impact of Immigration on the Wages of Native Workers in Denmark"), I examine the impact of immigration on wages of native workers, using the Danish register data which contain more information on immigration than the German dataset applied for Chapter 1. My first motivation of this chapter is that Denmark has witnessed a substantial increase in the employment of immigrants since the early twenty-first century, especially after EU expansion (in the year 2004 and 2007). This fact makes it interesting to study the impact of the increasing immigrant population brought to the local labour market in Denmark. Secondly, Denmark is a place where individual-level and employer-employee linked labour market data for the full population are available. With the additional data for immigrants, I am able to investigate the wage effect of immigrants on natives empirically. The third motivation is that even though a vast amount of research projects have been carried out on this topic, few consistent empirical evidence has been found, particularly, in Denmark.

In Chapter 2, I briefly present reviews of selected studies on the impact of immigration on wages, both theoretically and empirically. As regards those empirical papers with statistically significant results, neither the U.S. nor European literature has reached a clear consensus. Some studies have suggested a positive impact of immigrants on the wages of native workers (Ottaviano and Peri, 2006; Ottaviano and Peri, 2012; Fogged and Peri, 2016, and etc.), while others have indicated a negative effect (Card, 2001); Ortega and Verdugo, 2016; Malchow-Møller et al., 2012, and etc.). Moreover, less empirical evidence has been provided on high-skilled immigrants. Therefore, I follow and extend the approaches used by Malchow-Møller et al. (2012). Apart from OLS, FE and IV models, I add a quantile regression model and

apply the Danish register data. Chapter 2 contributes to implementing the existing knowledge on the impact of immigration on the wages of native workers in Denmark, during the period (2004-2013) when immigrants increased rapidly. Particularly, this study provides empirical evidence under several wage quantiles as well as within each skill level in Denmark.

Not surprisingly, although immigrants only account for a minor share of the population in most countries, they have attracted increasing attention within both academia and politics. After the empirical study on the impact of immigration on the wages of native workers in Chapter 2, I turn my view on the aspect of immigrants. Following Chapter 2, Chapter 3 (entitled "Analysis on Native-immigrant Wage Gap in Denmark") analyses the native-immigrant wage gap, as well as the existence of wage discrimination against male immigrants, in different skill levels, ethnic groups and wage quantiles for the labour market in Denmark. The years after EU enlargement and the free movement in the labour market are of the interest (2004, 2007, 2009, 2010, and 2013).

Numerous theoretical and empirical studies have investigated into the native-immigrant wage gap for decades, which include Chiswick (1978), Kee (1995), Lehmer and Ludsteck (2011), and Hofer et al. (2017) etc. Several studies have empirically analysed income inequality in Denmark (e.g. Nielsen et al., 2004; Nielsen, 2011). However, except for Nielsen et al. (2004), empirical evidence on wage differentials by migration status is very scarce, not to mention empirical evidence for the period after EU enlargement. This research gap provided an incentive to conduct such an empirical study for the period from 2004 to 2013. Similar to Chapter 2, this study is also enabled by the rich administrative data and a rapidly growing immigrant population in Denmark during the year 2004 to 2013. Denmark is a case worthy of further analysis in terms of how the native-immigrant wage gap differs depending on skill level and nationality and whether potential discrimination plays a role in the wage gap.

Chapter 3 provides an overall summary of changes in the population and wage distributions for both of the native and immigrant groups in the labour market in Denmark for the period of 2004-2013. I apply and compare two decomposition frameworks—Oaxaca-Blinder (1973)

and Melly (2005)—for my empirical studies on the native-immigrant wage gap. I mainly focus on changes in the potential discrimination within skill levels and groups of origin over the period following EU enlargement in 2004 and 2007, and the introduction of free movement in Denmark in 2009. The findings from the empirical analysis contribute to the wage inequality literature in labour market in Denmark. Moreover, empirical evidence obtained in Chapter 3 provides comprehensive insights into native-immigrant wage gap and potential discrimination against immigrants, under different skill levels as well as within various ethnic groups.

Bibliography

Arni, P., Caliendo, M., Künn, S. & Mahlstedt, R. (2014), Predicting the risk of long-term unemployment: What can we learn from personality traits, beliefs and other behavioral variables, Technical report, Working Paper.

Blinder, A. S. (1973), 'Wage discrimination: reduced form and structural estimates', *Journal of Human resources* pp. 436–455.

Caliendo, M., Mahlstedt, R. & Mitnik, O. A. (2014), 'Unobservable, but unimportant? the influence of personality traits (and other usually unobserved variables) for the evaluation of labor market policies'.

Card, D. (2001), 'Immigrant inflows, native outflows, and the local labor market impacts of higher immigration', *Journal of Labor Economics* **19**(1), 22–64.

Chiswick, B. R. (1978), 'The effect of americanization on the earnings of foreign-born men', *Journal of political Economy* **86**(5), 897–921.

Foged, M. & Peri, G. (2016), 'Immigrants' effect on native workers: New analysis on longitudinal data', *American Economic Journal: Applied Economics* **8**(2), 1–34.

Hofer, H., Titelbach, G., Winter-Ebmer, R. & Ahammer, A. (2017), 'Wage discrimination against immigrants in austria?', *Labour* **31**(2), 105–126.

- Kee, P. (1995), 'Native-immigrant wage differentials in the netherlands: discrimination?', *Oxford Economic Papers* pp. 302–317.
- Lehmer, F. & Ludsteck, J. (2011), 'The immigrant wage gap in germany: Are east europeans worse off?', *International migration review* **45**(4), 872–906.
- Lechner, M. & Wunsch, C. (2013), 'Sensitivity of matching-based program evaluations to the availability of control variables', *Labour Economics* **21**, 111–121.
- Malchow-Møller, N., Munch, J. R. & Skaksen, J. R. (2012), 'Do immigrants affect firm-specific wages?', *The Scandinavian Journal of Economics* **114**(4), 1267–1295.
- Melly, B. (2005), 'Decomposition of differences in distribution using quantile regression', *Labour economics* **12**(4), 577–590.
- Nielsen, C. P. (2011), 'Immigrant over-education: evidence from denmark', *Journal of Population Economics* **24**(2), 499–520.
- Nielsen, H. S., Rosholm, M., Smith, N. & Husted, L. (2004), 'Qualifications, discrimination, or assimilation? an extended framework for analysing immigrant wage gaps', *Empirical Economics* **29**(4), 855–883.
- Oaxaca, R. (1973), 'Male-female wage differentials in urban labor markets', *International economic review* pp. 693–709.
- Ortega, J. & Verdugo, G. (2016), Moving up or down? immigration and the selection of natives across occupations and locations, Technical report, IZA.
- Ottaviano, G. I. & Peri, G. (2006), 'The economic value of cultural diversity: evidence from us cities', *Journal of Economic geography* **6**(1), 9–44.
- Ottaviano, G. I. & Peri, G. (2012), 'Rethinking the effect of immigration on wages', *Journal of the European economic association* **10**(1), 152–197.

Chapter 1

ON OMITTED VARIABLES, PROXIES AND UNOBSERVED EFFECTS IN ANALYSIS OF ADMINISTRATIVE LABOUR MARKET DATA

1

Author

Shihan Du²; Pia Homrighausen³, Ralf A. Wilke⁴

¹We thank the DIM unit of the IAB for providing the data and Arne Bethmann for his support with the PASS data.

²Copenhagen Business School, Department of Economics, Porcelaenshaven 16A, 2000 Frederiksberg, Denmark, E-mail: sd.eco@cbs.dk

³Institute for Employment Research (IAB), E-mail: pia.homrighausen@iab.de.

⁴Corresponding author: Copenhagen Business School, Department of Economics, Porcelaenshaven 16A, 2000 Frederiksberg, Denmark, Phone: +4538155648, E-mail:rw.eco@cbs.dk

Abstract: Empirical research addresses omitted variable bias in regression analysis by means of various approaches. We present a framework that nests some of them and put it to German linked administrative labour market data. We find evidence for sizeable omitted variable bias in a wage regression, while a labour market transition model appears to be less affected. Additional survey variables contribute only to the wage model, while the use of work history variables and panel models lead to changes in coefficients in the two models. Overall, panel data models with a restricted regressor set are found to control for more unobserved information than cross-sectional analysis with an extended variable set.

Keywords: linked survey-administrative data, statistical regularisation

1.1 Introduction

The problem of omitted variable bias is known as one of the classical issues in statistics. It occurs in estimation results when one or more relevant variables are missing in a model. The model attributes the effect of the missing variables to the estimated effects of the included variables, causing bias on estimation results. To mitigate the omitted variable bias, empirical research often includes the use of proxies or instrumental variables in an attempt to reduce omitted variable bias in multivariate statistical regression analysis.

In practice, the problem of missing crucial variables in an estimation model spans almost every empirical analysis within the field of labour economics. In empirical studies, based on labour market data, researchers often use constructed variables from the work history of individuals, add survey-based variables to the administrative data, or use available panel data to mitigate omitted variable bias. Despite the widespread use of work history and survey-based variables, little systematic research has been conducted to assess how they contribute to the estimation of the models. Motivated by this research gap, we conduct a study to examine how such additional information from labour market data contribute to the omitted variable bias in estimation.

In this paper, theoretically, we present a unified framework that nests various approaches that aim to reduce omitted variable bias in linear regression analysis. We then apply our approach to two widely studied empirical applications – wage regression and a linear probability model for labour market transition analysis, which are based on linked German administrative labour market data. Empirically, we provide evidence on to what extent does the additional information to reduce omitted variable bias contribute to the quality of results in our two applications. Moreover, in many countries, the use of administrative data and the addition of variables requires a well-justified research plan. The findings in our paper can be used as a guide.

This paper is organised as follows: in this section, introduction and background information of this paper are presented. Then in Section 1.2, the econometric problem is outlined. Section

1.3 describes the data and Section 1.4 shows the empirical findings. Finally, the last section concludes.

1.1.1 Linked administrative and survey data

Linked administrative data is increasingly used for empirical research in economics, social sciences and related disciplines. Their main advantages over survey-based data sources are bigger sample sizes and higher precision of key variables. Administrative data cover the population; hence its availability is not restricted to smaller and possibly non-random samples. Key variables are generated through operations in firms and public services. They are less prone to be misclassified due to few responding recall errors.

However, administrative data also have disadvantages over survey data. The variable set is restricted to information generated through operations. Thus, there is often a systematic lack of information on everything that exceeds the operational processes. This includes, for example, the motivation of individuals, their personality traits, the size of social networks and working climate in firms among many other things. Indeed, a number of studies based on survey data have shown that such additional variables contribute to the estimation model. Besides that, their availability enables the researcher to analyse problems which could not be analysed with administrative data. Examples include Nyhus and Pons (2005), Mueller and Plug (2006), and Heineck and Anger (2010) who use survey data with information about personality traits to analyse individual labour market outcomes.

The existence of administrative data does not directly imply that all information collected is indeed accessible to the researcher. In particular, not all variables may be available due to a lack of data linkage between administrative registers. Moreover, in common practice, due data confidentiality restrictions, data providers usually only give access to a random sample of the population data, and only to a restricted set of variables. Therefore, typical research based on administrative data is far away from using complete information about the population with all variables collected in administrative processes.

A sizeable random sample should not raise too many concerns for making inference with these data. However, the unavailability of important variables casts concerns for the consistency of estimates. There is extensive literature that considers the problem of omission of variables in regression analysis. For example, Gelbach (2016) suggests a variable selection approach that takes into account how much the omission of an available variable induces a bias for the coefficients on the other still included variables. In the case where the variables are missing due to the unavailable excess, Oster (2017) presents a comprehensive treatment on the omitted variables and suggests approaches on how to approximate the size of corresponding bias under restrictions.

1.1.2 Overview on empirical approaches for reducing omitted variable bias

One general empirical approach for reducing omitted variable bias is to include constructed variables from the individual work history. Examples include Kauhanen and Napari (2012) who use linked employer-employee data to study career and wage dynamics within and between firms in Finland. Fernández-Kranz and Rodríguez-Planas (2011) investigate the earnings effect of women who switch to part-time work under different types of contracts in Spain. Their study is based on Spanish longitudinal data from social security records. Baptista et al. (2012) obtain new insights into career mobility using Portuguese longitudinal matched employer-employee data. Using German administrative data, Biewen et al. (2014) conduct an analysis of the treatment effects of labour market programmes.

Work history variables may directly belong to the population model, or they might be proxies for otherwise unobserved variables such as performance. A prominent example in labour economics is that human capital is difficult to measure and usually unobserved. However, human capital is supposed to be an important variable in wage regression models. Therefore, researchers use test scores, such as the IQ, as proxies for human capital (compare Neal and Johnson, 1996; Bollinger, 2003). While the use of proxies is practically appealing, except for some cases under certain assumptions, there is no guarantee that their use leads to a bias

reduction or consistent estimation.

Another approach to mitigate the omission of variables is adding survey-based variables to the administrative data, especially information on personal traits. While adding variables is appealing, the generation of survey data is typically costly and time-consuming. Moreover, the question arises to what extent these variables indeed reduce the omitted variable bias in the model. The third approach is to use panel data instead of additional variables. The availability of panel data makes it possible to control for correlated unobserved time-invariant effects, reducing the need to control for as many variables as possible compared to cross-sectional analysis.

1.1.3 Motivation and contribution

The use of work history and survey-based variables to reduce omitted variable bias has been regarded as a common approach in the empirical literature. However, there is limited systematic research carried out to evaluate how those variables contribute to the estimation of the models. Attempts to investigate the role of work history and survey-based variables are so far restricted to sensitivity analysis, which is how the additional variables included in the model affects estimation results. For example, Lechner and Wunsch (2013), Arni et al. (2014) and Caliendo et al. (2014) investigate whether estimated treatment effects of labour market programmes on labour market outcomes are sensitive with respect to the inclusion of additional variables.

Our study exceeds a sensitivity analysis, as it tests a number of relationships and restrictions that can be partly derived from a panel model. This is helpful in obtaining a deeper understanding of the viability of the different approaches. We suggest a statistical framework that allows us to test the conditions for the work history variables to be feasible proxy variables. Moreover, we relate the results of the cross-sectional analysis with those of panel analysis, to investigate to what extent additional cross-sectional variables explain the variation in unobserved individual time-invariant effects. In our analysis, we do exemplary wage regressions and an analysis of labour market transitions. Our results suggest that additional

cross-sectional variables control for considerably less relevant information than fixed effects in panel analysis. Panel data analysis is found to give significantly different results, particularly, in the wage regression model. The endogeneity of a number of regressors in the cross-sectional models is confirmed. Our results are important for both empirical researchers and data providers.

This paper addresses research gap on the evaluation of how additional information contribute to the estimation of the models in labour market research as follows: Our starting point is a widely used administrative data product with only a limited number of variables. We use a sample of linked administrative data which are linked to extensive survey data from Germany. In particular, we use the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB), which is linked with the Panel Study "Labour Market and Social Security" (PASS). The PASS survey was funded by the German government to provide a more comprehensive database for the evaluation of the effects of the so-called Hartz reforms during the 2000s. Our data, therefore, contains many non-operations based variables which are not available in administrative data. Centred around this scenario, we provide a formal framework for estimation bias. The bias is due to the omission of important variables or the use of imperfect proxy variables. We then assess the contribution of additional survey-based non-operations related variables and work history variables to the model, as well as evaluate to what extent the variables change the estimation results.

1.2 The model

We consider the situation where a researcher has access to some standard administrative data product, with only a smaller number of administrative registers linked. Therefore, the set of available variables is restricted to some core variables. We restrict ourselves to the linear regression model. The population model is assumed to be:

$$y = X\beta + W\gamma + v, \tag{1.1}$$

where β ($J \times 1$) and γ ($L \times 1$) are unknown parameters, and the set of β is the one we are interested in this study. X ($1 \times J$) are observable regressors (with the first element being a constant) and W ($1 \times L$) are unobserved regressors. We will later relax this to some of the components of W being observed. We assume that the components of X and W are not perfectly multicollinear. y is observed and v is unobserved. We assume $E(v|X, W) = 0$.

1.2.1 General case of omitted variable bias

Because W is unobserved, the model in (1.1) cannot be directly estimated. Instead, one could choose ignore the unobserved variables and use OLS to estimate the following model (Equation 1.2):

$$y = X\beta + u, \quad (1.2)$$

where $u = W\gamma + v$. This is what is typically estimated in applications. It is well known that if $\text{cov}(x_j, u) \neq 0$ for some j causes $\hat{\beta}$, the OLS estimator for β , to be inconsistent. We focus here on a model with an unknown number(L) of omitted variables as this is the most realistic scenario in applications. When there are more than one omitted variables, the L linear projections of W onto the observable regressors are

$$W = X\delta + R,$$

with δ is $J \times L$ and R is $1 \times L$. Let r_l be the l 'th component of R . By definition $E(r_l) = 0$ and $\text{cov}(x_j, r_l) = 0$ for $j = 1, \dots, J$ and $l = 1, \dots, L$. When plugging W into (1.1) we obtain

$$y = X(\beta + \delta\gamma) + R\gamma + v.$$

In this model, we assume $\text{Cov}(X, \gamma) = 0$. All regressors are uncorrelated with the composite error, i.e. $E(v|X, W) = 0$, and therefore, the probability limit of the OLS estimator $\hat{\beta}$ for model (1.2) is

$$\text{plim}\hat{\beta} = \beta + \delta\gamma. \quad (1.3)$$

Equation 1.3 is the well known omitted variables bias and its size depends on the strength of the partial correlation between W and X , and the size of the elements of γ , i.e. the relevance of the omitted variables in the population model (1.1).

Since W is not observed, the size and direction of the bias are unknown in an application. For this reason, the approach developed in Gelbach (2016) that focuses on variable selection cannot be applied in our case. Although Gelbach has used omitted variables bias formula to construct a conditional decomposition that accounts for various covariates' role, in moving base regressors' coefficients. There is a limitation of his decomposition. In his approach, it is generally required that the regression function can be correctly written as a linear function of X and W . In order to make his framework valid in our case, X or W should not be endogenous, nor mismeasured.

Neither the approach developed by Altonji et al. (2005), for using the degree of selection on observables to investigate bias from the selection on unobservables, can apply in our case. There are strong assumptions in their approach, such as the number of observed X and unobserved W is large enough in order to avoid that any part dominates the distribution of the outcome y . The size of W is unknown thus it is difficult to judge whether it has a similar effect as the observed X on y . Therefore, the method in Altonji et al. (2005) is not suitable to be applied in our study. We also focus on alternative approaches aiming at reducing the omitted variable bias. However, none of these approaches is able to entirely remove bias or reveal the size of the bias in the absence of additional restrictions.

We then looked at the method developed by Oster (2017), because it is a new approach to estimate the omitted variable bias and would have therefore fit very well in our analysis. Oster provides an in-depth analysis of omitted variable bias. She shows that a consistent, closed-form estimator for omitted variable bias is possible to be constructed under less restrictive assumptions, e.g. without observing one or multiple W . In particular, her model considers the case of one component of X being related to W and requires that the components of W to be uncorrelated. The restrictions rely on the relationship between X and the omitted factors (proportional selection relationship), and knowledge of the R^2 of the population model.

We apply Oster's method in our empirical application, and a brief presentation of Oster's method is presented in Appendix II. We use R_{max} which is developed by Oster to test to what

extent additional information used to reduce omitted variable bias will contribute to the quality of results. It is found that the sign and magnitude of the estimated proportional selection relationship jumped strongly across variables. Given the coefficient instability and that the restrictions on the models in Oster (2017) exceed what we assume in our model, we only apply her method to our problem using information on some of the components of W (Z and W_1).

In this section, we suggest statistical frameworks on three common approaches to mitigate omitted variable bias: Add work history variables Z , add linked survey variables W_1 (a subset of W), and perform panel analysis with unobserved effects. Moreover, we present mechanisms for additional tests on several restrictions.

1.2.2 Add work history variables

One approach to mitigate omitted variable bias is to plug in constructed variables from the observable history of cross section units. In labour market research these are for example variables that characterise the work history of an individual and not simply lagged observable variables. These are denoted as Z ($1 \times P$). We assume that none of the components of X and Z are highly correlated or perfectly multicollinear in an application. In most applications P is a small integer and $P \leq L$. This means there are fewer constructed variables than omitted variables.

The role of Z requires some discussion. A special case is attained if a z_j is a proxy variable for one unobserved w_l , i.e. $z_j = w_l + \text{error}$ with $E(\text{error}) = 0$. However, more generally z_j can be related to any W , i.e. $z_j = \theta_0 + W\theta_j + m_j$ with $E(m_j|W) = 0$ for all j . θ_0 (1×1) and θ_j ($L \times 1$) are unknown to the researcher. If z_j is a proxy for w_l , then only the l 'th element of θ_j is non-zero. This is the case that is typically considered by the proxy variable literature (Lubotsky and Wittenberg, 2006, Bollinger and Minier, 2015). Using Z instead of W can be also interpreted as a measurement error problem. Here any deviation from the linear combination $W\theta_j$, which is m_j , is the measurement error. Alternatively, one could think of $z_j \in W$. In this case the constructed variable would directly belong to the population

model. Then $m_j = 0$, one component of θ_j is 1 and the others are 0. Lastly, z_j may not be correlated with any component of W . In this case $\theta_j = 0$ and z_j should not be included at all. A researcher normally faces the problem of not knowing the exact role of the components of Z . In any case it depends on the statistical relationship between X , W and the m_j s, whether the inclusion of Z mitigates or increases the omitted variable bias.

Given that W and L are unknown, it is more convenient to write the linear projection on the linear combination of W s, i.e, $W\gamma = \alpha + Z\lambda + e$ with $E(e|Z) = 0$ and parameters α (1×1) and λ ($P \times 1$). α is the intercept. e can be interpreted as the measurement or approximation error between $W\gamma$ and $Z\lambda$, which is the variation in the linear combination of unobserved variables that is not explained by the linear combination of constructed and included variables in Z . Therefore,

$$\begin{aligned} y &= X\beta + W\gamma + v \\ &= X\beta + Z\lambda + \alpha + e + v. \end{aligned} \tag{1.4}$$

For β in model (1.4) to be consistently estimated by OLS, it is additionally required that e is uncorrelated with X and v with Z . The former is not the case if X plays a role in the linear projection of Z and X on $W\gamma$, so it is required $E(W\gamma|X, Z) = E(W\gamma|Z)$. The latter requires $E(y|X, W, Z) = E(y|X, W)$, i.e. the redundancy of Z in the population model. Whether the bias in $\hat{\beta}$ in model (1.4) is smaller or greater than in model (1.2) is an empirical question. This depends on whether the correlations between the components of X and $W\gamma$ are greater or smaller than the correlations between the components of X and e , respectively. If for example the size of the components of δ are zero or very small, the inclusion of Z will increase the bias in $\hat{\beta}$ if there is correlation between Z and both X and v . Evidently, the better the fit of the model for $W\gamma$ on Z , the more likely plugging in Z leads to bias reduction in $\hat{\beta}$. This is because e becomes smaller in magnitude which reduces its covariance with X . It is remarked that λ has the interpretation of parameters of the linear projection on $W\gamma$ and we ignore the identifiability of α and the first component of β because the intercept is assumed to be not of interest. Bias in $\hat{\beta}$ is the focus in our application.

1.2.3 Add linked survey variables

Another approach to mitigate omitted variable bias is to enhance the regressor set by conducting a survey or by using additional administrative variables that are normally not accessible. Suppose that a subset W_1 of W , by assumption the first $L1$ variables of W , is observable in some random sample of the population. The idea is to do an analysis with a richer variable set. For direct comparability of the results across models we always restrict the analysis to the cross section units for which we have information on W_1 . Thus, we ignore the potential loss in precision and focus on asymptotic bias only. We consider the case, where the researcher is primarily interested in estimating the partial relationship between y and elements of X , rather than between y and elements W_1 , although the latter will be typically also of interest. W_2 is $1 \times L2$ and comprises of the last $L2$ elements of W with $L1 + L2 = L$. W_2 , the remaining unobservable variables, may be correlated with X and W_1 . Therefore, their omission induces a bias for estimated β and $\gamma_{(1)}$ in the regression of y on X and W_1 :

$$y = X\beta + W_1\gamma_{(1)} + u_2, \quad (1.5)$$

where $\gamma_{(1)}$ contains the first $L1$ elements of γ and $u_2 = W_2\gamma_{(2)} + v$, where $\gamma_{(2)}$ consists of the last $L2$ elements of γ . Unfortunately, there is no guarantee that including more variables indeed reduces the bias, but in practice one should expect this. The reason is that the number of summands in the bias term in equation (1.3) decreases from L to $L2$, when reducing the number of omitted variables. However, this may not lead to a reduction in the bias as the magnitude and sign of the various components of δ and γ are not restricted.

1.2.4 Panel analysis with unobserved effects

Instead of enhancing the set of observable variables, one can exploit the availability of longitudinal information, i.e. panel data, to mitigate the bias from the omission of W . y , X and Z are observed in periods $t = 1, \dots, T$ with $T \geq 2$ and observations are denoted as y_{it} , X_{it} and Z_{it} , respectively, for units $i = 1, \dots, N$. W_1 is assumed to be observed in one period only and W_2 is never observed, thus, W has to be omitted from the model. In order to relax the

exogeneity restrictions on X , we consider a fixed effects model:

$$y_{it} = X_{it}\beta + a_i + q_{it}$$

with $a_i + q_{it} = u_{it}$. a_i is assumed to be time invariant (the so called fixed effect) and q_{it} is a time varying error. Though, X is allowed to be correlated with a , the fixed effects estimator will only consistently estimate β if $E(q_{it}|X_i, a_i) = 0$ with $X_i = (X'_{i1}, \dots, X'_{iT})'$. However, this depends on the relationship between W and X because

$$\begin{aligned} y_{it} &= X_{it}\beta + W_{it}\gamma + v_{it} \\ &= X_{it}\beta + (\bar{W}_i + C_{it})\gamma + v_{it} \\ &= X_{it}\beta + a_i + q_{it} \end{aligned} \tag{1.6}$$

with $\bar{W}_i = \sum_{t=1}^T W_{it}/T$ and $q_{it} = C_{it}\gamma + v_{it}$. a_i therefore corresponds to the time constant part of $W_{it}\gamma$, which is not only the time constant variables in W but also the time average of the time varying components of W . $E(C_{it}\gamma|X_i, \bar{W}_i\gamma) = 0$ is required for consistent estimation by means of a fixed effects panel data model provided that v is idiosyncratic. It is also insightful to consider the role of Z when used in the fixed effects model. As discussed above, $W\gamma$ can be expressed as a linear combination of the Z plus a measurement error. In terms of the panel model this is $W_{it}\gamma = Z_{it}\lambda + b_i + s_{it}$. This linear projection decomposes the measurement error into a time constant part (b_i) and a time varying part (s_{it}). Then, for the main model we have

$$\begin{aligned} y_{it} &= X_{it}\beta + W_{it}\gamma + v_{it} \\ &= X_{it}\beta + Z_{it}\lambda + b_i + s_{it} + v_{it}. \end{aligned} \tag{1.7}$$

In order to consistently estimate β by means of a fixed effects model, b_i is allowed to be correlated with X_{it} and Z_{it} , but we need $E(s_{it}|X_i, Z_i, b_i) = 0$ and $E(v_{it}|X_i, Z_i, b_i) = 0$ with $Z_i = (Z'_{i1}, \dots, Z'_{iT})'$. The latter is again satisfied if Z does not play a role in the population model. The former, however, requires some discussion. b_i captures all time constant features of W which are not being absorbed by Z . The more of the time varying information of W is

captured by Z , the smaller is s_{it} . If the time varying information in Z_{it} is related to the time varying part of W_{it} , s_{it} is smaller in size than $C_{it}\gamma$. Then the inconsistency of the estimated β compared to model (1.6) is smaller. If the measurement error is time constant, i.e. $s_{it} = 0$, the fixed effects estimator for model (1.7) is consistent (Wooldridge, 2010). A roughly time constant measurement error (i.e. $s_{it} \approx 0$) may not be implausible in applications if Z_{it} has the interpretation of containing proxies.

1.2.5 Testable restrictions

In our empirical analysis, we do a comparative estimation of the various approaches aiming to reduce omitted variable bias. In order to examine to what extent the results are sensitive, we relate the result patterns to the theoretical considerations outlined in this section. Such analysis exceeds a sensitivity analysis which has been done in previous empirical studies. The underlying theory and the availability of W_1 as well as the estimated fixed effects provide a starting point for testing and checking the following restrictions:

- I. The role of Z .
- II. How the considered approaches relate, in terms of the ability to control for parts of $W\gamma$.
- III. Which of the X and Z show evidence of endogeneity.

Testable restrictions I The availability of W_1 makes it possible to get some ideas of how usually omitted variables are related to Z . In particular, one can estimate the strength of the relationship between $W_1\gamma_{(1)}$ and the Z . This shows which of the Z variables are related with unobservables and how much the variation in Z is able to explain the variation in $W_1\gamma_{(1)}$. A high R^2 would point to small measurement error. One can also test restrictions required for Z being a set of valid proxy variables. However, valid inference requires that a model without the omitted W_2 can be consistently estimated, i.e. W_2 is uncorrelated with all included variables. Testable restrictions are $E(W_1\gamma_{(1)}|X, Z) = E(W_1\gamma_{(1)}|Z)$ and $E(y|X, W_1, Z) = E(y|X, W_1)$, which have been motivated above. However, any correlations between (X, Z) and W_2 invalidate the inference.

Testable restrictions II Once panel models (1.6) and (1.7) have been estimated, one can check to what extent the survey variables W_1 explain the components of these models that control for the omitted W . One can test this by relating the estimated fixed effects to W_1 and Z in a cross-sectional model.

In our model settings, W_1 is observed and is a subset of W , and W_2 is never observed. Been discussed in the cross-sectional models, W_2 is $1 \times L2$ and comprises of the last $L2$ elements of W with $L1 + L2 = L$. W_2 , the remaining unobservable variables, may be correlated with X and W_1 . Similarly to the framework in model (1.5), a panel regression model of y_{it} on X_{it} and W_{1it} can be written as:

$$y_{it} = X_{it}\beta + W_{1it}\gamma_{(1)} + u_{2it}, \quad (1.8)$$

Where $u_{2it} = W_{2it}\gamma_{(2)} + v_{it}$, since the omission of W_{2it} in panel models induces a bias for estimated β and γ_1 in model (1.8).

In the case where W_{2it} is omitted, the effect of is included in an error term, as presented in model (1.8). We are not sure whether and how much W_{2it} can be captured by time consistent part. It has been shown that in panel models that: $a_i = \bar{W}_i\gamma$ in model (1.6), and $b_i = W_{it}\gamma - Z_{it}\lambda - s_{it}$ in model (1.7).

Therefore, we are able to test how much the estimated fixed effects can explain W_{1it} by relating the estimation result a_i and b_i to W_{1it} . After panel models (1.6) and (1.7) have been estimated, we can perform above test in a cross sectional model. Given that only W_1 is observed in one period, the following linear projections are suggested:

$$\hat{a} = W_1\rho + d \quad (1.9)$$

$$\hat{b} + Z\hat{\lambda} = W_1\varrho + f, \quad (1.10)$$

Where d and f in model (1.9) and (1.10) include the effect of W_2 . d and f are unobserved and uncorrelated with W_1 , and $E(d) = E(f) = 0$. The dependent variables in these models are

the estimated components of the panel models (1.6) and (1.7) that are supposed to control for the omitted W .

These regressions in model (1.9) and (1.10) can test two things : Firstly, the regression results can reveal which components of W_1 are indeed at least controlled for to some extent. This can be indicated through the test on whether there is a linear partial relationship between the components of W_1 and the dependent variables.

Secondly, the R^2 of these models shows us how much the variation in W_1 explains the variation of the components that control for W . A low R^2 would point to that the panel models mainly control for information that is not in W_1 and Z , thus the fixed effects a_i and b_i captures the information in d , which includes W_2 . This result would suggest that a panel analysis using a reduced regressor set is expected to be the more fruitful empirical approach than a cross sectional analysis with an expanded regressor set. In contrast, if the R^2 was high, the reverse applies. And this would suggest that the fixed effects capture only little time constant information of W_2 , meaning that a fixed effects panel analysis does not control for much more than what is in W_1 . It is remarked that the R^2 of models (1.9) and (1.10) increases with $L1$ and approaches 1 if the entire W was used. Moreover, the models use W_1 at one time point and not the time constant part of W_1 which is expected to result in a lower R^2 . However, the more important the cross sectional variation in W_1 than the longitudinal variation, the smaller the expected effect on the R^2 .

Testable restrictions III Finally, simple regression based tests of the endogeneity of X and Z can be conducted once fixed effects have been estimated. The idea is here to regress \hat{a} or \hat{b} on X or (X, Z) , respectively. Any significant relationship points to that the fixed effects are partially correlated with the observables, thus leading to inconsistencies of OLS estimates for β for models (1.2) or (1.4). These tests will also reveal which variables or groups of variables possess these patterns.

1.3 German Administrative Data linked with Survey Data

For our analyses, we use the Integrated Employment Biographies (IEB) of the IAB. These administrative registers contain information for every German once employed in a job subject to social insurance contributions since 1973. This information includes socio-demographic characteristics as well as daily records on employment and job seeking periods, receipt of unemployment benefits and information about participation in active labour market policy programs.

Usually, access to these data is restricted to random samples and a subset of variables due to data confidentiality reasons. In our application, we mimic the situation of a researcher working with a standard administrative data set, which is accessible to a wider group of data users. In particular, we focus on the widely used scientific-use-file version of the "Sample of Integrated Labour Market Biographies" (SIAB, cf. vom Berge et al. 2013). The SIAB is a 2 percent random sample drawn from the IEB (approximately 1.6M individuals) and provides restricted access to variables available in the IEB records. The SIAB is available as a standard data set through the Research Data Center (FDZ) of IAB (<http://fdz.iab.de/>).

We enrich the administrative data by linking it with comprehensive survey data on the individual level, with the household panel study "Labour Market and Social Security" (PASS, cf. Berg et al. 2012). The PASS survey was implemented in 2006 to gain more insights into the living conditions of (means-tested) unemployment benefit recipients in the household context. Since then, the PASS survey, in general, provides several waves of survey data from household and individual interviews on a wide variety of issues relating to the socio-economic situation. About 80 percent of the individuals interviewed in the PASS survey agreed to link the PASS survey data to the administrative records (approximately 22,000 individuals). A very similar linked dataset is the "PASS survey data linked to administrative data of the IAB" (PASS-ADIAB) that is also available through the Research Data Center (FDZ) of IAB. For more information on these data see Antoni and Bethmann (2014).

Table 1.1: Data sources

	Size	IEB Variables	SIAB Variables (X)	PASS survey variables (W_1)
Integrated Employment Biographies (IEB)	100% of the population	x	x	
Sample of Integrated Labour Market Biographies (SIAB)	2% of IEB		x	
Panel Study "Labour Market and Social Security" linked with IEB (PASS-ADIAB)	0.03% of IEB		x	x

For our comparative analysis, we restrict the sample to individuals aged 16 to 64 of different households who have participated in the 5th wave of the PASS survey in 2011. This leaves us with approximately 9,700 individuals. Since it is a common situation that survey data is only available for one period, we do not use further waves of the PASS survey. We restrict the analysis to the 5th wave to have information on personality traits that are not available in prior waves. Using both, the restricted IEB data as well as information from the PASS data, our sample contains variables from administrative registers available in the SIAB (X), generated work history variables (Z) as well as additional survey-based variables from PASS (W_1). Information on the size of each dataset and how they are linked to our empirical applications are shown in Table 1.1.

With these data, we perform two exemplary applications: one wage regression and one labour market transition analysis. Focusing only on individuals who are observed at least three years in the administrative data, the sample of the wage regression consists of 2,435 persons employed during the interview months. The sample of the transition regression consists of 1,484 persons who once have been registered as unemployed during the interview year and are observed at least for three years in the administrative data. The dependent variable y of the wage regression is the logarithmized average daily gross wage at the time of the interview. X includes socio-demographic and employment-related variables such as gender, age, trainee status, education, nationality, and industrial sector. The dependent variable y in

the transition analysis is a dummy variable indicating whether an unemployed individual left unemployment within 12 months ($y = 1$) or not ($y = 0$). As regressors, we use a subset of the variables of the wage regression as well as dummies of unemployment related registers such as the receipt of unemployment insurance benefits (German: Arbeitslosengeld, ALG I) and means-tested unemployment benefits (German: Arbeitslosengeld II, ALG II). Table 1.9 in Appendix III presents the full set of regressors used in the wage and transition regression as well as their descriptive statistics.

The survey-based variables constituting W_1 are linked PASS data. Among the survey variables, those supposed to have an impact on wage levels and/or labour market transitions, are of special interest. While the survey incorporates a wide array of topics, we mainly focus on labour market-related information. This includes information on personality traits and attitudes (Big Five), job search, working hours and other social factors. Table 1.9 in Appendix III presents the full set of survey variables used as well as their descriptive statistics.⁵ Despite that we use a rich set W_1 variables, there may well be further important variables in the population models that are unobservable to us (and thus in W_2).

Variables Z are constructed from individual (un-) employment histories. Thus, they are computed from past administrative records on employment and unemployment among other past labour market outcomes. We construct four variables for the wage regression: length of job tenure, the share of time employed over a total length of recorded labour market history, past unemployment history, and working experience. For the transition analysis, we construct five variables: past unemployment history at the time of transition, duration of current unemployment episode, recall history, past long-term unemployment (i.e. last unemployment episode longer than 12 months), and participation in active labour market programmes within the last three years.

⁵See www.fdz.iab.de for a full list of variables available in the PASS survey data.

1.4 Empirical Analysis

Due to the limited size of the survey population, we restrict ourselves to two exemplary linear regression models: A wage regression and a linear probability transition model. For both models, we do the analysis steps as outlined in Section 1.2. From our findings, we derive some general guidance for empirical researchers who work with these or similar data.

The idea behind using work history variables Z in wage or transition models is twofold: These variables capture otherwise unobserved individual features related to past labour market performance and therefore they can be interpreted as proxy variables. In our application Z include among others the unemployment history and tenure with the current employer. While past unemployment experiences should be related to work motivation and performance, the tenure in a job should reflect job specific skills. Thus, these variables are correlated with something that is typically not observable.

However, work history variables may actually belong to the population model. This is for example if past unemployment experiences play a direct role in hiring decisions and therefore for the probability of starting a new job. Similarly, job safety or the collective wage bargaining process can be direct functions of tenure due to legal restrictions. In many countries, dismissal protection is stronger for long-time employees and recently hired employees usually are not entitled to wage increases. However, if a component of Z belonged to the model, it is correlated with the unobservables for the reasons mentioned above. Thus, it is endogenous. This is why adding additional variables W_1 to the model is expected to not only uncover endogeneity of X but in particular of Z . The PASS data provide a large number of additional variables. Similar to the method mentioned in Belloni et al. (2014), We apply the Post-LASSO and an elastic net (see Appendix I) as tools for the selection of relevant variables in the two models. While for the wage regression 35 variables are selected as the set of relevant W_1 variables (see Table 1.8 in Appendix III), none of the survey variables appears to be relevant in the transition model.

1.4.1 Wage Regression

We consider a standard Mincer type wage equation with y is the log of the average daily gross wage. As regressors X we use individual level and firm level data such as age, gender, education and industry. As Z we use work history variables related to previous working experience, tenure and previous unemployment experiences. For a complete list of variables in this model see Table 1.9 in Appendix III.

As the first step we apply ordinary least squares to estimate linear models for $E(y|X)$, $E(y|X, Z)$, $E(y|X, W_1)$ and $E(y|X, Z, W_1)$. Table 1.2 contains the main estimation results for these models, denoted by W.A-W.D. The coefficients on W_1 are for completeness reported in Table 1.10 in Appendix III. The R^2 increases from 0.32 in Model W.A to 0.41 in Model W.B and to 0.50 in Model W.C. It increases further to 0.57 in Model W.D, pointing to that the set of variables individually and jointly contribute to explaining variation in the dependent variable. It is found that a number of coefficients on X differ considerably across the models, pointing to omitted variable bias. For example the coefficient on gender decreases from 0.499 in Model W.A to 0.148 in Model W.D. This suggests that the estimated gender wage gap is much smaller (only around 14% compared to 39%) when non-operations based variables are included. Although still highly significant, this is an economically relevant reduction. In contrast, other coefficients such as nationality and several business sectors are invariant across models W.A-W.D.

Table 1.2: Wage regression: Dependent variable $\log(\text{wage})$

	W.A	W.B	W.C	W.D	W.E	W.F
	$E(y X)$ coef. / (SE)	$E(y X, Z)$ coef. / (SE)	$E(y X, W_1)$ coef. / (SE)	$E(y X, Z, W_1)$ coef. / (SE)	$E(y X, Z, W_1)$ coef. / (SE)	$E(y_{it} X_{it}, Z_{it})$ coef. / (SE)
Gender (male=1)	0.499*** (0.026)	0.443*** (0.024)	0.174*** (0.027)	0.148*** (0.026)	6.537* (3.564)	6.199* (3.498)
Age	0.006*** (0.001)	-0.001 (0.001)	0.011*** (0.001)	0.002 (0.001)	-0.055 (0.084)	-0.047 (0.082)
Dummy: trainee	-0.437 (0.369)	-0.342 (0.305)	-0.628** (0.285)	-0.518** (0.246)	-0.494*** (0.166)	-0.515*** (0.164)
Missing information on education	-0.528*** (0.163)	-0.438*** (0.154)	-0.394*** (0.141)	-0.300* (0.154)	0.060 (0.180)	0.055 (0.175)
No formal degree	-0.264** (0.117)	-0.215** (0.107)	-0.146 (0.090)	-0.112 (0.081)	-0.053 (0.146)	-0.040 (0.141)
Vocational training	0.030 (0.113)	-0.008 (0.103)	0.045 (0.085)	0.022 (0.076)	0.062 (0.133)	0.080 (0.127)
Higher Education	0.522*** (0.117)	0.483*** (0.107)	0.434*** (0.088)	0.417*** (0.079)	0.050 (0.130)	0.056 (0.125)
Dummy: German nationality	0.030 (0.058)	-0.057 (0.055)	0.001 (0.048)	-0.072 (0.045)	-0.048 (0.130)	-0.030 (0.127)
Agriculture	-0.627*** (0.097)	-0.443*** (0.093)	-0.558*** (0.077)	-0.404*** (0.080)	-0.189 (0.561)	-0.174 (0.561)
Hotel and restaurant	-0.543*** (0.076)	-0.368*** (0.074)	-0.571*** (0.063)	-0.436*** (0.059)	-0.309 (0.200)	-0.328* (0.197)
Construction	-0.310*** (0.055)	-0.211*** (0.053)	-0.284*** (0.049)	-0.200*** (0.047)	0.139 (0.155)	0.130 (0.155)
Trade	-0.249*** (0.038)	-0.175*** (0.035)	-0.193*** (0.034)	-0.135*** (0.031)	-0.025 (0.087)	-0.022 (0.086)
Services	-0.232*** (0.034)	-0.124*** (0.032)	-0.198*** (0.032)	-0.115*** (0.030)	-0.108* (0.064)	-0.115* (0.063)
Education and social health	-0.136*** (0.037)	-0.054 (0.035)	-0.092*** (0.032)	-0.033 (0.030)	-0.066 (0.101)	-0.073 (0.099)
Public institutions	0.082* (0.045)	0.070 (0.044)	0.086** (0.037)	0.076** (0.036)	0.158 (0.179)	0.153 (0.177)

Continued on next page

... table 1.2 continued

	W.A	W.B	W.C	W.D	W.E	W.F
Other sectors	-0.083 (0.074)	-0.010 (0.061)	-0.015 (0.062)	0.038 (0.050)	-0.514** (0.226)	-0.502** (0.218)
Tenure (in years)		0.019*** (0.002)		0.018*** (0.002)		0.005 (0.006)
Share of working experience over total observation time		0.217*** (0.047)		0.118*** (0.043)		-0.123 (0.106)
Additional working experience (in years)		0.011*** (0.002)		0.013*** (0.002)		0.008 (0.007)
Dummy: unemployment history in the past		-0.492*** (0.032)		-0.427*** (0.029)		0.101 (0.145)
Constant	3.770*** (0.131)	4.200*** (0.125)	2.506*** (0.189)	2.996*** (0.178)		
N	2435	2435	2435	2435	3 × 2435	3 × 2435
R ²	0.319	0.412	0.502	0.570	0.997	0.997

Robust standard errors of model W.A-W.D and clustered standard errors of model W.E-W.F in parentheses.

* p<0.10, ** p<0.05, *** p<0.010

The coefficients on several components of X , such as gender and higher education, change monotonically from Models W.A to W.D. This could be interpreted as an improvement of the estimates and a reduction in the omitted variable bias as the model R^2 increases. As outlined in Section 1.2, however, there is no theoretical foundation that this is always true. For some X , such as vocational training and nationality, the change is small and not statistically significant. For other variables in X , such as trainee, the coefficients do not change monotonically (although not significantly) but they gain in precision and become statistically significant. As all Z variables are individually significant in Model W.D, the restriction $E(y|X, W_1, Z) = E(y|X, W_1)$ is violated. It can be seen from Table 1.3 that all but one component of Z are individually significant in the linear projection on $W_1\hat{\gamma}_{(1)}$. This suggests that there is a statistical partial relationship between the linear combination of Z and the linear combination of W . However, the R^2 of only 0.04 points to that the variation in Z only very little explains the variation in $W_1\hat{\gamma}$ and therefore Z are poor proxies for W_1 . This is also confirmed by a rejection of the restriction $E(W_1\gamma_{(1)}|X, Z) = E(W_1\gamma_{(1)}|Z)$ with a P-value of virtually 0. Moreover, the coefficients on Z are mainly unchanged between Models W.B and W.D., which also suggest that the endogeneity of Z is not removed by adding W_1 . If anything, these observations suggest that the Z variables are either components of W_2 or they proxy for components in W_2 . This would be in line with the increase in the R^2 when we go from Model W.C to W.D.

Table 1.3: Wage regression: Test restrictions for Z being feasible proxy variables

	$E(W_1\hat{\gamma}_{(1)} Z)$ coef. / (SE)
Tenure (in years)	-0.000 (0.001)
Share of working experience over total observation time	0.241*** (0.028)
Additional working experience (in years)	-0.006*** (0.001)
Dummy: unemployment history in the past	-0.175*** (0.019)
N	2435
R^2	0.042

Robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

In order to shed more light on the role of W_1 and Z in the previous models we estimate panel data regression (1.6) and (1.7) with 3 periods for the same individuals as for the other models. We include period interactions for all regressors and only report the coefficients for the period that is used in the cross sectional models. In order to obtain coefficients on the time constant variables, we estimate a dummy variable regression model with 2,435 individual specific dummy variables. The results - without the estimated α - are displayed in Table 1.2 as Models W.E and W.F, respectively. It is evident that the coefficients on several of the X and Z variables change considerably when using a panel model that allows for correlation between (X, Z) and the time constant part of the error. This points to violations of the stronger exogeneity restrictions in cross sectional analysis. For example, the coefficient on higher education drops sharply from 0.483 in Model W.B to 0.056 in Model W.F. A similar pattern can be observed for several of the business sectors, while other previously strongly significant coefficients become weakly or insignificant in the panel analysis (e.g. gender). The multicollinearity pattern driving this result is briefly discussed at the end of this subsection. But there are also variables, such as trainee, for which precision increases. The coefficients on the Z variables decrease in magnitude and these variables become considerably less individually significant. A robust test whether the components of Z are jointly significant in Model W.F has a p-value of 0.704. This observation and given that the R^2 of Model W.F is not higher than that of Model W.E suggest that Z does not additionally contribute to the model. The relevance of Z in Models W.B and W.D is therefore more likely due to correlation with W_2 rather than because Z directly belongs to the population model.

In the following we shed light on two more questions: First, to what extent do the variables in W_1 explain the variation of the estimated part of the panel model that is supposed to capture the omitted W ? Second, to what extent are the estimated fixed effects statistically related to the included X and Z ? Any relationship suggests endogeneity of the latter in a cross sectional regression.

Table 1.4: The statistical relationship between the estimated component of the panel model that controls for omitted W and the observable W_1

	$E(\hat{a} W_1)$ coef. / (SE)	$E(\hat{b} + Z\hat{\lambda} W_1)$ coef. / (SE)
Big Five: I am rather cautious, reserved	0.036 (0.047)	0.029 (0.057)
Big Five: I tend to criticise people	0.000 (0.041)	-0.013 (0.051)
Big Five: I attend to all my assignments with precision	0.044 (0.066)	0.045 (0.080)
Big Five: I have versatile interests	-0.132** (0.060)	-0.171** (0.073)
Big Five: I am inspirable and can inspire other people	0.027 (0.052)	0.032 (0.064)
Big Five: I easily trust in people and believe in the good in humans	0.070* (0.041)	0.094* (0.050)
Big Five: I tend to be lazy	-0.203*** (0.043)	-0.250*** (0.053)
Big Five: I am profound and like to think about things	-0.115** (0.045)	-0.139** (0.055)
Big Five: I am rather quiet, introverted	-0.292*** (0.046)	-0.375*** (0.056)
Big Five: I can act cold and distant	0.004 (0.040)	0.017 (0.049)
Big Five: I am industrious and work hard	0.200*** (0.076)	0.282*** (0.092)
Big Five: I worry a lot	0.225*** (0.041)	0.291*** (0.051)
Big Five: I have a vivid imagination and have a lot of phantasy	-0.188*** (0.052)	-0.225*** (0.063)
Big Five: I am outgoing and like company	-0.005 (0.054)	0.021 (0.066)
Big Five: I can be gruff and repellend towards other people	-0.117*** (0.043)	-0.140*** (0.053)
Big Five: I make plans and carry them out	-0.025 (0.057)	-0.038 (0.070)
Big Five: I easily get nervous and insecure	0.141*** (0.048)	0.200*** (0.058)
Big Five: I treasure artistic and aesthetic impressions	0.219*** (0.047)	0.248*** (0.057)
Big Five: I am not very interested in art	-0.152*** (0.044)	-0.176*** (0.053)
Dummy: satisfied with one?s life in general	0.389*** (0.142)	0.410** (0.173)

Continued on next page...

... table 1.4 continued

	$E(\hat{a} W_1)$ coef. / (SE)	$E(\hat{b} + Z\hat{\lambda} W_1)$ coef. / (SE)
Dummy: was looking for a new job	-0.458*** (0.165)	-0.402** (0.202)
Dummy: was looking for an additional job	-0.708* (0.379)	-0.808* (0.463)
Dummy: was looking for a new and an additional job	0.170 (0.881)	0.402 (1.078)
strength of connection to place of residence	-0.031 (0.046)	-0.033 (0.057)
Frequency of misunderstandings, tensions or conflicts	-0.108** (0.049)	-0.168*** (0.059)
Number of children in total (within and outside the household)	0.212*** (0.056)	0.090 (0.068)
Number of children in household	0.295*** (0.084)	0.415*** (0.103)
Dummy: none of parents has a HE degree	0.056 (0.092)	0.019 (0.112)
Dummy: one parent has a HE degree	0.032 (0.173)	-0.016 (0.212)
Current contract working time, total, without mini-job	-0.042*** (0.008)	-0.055*** (0.010)
Current actual working time, main occupation, without mini-job	-0.074*** (0.014)	-0.095*** (0.017)
Current actual working time, total, without mini-job	0.024* (0.014)	0.029* (0.017)
Dummy: none of parents with migrational background	-0.369** (0.182)	-0.314 (0.223)
Size of household	-0.753*** (0.064)	-0.844*** (0.078)
Constant	9.482*** (0.639)	9.667*** (0.782)
N	2435	2435
R^2	0.327	0.334

Robust standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.010

Table 1.4 displays the results of the linear projections of W_1 on the estimated components of the panel models that capture the unobserved W as given by (1.9) and (1.10) for the cross sectional data. In the case of Model (1.6) this is simply the estimated fixed effects \hat{a} . In the case of Model (1.7), this is the estimated fixed effect plus the estimated component related

to Z , i.e. $\hat{b} + Z\hat{\lambda}$. The estimated coefficients are from the panel regressions. Given that the two regressions in Table 1.4 have different dependent variables with different variation, the estimated coefficients and the R^2 are not directly comparable. However, they show that the variation in W_1 explains around one third of the variation of the dependent variables. They also show that a number of W_1 variables is partially related with the dependent variables. This is evidence that for the panel models effectively controlling for information in W_1 without directly using it. However, the remaining $2/3$ of the variation must be due to W_2 . This suggests that the panel models also effectively control for additional unobservables.

Table 1.5: Wage regression: Regression based endogeneity test for components of X and Z

	$E(\hat{a} X)$ coef. / (SE)	$E(\hat{b} X, Z)$ coef. / (SE)
Gender (male=1)	-4.867*** (0.026)	-6.249*** (0.024)
Age	0.060*** (0.001)	0.034*** (0.001)
Dummy: trainee	-0.150 (0.375)	-0.033 (0.299)
Missing information on education	-0.582*** (0.158)	-0.483*** (0.155)
No formal degree	-0.188* (0.107)	-0.126 (0.098)
Vocational training	-0.050 (0.103)	-0.082 (0.094)
Higher Education	0.329*** (0.106)	0.304*** (0.097)
Dummy: German nationality	0.066 (0.059)	-0.037 (0.055)
Agriculture	-0.473*** (0.094)	-0.285*** (0.095)
Hotel and restaurant	-0.132 (0.080)	0.056 (0.075)
Construction	-0.285*** (0.059)	-0.179*** (0.058)
Trade	-0.147*** (0.039)	-0.076** (0.036)
Services	-0.145*** (0.035)	-0.030 (0.033)
Education and social health	-0.185*** (0.037)	-0.102*** (0.034)
Public institutions	-0.192***	-0.196***

Continued on next page...

... table 1.5 continued

	$E(\hat{a} X)$ coef. / (SE)	$E(\hat{b} X, Z)$ coef. / (SE)
	(0.045)	(0.044)
Other sectors	0.223*** (0.080)	0.293*** (0.066)
Tenure (in years)		0.017*** (0.002)
Share of working experience over total observation time		0.327*** (0.048)
Additional working experience (in years)		0.004* (0.002)
Dummy: unemployment history in the past		-0.479*** (0.034)
Constant	3.735*** (0.124)	4.172*** (0.120)
N	2435	2435
R^2	0.954	0.974

Robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 1.5 reports the results for regressions of the estimated fixed effects from the panel analysis on the included regressors in the two models using the cross sectional data. It is apparent that a large number of the coefficients differ significantly from 0. This points to partial correlation between fixed effects and regressors and thus to endogeneity of the latter in the cross sectional models of Table 1.2. This means there is significant bias in many of the estimated coefficients of the cross sectional models W.A and W.B in Table 1.2. The large values of the R^2 for the two models in Table 1.5 reveal that the included regressors nearly entirely explain the variation in estimated fixed effects. This causes a strong multicollinearity pattern between some of the variables in the panel models of Table 1.2, which is reflected by the partly huge standard errors in Models W.E and W.F., e.g. for the coefficient on gender. A solution to mitigate this pattern would be to use information from additional periods but this would then lead to an unbalanced panel.

1.4.2 Transition Analysis

In this subsection we repeat the analysis by considering a probability model for leaving unemployment. The framework In Section 2 is applied in a linear probability model. In particular, the dependent variable " y " is now binary, and takes on the value one if the individual has left unemployment within 12 months since the time of the interview. We apply the linear probability model($P(y = 1|X)$, $P(y = 1|X, Z)$, $P(y_{it} = 1|X_{it})$ and $P(y_{it} = 1|X_{it}, Z_{it})$) to estimate the partial relationship between various observables and the transition probability. As this analysis is restricted to unemployed job seekers, the sample conditions on those being in the job seekers register. For this reason, firm level variables are no longer available but other variables such as claiming unemployment benefits. The full set of variables is again provided in Table 1.9 in Appendix III. Also, the set of work history variable Z changes and becomes related to past unemployment experiences and participation in active labour market policy programs. As mentioned above, none of the survey variables have been selected by the Post-Lasso and elastic net, thus the set W_1 is empty. This suggests that the survey does not contribute relevant information to the analysis.

Table 1.6: Transition analysis: Binary dependent variable (Dummy: left unemployment within 12 months)

	T.A $E(y X)$ coef. / (SE)	T.B $E(y X, Z)$ coef. / (SE)	T.C $E(y_{it} X_{it})$ coef. / (SE)	T.D $E(y_{it} X_{it}, Z_{it})$ coef. / (SE)
Gender (male=1)	0.043* (0.022)	0.034 (0.021)	0.111 (0.879)	-0.833 (0.966)
Age	-0.003*** (0.001)	-0.002* (0.001)	0.016 (0.023)	0.033 (0.026)
Missing information on education	-0.007 (0.104)	0.006 (0.099)	0.048 (0.183)	0.013 (0.143)
No formal degree	-0.010 (0.095)	-0.030 (0.090)	-0.003 (0.172)	-0.033 (0.133)
Vocational training	0.025 (0.094)	0.005 (0.089)	0.030 (0.168)	-0.003 (0.127)
Higher Education	0.043 (0.128)	-0.021 (0.127)	0.094 (0.186)	0.053 (0.151)
Dummy: German nationality	0.002 (0.037)	0.003 (0.034)	-0.077 (0.074)	-0.058 (0.068)
Dummy: receiving	0.121***	0.034	0.322***	0.201***

Continued on next page...

... table 1.6 continued

	T.A	T.B	T.C	T.D
	coef. / (SE)	coef. / (SE)	coef. / (SE)	coef. / (SE)
unemployment insurance benefits	(0.026)	(0.025)	(0.097)	(0.073)
Dummy: receiving mean-tested unemployment benefits	-0.163*** (0.031)	-0.088** (0.036)	0.003 (0.144)	0.092 (0.121)
Dummy: West Germany	-0.006 (0.023)	-0.023 (0.023)	0.027 (0.097)	0.024 (0.095)
Dummy: left unemployment in the past		0.510*** (0.044)		0.537*** (0.063)
Unemployment duration (in months)		-0.003*** (0.000)		-0.020*** (0.002)
Dummy: left long-term unemployment (>12 months) in the past		-0.009 (0.032)		0.293*** (0.043)
Dummy: be recalled in the past		0.022 (0.025)		-0.004 (0.039)
Dummy: participation in active labour market programmes in the past 3 years		0.035 (0.022)		-0.000 (0.026)
Constant	0.949*** (0.119)	0.555*** (0.118)		
N	1484	1484	3 × 1484	3 × 1484
R ²	0.036	0.151	0.925	0.936
Percent correctly predicted (PCP)	0.763	0.787		

Robust standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.010

Table 1.6 shows the estimation results for the models $P(y = 1|X)$, $P(y = 1|X, Z)$, $P(y_{it} = 1|X_{it})$ and $P(y_{it} = 1|X_{it}, Z_{it})$, which are denoted as T.A - T.D, respectively. It is apparent that the estimated coefficients on the X variables are often similar and statistically not different across the regressions T.A and T.B, except for the benefit claim related variables, which both decrease in magnitude. The R^2 increases from 0.036 to 0.151, which shows that the work history variables contribute to the model, though, the Percent Correctly Predicted (PCP) after the inclusion of Z only very marginally increases from 0.763 to 0.787. The results for the panel models T.C and T.D in Table 1.6 also only show evidence for a small number of coefficients (X, Z) to be sizably different in comparison to Models T.A and T.B. The coefficient on receiving unemployment insurance benefits increases considerably, while the coefficient on receiving mean-tested unemployment benefits changes sign but loses statistical significance. Among

the Z variables, only the coefficients on unemployment duration and on having left long-term unemployment change. In particular, they increase strongly in magnitude in model T.D. As the latter is the interaction of having been long term unemployed and having left unemployment in the past, these results suggest that past successes play an important role in explaining future successes. While the W_1 variables turned out to be irrelevant, the Z variables appear to be the most important variables in the transition model and they remain relevant after controlling for unobserved fixed effects. Thus, they may be relevant in the population model or may be related to time varying information in W_2 that has not been captured by the panel models.

Table 1.7: Transition sample: Regression based endogeneity test for components of X and Z

	(1) $E(a X)$ $\beta / (SE)$	(2) $E(b X, Z)$ $\beta / (SE)$
Gender (male=1)	-2.844*** (0.018)	-2.020*** (0.018)
Age	0.151*** (0.001)	0.110*** (0.001)
Missing information on education	-0.024 (0.073)	0.024 (0.077)
No formal degree	-0.005 (0.067)	-0.013 (0.070)
Vocational training	0.008 (0.066)	0.014 (0.069)
Higher Education	-0.027 (0.098)	-0.032 (0.097)
Dummy: German nationality	0.054* (0.030)	0.051* (0.027)
Dummy: receiving unemployment insurance benefits	-0.232*** (0.021)	-0.196*** (0.020)
Dummy: receiving mean-tested unemployment benefits	-0.117*** (0.028)	-0.117*** (0.034)
Dummy: West Germany	-0.062*** (0.019)	-0.077*** (0.019)
Dummy: left unemployment in the past		0.040 (0.039)
Unemployment duration (in months)		0.017*** (0.000)
Dummy: left long-term unemployment (>12 months) in the past		-0.274*** (0.026)
Dummy: be recalled in the past		0.031 (0.020)

Continued on next page...

... table 1.7 continued

	(1) $E(a X)$ $\beta / (SE)$	(2) $E(b X, Z)$ $\beta / (SE)$
Dummy: participation in active labour market programmes in the past 3 years		0.041** (0.017)
Constant	0.953*** (0.089)	0.565*** (0.096)
N	1484	1484
R^2	0.976	0.968

Robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 1.7 presents the results for the linear projection of X and X, Z on the estimated fixed effects. Similar to the wage regressions, there is evidence for a number of variables being endogenous in the cross sectional analysis of Models T.A and T.B. However, as already discussed, our results suggest that the omitted variable bias is limited for most variables in the transition model. Table 1.7 also shows that the included regressors almost perfectly explain the variation in estimated fixed effects, which suggests again a multicollinearity pattern for a subset of the regressors in Table 1.6.

Before finishing this section we pay some special focus on the variable "participation in an active labour market policy program" as participation in an active labour market policy program, such as training, is a policy relevant variable that has received a lot of attention in empirical labour market research. We do not find economically, nor statistically relevant changes when comparing Models T.B and T.D. Lechner and Wunsch (2013) and Caliendo et al. (2014), who focus among other things on the estimation of treatment effects on labour market transitions, add more operations based administrative or interview based survey variables to their models to check sensitivity of results. Our findings confirm their findings that the estimated treatment effects are stable and thus may not be affected by omitted variables.

Despite these findings the endogeneity tests in Table 1.7 provide some evidence for this variable being endogenous. In order to tackle endogeneity of labour market treatment variables,

the academic literature typically applies instrumental variable techniques. In the future research, we can construct an additional variable, e.g. the regional treatment intensity, similar to the one used in Frölich and Lechner (2010) and Bookmann et al. (2014), as a candidate for an instrument.

1.5 Conclusion and Discussion

In virtually any empirical regression analysis, there is limited availability of observed variables and limited prior knowledge on which variables belong to the model. This paper provides a unified framework that nests various approaches aiming at reducing omitted variable bias in linear regression analysis. We work out the mechanisms driving the size of the bias and how various models with different regressor sets or unobserved effects relate. Without imposing restrictions on the relationship and role of the variables, it is, however, not possible to derive model rankings that are valid in every application.

In our applications, we find evidence for sizeable omitted variable bias for a number of variables in the wage regression, while only a small number of coefficients is systematically affected in the transition analysis. While the use of work history and survey variables in the transition analysis hardly changes the results, they seem to contribute to a reduction in omitted variable bias in the wage regression as by including more and more variables the coefficients often converge to their values in the most comprehensive panel data model. In particular, key socio-demographic variables appear to move closer to the results of a panel analysis. When exploiting the availability of panel data, we obtain evidence for cross-sectional results being biased due to correlations with unobserved effects. Our results suggest that panel analysis is expected to capture more relevant unobserved model components than an expanded regressor set at one point of time. Beside asymptotic bias considerations, an analysis based on administrative data only should also benefit from a higher precision due to the larger sample size, if for example, survey-based variables were only available for a small subset of the population.

Our results are not only crucial for empirical researchers but also for data providers. Due to cost and data confidentiality constraints, data providers aim at supplying a maximum amount of relevant information but a minimum of irrelevant information. Given our findings, the availability of longitudinal information for key variables appears to add more to the analysis than a greatly but possibly unfocused set of additional (survey) variables at one time point. Further, simulation studies on the case in this paper can be conducted in the future to support our findings.

Bibliography

- Altonji, J. G., Elder, T. E. & Taber, C. R. (2005), 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of political economy* **113**(1), 151–184.
- Antoni, M., Bethmann, A. et al. (2014), 'Pass-befragungsdaten verknüpft mit administrativen daten des iab (pass-adiab) 1975–2011', *FDZ-Datenreport* **3**, 2014.
- Arni, P., Caliendo, M., Künn, S. & Mahlstedt, R. (2014), Predicting the risk of long-term unemployment: What can we learn from personality traits, beliefs and other behavioral variables, Technical report, Working Paper.
- Baptista, R., Lima, F. & Preto, M. T. (2012), 'How former business owners fare in the labor market? job assignment and earnings', *European Economic Review* **56**(2), 263–276.
- Belloni, A., Chernozhukov, V. & Hansen, C. (2014), 'High-dimensional methods and inference on structural and treatment effects', *Journal of Economic Perspectives* **28**(2), 29–50.
- Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, M., Bethmann, A., Fuchs, B., Trappmann, M. & Wurdack, A. (2012), Codebook and documentation of the panel study 'labour market and social security' (pass): Datenreport wave 5, Technical report, FDZ-Datenreport, Nürnberg.

- Biewen, M., Fitzenberger, B., Osikominu, A. & Paul, M. (2014), 'The effectiveness of public-sponsored training revisited: The importance of data and methodological choices', *Journal of Labor Economics* **32**(4), 837–897.
- Bollinger, C. R. (2003), 'Measurement error in human capital and the black-white wage gap', *Review of Economics and Statistics* **85**(3), 578–585.
- Bollinger, C. R. & Minier, J. (2015), 'On the robustness of coefficient estimates to the inclusion of proxy variables', *Journal of Econometric Methods* **4**(1), 101–122.
- Boockmann, B., Thomsen, S. L. & Walter, T. (2014), 'Intensifying the use of benefit sanctions: an effective tool to increase employment?', *IZA Journal of Labor Policy* **3**(1), 21.
- Caliendo, M., Mahlstedt, R. & Mitnik, O. A. (2014), 'Unobservable, but unimportant? the influence of personality traits (and other usually unobserved variables) for the evaluation of labor market policies'.
- Fernández-Kranz, D. & Rodríguez-Planas, N. (2011), 'The part-time pay penalty in a segmented labor market', *Labour Economics* **18**(5), 591–606.
- Friedman, J., Hastie, T. & Tibshirani, R. (2010), 'Regularization paths for generalized linear models via coordinate descent', *Journal of statistical software* **33**(1), 1.
- Frölich, M. & Lechner, M. (2010), 'Exploiting regional treatment intensity for the evaluation of labor market policies', *Journal of the American Statistical Association* **105**(491), 1014–1029.
- Gelbach, J. B. (2016), 'When do covariates matter? and which ones, and how much?', *Journal of Labor Economics* **34**(2), 509–543.
- Heineck, G. & Anger, S. (2010), 'The returns to cognitive abilities and personality traits in germany', *Labour economics* **17**(3), 535–546.

- Kauhanen, A. & Napari, S. (2012), Career and wage dynamics: Evidence from linked employer-employee data, *in* 'Research in labor economics', Emerald Group Publishing Limited, pp. 35–76.
- Lechner, M. & Wunsch, C. (2013), 'Sensitivity of matching-based program evaluations to the availability of control variables', *Labour Economics* **21**, 111–121.
- Lubotsky, D. & Wittenberg, M. (2006), 'Interpretation of regressions with multiple proxies', *The Review of Economics and Statistics* **88**(3), 549–562.
- Mueller, G. & Plug, E. (2006), 'Estimating the effect of personality on male and female earnings', *ILR Review* **60**(1), 3–22.
- Neal, D. A. & Johnson, W. R. (1996), 'The role of premarket factors in black-white wage differences', *Journal of political Economy* **104**(5), 869–895.
- Nyhus, E. K. & Pons, E. (2005), 'The effects of personality on earnings', *Journal of Economic Psychology* **26**(3), 363–384.
- Oster, E. (2017), 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business & Economic Statistics* pp. 1–18.
- Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 267–288.
- Townsend, W. (2017), 'Elasticregress: Stata module to perform elastic net regression, lasso regression, ridge regression'.
- Vom Berge, P., König, M., Seth, S. et al. (2013), 'Sample of integrated labour market biographies (siab) 1975-2010', *FDZ Datenreport* **1**, 2013.
- Wooldridge, J. M. (2010), *Econometric analysis of cross section and panel data*, MIT press.

Appendices

Appendix I: Statistical regularisation and variable selection

The survey data contains a large number of variables (around 40). But it is unknown which of them actually belong to the regression model. In an overfitted model with many variables, estimated coefficients may become implausibly large, while not contributing much to the precision of the model fit. Moreover, the reporting of results is more convenient if irrelevant variables are excluded. We apply the Lasso (least absolute shrinkage and selection operator) as a numerical procedure to eliminate variables that do not or very little contribute to the model. Beside their elimination, the model constraints the sum of the parameters on the regressors, making their interpretation easier. The objective is minimising the usual sum of squared residuals subject to a linear inequality constraint

$$(\hat{\beta}, \hat{\gamma}) \arg \min \sum_{i=1}^N (y_i - X_i\beta - W_i\gamma) \text{ subject to } \sum_j \gamma_j \leq \lambda$$

with λ being the regularisation parameter (Tibshirani, 1996) and W is a regressor set that contains the eventually chosen W_1 but does not include W_2 . The linear inequality constraint leads the Lasso to un-select variables, i.e. $\hat{\gamma}_j = 0$, if their coefficient is small in magnitude. It also leads to the selection of one variable in the case of a group of highly correlated regressors. In order to find $(\hat{\beta}, \hat{\gamma})$ we apply the algorithm suggested by Friedman et al. (2010). λ is determined by cross-validation such that it minimises the mean squared error. We use the STATA package `elasticregress` (Townsend, 2017). As for the transition model the Lasso does not select any of the variables, we have also applied an Elastic Net which combines the Lasso with a Ridge regression. The Elastic Net did not select additional variables for the transition model. Moreover, in this paper, Post-LASSO procedure is applied (Similar to the method used in Belloni and Chernozhukov(2013)). We firstly apply LASSO to determine which variables to drop, and then the selected variables are estimated using restricted model.

Appendix II: A Brief explanation on the method using in Oster (2017)

In this paper, we apply method from Oster (2017) in part of our empirical strategy to test to what extent additional information used to reduce omitted variable bias will contribute to the quality of results. The following presents a brief explanation of Oster's method.

Oster starts with a simple setup, and this is a very similar situation as discussed in our paper. Assume that the model that determines wages is given by equation 1.11:

$$Y = \beta X + W + C \quad (1.11)$$

Where W and C are two orthogonal components of unobserved "ability" and X is education. Assume that both W and C are related to X , but W has a much larger variance than C has. Oster also assumes that if one regress X on W and C , same coefficients will be obtained. However, when one observes W (the high variance control), and the other observes C (the lower variance control), the case will be different. The coefficient will appear stable when C is included. This is not because of the small bias, but just because of the fact that the control is less important in explaining wages.

Then Oster develops the case in model 1.11 to the regression model in equation 1.12, now C in equation 1.11 can be observed:

$$Y = \beta X + \psi \omega^0 + W_2 + \epsilon \quad (1.12)$$

where X is the (scalar) treatment, and ω^0 is a vector of the observed controls, $\omega_1^0, \dots, \omega_J^0$. W_2 is not observed. Oster draws the setup of ϵ from AET who assumes that $\epsilon = 0$ and that W_2 contains some error unrelated to X . Moreover, the proportional selection relationship is defined as $\delta \frac{\sigma_{1X}}{\sigma_1^2} = \frac{\sigma_{2X}}{\sigma_2^2}$, $\sigma_{iX} = \text{cov}(W_i, X)$ and $\sigma_i^2 = \text{var}(W_i)$ for $i \in \{1, 2\}$. δ is the coefficient of proportionality.

Define $W_1 = \psi \omega^0$. All elements of ω^0 are assumed to be orthogonal to W_2 , thus W_1 and W_2 are orthogonal. Let $\hat{\beta}$ be the coefficient from the short regression of Y on X , and \hat{R} be the R-squared from that short regression. Let $\tilde{\beta}$ be the coefficient from the intermediate regression of Y on X and ω^0 , and the R-squared denoted as \tilde{R} .

Then, by performing the following auxiliary regressions, the omitted variable bias of the OLS estimates $\mathring{\beta}$ and $\tilde{\beta}$ are determined:

(1) Regress each value $\omega_1^0, \dots, \omega_J^0$ on X . $\hat{\lambda}_{\omega_i^0|X}$ is the in-sample estimated coefficient on X from these regressions in (1).

(2) Regress W_2 on X . $\hat{\lambda}_{W_2|X}$ is the (unobservable) in-sample estimated coefficient on X from the regressions in (2).

(3) Regress W_2 on X and ω^0 . $\hat{\lambda}_{W_2|X, \omega^0}$ is the coefficient on X from the regressions in (3).

Denote the population analogs of these values $\lambda_{\omega_i^0|X}$, $\lambda_{W_2|X}$, and $\lambda_{W_2|X, \omega^0}$.

Finally, R_{max} is defined. R_{max} is based on the R-squared from the hypothetical regressions of Y on X , ω^0 and W_2 . Oster points out that these R^2 are values that are estimated in-sample based on the available data, while R_{max} is a (theoretical) population value.

Under restrict assumptions, the probability limits of the short and intermediate regression coefficients in terms of the auxiliary regression coefficients can be written in equation 1.13:

$$\beta^* = \tilde{\beta} - \left[\mathring{\beta} - \tilde{\beta} \right] \frac{R_{max} - \tilde{R}}{\tilde{R} - \mathring{R}}, \text{ and } \beta^* \xrightarrow{P} \beta \quad (1.13)$$

It has also been proved under less restricted assumptions, that the probability limits of the short and intermediate regression coefficients in terms of the auxiliary regression coefficients can be written as follows:

$$\tilde{\beta} \xrightarrow{P} \beta + \lambda_{W_2|X, \omega^0}$$

Further, Oster argues that in practical empirical applications, the outcome cannot be fully explained even if the full control set is included, due to measurement error; hence an idea is proposed to construct R_{max} by using the information on the measurement error or expected idiosyncratic variation in the outcome. In our case, R_{max} equals to one if W can be fully explained by the full control set. And in theory, R_{max} should reflect how much of the variation of in W_2 could be explained if we had full controls for W_1 .

Once a dataset with a full set of controls is accessible, one can then explore coefficient bias

when various sets of controls are excluded. Therefore, a test of whether the proportional selection relationship would lead to better inference in the model setting becomes possible.

Appendix III: Tables

Table 1.8: Variable selection by LASSO and elastic net

Variable Names
<i>Wage sample:</i>
Big Five: I am rather cautious, reserved
Big Five: I tend to criticise people
Big Five: I attend to all my assignments with precision
Big Five: I have versatile interests
Big Five: I am inspirable and can inspire other people
Big Five: I easily trust in people and believe in the good in humans
Big Five: I tend to be lazy
Big Five: I am profound and like to think about things
Big Five: I am rather quiet, introverted
Big Five: I can act cold and distant
Big Five: I am industrious and work hard
Big Five: I worry a lot
Big Five: I have a vivid imagination and have a lot of phantasy
Big Five: I am outgoing and like company
Big Five: I can be gruff and repellend towards other people
Big Five: I make plans and carry them out
Big Five: I easily get nervous and insecure
Big Five: I treasure artistic and aesthetic impressions
Big Five: I am not very interested in art
Dummy: satisfied with one?s life in general
Dummy: was looking for a new job
Dummy: was looking for an additional job
Dummy: was looking for a new and an additional job
strength of connection to place of residence

Continued on next page...

... table 1.8 continued

Variable Names

Frequency of misunderstandings, tensions or conflicts

Number of children in total (within and outside the household)

Number of children in household

Dummy: none of parents has a HE degree

Dummy: one parent has a HE degree

Current contract working time, total, without mini-job

Current actual working time, main occupation, without mini-job

Current actual working time, total, without mini-job

Dummy: none of parents with migrational background

Size of household

Not Selected by LASSO:

Big Five: I tend to be depressed, crestfallen

Big Five: I am relaxed and don't let stress get to me

Dummy: satisfied with health

Working even without being dependent on wage

Number of real close friends/family members outside the household

Transition sample:

No variable is selected by LASSO and elastic net

Table 1.9: Descriptive statistics

Variable Names	Wage			Transition		
	mean	sd	min	max	mean	sd
<i>y variables:</i>						
log(average daily gross wage)	4.15	0.66	1.21	5.44	-	-
Dummy: left unemployment within 12 months	-	-	-	-	0.76	0.42
<i>X variables:</i>						
Gender (male=1)	0.47	0.50	0.00	1.00	0.40	0.49
Age	42.87	10.19	18.00	64.00	39.74	11.30
Dummy: trainee	0.00	0.04	0.00	1.00	-	-
Missing information on education	0.00	0.07	0.00	1.00	0.06	0.25
No formal degree	0.12	0.33	0.00	1.00	0.31	0.46
Vocational training	0.75	0.43	0.00	1.00	0.60	0.49
Higher Education	0.11	0.31	0.00	1.00	0.01	0.12
Dummy: German nationality	0.96	0.19	0.00	1.00	0.89	0.32
Dummy: West Germany	-	-	-	-	0.63	0.48
Agriculture	0.01	0.09	0.00	1.00	-	-
Hotel and restaurant	0.03	0.18	0.00	1.00	-	-
Construction	0.05	0.22	0.00	1.00	-	-
Trade	0.14	0.35	0.00	1.00	-	-
Services	0.29	0.45	0.00	1.00	-	-
Education and social health	0.20	0.40	0.00	1.00	-	-

Continued on next page...

... table 1.9 continued

Variable Names	Wage			Transition		
	mean	sd	min	max	mean	sd
Public institutions	0.07	0.25	0.00	1.00	-	-
Other sectors	0.02	0.14	0.00	1.00	-	-
Dummy: receiving unemployment insurance benefits	-	-	-	-	0.67	0.47
Dummy: receiving mean-tested unemployment assistance	-	-	-	-	0.37	0.48
<i>Z variables:</i>						
Tenure (in years)	4.85	5.99	0.00	36.59	-	-
Share of working experience over total observation time	0.53	0.34	0.00	1.00	-	-
Additional working experience (in years)	8.31	8.03	0.00	36.14	-	-
Dummy: unemployment history in the past	0.74	0.44	0.00	1.00	-	-
Dummy: left unemployment in the past	-	-	-	-	0.89	0.31
Unemployment duration (in months)	-	-	-	-	40.19	30.20
Dummy: left long-term unemployment(>12 months) in the past	-	-	-	-	0.22	0.41
Dummy: be recalled in the past	-	-	-	-	0.73	0.44
Dummy: participation in an active labour market policy program in the past 3 years	-	-	-	-	0.39	0.49
<i>W1 variables:</i>						
Big Five: I am rather cautious, reserved	2.82	1.16	1.00	5.00	-	-
Big Five: I tend to criticise people	2.73	1.12	1.00	5.00	-	-
Big Five: I attend to all my assignments with precision	4.41	0.71	1.00	5.00	-	-
Big Five: I have versatile interests	4.22	0.83	1.00	5.00	-	-
Big Five: I am inspiring and can inspire other people	3.80	1.02	1.00	5.00	-	-

Continued on next page...

... table 1.9 continued

Variable Names	Wage			Transition		
	mean	sd	min	max	mean	sd
Big Five: I easily trust in people and believe in the good in humans	3.57	1.12	1.00	5.00	-	-
Big Five: I tend to be lazy	2.21	1.10	1.00	5.00	-	-
Big Five: I am profound and like to think about things	3.70	1.05	1.00	5.00	-	-
Big Five: I am rather quiet, introverted	2.52	1.21	1.00	5.00	-	-
Big Five: I can act cold and distant	3.10	1.24	1.00	5.00	-	-
Big Five: I am industrious and work hard	4.31	0.65	1.00	5.00	-	-
Big Five: I worry a lot	3.23	1.18	1.00	5.00	-	-
Big Five: I have a vivid imagination and have a lot of phantasy	3.81	0.96	1.00	5.00	-	-
Big Five: I am outgoing and like company	3.71	1.01	1.00	5.00	-	-
Big Five: I can be gruff and repellend towards other people	3.01	1.18	1.00	5.00	-	-
Big Five: I make plans and carry them out	3.98	0.85	1.00	5.00	-	-
Big Five: I easily get nervous and insecure	2.43	1.03	1.00	5.00	-	-
Big Five: I treasure artistic and aesthetic impressions	3.34	1.19	1.00	5.00	-	-
Big Five: I am not very interested in art	2.76	1.26	1.00	5.00	-	-
Dummy: satisfied with one's life in general	0.88	0.32	0.00	1.00	-	-
Dummy: was looking for a new job	0.08	0.27	0.00	1.00	-	-
Dummy: was looking for an additional job	0.01	0.12	0.00	1.00	-	-
Dummy: was not looking for a new job	0.91	0.29	0.00	1.00	-	-
Dummy: was looking for a new and an additional job	0.00	0.05	0.00	1.00	-	-
strength of connection to place of residence	1.98	0.96	1.00	5.00	-	-

Continued on next page...

... table 1.9 continued

Variable Names	Wage			Transition		
	mean	sd	min	max	mean	sd
Frequency of misunderstandings, tensions or conflicts	3.55	0.95	1.00	5.00	-	-
Number of children in total (within and outside the household)	1.53	1.13	0.00	7.00	-	-
Number of children in household	1.07	1.02	0.00	7.00	-	-
Dummy: none of parents has a HE degree	0.53	0.50	0.00	1.00	-	-
Dummy: one parent has a HE degree	0.08	0.27	0.00	1.00	-	-
Current contract working time, total, without mini-job	33.83	9.17	0.00	80.00	-	-
Current actual working time, main occupation, without mini-job	37.49	11.37	0.00	80.00	-	-
Current actual working time, total, without mini-job	37.85	11.87	0.00	120.00	-	-
Dummy: none of parents with migrational background	0.06	0.24	0.00	1.00	-	-
Size of household	3.11	1.09	2.00	10.00	-	-

Table 1.10: Appendix: Estimated coefficients on W_1 variables in Models W.C and W.D (continued from Table 1.2)

	$E(y X, W_1)$	$E(y X, Z, W_1)$
	coef. / (SE)	coef. / (SE)
Big Five: I am rather cautious, reserved	-0.017 (0.011)	-0.017* (0.010)
Big Five: I tend to criticise people	0.025** (0.010)	0.026*** (0.009)
Big Five: I attend to all my assignments with precision	-0.019 (0.014)	-0.016 (0.013)
Big Five: I have versatile interests	0.001 (0.013)	0.006 (0.012)
Big Five: I am inspirable and can inspire other people	-0.009 (0.011)	-0.009 (0.011)
Big Five: I easily trust in people and believe in the good in humans	0.002 (0.009)	0.007 (0.008)
Big Five: I tend to be lazy	0.036*** (0.010)	0.036*** (0.009)
Big Five: I am profound and like to think about things	0.018* (0.010)	0.018* (0.009)
Big Five: I am rather quiet, introverted	-0.009 (0.010)	-0.010 (0.009)
Big Five: I can act cold and distant	-0.005 (0.008)	-0.005 (0.008)
Big Five: I am industrious and work hard	-0.008 (0.018)	-0.002 (0.017)

Continued on next page...

... table 1.10 continued

	$E(y X, W1)$	$E(y X, Z, W1)$
	coef. / (SE)	coef. / (SE)
Big Five: I worry a lot	-0.023** (0.009)	-0.018** (0.009)
Big Five: I have a vivid imagination and have a lot of phantasy	0.004 (0.011)	-0.000 (0.010)
Big Five: I am outgoing and like company	-0.028** (0.011)	-0.029*** (0.010)
Big Five: I can be gruff and repellend towards other people	-0.007 (0.010)	-0.000 (0.009)
Big Five: I make plans and carry them out	0.059*** (0.013)	0.050*** (0.012)
Big Five: I easily get nervous and insecure	-0.023** (0.011)	-0.019* (0.011)
Big Five: I treasure artistic and aesthetic impressions	0.005 (0.010)	0.009 (0.009)
Big Five: I am not very interested in art	0.000 (0.010)	-0.004 (0.009)
Dummy: satisfied with one?s life in general	0.164*** (0.031)	0.122*** (0.029)
Dummy: was looking for a new job	-0.180*** (0.035)	-0.128*** (0.033)
Dummy: was looking for an additional job	-0.028 (0.087)	-0.001 (0.079)
Dummy: was looking for a new and an additional job	-0.093 (0.171)	-0.030 (0.179)
strength of connection to place of residence	0.017* (0.008)	0.024*** (0.008)

Continued on next page...

... table 1.10 continued

	$E(y X, W1)$	$E(y X, Z, W1)$
	coef. / (SE)	coef. / (SE)
	(0.010)	(0.009)
Frequency of misunderstandings, tensions or conflicts	-0.024**	-0.023**
	(0.011)	(0.011)
Number of children in total (within and outside the household)	-0.060***	-0.027**
	(0.014)	(0.013)
Number of children in household	0.050***	0.049***
	(0.017)	(0.017)
Dummy: none of parents has a HE degree	0.033	0.013
	(0.020)	(0.019)
Dummy: one parent has a HE degree	0.105***	0.100***
	(0.036)	(0.034)
Current contract working time,total, without mini-job	0.021***	0.020***
	(0.002)	(0.002)
Current actual working time, main occupation, without mini-job	0.022***	0.020***
	(0.003)	(0.003)
Current actual working time,total, without mini-job	-0.011***	-0.010***
	(0.004)	(0.003)
Dummy: none of parents with migrational background	0.089**	0.108***
	(0.040)	(0.039)
Size of household	0.011	-0.015
	(0.013)	(0.013)
N	2435	2435
R^2	0.502	0.570

* p<0.10, ** p<0.05, *** p<0.010

Chapter 2

IMPACT OF IMMIGRATION ON THE WAGES OF NATIVE WORKERS IN DENMARK

1

Author

Shihan Du²

¹Acknowledgement: Great thanks to my PhD supervisors Ralf Wilke and Birthe Larsen for giving many useful comments for this paper. I also thank Jakob Munch, David Jenkins, Fane Groes, Dario Pozzoli and other professors for the suggestions and comments provided.

²Copenhagen Business School, Department of Economics, Porcelaenshaven 16A, 2000 Frederiksberg, Denmark, E-mail: sd.eco@cbs.dk

Abstract: In this paper, I analyse the impact of immigrants with different skill levels on the wages of natives in Denmark. I use administrative data on full population in Denmark for the period 2004 to 2013. I follow and extend the paper by Malchow-Møller et al. (2012) in our analysis. I find empirical evidence that an increase in the share of high-skilled immigrants is associated with a higher wage of native workers, especially for high-wage earners. Low-skilled immigrants are found to have a positive impact on most wage quantiles of natives, while medium-skilled immigrants tend to bring negative wage effect. I further argue that the empirical evidence I have obtained actively support the hypothesised mechanisms in this paper.

Keywords: Immigration; Wage; Natives; Quantile regressions; Wage efficiency

2.1 Introduction

Many developed countries have witnessed an immigration boom in recent years. In the context of immigration into European countries, the recent literature agrees that a large inflow of immigrants has been entering Western countries during recent decades. Attitudes towards the increased immigrant numbers vary both at the national level and among ordinary citizens. Debates on whether to open the door for more immigrants have become more fierce in most European countries.

One of the major concerns raised by most parties is the potential impact of immigration on employment in the local labour market. A vast amount of research projects have been carried out on this topic, and wage, being an important part of labour economics, has been discussed in numerous relevant works. Denmark has witnessed a strong increase in the employment of immigrants since the early twenty-first century, and it is a place where individual-level and employer-employee linked labour market data for the full population are available. This makes it possible to conduct an empirical study on the wage effect of immigrants on natives.

There are a number of studies that have focused on immigration to Denmark, but they have mainly focused on refugees and low-skilled immigrant groups. Less empirical evidence has been presented on high-skilled immigrants, though Denmark is regarded as one of the EU's most active members in trying to recruit skilled workers from around the world. Moreover, the conclusions of those empirical studies have varied. Therefore, I conducted a study based on Danish register data in an attempt to obtain further relevant empirical evidence from different econometric models. Particularly, in this paper, I am aiming to answer: Firstly, whether immigration has any impact on the wages of natives in Denmark, after the EU enlargement in the year 2004. Secondly, how the impact from immigrants within each skill levels differs. Thirdly, how the impact differs across different wage quantiles of native workers.

This paper contributes to gaining a more comprehensive view of such impact on wages of the native workers in Denmark during the period (2004-2013) when immigrants increased rapidly.

Moreover, this study provides empirical evidence within several wage quantiles as well as different skill levels in Denmark. The paper is organised as follows: Section 2.2 describes the background and motivation of the paper. Section 2.3 presents our framework setting based on ordinary least squares (OLS), fixed-effects (FE), instrumental variable (IV), and quantile regression models. Information about the datasets and the variables used in the analysis is given in Section 2.4. Section 2.5 compares the different estimation models, and then Section 2.6, further discusses what the empirical results suggest the impact of immigrants on the wages of natives. Finally, conclusions and a further discussion appear in Section 2.7.

2.2 Background and motivation

Many observers have noted that increased immigration is likely to be part of any strategy to keep European social security systems solvent. At the same time, the rise in immigration has been associated with high levels of anti-foreigner sentiment, and the view that immigrants take jobs from natives is widespread. Attitudes towards immigrants depend on the native group. One major concern of natives is the impact of immigration on their wages. Some are quite open to immigrants since they believe that all of society will gain from the new technologies the immigrants bring into the country. With increased firm productivity, the wages of natives could rise. In contrast, those who are not in favour of more immigration propose that it could lower their wages. More immigrants do mean a larger labour supply, which might, in turn, lowers average wages. Moreover, low-skill immigrants could ask for a lower wage because their countries of origin might have a lower wage level than that of native Danish workers. Hence, firms stand to gain more financially by employing those low-skilled immigrants than by employing native Danes. To compete with the increasing number of low-skilled immigrants, native workers could compromise by requesting a lower wage to avoid losing their jobs.

2.2.1 Background and literature

There are classical theoretical models on the impact of immigration on the wages of native workers. Supply-and-demand theory predicts that the real wage rate for all workers will fall

after the arrival of immigrants. Moreover, based on the bargaining model (McDonald and Solow, 1981), increased "cheap labour" performed by immigrants could lower the bargaining power of trade unions or workers in general and could increase firms' alternative payback, thereby decreasing wages. The model also implies that hiring more immigrants signals that a firm has better alternative options; thus, native wages are additionally affected when this occurs. On the other hand, according to the O-ring theory (Kremer, 1993), wages are often an indirect measure of firm productivity, and an increase in one's productivity increases the productivity thus the wages of his or her co-workers. As high-skilled immigrants enter the receiving country, their native co-workers' wages could rise through this channel.

Later, as richer datasets became available for research, more empirical discussions began to take place on the impact of immigrants on native workers in the local labour market in the receiving country. Angrist and Kugler (2003) have reported small, mostly negative immigration effects for European countries. They also apply an IV strategy based on the data of immigrants from former Yugoslavia and find larger, though mostly insignificant, negative estimates. The estimates typically demonstrate that a 10% increase in the foreign share would reduce native employment rates by 0.2-0.7 of a percentage point. Ordinary least squares estimates are at the low end of this scale, while IV estimates using the Balkan Wars are mostly larger than the corresponding OLS estimates, implying the substantial displacement of native workers by immigrants.

I narrow down the field of impact to the wages of native workers to focus on our primary research question: Do immigration affect the wages of native workers? Wages are an important topic in labour economics and immigration studies. There is a rich empirical literature stream, and more specific discussions have studied the impact of immigrants on the wages of native workers. Dustmann et al. (2005) briefly reviewed the simplest theoretical model that helps to explain the effects of immigration on the economic outcomes of native workers in the receiving country. They concluded that the same immigrant inflow might affect different recipient countries in divergent ways, depending on the skill structure of the native workforce. Studying a

range of empirical evidence on the UK and other countries allowed them also to point out the difficulties an analyst faces when attempting to assess the effects of immigration on wages empirically. To date, at least in the empirical literature, a universal agreement has not been reached regarding the impact of immigration on the wages of native workers.

Indeed, many recent empirical papers adopting different identification strategies have failed to find a negative impact of immigration on natives' wages. Manacorda et al. (2012) used a pooled time series of British cross-sectional micro-data on male wages and employment from the mid-1970s to the mid-2000s. They illustrated that immigration has primarily reduced the wages of immigrants, but with little discernible effect on the wages of the native-born. Using Spanish data, Carrasco et al. (2008) explored the effects of immigration on employment and the wages of native workers using three samples varying in coverage of legal and illegal immigration and in the dimensions along which labour market segments can be constructed. Ordinary least squares and IV approaches were adopted, but overall, they did not find any significant negative effect of immigration on either the employment rate or the wages of native workers.

As regards those empirical papers with statistically significant results, neither the U.S. nor European literature has reached a clear consensus. Some studies have suggested a positive impact of immigrants on the wages of native workers, while others have indicated a negative effect. Since the situations in the U.S. and Europe are not comparable due to socio-backgrounds and other factors, I first review empirical evidence from the U.S.

Various studies have reported a positive effect using U.S. data, and more of a focus on high-skilled immigrants has emerged in recent decades. Ottaviano and Peri (2006) have demonstrated that from 1990 to 2004, immigration increased the average wages of native U.S. workers. This positive effect arose from a relatively large positive effect on the wages of more highly educated native workers and a slight negative effect on the wages of native high-school drop-outs. Ottaviano and Peri (2012) reported similar findings. They calculated the effects of immigration on the wages of native U.S. workers of various skill levels, and they

found that from 1990 to 2006, immigration had a small effect on the wages of native workers with no high school degree (between +0.6% and +1.7%). It also had a small positive effect on average native wages (+0.6%) and a substantial negative effect (-6.7%) on the wages of previous immigrants in the long run.

At the same time, empirical evidence also supports the hypothesis that the presence of immigrants lowers the wages of native workers. Orrenius and Zavidny (2007) were the first to use U.S. INS data on new recipients of legal permanent resident status to examine the effects of immigration on wages. They used occupation as a proxy for skill and found that an increase in the fraction of foreign-born workers tended to lower the wages of natives in blue-collar occupations, but did not have a statistically significant negative effect among natives in skilled occupations. Using 1990 census data, Card (2001) likewise reported that immigration inflows over the 1980s reduced the wages of low-skilled natives in traditional gateway cities in the U.S.

Scant empirical evidence has been found in Europe supporting the hypothesis that immigrants have a positive effect on the wages of native workers. Malchow-Møller et al. (2011) used a dataset covering the entire Danish population from 1995 to 2007, and they reported positive effects of foreign experts on the wages of local high-skilled workers in Denmark, but the effects on low-skilled natives were not significant.

Some empirical evidence gained based on European data indicates that immigrants, particularly low-skilled immigrants, have a negative impact on the wages of native workers. For example, Ortega and Verdugo (2016) exploited a large French panel for 1976-2007 and applied an FE model to examine the impact of low-educated immigration on the labour market outcomes of blue-collar natives initially in jobs where immigrants had become overrepresented in recent decades. Low-educated immigration generally lowers the wages of blue-collar workers, but its impact is heterogeneous across sectors. More evidence on the overall impact of immigration on natives' wages in France for the period from 1990 to 2010 is quantified in Edo and Torbal (2015). Their short-run simulations indicated that immigration had reduced native

wages by 0.6%.

Returning to evidence from Denmark, taking advantage of the rich Danish dataset, Malchow-Møller et al. (2012) tested the hypothesis that the employment of immigrants affects firm-specific wages by using linked employer-employee data for 1993-2004. Denmark has experienced a particularly pronounced increase in the employment of immigrants, and the researchers found that the increased use of low-skilled immigrant workers had a significantly negative effect on the wages of native workers at the same workplace. However, a more recent paper by Foged and Peri (2016) provided a different conclusion: Immigration has had positive effects on native unskilled wages, employment, and occupational mobility. They used longitudinal data on the universe of workers in Denmark from 1991 to 2008, and they tracked the labour market outcomes of low-skilled natives in response to an exogenous inflow of low-skilled immigrants. Hence, even studies using the same dataset can come to different conclusions.

Therefore, in order to analyse the impact of immigration on the wages of native workers in Denmark, I start by comparing our empirical results with the contradictory findings of Malchow-Møller et al. (2012) and Foged and Peri (2016), based on the individual-level dataset from Denmark. Malchow-Møller et al. (2012) use FE and IV models to study the impact from immigrants with different skill levels, while Foged and Peri (2016) focus on Difference-in-difference, and they are more interested in the refugee population. The period covering for the studies are also slightly different—Malchow-Møller et al. (2012) only covers the period until the year 2004, and Foged and Peri (2016) studies on a longer duration (covers the period 2004-2007, when EU enlargement took place). This is an indication that the event of EU expansion could have played a role in the impact of immigration. Hence, I am then motivated to look into the period after the year 2004. The following paragraphs introduce two major changes: the eastern enlargement of the EU, which took place in 2004, and the increasingly important role of high-skilled immigrants in immigration studies.

2.2.2 Recent changes in the Danish immigration environment

The recent changes in the Danish immigration environment, which motivate this study, are briefly summarised in the following two aspects:

Immigrant flows to Denmark after 2004 The eastern enlargement of the EU took place in 2004, and from 2004 to 2007, the labour inflow to Denmark from Eastern European countries tripled, rising from 10,000 individuals in 2003 to nearly 30,000 in 2007. Restrictions were eased in 2007 and phased out in 2009. There are three different ways that people from new EU member states can come to Denmark to work: they can be employed by Danish companies, they can establish their own companies in Denmark, or they can be stationed in Denmark by employers based overseas. Among the Eastern European immigrants living and working in Denmark, Poles are clearly the dominant group. They comprised nearly 60% of the Eastern European employees working in Denmark in 2006.

Highly skilled immigrants Highly educated immigrants or high-skilled workers are among the populations that are increasingly welcomed by most countries. One famous case is the H1B visa scheme in the U.S. The H1B visa policy was introduced in 1990. It is a non-immigrant U.S. visa that allows U.S. employers to hire foreign workers in speciality occupations temporarily. It was the main channel of entry for foreign science, technology, engineering, and mathematics (STEM) workers in the U.S. In fact, 70%-80% of H1B visa holders worked in STEM jobs.

The U.S. states that saw a large inflow of highly educated foreign-born workers experienced faster growth in patenting per person (Hunt and Gauthier-Loiselle, 2010). Peri et al. (2015) estimated the equilibrium elasticity of aggregate employment and the wages of different groups to STEM across cities and years. They found that an exogenous increase in STEM workers totalling 1% of total employment increased the wages of college-educated workers and non-college educated workers by 7%-8% and 3%-4%, respectively. However, the effects on employment were not significant for the 1990-2010 period. Various studies have discussed

whether the effects of highly educated immigrants on innovation translate into productivity and wages. The early focus was on the key role of technological growth in productivity growth (Jones, 2002). Later, some researchers started examining local human capital spillovers and local multiplier effects of high-tech (STEM) workers (Moretti, 2004, 2012 and Glaeser, 2011).

In the context of Europe, Markusen and Trofimenko (2009) considered the special role of foreign experts in the productivity of domestic firms at the micro level. Foreign experts can increase productivity by teaching local workers 'new tricks', and this phenomenon is more relevant in developing countries. The researchers used the employment of native experts as a proxy for an exogenous increase in productivity.

Europe started to focus on visa policies for high-skilled workers much later than the U.S. Denmark is one of the EU's most active members in trying to recruit skilled workers from around the world. Since 2002, Denmark has developed a selective system for labour immigration designed to ensure that migrants with desirable and in-demand competencies can gain access and residence. A new tax system for foreign experts was introduced in 2002. Under the new scheme, foreign experts can stay as long as they wish without paying additional taxes after the first three years, which had been required before 2002.

Highly skilled and educated migrants are in great demand, and Denmark is in competition with other immigration countries. Immigration policy changes were introduced between 2006 and 2007, and these aimed at attracting high-skilled migrants (OECD, 2008). Denmark introduced a new point-based green card scheme in 2008 for attracting skilled workers; points may be accumulated based on earnings, qualifications, and a shortage list. Under this scheme, foreign workers with or without a concrete job offer can come to Denmark to seek work. They are granted a residence permit solely on the basis of their qualifications. The majority of green card holders originate from Pakistan, India, China, Iran, and Bangladesh. The green card scheme was abolished in June 2016, but another programme, called 'the establishment card',

has a similar function.

2.2.3 Motivation of the paper

I explore the situation discussed above further via Figure 2.1, the left compares the wage distributions of the total Danish population in 2004 and 2014, and the right separately displays the log wages of immigrants and natives from 2004 to 2013. Both wages and log wages were deflated according to the price level in the corresponding year. As both figures illustrate, a slight increase occurred in the mean wages of both working populations. Moreover, the distribution curve is steeper for 2004 than for 2014, as seen in the left of Figure 2.1; this could imply a change in the wage distribution. The right of Figure 2.1 indicates a gap between the wages of natives and immigrants during the ten years, and both groups experienced an increase in wages during the period.

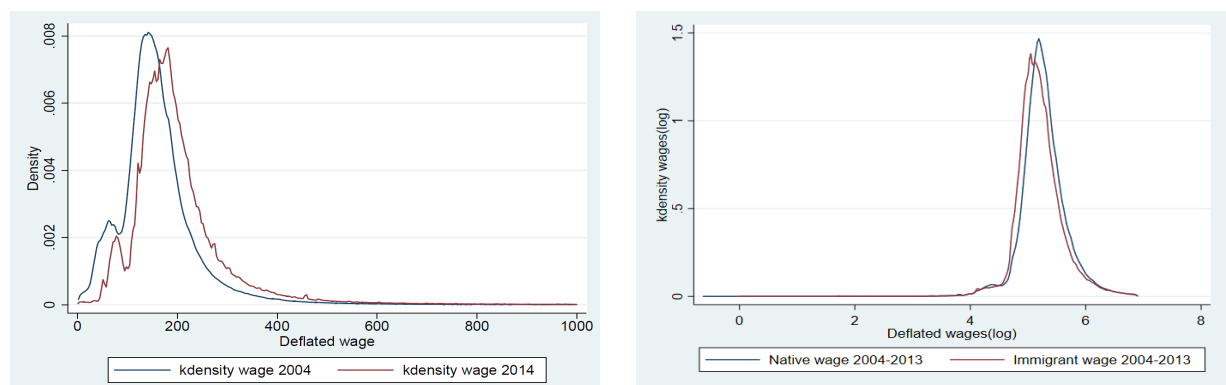


Figure 2.1: Comparisons on wage distributions

The large increase in immigration from Eastern Europe might have had an impact on the skill structure of the local labour market in Denmark, while the high-skilled immigrant group might be playing a more important role than in earlier years. Based on the empirical evidence pre-

sented in most of the relevant literature from the U.S., which has already been discussed, high-skilled immigrants could have a positive effect on the wages of native high-skilled workers, and, in certain cases, the effect could be large enough to mitigate the negative impact of low-skilled immigrants on native workers. This could be one explanation underlying the evidence supporting the positive impact of immigrants on the wages of native workers. Therefore, it is necessary to analyse the impact within and across different skill groups, and the contribution of high-skilled immigrants cannot be ignored.

Earlier, Dustmann et al. (2013) used models similar to those of Malchow-Møller et al. (2012) (IV and OLS). Dustmann et al. (2013) employed UK Labour Force Survey data for 1997-2005, and they calculated results for the different percentiles of the wage distribution. They found that immigration depresses wages below the 20th percentile of the wage distribution, but leads to slight wage increases in the upper part of the wage distribution. Dustmann et al. (2013) have not focused on the impact of different immigrant skill groups, which Malchow-Møller et al. (2012) did consider. Malchow-Møller et al. (2012) did not, however, examine different quantiles of the wage distribution to analyse the impact of immigrants on the wages of native workers.

There has not been a study that combined analysis of both different quantiles of the wage distribution and different skill groups. Therefore, motivated by this research gap as well as the two above-mentioned major changes: (1) the large inflow of the immigrants from the European Economic Area to the local labour market in Denmark, and (2) the increasingly attractive policies applicable to highly skilled immigrants in Denmark. It is worthwhile to pose the following question:

1. Whether immigration has any impact on the wages of natives in Denmark, after major changes that took place from 2004 to 2013.
2. How the impact from immigration within each skill levels differs. After the enlargement of

the EU, have low-skilled immigrants still had a significant negative impact on the wages of native workers, as supported by the previous empirical literature? Moreover, as high-skilled immigrants play increasingly important roles in local labour markets, do they have a significant impact on the wages of native workers in the workplace in Denmark?

3. How the impact of immigration differs across different wage quantiles of native workers.

This paper applies different econometric approaches to explore the above research questions.

Therefore, in this study, I further explored the full wage distribution, and, to obtain more comprehensive estimations for different percentiles of wages, I used IV models and the conditional quantile regression approach instead of OLS, to further investigate the case in Denmark. I examined the impact of immigration on the wages of native workers from different skill groups for various quantiles of the wage distribution.

Furthermore, I decided to use the newest dataset based on the above two major changes in the Danish immigration market, which took place after the period Malchow-Møller et al. (2012) used. In other words, the analysis in this paper was based on data covering the ten years following the sample period of Malchow-Møller et al. (2012). One purpose to determine whether the immigration has an impact on the wages of native workers after EU enlargement. At the same time, I can further verify the findings by Malchow-Møller et al. (2012) and Foged and Peri (2016).

2.2.4 Hypothesis

Based on the classical theoretical models of wage efficiency theory and demand-supply model which is discussed in Section 2.2, I propose the following hypothesis for the impact of immigration within different skill levels on wages of natives:

Hypothesis 1(H1): Positive wage effect might be brought by low-skilled immigrants, due to the occupation upgrading behaviour of natives. Wage efficiency theory also support the possible positive wage effect by low-skilled immigrants.

Hypothesis 2(H2): The increased natives and immigrants could squeeze in the medium-skilled class, causing more fierce competition in this market as the share of medium-skilled immigrants increases, and worsening the wages of natives.

Hypothesis 3(H3): Wage efficiency theory together with the O-Ring theory predict a positive impact of high-skilled immigrants on the wages of natives.

The three hypothesis are discussed in detail as follows:

Low-skilled immigrants Task complexity for lower-skilled jobs usually is low; thus, low-skilled workers often perform routine or manual work requiring fewer communication skills and other specific abilities. These positions are the first choice for immigrants with low educational levels, especially during their first year in the local labour market. With less bargaining power, those low-skilled immigrants acquire a job by lowering the starting wage, which could potentially reduce wages for natives with the same skill level.

However, if taking the occupation upgrading behaviour of natives into consideration, low-skilled immigrants will not reduce the wages of natives. Occupation upgrading on the part of native workers can occur as the local labour market seeks to protect labour from immigration (D'Amuri and Peri, 2014). In particular, natives can upgrade their skills to avoid competing with low-skilled immigrants(Peri and Sparber, 2009), and this phenomenon is also supported by empirical evidence based on European Labour Force survey data.

Besides, according to wage efficiency theory (Akerlof, 1982), the wage could exceed the market-clearing level in industries where the cost of replacing labour is high if managers want to increase a firm's productivity or efficiency or to decrease costs related to labour turnover. In Denmark, labour turnover costs are high. Managers can increase the wage to motivate the natives to stay for this reason. The comparison on the average experience of the low-skilled workers within different types of firms shown in Table 2.1 confirmed the wage efficiency assumption. As I can see from the figure, within firms of both natives and immigrants, the average experience of the worker in the low-skilled group is longer than it is in the firms with only natives. In other words, with the share of immigrants within the firm increases, the average experience of the workers increases. One possible explanation is that more native workers have received the incentives to stay longer, and longer experience can also be associated with higher pay. Therefore, immigrants will bring a positive impact on the wages of native workers for the above reasons.

Table 2.1: Average experience in the firm of natives, mixed of natives and immigrants, and immigrants

Experience (years)	Natives only	Natives and Immigrants	Immigrants only
All	9.93	9.80	6.98
Low-skilled	9.16	9.32	6.58
Medium-skilled	10.3	10.23	7.98
High-skilled	9.36	9.09	6.84

Medium-skilled immigrants Regarding the medium-skilled group, a negative impact of immigrants can be placed on native workers because of the potential occupational upgrading behaviour of the native workers. The increased native and immigrant workers could squeeze in the medium-skilled class, causing more fierce competition in this market as the share of medium-skilled immigrants increases, and worsening the wages of natives. Since the supply of medium-skilled workers would have been more than actual demand, more immigrants

would have meant more competition for positions. The wages of natives, especially in the low-skilled and medium-skilled markets, could plausibly have been worsened by the greater number of medium-skilled immigrants holding less bargaining power and asking for lower wages for the same job position.

High-skilled immigrants According to the O-Ring theory (Kremer, 1993), high-skilled immigrants can increase their native co-workers' productivity by teaching new skills; they thus increase a firm's overall productivity. At the same time, firm productivity is often an indirect measure of wages. Therefore, immigrants can increase the wages of natives by boosting a firm's total productivity.

Moreover, wage efficiency theory also supports our hypothesis for the high-skilled group. Another incentive for managers to pay more is to hire more productive workers; thus, a selection effect cannot be ignored, particularly among high-skilled workers. High-skilled workers only select high-wage positions because their reservation wage is too high for low-wage firms, but high-wage positions attract applicants with various skill levels. Moreover, high-wage firms prefer to recruit more productive workers to achieve an efficiency wage (i.e., higher pay results typically in a higher return; e.g., Malcolmson, 1981). With more high-skilled immigrants arriving in Denmark, managers now have more applicants in the selection 'pool'. Higher wages attract more applicants, especially highly qualified ones (Schlicht, 2005). The managers of a high-wage firm tend to tighten hiring standards and to provide higher wages to attract more productive workers to enlarge the potential applicant pool; in this way, they recruit more competitive employees to enhance the firm's productivity. Again, wages for all high-skilled and high-wage earners are pushed upwards by this selection effect, and the increased number of high-skilled immigrants facilitates this outcome.

2.3 Econometric frameworks and estimation strategies

I tested the hypothesis that the employment of immigrants affects firm-specific wages by using linked employer-employee data from 2004 to 2013 from a developed country (Denmark) which has experienced a particularly pronounced increase in the employment of immigrants. I started by implementing the test strategy adopted by Malchow-Møller et al. (2012), who conducted similar research. In this paper, the increased use of low-skilled immigrant workers has a significantly negative effect on the wages of native workers in the workplace, when controlling for the potential endogeneity of the immigrant share using both FEs and instrumental variables. Next, I extended their empirical approach by applying FE quantile regressions.

2.3.1 OLS and FE

To examine whether the increased use of low-skilled immigrant workers following the eastern enlargement of the EU had a significant negative effect on the wages of native workers, I employed the identification approach for the OLS and the FE models used by Malchow-Møller et al. (2012). I compared the effects within and across skill groups and especially focused on the high-skilled immigrant group. Next, I compared the results from the OLS estimation and FE models.

In this paper, I assume the wage equation of native workers is as follows:

$$w_{ijt} = \beta_0 + \beta_1 S_{jt} + \beta_2 C_{jt} + e_{ijt}, \quad (2.1)$$

Where, w_{ijt} is the (log) wage of individual i in workplace j in the period t . S_{jt} is the variables of interested in this paper, it is the share of immigrants in the employment at workplace j at time t , and C_{jt} is the control variable, it is a set of a time-variant observable individual and workplace characteristics that can affect individual i at the workplace j at time t . In the empirical test, I split the share of immigrant S_{jt} into three categories: low-skilled, medium-skilled, and high-skilled. Hence, I could separately estimate the effects for each group of

immigrants.

The basic OLS estimation of equation 2.1 could be biased due to shocks in a specific year, characteristics of particular industries, or regional differences. Therefore, I firstly included $\text{year} \times \text{region}$ to exclude the influence of year, industry, and region. The results of the estimation including these $\text{year} \times \text{region} \times \text{industries}$ FEs are reported as the OLS model estimation in the empirical analysis.

Then, I introduced a set of FEs for each combination of workplace and worker, called job-spell FEs, to control for shocks from the unobserved workplace and worker characteristics. Hence, the OLS model incorporated the job-spell FEs, and a classic AKM model, presented in equation 2.2, for wage analysis was applied for our empirical test as an FE model:

$$w_{ijt} = \beta_0 + \beta_1 S_{jt} + \beta_2 C_{jt} + \theta_{ij} + \epsilon_{ijt}. \quad (2.2)$$

Where θ_{ij} is the job-spell effect, and other regressors are the same as in the OLS models. Nevertheless, there could still be potential upward bias for β_2 in the estimates due to positive demand shocks during a specified period. When firms pay higher wages because of an increased need for workers, more immigrants can be attracted and hired thanks to their potentially lower bargaining power and higher flexibility. Another upward bias could be caused by the different attitudes of natives towards immigrants. Natives' responses, as mentioned by Malchow-Møller et al. (2012), can affect β_2 in some cases. For example, if natives are forced to leave the workplace because of the increased number of immigrants at a firm, the average wages at the firm may decrease due to the reduced bargaining power of immigrants and the increased share of immigrants in the firm. This type of endogeneity cannot be mitigated by job-spell FEs, nor could I assess endogeneity with the datasets at hand.

2.3.2 The IV approach

Various studies have proposed using an IV approach to overcome the endogeneity issue mentioned above. However, most empirical evidence obtaining using such approaches in immigration studies is only significant within the low-skilled group (e.g. Foged and Peri, 2016 and Malchow-Møller et al. 2012), and certain researchers have even reported insignificant results for the IV approach. I applied an approach widely used in similar studies to include the increase in the share of immigrants at the regional level to construct the instrument for the share of immigrants in the firm.

To be more specific, for the low-skilled immigrant group, I instrumented the share of low-skilled immigrants in the workplace j in the year t S_{jt} as z . I chose the shares of low-skilled immigrants in the workplace j in 1999 (five years prior to the first year of our sample) for four groups of countries of origin for immigrants: (1) the EU-15 countries, plus Norway and Iceland; (2) the EU-12 countries (those that joined the EU in 2004 and 2007); (3) all remaining developed countries; and (4) developing countries and other observations with unknown origins. Then, I multiplied them by regional increases between 1999 and year t in the employment of low-skilled immigrants from each group of origin. In this way, I hence obtained the "new" share of low-skilled immigrants in the workplace in the year t for each particular ethnic group. Finally, I added the "new" shares of low-skilled immigrants from each group of origin together to form the instrument for our IV (2SLS) regression model. I constructed the instruments for the other two skill groups in the same way. Instrument z is described as follows:

$$z = \sum_{g=1}^4 S_{jg,1999}(\bar{S}_{gt}/\bar{S}_{g,1999}) \quad (2.3)$$

Here, g represents the origin group. The shares of low-skilled immigrants from origin group g within overall regional employment in the year t and 1999 are expressed as the terms \bar{S}_{gt} and $\bar{S}_{g,1999}$, respectively.

According to the classical assumptions for the IV approach outlined by the Wooldridge text-

book, in order for the variable, z to serve as a valid instrument for S_{jt} , the following requirements must be satisfied: Firstly, the instrument must be exogenous—in other words, $Cov(z, \epsilon) = 0$. The regional development in immigrant employment can be assumed to be exogenous at the workplace level; thus the first condition is normally satisfied. Secondly, the instrument must be correlated with the endogenous explanatory variable S_{jt} . That is, $Cov(z, S_{jt}) \neq 0$. Networking effects among immigrants have been proven to have effects on immigrants' location choices during the job-search process (Hellerstein et al., 2008). In other words, immigrants are more likely to be employed in workplaces that also employ immigrants from similar ethnic groups. Therefore, the regional increase should theoretically be correlated with an increased share of immigrants within the workplace. I verified this in the first stage of the IV regression analysis. Similar instruments have been used for other empirical studies, such as Card (2001), Ottaviano and Peri (2012) and Peri et al. (2015), and they have produced relatively effective results.

Later, I combined the IV and FE models in our empirical analysis. In short, I included $\text{year} \times \text{region} \times \text{industry}$ FEs and job-spell FEs in the IV (2SLS) models. These two FEs are non-nested. In order to handle multi-way FEs and high-dimensional models in empirical applications, I choose the STATA package `reghdfe` which implements the estimator described in Correia (2016) to perform regression on FE and FE-IV models in this paper.

2.3.3 Quantile regression

All methods used by Malchow-Møller et al. (2012) are linear regression models that only based on the regression on the mean of wages. OLS, FE and IV approaches have been applied in their paper, and earlier in this section, I also presented how the models are fitted into our application. Then, we implemented the Malchow-Møller et al. (2012) study by applying quantile regression to the overall immigrant sample (different skill groups) to obtain more comprehensive estimations for different wage quantiles.

The traditional linear regression model describes the process by which the dependent variable is affected by the independent variable. OLS is one of the widely used linear regression models for estimating the coefficient. If the random error term of the model is from a distribution with a zero mean and is homogeneous, then the least squares estimation is unbiased and effective. Practically, the assumption of homogeneity is usually not met. For example, when there is severe heteroscedasticity in the data, or in the case of fat tails and leptokurtosis, the estimation of OLS will no longer be unbiased. To compensate for these shortcomings of the OLS regression analysis, Laplace proposed a median regression (minimum absolute deviation estimate), then on this basis, Koenker and Baset (1978) proposed a quantile regression approach. The quantile regression is used to estimate a model by minimising the asymmetrically weighted sum of absolute errors, instead of the sum of the squared residuals, as in OLS (Koenker and Hallock, 2001).

Regarding the potential heterogeneity issue in my empirical application, different groups at each wage level could consist of different combinations of workers with diverse nationalities and skill levels, and an estimation bias could result from this variety. Such heterogeneity is allowed in this paper, though this will to some extent lead to biased estimation results in the least squares models (e.g. OLS, FE and IV). I will not place too much emphasis on this issue in the framework since it is a common problem from the data. While the quantile regression does not restrict the distribution of the error term, it takes the heteroscedasticity of the error term into consideration. For this reason, approach Malchow-Møller et al. (2012) study can be implemented by adapting quantile regressions to reduce corresponding potential heterogeneity issue arising from OLS, FE and IV models.

Compared to the least squares models, such as FE and IV, the quantile regression model not only has more relaxed restrictions, e.g. on the assumption of distributions, but also can obtain more abundant information. The independent variable is estimated according to the conditional quantile of the dependent variable, the regression models under all quantiles are

then obtained. Therefore, it can provide a more comprehensive analysis of the relationship between the variables, and it reflects the different influences of the various quantiles of the independent variables on the dependent variable. The model includes the extreme values, rather than just describing its conditional expectation (the mean value), as is in OLS, FE and IV models. In other words, the quantile regression approach provides more robust estimations to the outliers in the response measurements. The quantile regression approach provides an opportunity to reveal the whole picture of the distribution of the wages in my empirical applications. Therefore, we adopt the quantile regression method to study separately considered the influence of each on the wages of native workers.

In the econometric setting, I supposed the τ^{th} ($0 < \tau < 1$) conditional quantile function of wage (log) is as follows:

$$Q_\tau(W|X) = X\beta_\tau \quad (2.4)$$

Given the distribution function of W : $W = X\beta_\tau + \mu_\tau$, where $X = S\beta_{1\tau} + C\beta_{2\tau}$.

β_τ can be obtained by solving:

$$\beta_\tau = \arg \min_{\beta \in R^k} E(\rho_\tau(W - X\beta)) \quad (2.5)$$

$$\hat{\beta}_\tau = \arg \min_{\beta \in R^k} \sum_{i=1}^n (\rho_\tau(W_i - X\beta)) \quad (2.6)$$

Where ρ_τ is the checkpoint function, defined as: $\rho_\tau = \tau\gamma$ if $\gamma \geq 0$, or $\rho_\tau(\gamma) = (\tau - 1)\gamma$ if $\gamma < 0$. Further, β^τ is the vector of parameters. X is the vector of exogenous variables; it includes the variables of interest for this paper S and the other control variables C . $Q(W|X)$ denotes the τ^{th} conditional quantile of W , given X . In our empirical setting of the model, W is the logarithm of the hourly wage of native workers. The set of the share of immigrant within the firm is included in S . The set of exogenous variables in C includes exogenous individual characteristics such as age group, gender, marital status, education, experience, and tenure; these are defined as in the OLS and FE models. Using the estimation results

for each τ th quantile of wages, I will be able to analyse the different impacts of immigrants on natives with the same wage levels. Similar to the OLS model, the interaction terms of $\text{year} \times \text{region} \times \text{industries}$ are included in the quantile regression model, and this makes the estimation results comparable. In terms of individual fixed effects, I exclude them from this quantile model at this stage, to avoid the incidental parameters problem (Neyman and Scott, 1948) arising from the individual-specific (fixed) effect panel quantile regression.

2.4 Data

Our empirical analysis was based on administrative data from Statistic Denmark (DST). As the data were register based, measurement problems were negligible. I used individual data as well as linked employer-employee data from the Integrated Database for Labour Market Research (IDA) to perform analyses using the OLS, FE, IV 2SLS, and quantile regression models. Thanks to the rich data from DST, our sample included 31,685,350 observations (50.36% males and 49.64% females) for all individuals of working age for the period from 2004 to 2013.

One previous relevant study (Malchow-Møller et al., 2012) only covered the period from 1993 to 2004, which was before the immigration boom from Eastern Europe, which took place from 2004 onwards. In our analysis, I used a more recent dataset to study the period from 2004 to 2013—the period after the labour inflow from Eastern European countries—to assess whether the negative effect on the wages of native workers persisted after 2004.

I used observations on all individuals, both natives and immigrants, in the Danish labour market to construct very detailed measures of the employment of immigrants at the workplace (establishment) level. By examining the full population, I was able to compare the results with previous empirical evidence from similar studies on low-skilled immigrants, and I also analysed the impact of high-skill immigrants on the local labour market.

I considered as workers all individuals of working age, defined as 18-65 years, residing in Denmark. Any wage model based on worker and firm FEs is identified by worker moves between firms. I applied OLS, FE, IV, and quantile regression models to estimate the impact of immigrants on natives wages W ; Hence, I used the linked employer-employee data (IDA) for the full population; this was particularly essential for small firms and cases in Denmark. Without the full population, I could have easily run out of observations after several controls on particular variables.

The (log) hourly wage rate was the dependent variable for all four models. I obtained this value from the DST registers IDAN (2004-2010) and BFL (2011-2013). Wages for all years were deflated by the average CPI of the year. Only full-time workers working in the private sector in Denmark were included. After implementing the above controls, I had 9,599,171 observations for our estimation sample, which was sufficient for our empirical analysis.

Data on workers' age, gender, marital status, labour market experience, tenure, education, and family, and other characteristics³ were included in the wage models. I obtained most of this information from the DST registers BEF and IDAN. Education is a highly important identifier for classifying an employee's skill level. I split education into three levels—basic education (low skilled), vocational education (medium skilled), and higher education (high skilled)—according to International Standard Classification of Education codes. Table 2.2 presents descriptive statistics for the individual characteristics for all population members, including both natives and immigrants. The table suggests that the majority of workers had a vocational education level (i.e., medium skilled), and only 4.7% workers were high-skilled.

Moreover, I used data from the BEF register to construct the immigrant group and native group and to further examine the impact of immigrants on native workers. In this study, I defined immigrants as individuals born outside Denmark with non-Danish parents. Moreover,

³Full list of variables used is provided in Section 2.5, and they are also explained in the appendix.

Table 2.2: Summary statistics: Individual characteristics

Variable	Mean	Std. Dev.	Min	Max
Age 18 to 24	0.0697		0	1
Age 25 to 29	0.0955		0	1
Age 30 to 39	0.2720		0	1
Age 40 to 49	0.2811		0	1
Age 50 to 59	0.2228		0	1
Age 60 to 65	0.0498		0	1
Female	0.4470		0	1
Married	0.5501		0	1
Children 0-6 years	0.2234		0	1
Experience	9.7476	7.5652	0	36
Tenure	4.9812	6.3245	0	21
Union member	0.6944		0	1
Basic education	0.2881		0	1
Vocational education	0.6499		0	1
Higher education	0.0468		0	1
Copenhagen	0.1124		0	1
Large city	0.4729		0	1
Small city	0.4189		0	1
Number of observations	9,599,171			

Note: Classification of size of cities is based on the population of the city, and the term 'Copenhagen' is referred to the Copenhagen city instead of Greater Copenhagen Area.

Table 2.3: Immigrant characteristics at individual level, selected years

	Mean			
	1999	2004	2008	2013
Age 18-24	0.1119	0.0669	0.0604	0.0571
Age 25-29	0.1349	0.1241	0.1227	0.1495
Age 30-39	0.3008	0.3658	0.3259	0.3338
Age 40-49	0.2386	0.2686	0.2914	0.2678
Age 50-59	0.1861	0.1469	0.1673	0.1520
Age 60-65	0.0188	0.0238	0.0298	0.0335
Basic education(low-skilled)	0.6041	0.5778	0.5422	0.5138
Vocational education(media-skilled)	0.3475	0.3724	0.3378	0.3456
Higher education(high-skilled)	0.0325	0.0385	0.0809	0.1256
<i>Country of origin:</i>				
EU-15,Norway and Iceland	0.5489	0.2944	0.2671	0.2544
EU-12 new	0.0034	0.0721	0.0943	0.1771
Remaining developed countries	0.0032	0.0576	0.0513	0.049
Less developed countries	0.4445	0.5759	0.5873	0.5190

an individual born abroad was regarded as an immigrant if no information on his or her parents was available. According to this definition, our sample of working-age individuals was comprised of 90.55% native workers and 9.45% immigrants. I then used the information on countries of origin from the DST register BEF and the employer-employee linked IDA data to construct the instrument variables. The DST defines the country of origin of an immigrant based on both parents' countries of birth. When none of the parents is known, the country of origin is defined based on the person's own information. If the person is immigrant, it is assumed that the country of origin is equal to the birth country. When only one of parents' information is known, the country of origin is defined based on the known parent. When both parents are known, the country of origin is defined based on the mother's country of birth or the country of citizenship respectively. As mentioned earlier, I followed the classification used by Malchow-Møller et al. (2012) and grouped the immigrants into four categories by country of origin: (1) classical EU-15 countries, plus Norway and Iceland; (2) new EU-12 countries (those that joined EU in 2004 and 2007); (3) remaining developed countries (according to UN classification); and (4) developing countries, including unknown countries of origin.

Table 2.3 provides information on immigrant characteristics at the individual level for the years 1999, 2004, 2008, and 2013. The table clearly displays trends in the distribution of countries of origin. As discussed in Section 2.2, with the enlargement of the EU in 2004 and 2007 and the full opening of the Danish labour market to Eastern EU workers in 2009, the "new EU-12" share almost doubled from 2008 to 2013. At the same time, the proportion of EU-15 immigrants decreased over the last decades, and the share of immigrants from less developed and developing countries remained relatively steady; a slight increase was found before 2008, and the figure fell in 2013. All evidence indicates that Denmark is becoming more international. When I turn to the distribution of immigrant skill levels, I can see from Table 2.3 that the share of high-skilled immigrants increased from 2004 to 2013. Various immigration policies aiming to attract high-skilled immigrants may have succeeded to a certain extent. High-skilled immigrants doubled within the ten years from 2004 to 2007.

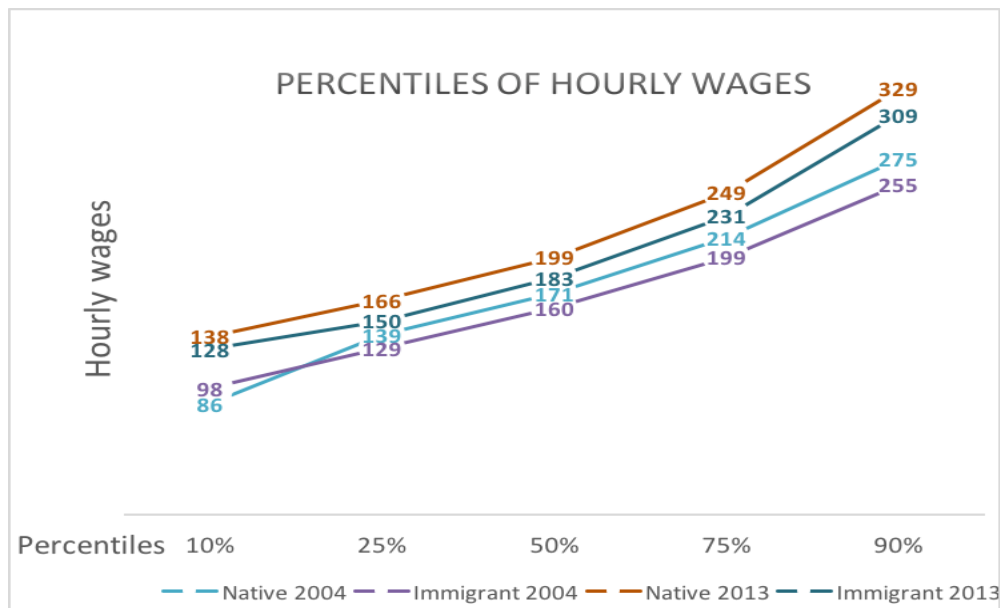


Figure 2.2: Wage Percentile

Figure 2.2 compares the hourly wages of natives and immigrants in different percentiles for both 2004 and 2013. Immigrants belonging to most percentiles earned less than natives in both 2004 and 2013, except for the lowest percentile in 2004. Furthermore, the average wages for both groups increased from 2004 to 2013, and this finding confirms the information found in Figures 2.1 in Section 2.2.

A more detailed comparison of the wages of natives and immigrants is in Table 2.4, which compares natives and immigrants within similar groups. Not surprisingly, immigrants in most groups earned less than natives. Of course, different factors might have contributed to this situation, and job mismatches for immigrant could potentially be one cause. I will not further discuss this issue here, and I leave it to Section 2.6. For the "abnormal" part in the table, however, it is interesting to note that in both 2004 and 2013, for the youngest group (aged 18-24), immigrants earned more than natives. Very old (aged 60-65) immigrants apparently received higher wages than natives in 2004, but the situation reversed in 2013. The situation in 2004 was also described by Malchow-Møller et al. (2012). They deemed selection and

other characteristics responsible. Finally, the wages of immigrants from the Eastern EU and less developed countries were much lower than those of immigrants from other countries; education and skill level could play a role in this gap.

Regarding the workplace-level data, I used variables from the BFL, FIRM, IDAN, IDAP, and IDAS DST registers to construct relevant variables for the estimation. I measured the size of the workplace by the number of employees. I then summarised and compared the share of immigrants with different skill levels in each workplace and other employee traits (e.g., age, gender, and other controlled factors). Those summary statistics for the workplace-level variables are presented in Table 2.5.

Being a country that has witnessed a substantial increase in the employment of immigrants, Denmark is regarded as an interesting case for immigration studies. As seen from Table 2.5, the number of workplaces with immigrants doubled from 2004 to 2013, and the overall share of immigrants in the workplace slightly increased during the ten years. However, the employment of high-skilled workers in workplaces with immigrants was three times what it was in 2004, while the employment of high-skilled workers in firms with only natives experienced a minor increase during the same period and the percentage of high-skilled workers was far less than that at firms that recruited immigrants. Additionally, the average size of firms with immigrants was much larger than firms with only natives, and the difference was much larger in 2013. In fact, the average firm with immigrants was almost 20 times larger than the average firms without immigrants in 2013, while the former was only three times larger than the latter in 2004.

Table 2.4: Hourly wage of natives and immigrants

	2004			2013		
	Natives	Immigrants	Natives	Immigrants	Natives	Immigrants
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
All	180.220	93.483	170.535	89.696	222.688	102.345
Age 18-24	112.654	66.658	122.407	67.757	141.723	63.715
Age 25-29	158.662	75.938	154.179	75.294	186.522	72.240
Age 30-39	187.392	92.383	172.584	88.109	223.871	93.160
Age 40-49	196.376	99.367	176.517	95.341	243.633	109.626
Age 50-59	189.415	89.264	185.595	91.337	234.152	102.982
Age 60-65	195.131	95.664	203.468	89.303	240.465	105.420
Low-skilled	154.251	79.429	146.473	73.732	185.751	88.247
Medium-skilled	188.847	93.773	169.844	86.543	226.500	98.969
High skilled	241.574	129.573	215.173	120.970	279.387	120.279
<i>Country of origin:</i>						
EU-15,Norway and Iceland			197.383	108.946	239.542	123.501
EU-12 new			174.208	75.611	187.789	84.757
Remaining developed countries			204.281	114.184	246.427	126.335
Less developed countries			152.981	71.370	195.730	87.540

Table 2.5: Summary statistics:workplace level-updatd

	Workplaces without immigrants				Workplaces with immigrants			
	2004		2013		2004		2013	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Nr. of employee	28.642	38.272	30.903	26.002	103.545	41.719	322.683	624.565
Share of female	0.3329	0.2687	0.3565	0.2777	0.3315	0.2198	0.4777	0.2710
Share of low skilled	0.2833	0.1807	0.2432	0.1757	0.6091	0.1656	0.6474	0.1827
Share of medium-skilled	0.695	0.2591	0.7111	0.2211	0.3461	0.1426	0.2292	0.1459
Share of high-skilled	0.0217	0.1211	0.0457	0.1123	0.0440	0.1009	0.1234	0.1638
Share of aged 40-65	0.4638	0.2151	0.5198	0.2212	0.4643	0.1722	0.5626	0.1773
Share of Union members	0.5782	0.1917	0.6367	0.1883	0.6690	0.1304	0.7103	0.1316
Share of immigrant	0	0	0	0	0.0723	0.0797	0.0904	0.0983
Nr. Of workplaces	10,387		14,157		6,714		16,583	

Table 2.6: Distribution of share of immigrant and share of different skill levels, workplace level-updated

	Share of immigrant with:		Share of all workers with:	
	Mean	Std. dev.	Mean	Std. dev.
Low-skilled, year 1999	0.0300	0.0576	0.3478	0.1902
Low-skilled, year 2004	0.0336	0.0542	0.3122	0.1543
Low-skilled, year 2008	0.0320	0.0440	0.2957	0.1318
Low-skilled, year 2013	0.0390	0.0505	0.2308	0.1525
Medium-skilled, year 1999	0.0051	0.0193	0.6232	0.1980
Medium-skilled, year 2004	0.0095	0.0150	0.6520	0.1827
Medium-skilled, year 2008	0.0128	0.0129	0.6451	0.1713
Medium-skilled, year 2013	0.0246	0.0156	0.6604	0.1903
High-skilled, year 1999	0.0028	0.0123	0.0290	0.0831
High-skilled, year 2004	0.0032	0.0110	0.0358	0.0918
High-skilled, year 2008	0.0045	0.0106	0.0592	0.1087
High-skilled, year 2013	0.0095	0.0184	0.1088	0.1584

Table 2.6 demonstrates the distribution of the share of immigrants across workplaces. I include the share of all workers in the workplace with the same skill level for comparison purposes in the right-hand column. The share of low-skilled immigrants increased from 1999 to 2013; this could be the result of EU enlargement, the more open labour markets in Denmark after 2009, and the arrival of more refugees who, on average, had less education. The majority of the recently arrived workers from the Eastern EU may also belong to the low-skilled group since their wages are far below the average for medium-skilled workers. Meanwhile, the average share of employment of all low-skilled workers decreased over the last 13 years. The mean share of medium-skilled immigrants remained stable during the same period, but the share of all workers increased. This could be the result of a possible response from natives

seeking to protect themselves from the effect of immigration by upgrading their occupation to a higher wage level. I discuss this effect in the discussion in Section 2.6. Moreover, the share of both high-skilled immigrants and all high-skilled workers in the firm increased from 1999 to 2013, revealing the demand for high-skilled employees in Danish labour markets.

2.5 Empirical Analysis

In this section, I perform the OLS, FE, IV, and quantile regressions of the wages of native workers on the share of immigrants with different skill levels in the workplace; the models also included control variables, such as age group, tenure, and experience. I discuss the results for the OLS, FE, and IV regressions in the following baseline specification. Then, I perform the quantile regression and explore in detail the effects of immigrants of different skill levels on the various wage percentiles. Robustness checks and other additional information are presented in the appendix.

2.5.1 Baseline specification

For all regression models (OLS, FE, and IV), both individual- and workplace-level control variables were included. The individual characteristics were dummy variables for age group, education level, gender, marital status, number of children aged 0-6 years old, and the labour market-related variables included tenure, tenure squared, experience and experience squared. At the workplace level, I particularly focused on the share of low-skilled, medium-skilled, and high-skilled immigrants, and I also included the share of workers over 40 years old, the share of female workers, the share of union members, the share of individuals with a vocational education, the share of individuals with a higher education, and the logarithm of the number of employees. Just as discussed in the framework setting in Section 2.3, $\text{year} \times \text{region} \times \text{industries}$ FEs were included in the OLS regression. I had 470 such FEs controlling for the particular changes resulting from effects of specific regions and industries.

Additionally, I included job-spell FEs in the FE regressions to control for the shocks from specific time-invariant worker-workplace characteristics. All standard errors were clustered at workplace-year level.

I present estimation results for the OLS and FE in Table 2.7.⁴ Full results are presented in the Appendix. I only considered the model jointly containing all shares of low-skilled, medium-skilled, and high-skilled immigrants. I did not analyse any single share of immigrants of a particular skill level (e.g., only the share of low-skilled immigrants in the workplace in a single model) because a bias could have emerged from potentially omitting critical variables. I present the additional specifications and further robustness checks in the appendix. As illustrated in column (1) in Table 2.7, all results from the OLS regressions were significant. Both low-skilled and medium-skilled immigrants had a negative wage effect on native co-workers in the firm, while the coefficient of high-skilled immigrants indicated a positive wage effect; this figure was much larger in absolute value (0.34) compared to the effects of low-skilled and medium-skilled immigrants (-0.09 and -0.18, respectively). The contribution from the low-skilled immigrant group was relatively small, but consistent with most previous empirical studies (e.g., Malchow-Møller et al., 2012 and Ortega and Verdugo, 2016).

However, after adding a job-spell FE to the model, the estimation results significantly changed for the low-skilled immigrant group. The results from the FE regressions are illustrated in column (2) in Table 2.7. Again, all estimations were significant. As I can conclude from the

⁴Notes: Estimations are based on a panel from 2004-2013. Full population include full-time native workers aged 18-65 from workplaces with 10+employees. I included but not reported following control variables: 1) individual-level characteristics: experience, experience², tenure, tenure², six age groups(dummies), marital status, children with 0-6 years, city size. 2) Workplace characteristics: share of female, number of employees(log), the share of medium educated, the share of highly-educated, the share of union members and share of aged 40+. In fixed effects estimation, time-invariant variables drop out. Standard errors are robust in OLS, and are clustered at the workplace-year level in FE

Table 2.7: Share of Immigrants at workplace and the wage of native workers, OLS and FE results

	Dependent variable: log(hourly wage)	
	(1)	(2)
	OLS	FE
Share in the workplace of:		
Low-skilled immigrants	-0.097*** (0.032)	0.156*** (0.041)
Medium-skilled immigrants	-0.183*** (0.016)	-0.231*** (0.015)
High-skilled immigrants	0.340*** (0.044)	0.259*** (0.054)
Job-spell fixed effects	No	Yes
Region×time×industry fixed effects	Yes	Yes
Observations	8,216,496	8,216,496
R-squared	0.215	0.603

Standard errors in parentheses:

Robust standard errors are reported for OLS in column (1).

Standard errors for FE in column (2) are clustered in workplace-year level.

***, ** and *: $H_0 : \beta_j = 0$ rejected at 1%, 5% and 10% significance level respectively.

table, the estimations for the medium-skilled immigrant group and the high-skilled immigrant group were consistent with the results of the OLS regression. The absolute value of the coefficient of the medium-skilled immigrant group increased. Meanwhile, the coefficient for the high-skilled immigrant group was smaller than the OLS estimation, which is probably because the positive effect was shared by the low-skilled immigrant group. In the FE regression, the effects from the low-skilled immigrant group became positive and were larger as compared to the OLS regression coefficient.

To be more specific, as regards the OLS and FE results, consistent estimations indicate that an one percentage point increase in the share of high-skilled immigrants in the workplace can contribute to 0.26 percentage point wage increase for native co-workers in the firm. The impact of medium-skilled immigrants on native workers was significantly negative. The effect from the low-skilled immigrant group was not clear, although after the job-spell FE was added to the model, the impact of that group on the wages of native workers turned positive.

To solve the potential endogeneity problem, I then performed IV regressions with FEs on $\text{year} \times \text{region} \times \text{industry}$ and job-spell FEs. The results are reported in the Appendix (see Table 2.12). All instruments chosen passed the F-test and were positively correlated with the potential endogenous variables, but all coefficients were small. Besides, I obtained much larger coefficients from the IV (2SLS) regressions than from the OLS and FE regressions. This is consistent with the empirical evidence from Malchow-Møller et al. (2012) except for the positive sign I found. In their paper, larger negative coefficients for significant estimations were found in IV estimation for the low-skilled immigrant group. Considering the comparableness of these results, I choose not to discuss them further.

To conclude, according to the OLS, FE, and IV regressions, positive effects of the high-skilled immigrant group on the wages of native workers were confirmed by all estimations. However, I cannot draw firm conclusions regarding the other two immigrant groups at this stage. As

explained earlier, the effects of differently skilled immigrants could be felt by different quantiles of wage earners. I examined those effects in detail through quantile regressions, illustrating the impact of immigrants of different skill levels on the various percentiles of wage earners.

2.5.2 Additional empirical evidence from the quantile regression analysis

By applying quantile regressions, I explored the impact of immigrants belonging to different wage quantiles to gain a more comprehensive picture of the impact of immigrants of different skill groups on natives. The same set of regressors employed in the OLS, FE, and IV analyses was used for our quantile regression model, and a set of region×year×industry FEs was included, same as what was used for OLS model. I used the 0.1th, 0.25th, 0.50th, 0.75th and 0.9th wage quantiles for our analysis. I investigated the impact of key variables on different wage quantiles separately. The estimation results together with bootstrapped standard errors are provided in Table 2.8, and columns 1 to 5 present the estimation for each quantile separately.⁵ Full results are presented in the Appendix.

⁵Notes for Table 2.8: Same data set is used in quantile regressions. All control variables are included but are not reported in Table 2.8, please refer to the explanations for Table 2.7.

Table 2.8: Share of Immigrants at workplace and the wage of native workers, Quantile regression results

	Dependent variable:log(hourly wage)				
	(1)	(2)	(3)	(4)	(5)
	q10	q25	q50	q75	q90
Share in the workplace of:					
Low-skilled immigrants	0.235*** (0.0117)	0.265*** (0.0083)	0.162*** (0.0100)	-0.031*** (0.0115)	-0.147*** (0.0197)
Medium-skilled immigrants	-0.089*** (0.0036)	-0.154*** (0.0025)	-0.205*** (0.0039)	-0.188*** (0.0044)	-0.235*** (0.0074)
High-skilled immigrants	0.300*** (0.0104)	0.094*** (0.0110)	0.097 *** (0.0122)	0.300*** (0.0157)	0.540*** (0.0242)
Job-spell fixed effects	No	No	No	No	No
Region×time×industry fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	8,216,496	8,216,496	8,216,496	8,216,496	8,216,496

Standard errors in parentheses.

***, ** and *: $H_0 : \beta_j = 0$ rejected at 1%, 5% and 10% significance level respectively.

The results in Table 2.8 make clear that low-skilled immigrants had positive effects on the wages of natives among most wage quantiles, while the impact of medium-skilled immigrants was significantly negative. High-skilled immigrants had significant estimations for higher wage quantiles. I then focused on different wage quantiles within each skilled immigrant group.

For the low-skilled immigrants' group, the estimations for low and middle quantiles were consistent with the FE and FE-IV results, while the estimations for higher quantiles (q75 and q90) were similar to OLS results in Table 2.7. The estimation coefficients showed a decreasing trend when moving from lower to higher quantiles (q10-q90). The positive wage effect on low-wage and middle-wage was larger in absolute value than it was on natives in high-wage quantiles. In other words, as the share of low-skilled immigrants increased, natives who earned low wages experienced more positive growth in wages than natives with wages falling in the middle, and natives who earn relatively high wages received negative impact from the increase of low-skilled immigrants on their wages. This results indicated that the negative results obtained OLS could be from the impact on higher wage quantiles. The evidence showed that a positive impact from immigration on the wages of native workers dominated in the low and medium wage quantiles, and it will be used for later interpretations.

All coefficients for medium-skilled immigrants were significantly negative, and the impact showed an increasing trend as the wage quantile rose from the low of 10% to the high of 90%. The effect was the largest among the three skill groups in 0.5th wage quantiles. The estimations were also consistent with the OLS and FE results.

The results for the high-skilled immigrant group were consistent with all results obtained using the previous OLS, FE, and IV models. Besides, the estimation results from the quantile regression for this group were more informative. As indicated in columns 2 to 3, which represent low- and medium-wage quantiles, the estimated coefficient for the high-skilled group for this wage range was numerically much less than that for the higher wage quantile. This indicates

that high-skilled immigrants tend to have a more positive impact on high-wage natives in the statistical aspect.

From these results, I can conclude that low-skilled immigrants have a positive effect on the wages of natives, while medium-skilled immigrants have a negative effect on natives, and especially on medium-wage natives. Negative effects were dominant among medium-wage natives. Finally, high-skilled immigrants tended to have a positive impact on natives who earned higher wages, and their influence on other wage quantiles was not statistically significant. The cross-quantile differences did not seem to be very large, except for high-skill immigrants. Here, the top quantiles were much higher. Perhaps the executives and managers at large international Danish firms are paid much more than the executives and managers at firms with only Danish managers (and that possibly use Danish as the internal language).

2.6 Discussion on the impact of immigrants of different skill levels

Based on the empirical evidence outlined in Section 2.5, I can now draw a clearer picture of the impact of immigrants of different skill levels on the wages of native workers. In this section, the above empirical analysis is discussed in the light of economic theories. Apart from the low-skilled and medium-skilled immigrant groups, which have been studied by numerous researchers, I additionally focus on high-skilled immigrants, who have more influence today than in the past.

2.6.1 Low-skilled immigrants

Our empirical results from the OLS regression demonstrate minor negative effects of lower-skilled immigrants on the wages of natives, and it could be biased by job/firm-specific shocks.

The results from corresponding quantile regressions further indicate such shocks, since the impact was only negative in high wage quantiles. In other words, low-skilled immigrants do not reduce the wages of those natives. This is in accordance with the hypothesis H1 in Section 2.2 that positive wage effect might be brought by low-skilled immigrants, due to the occupation upgrading behaviour of natives. Also, the wage efficiency theory predicts the positive wage effect by low-skilled immigrants.

As discussed in the empirical analysis, the increase in low-skilled immigrants led to higher wages among natives, especially in the low-wage quantile. This also confirms the occupation upgrading theory discussed in Section 2.2. Low-skilled immigrants are increasing while low-skilled workers' share of total employment is indeed decreasing, as our data have demonstrated in Section 2.4. The increase could be due to the large inflow of immigrants from Eastern European countries; these individuals are mainly low-skilled workers. With better unemployment benefits and social welfare coverage than immigrants, natives find it easier than immigrants to return to college or continue with technical training to gain new skills and to upgrade to new types of jobs where language skills and networking play a greater role. Therefore, those natives are pushed to occupations paying higher wages. As Foged and Peri (2016) have agreed, natives and immigrants are not perfect substitutes in the low-skilled labour market; the authors also found positive impacts of low-skilled immigrants on native wages, particularly for low-skilled natives. These results are aligned with our findings from the FE, IV, and quantile regressions.

Moreover, our empirical evidence strongly supports the wage efficiency theory. The wage could exceed the market-clearing level in industries where the cost of replacing labour is high if managers want to increase a firm's productivity or efficiency or to decrease costs related to labour turnover. The empirical evidence I obtained indicated that managers could increase the wage to motivate the natives to stay in order to reduce labour turnover costs. With more low-skilled immigrants entered the Danish labour market, the increase of wages of natives

increases may partly resulting from managers' motivation to reduce labour turnover costs within the firm.

2.6.2 Medium-skilled immigrants

The hypothesis H2 mentioned in Section 2.2—The increased natives and immigrants could squeeze in the medium-skilled class, causing more fierce competition in this market as the share of medium-skilled immigrants increases, and worsening the wages of natives—is firstly confirmed by the descriptive data in Section 2.4 indicating that more natives moved from the low-skilled group to medium-skilled occupations from 2004 to 2013. Then, as strongly confirmed by the empirical evidence from the OLS, FE, and conditional quantile regressions, medium-skilled immigrants have a negative impact on the wages of native workers. Basic labour demand-and-supply theory could further explain this evidence. The share of medium-skilled workers at native firms increased significantly from 2004 to 2013. One explanation for this increase may be lower-skilled natives moving to medium-skilled types of work. Another route could be the share of immigrants. Since the supply of medium-skilled workers would have been more than actual demand, more immigrants would have meant more competition for positions.

2.6.3 High-skilled immigrants

The estimation results from all regression models indicated a significant positive impact of high-skilled immigrants. Only limited empirical results have previously been reported for this group. According to the summary statistics displayed in Section 2.4, the proportion of high-skilled immigrants increased dramatically in the local labour market in Denmark. The evidence from our empirical tests indicates that high-skilled immigrants are influencing more local employees than ever, while Denmark has witnessed a boom in high-skilled immigrants. The two mechanisms predicting the positive wage effect on natives—especially those in higher wage quantiles' of high-skilled immigrants are discussed in Section 2.2. Our empirical evidence

strongly supports the hypothesis H3 in Section 2.2—Wage efficiency theory together with the O-Ring theory predict a positive impact of high-skilled immigrants on the wages of natives.

Our empirical evidence strongly supports the O-Ring theory (Kremer, 1993), mentioned in Section 2.2. Immigrants can increase the wages of natives by boosting a firm's total productivity. The literature on skill transfer and technology spillover also confirms our results (Peri et al., 2014; Peri, 2012; and Malchow-Møller, 2011 etc.).

Furthermore, our empirical analysis of high-skilled immigrant groups is also well aligned with the wage efficiency theory. The efficiency wage hypothesis in Section 2.2 argues that usually, wages are not always market clearing. To be more specific, the wage could exceed the market-clearing level in industries where the cost of replacing labour is high if managers want to increase a firm's productivity or efficiency or to decrease costs related to labour turnover.

Mobility for markets where high-skilled workers usually work is relatively high compared to those comprised of other groups (Tremblay, 2005), as those high-skilled workers can more easily find a position elsewhere due to their competency in specific skills. Moreover, as the earlier summary statistics illustrated, firms with immigrants have recruited many more high-skilled workers during recent decades, while shares of high-skilled workers at firms with only natives have remained stable, and the share of high-skilled workers across all workplaces has increased significantly. The arrival of more high-skilled immigrants could be a signal of greater labour demand in high-skilled sectors. To avoid high costs due to the labour turnover of high-skilled employees (e.g., hiring and training), firms pay a higher wage to keep their skilled workers. Hence, the wages of all high-skilled workers, including natives, increase. High-skilled immigrants play a role in this mechanism as an indirect measure of labour demand.

Further, according to the wage efficiency theory, another incentive for managers to pay more is to hire more productive workers; thus, a selection effect cannot be ignored, particularly among

high-skilled workers. As discussed in Section 2.2, high-skilled workers only select high-wage positions because their reservation wage is too high for low-wage firms, but high-wage positions attract applicants with various skill levels. Moreover, high-wage firms prefer to recruit more productive workers to achieve an efficiency wage. With more high-skilled immigrants arriving in Denmark, the managers of a high-wage firm tend to tighten hiring standards and to provide higher wages to attract more productive workers to enlarge the potential applicant pool; in this way, they recruit more competitive employees to enhance the firm's productivity. Again, wages for all high-skilled and high-wage earners are pushed upwards by this selection effect, and the increased number of high-skilled immigrants facilitates this outcome. Our empirical evidence clearly confirmed our hypothesis.

2.7 Conclusion

In this paper, I have focused on an increase in the number of immigrants, especially high-skilled groups, in a local labour market. I tested the impact of immigrants on the wages of natives by using administrative data from Statistics Denmark on the full population in Denmark for the period from 2004 to 2013. I found evidence that an increase in the share of high-skilled immigrants leads to higher wages for native workers, especially high-wage earners. Empirical evidence on low-skilled and medium-skilled immigrant groups has also been discussed in the paper.

The empirical tests applied OLS, FE, IV (2SLS), and quantile regression models. High-skilled immigrants were found to have a positive impact on natives, as confirmed by all estimation models. Moreover, evidence from the quantile regression indicates that the positive wage effect from high-skilled immigrants is mainly on natives who earn higher wages. In addition, according to the estimation results from the FE, FE-IV, and quantile regressions, low-skilled immigrants also have a positive effect on the wages of natives, and they have a more positive impact on low-wage natives. As confirmed by the OLS, FE, and quantile estimations, medium-

skilled immigrants have negative wage effects on, and the negative effects dominants for the medium-wage native group.

I have also discussed the empirical findings based on the hypothesised mechanisms. I argued that the positive wage effects of low-skilled immigrants could be due to wage efficiency theory and the occupation upgrading behaviour of natives seeking to protect themselves from the competition created by increased immigration. Then, the increase in medium-skilled natives and immigrants generates even fiercer competition in this market as the share of medium-skilled immigrants increases; the outcome is reduced wages for natives. Finally, the wage efficiency theory explains the positive impact of high-skilled immigrants on the wages of natives, especially high-wage earners, in terms of labour turnover and selection theory. Hiring standards and wage levels both rise due to high-wage firms' motivation to reduce costs associated with labour replacement and to hire more productive workers.

2.7.1 Future discussions

As our research question examining the impact of immigrants on the wages of native workers is part of the broader wage inequality literature on immigration, it was of interest to further study changes in the distribution of immigrant wages to gain a full picture of the influence of increased immigration on the local labour market in Denmark. According to our descriptive data, potential wage decimation exists over immigrants. A more in-depth analysis could also be performed on the skill mismatch between natives and immigrants to obtain more insight into wage inequality issues.

Bibliography

Akerlof, G. A. (1982), 'Labor contracts as partial gift exchange', *The quarterly journal of economics* **97**(4), 543–569.

Angrist, J. D. & Kugler, A. D. (2003), 'Protective or counter-productive? labour market institutions and the effect of immigration on eu natives', *The Economic Journal* **113**(488), F302–F331.

Card, D. (2001), 'Immigrant inflows, native outflows, and the local labor market impacts of higher immigration', *Journal of Labor Economics* **19**(1), 22–64.

Carrasco, R., Jimeno, J. F. & Ortega, A. C. (2008), 'The effect of immigration on the labor market performance of native-born workers: some evidence for spain', *Journal of Population Economics* **21**(3), 627–648.

Correia, S. (2016), Linear models with high-dimensional fixed effects: An efficient and feasible estimator, Technical report, Duke University. Working Paper.

D'Amuri, F. & Peri, G. (2014), 'Immigration, jobs, and employment protection: evidence from

- Europe before and during the great recession', *Journal of the European Economic Association* **12**(2), 432–464.
- Dustmann, C., Fabbri, F. & Preston, I. (2005), 'The impact of immigration on the British labour market', *The Economic Journal* **115**(507).
- Dustmann, C., Frattini, T. & Preston, I. P. (2013), 'The effect of immigration along the distribution of wages', *The Review of Economic Studies* **80**(1), 145–173.
- Edo, A. & Toubal, F. (2015), 'Selective immigration policies and wages inequality', *Review of International Economics* **23**(1), 160–187.
- Fogel, M. & Peri, G. (2016), 'Immigrants' effect on native workers: New analysis on longitudinal data', *American Economic Journal: Applied Economics* **8**(2), 1–34.
- Glaeser, E. (2011), *Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier*, Penguin.
- Hellerstein, J. K. & Neumark, D. (2008), 'Workplace segregation in the United States: Race, ethnicity, and skill', *The Review of Economics and Statistics* **90**(3), 459–477.
- Hunt, J. & Gauthier-Loiselle, M. (2010), 'How much does immigration boost innovation?', *American Economic Journal: Macroeconomics* **2**(2), 31–56.
- Jones, C. I. (2002), 'Sources of US economic growth in a world of ideas', *The American Economic Review* **92**(1), 220–239.
- Koenker, R. & Bassett Jr, G. (1978), 'Regression quantiles', *Econometrica: journal of the*

Econometric Society pp. 33–50.

Koenker, R. & Hallock, K. (2001), 'Quantile regression: An introduction', *Journal of Economic Perspectives* **15**(4), 43–56.

Kremer, M. (1993), 'The o-ring theory of economic development', *The Quarterly Journal of Economics* **108**(3), 551–575.

Malchow-Møller, N., Munch, J. R. & Skaksen, J. R. (2011), 'Do foreign experts increase the productivity of domestic firms?', *The Scandinavian Journal of Economics* .

Malchow-Møller, N., Munch, J. R. & Skaksen, J. R. (2012), 'Do immigrants affect firm-specific wages?', *The Scandinavian Journal of Economics* **114**(4), 1267–1295.

Malcomson, J. M. (1981), 'Unemployment and the efficiency wage hypothesis', *The Economic Journal* **91**(364), 848–866.

Manacorda, M., Manning, A. & Wadsworth, J. (2012), 'The impact of immigration on the structure of wages: theory and evidence from Britain', *Journal of the European Economic Association* **10**(1), 120–151.

Markusen, J. R. & Trefimienko, N. (2009), 'Teaching locals new tricks: Foreign experts as a channel of knowledge transfers', *Journal of Development Economics* **88**(1), 120–131.

McDonald, I. M. & Solow, R. M. (1981), 'Wage bargaining and employment', *The American Economic Review* **71**(5), 896–908.

Moretti, E. (2004), 'Workers' education, spillovers, and productivity: evidence from plant-level

- production functions', *The American Economic Review* **94**(3), 656–690.
- Moretti, E. (2012), *The new geography of jobs*, Houghton Mifflin Harcourt.
- Neyman, J. & Scott, E. L. (1948), 'Consistent estimates based on partially consistent observations', *Econometrica: Journal of the Econometric Society* pp. 1–32.
- OECD (2008), International migration outlook, Technical report, OECD, Paris.
- Orrenius, P. M. & Zavodny, M. (2007), 'Does immigration affect wages? a look at occupation-level evidence', *Labour Economics* **14**(5), 757 – 773.
URL: <http://www.sciencedirect.com/science/article/pii/S0927537106000674>
- Ortega, J. & Verdugo, G. (2016), Moving up or down? immigration and the selection of natives across occupations and locations, Technical report, IZA.
- Ottaviano, G. I. & Peri, G. (2006), 'The economic value of cultural diversity: evidence from us cities', *Journal of Economic geography* **6**(1), 9–44.
- Ottaviano, G. I. & Peri, G. (2012), 'Rethinking the effect of immigration on wages', *Journal of the European economic association* **10**(1), 152–197.
- Peri, G. (2012), 'The effect of immigration on productivity: Evidence from us states', *Review of Economics and Statistics* **94**(1), 348–358.
- Peri, G., Shih, K. & Sparber, C. (2015), 'Stem workers, h-1b visas, and productivity in us cities', *Journal of Labor Economics* **33**(S1), S225–S255.

Peri, G., Shih, K. Y. & Sparber, C. (2014), Foreign stem workers and native wages and employment in us cities, Technical report, National Bureau of Economic Research.

Peri, G. & Sparber, C. (2009), 'Task specialization, immigration, and wages', *American Economic Journal: Applied Economics* **1**(3), 135–169.

Schlicht, E. (2005), 'Hiring standards and labour market clearing', *Metroeconomica* **56**(2), 263–279.

Tremblay, K. (2005), 'Academic mobility and immigration', *Journal of Studies in International Education* **9**(3), 196–228.

Appendices

Appendix A: Notes for constructed variables for Table 2.2

In this paper, the following variables are constructed explicitly for the empirical analysis:

Hourly wage: It is calculated via dividing total annual wages including ATP by total hours worked in the year.

Age dummies: Age 18-24, Age 25-29, Age 30-39, Age 40-49, Age 50-59, Age 60-65.

Children 0-6 years: A dummy indicator of whether there is a child in the household aged 0-6 years.

Experience: In years. Sum of days employed in the labour market, divided by 365.

Tenure: In years. Time employed within the workplace.

Appendix B: Robustness checks

Appendix B.1: Robustness check for old workplace code

Table 2.9: Robustness check for old workplace code

Dependent variable: log(hourly wage)				
	(1)	(2)	(3)	(4)
Share in the workplace of:				
Low-skilled immigrants	-0.211*** (0.029)			-0.070** (0.029)
Medium-skilled immigrants		-0.177*** (0.011)		-0.188*** (0.011)
High-skilled immigrants			0.258*** (0.042)	0.366*** (0.042)
Region×time×industry fixed effects	Yes	Yes	Yes	Yes
Observations	8,268,580	8,268,580	8,268,580	8,268,580
R-squared	0.216	0.216	0.216	0.216

Appendix B.2: Robustness checks on data sample compared to Malchow-Møller et al.(2012)

Table 2.10: Comparison on Immigrant characteristics at individual level, selected years

	This paper		Malchow-Møller et al. (2012)	
	1999	2004	1999	2004
Age 18-24	0.1119	0.0669	0.090	0.067
Age 25-29	0.1349	0.1241	0.138	0.122
Age 30-39	0.3008	0.3658	0.364	0.356
Age 40-49	0.2386	0.2686	0.240	0.281
Age 50-59	0.1861	0.1469	0.150	0.147
Age 60-65	0.0188	0.0238	0.012	0.018
Basic education(low-skilled)	0.6041	0.5778	0.444	0.492
Vocational education(medium-skilled)	0.3475	0.3724	0.322	0.288
Higher education(high-skilled)	0.0325	0.0385	0.234	0.220
<i>Country of origin:</i>				
EU-15,Norway and Iceland	0.5489	0.2944	0.342	0.290
EU-12 new	0.034	0.0721	0.048	0.049
Remaining developed countries	0.032	0.0576	0.190	0.190
Less developed countries	0.4445	0.5759	0.419	0.470

Appendix C: Regression results from FE-IV model

Table 2.11: Results of 1st stage IV regression

	(1)	(2)	(3)
Instrument	share of Low-skilled immigrants	share of Medium-skilled immigrants	share of High-skilled immigrants
the immigrant group of:			
Low-skilled	0.027** (0.012)		
Medium-skilled		0.046***	
High-skilled			0.016*** (0.004)
Observations	8,216,496	8,216,496	8,216,496
R-squared	0.627	0.818	0.759

Standard errors in parentheses.

***, ** and *: $H_0 : \beta_j = 0$ rejected at 1%, 5% and 10% significance level respectively.

Table 2.12: Share of Immigrants at workplace and the wage of native workers, FE-IV results

Dependent variable: log(hourly wage)	
	(1)
	FE-IV
Share in the workplace of:	
Low-skilled immigrants	4.800*** (1.538)
Medium-skilled immigrants	-0.486 (0.476)
High-skilled immigrants	8.235*** (2.479)
Job-spell fixed effects	Yes
Region \times time \times industry fixed effects	Yes
Observations	8,216,496
R-squared	0.602

Standard errors clustered at workplace-year level in parentheses.

***, ** and *: $H_0 : \beta_j = 0$ rejected at 1%, 5% and 10% significance level respectively.

Appendix D: Full results from OLS regressions(From STATA log files)

*OLS Regression: Detailed results from STATA log file for colume (1) in Table 2.7

wage	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
share_of_lowskill	-.0969024	.0316977	-3.06	0.002	-.159029	-.0347758
share_of_mskill	-.183414	.0163114	-11.24	0.000	-.2153839	-.151444
share_of_highskill	.3397619	.0444052	7.65	0.000	.252729	.4267948
exp	.0176329	.0002615	67.44	0.000	.0171204	.0181454
exp2	-.0004862	7.78e-06	-62.52	0.000	-.0005014	-.000471
Tenure	.0018951	.0002872	6.60	0.000	.0013321	.002458
ten2	-.0000527	.0000105	-5.02	0.000	-.0000732	-.0000321
Age_18_24	.0930537	.0062727	14.83	0.000	.0807595	.1053479
Age_25_29	.3012516	.0063415	47.50	0.000	.2888225	.3136807
Age_30_39	.3889819	.0064013	60.77	0.000	.3764356	.4015282
Age_40_49	.4379847	.006452	67.88	0.000	.425339	.4506303
Age_50_59	.4187501	.0064093	65.33	0.000	.4061879	.4313122
Age_60_65	.4256956	.0065138	65.35	0.000	.4129288	.4384625
basic_edu	.0115519	.0021942	5.26	0.000	.0072513	.0158526
vocational_edu	.0841027	.0019157	43.90	0.000	.0803481	.0878573
higher_edu	.172492	.0030487	56.58	0.000	.1665166	.1784674
Married	.0541982	.0007795	69.53	0.000	.0526705	.055726
children_0_6_yrs	.0276824	.0012383	22.35	0.000	.0252553	.0301095
Copenhagen	.0188997	.0025571	7.39	0.000	.0138878	.0239115
large_city	.0033646	.0022449	1.50	0.134	-.0010353	.0077646
small_city	-.0029147	.0020405	-1.43	0.153	-.0069141	.0010848
lnemployee	.0189257	.0006287	30.10	0.000	.0176936	.0201579
share_female	-.1728413	.00281	-61.51	0.000	-.1783489	-.1673337
share_40_65	-.0326495	.004262	-7.66	0.000	-.0410029	-.0242962
share_highskill	.2717397	.0064374	42.21	0.000	.2591225	.2843568
share_mskill	.1924223	.006098	31.56	0.000	.1804704	.2043741
share_UM	.0127669	.0066649	1.92	0.055	-.0002962	.02583

Appendix E: Full Regression results from FE model and additional FE results on robust standard errors

Appendix E.1: Full regression results for FE model in Table 2.7

**FE regressions: Full results from STATA log file for Table 2.7
 ****FE include year*region*industry FE and job-spell FE,
 *****se clustered at workplace-year level

wage	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
share_of_lowskill	.1557325	.0347344	4.48	0.000	.0876542	.2238108
share_of_mskill	-.2310268	.0116161	-19.89	0.000	-.2537939	-.2082596
share_of_highskill	.2590338	.0390229	6.64	0.000	.1825503	.3355173
exp	.0195552	.0002479	78.87	0.000	.0190693	.0200412
exp2	-.0005395	7.65e-06	-70.56	0.000	-.0005545	-.0005246
Tenure	.0005858	.0002151	2.72	0.006	.0001642	.0010075
ten2	-.0000238	8.59e-06	-2.77	0.006	-.0000407	-6.99e-06
Age_18_24	.1023676	.0070085	14.61	0.000	.0886312	.116104
Age_25_29	.319842	.0069094	46.29	0.000	.3062998	.3333842
Age_30_39	.4044117	.0068542	59.00	0.000	.3909778	.4178456
Age_40_49	.4545789	.0068573	66.29	0.000	.4411388	.4680191
Age_50_59	.4283466	.0068712	62.34	0.000	.4148793	.4418138
Age_60_65	.4380757	.0070742	61.93	0.000	.4242105	.451941
basic_edu	-.0208581	.0022367	-9.33	0.000	-.0252421	-.0164742
vocational_edu	.0469668	.0022367	21.00	0.000	.042583	.0513506
higher_edu	.1366584	.0025717	53.14	0.000	.131618	.1416987
Married	.0547302	.0009518	57.50	0.000	.0528647	.0565956
children_0_6_yrs	.0271343	.0012251	22.15	0.000	.024733	.0295355
Copenhagen	.0433269	.007289	5.94	0.000	.0290408	.0576131
large_city	.0032354	.00744	0.43	0.664	-.0113467	.0178175
small_city	-.0079913	.007386	-1.08	0.279	-.0224675	.006485
lnemployee	.0188896	.0002884	65.49	0.000	.0183242	.0194549
share_female	-.191283	.0017341	-110.30	0.000	-.1946818	-.1878841
share_40_65	-.004697	.0028096	-1.67	0.095	-.0102038	.0008098
share_highskill	.2679309	.0046082	58.14	0.000	.258899	.2769628
share_mskill	.1180311	.0026884	43.90	0.000	.1127619	.1233002
share_UM	.1095753	.0034419	31.84	0.000	.1028292	.1163214

Appendix E.2: Additional regression results from using robust standard error in FE model in Table 2.7

* Fe regressions: Results from using robusted standard errors:

Statistics robust to heteroskedasticity	Prob > F	=	0.0000
	R-squared	=	0.6026
	Adj R-squared	=	0.2674
	Within R-sq.	=	0.1825
	Root MSE	=	0.3307

wage	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
share_of_lowskill	.1557325	.0347344	4.48	0.000	.0876542	.2238108
share_of_mskill	-.2310268	.0116161	-19.89	0.000	-.2537939	-.2082596
share_of_highskill	.2590338	.0390229	6.64	0.000	.1825503	.3355173
exp	.0195552	.0002479	78.87	0.000	.0190693	.0200412
exp2	-.0005395	7.65e-06	-70.56	0.000	-.0005545	-.0005246
Tenure	.0005858	.0002151	2.72	0.006	.0001642	.0010075
ten2	-.0000238	8.59e-06	-2.77	0.006	-.0000407	-6.99e-06
Age_18_24	.1023676	.0070085	14.61	0.000	.0886312	.116104
Age_25_29	.319842	.0069094	46.29	0.000	.3062998	.3333842
Age_30_39	.4044117	.0068542	59.00	0.000	.3909778	.4178456
Age_40_49	.4545789	.0068573	66.29	0.000	.4411388	.4680191
Age_50_59	.4283466	.0068712	62.34	0.000	.4148793	.4418138
Age_60_65	.4380757	.0070742	61.93	0.000	.4242105	.451941
basic_edu	-.0208581	.0022367	-9.33	0.000	-.0252421	-.0164742
vocational_edu	.0469668	.0022367	21.00	0.000	.042583	.0513506
higher_edu	.1366584	.0025717	53.14	0.000	.131618	.1416987
Married	.0547302	.0009518	57.50	0.000	.0528647	.0565956
children_0_6_yrs	.0271343	.0012251	22.15	0.000	.024733	.0295355
Copenhagen	.0433269	.007289	5.94	0.000	.0290408	.0576131
large_city	.0032354	.00744	0.43	0.664	-.0113467	.0178175
small_city	-.0079913	.007386	-1.08	0.279	-.0224675	.006485
lnemployee	.0188896	.0002884	65.49	0.000	.0183242	.0194549

share_female		-.191283	.0017341	-110.30	0.000	-.1946818	-.1878841
share_40_65		-.004697	.0028096	-1.67	0.095	-.0102038	.0008098
share_highskill		.2679309	.0046082	58.14	0.000	.258899	.2769628
share_mskill		.1180311	.0026884	43.90	0.000	.1127619	.1233002
share_UM		.1095753	.0034419	31.84	0.000	.1028292	.1163214

(est8 stored)

**Appendix F: Full results from QR regressions and tests for equality
across quantiles(From STATA log files)**

* Quantile regressions: Detailed results from STATA log file for Table 2.8 in Section 5.2.

Simultaneous quantile regression
bootstrap(20) SEs

Number of obs = 8216496
.10 Pseudo R2 = 0.1849
.25 Pseudo R2 = 0.1253
.50 Pseudo R2 = 0.0947
.75 Pseudo R2 = 0.0876
.90 Pseudo R2 = 0.0834

wage	Coef.	Bootstrap Std. Err.	t	P> t	[95% Conf. Interval]	
q10						
share_of_lowskill	.2345164	.0117355	19.98	0.000	.2115151	.2575176
share_of_mskill	-.08896	.0036303	-24.50	0.000	-.0960753	-.0818448
share_of_highskill	.2998738	.0104243	28.77	0.000	.2794426	.320305
exp	.0187423	.0000787	238.27	0.000	.0185882	.0188965
exp2	-.00052	2.31e-06	-225.19	0.000	-.0005245	-.0005154
Tenure	.0037093	.0000501	74.06	0.000	.0036111	.0038075
ten2	-.0000937	2.10e-06	-44.54	0.000	-.0000978	-.0000896
Age_18_24	.2237348	.0022839	97.96	0.000	.2192585	.2282111
Age_25_29	.6488959	.0021829	297.27	0.000	.6446176	.6531742
Age_30_39	.7206536	.0021559	334.27	0.000	.7164281	.7248791
Age_40_49	.7403253	.0020041	369.41	0.000	.7363974	.7442533
Age_50_59	.7289539	.0020954	347.88	0.000	.7248469	.7330609
Age_60_65	.7304621	.0021094	346.29	0.000	.7263277	.7345964
basic_edu	-.0503082	.0007076	-71.10	0.000	-.051695	-.0489214
vocational_edu	.0291858	.00055	53.07	0.000	.0281079	.0302637
higher_edu	.0901444	.0008737	103.17	0.000	.0884319	.0918568
Married	.0250025	.0002075	120.50	0.000	.0245958	.0254091
children_0_6_yrs	.0193784	.0004492	43.14	0.000	.0184979	.0202589
Copenhagen	.0611646	.0028376	21.56	0.000	.055603	.0667261

large_city	.0002235	.0028643	0.08	0.938	-.0053905	.0058375
small_city	-.0044001	.0028024	-1.57	0.116	-.0098928	.0010925
lnemployee	.0142614	.0000781	182.64	0.000	.0141084	.0144145
share_female	-.1557834	.0004584	-339.83	0.000	-.1566819	-.1548849
share_40_65	.0819471	.0007854	104.33	0.000	.0804077	.0834865
share_highskill	.2513881	.0011774	213.51	0.000	.2490805	.2536958
share_mskill	.1472006	.0007859	187.29	0.000	.1456602	.148741
share_UM	.0169437	.0011755	14.41	0.000	.0146396	.0192477
_cons	3.915779	.0037728	1037.90	0.000	3.908385	3.923174
<hr/>						
q25						
share_of_lowskill	.2652097	.0082959	31.97	0.000	.2489501	.2814693
share_of_mskill	-.1541295	.0024952	-61.77	0.000	-.1590199	-.149239
share_of_highskill	.0942061	.0109943	8.57	0.000	.0726576	.1157546
exp	.0172168	.0000738	233.44	0.000	.0170723	.0173614
exp2	-.0004749	2.29e-06	-207.74	0.000	-.0004794	-.0004704
Tenure	.0007328	.000054	13.58	0.000	.000627	.0008385
ten2	-3.35e-06	1.80e-06	-1.87	0.062	-6.87e-06	1.66e-07
Age_18_24	.3054428	.0025918	117.85	0.000	.300363	.3105225
Age_25_29	.6187723	.0022394	276.31	0.000	.6143832	.6231614
Age_30_39	.6861828	.0022139	309.94	0.000	.6818436	.690522
Age_40_49	.709859	.0022775	311.68	0.000	.7053951	.7143228
Age_50_59	.6981821	.0023278	299.94	0.000	.6936198	.7027444
Age_60_65	.7048431	.0022767	309.59	0.000	.7003809	.7093054
basic_edu	-.044347	.0005517	-80.39	0.000	-.0454283	-.0432658
vocational_edu	.0337831	.0004991	67.69	0.000	.0328049	.0347613
higher_edu	.1084028	.000745	145.50	0.000	.1069425	.109863
Married	.0288623	.0001767	163.30	0.000	.0285159	.0292087
children_0_6_yrs	.0225367	.0003246	69.42	0.000	.0219004	.0231729
Copenhagen	.0640417	.0020893	30.65	0.000	.0599468	.0681366
large_city	.001302	.0020041	0.65	0.516	-.002626	.0052301
small_city	-.0037447	.002028	-1.85	0.065	-.0077195	.0002301
lnemployee	.0141	.0000621	226.94	0.000	.0139783	.0142218
share_female	-.1656077	.0004613	-359.01	0.000	-.1665118	-.1647036
share_40_65	.0267267	.0008204	32.58	0.000	.0251187	.0283347

share_highskill	.2879096	.0012172	236.53	0.000	.2855239	.2902953
share_mskill	.1521054	.0006341	239.88	0.000	.1508626	.1533482
share_UM	.0173675	.0010568	16.43	0.000	.0152963	.0194388
_cons	4.125075	.002882	1431.34	0.000	4.119427	4.130724

q50						
share_of_lowskill	.1617596	.0100313	16.13	0.000	.1420986	.1814206
share_of_mskill	-.2048517	.0030741	-66.64	0.000	-.2108768	-.1988265
share_of_highskill	.0968769	.0121712	7.96	0.000	.0730217	.1207321
exp	.0173602	.0000701	247.71	0.000	.0172229	.0174976
exp2	-.0004774	1.90e-06	-251.12	0.000	-.0004812	-.0004737
Tenure	-.0012902	.0000551	-23.40	0.000	-.0013983	-.0011822
ten2	.0000524	2.01e-06	26.03	0.000	.0000485	.0000564
Age_18_24	.152593	.0067689	22.54	0.000	.1393263	.1658598
Age_25_29	.3178937	.0068596	46.34	0.000	.3044491	.3313384
Age_30_39	.3901314	.0067317	57.95	0.000	.3769374	.4033254
Age_40_49	.4219781	.0067688	62.34	0.000	.4087115	.4352448
Age_50_59	.4042836	.0066845	60.48	0.000	.3911823	.4173849
Age_60_65	.4195555	.0067514	62.14	0.000	.406323	.432788
basic_edu	-.017574	.0007609	-23.10	0.000	-.0190652	-.0160827
vocational_edu	.0447571	.0006573	68.09	0.000	.0434688	.0460455
higher_edu	.1325155	.0007696	172.19	0.000	.1310071	.1340239
Married	.0399547	.0002572	155.33	0.000	.0394505	.0404588
children_0_6_yrs	.0250412	.0003198	78.30	0.000	.0244144	.025668
Copenhagen	.0607857	.0017544	34.65	0.000	.0573472	.0642242
large_city	.0001576	.0018103	0.09	0.931	-.0033905	.0037058
small_city	-.0029957	.0017904	-1.67	0.094	-.0065048	.0005133
lnemployee	.01735	.0000735	235.90	0.000	.0172059	.0174942
share_female	-.188369	.0004297	-438.36	0.000	-.1892112	-.1875268
share_40_65	.0022015	.000677	3.25	0.001	.0008747	.0035283
share_highskill	.2735892	.0013182	207.54	0.000	.2710055	.2761729
share_mskill	.1025354	.0006151	166.70	0.000	.1013299	.103741
share_UM	.0757468	.0010034	75.49	0.000	.07378	.0777135
_cons	4.566655	.0067888	672.67	0.000	4.553349	4.579961

q75						
share_of_lowskill	-.0308245	.0115399	-2.67	0.008	-.0534423	-.0082067
share_of_mskill	-.1877681	.0044148	-42.53	0.000	-.1964209	-.1791153
share_of_highskill	.2996361	.0157383	19.04	0.000	.2687896	.3304827
exp	.0182017	.0000982	185.37	0.000	.0180092	.0183941
exp2	-.0005122	2.72e-06	-188.42	0.000	-.0005175	-.0005068
Tenure	-.004437	.0000965	-46.00	0.000	-.004626	-.004248
ten2	.0001313	3.58e-06	36.69	0.000	.0001243	.0001383
Age_18_24	-.1431733	.0027863	-51.38	0.000	-.1486344	-.1377121
Age_25_29	-.0122186	.0027704	-4.41	0.000	-.0176484	-.0067887
Age_30_39	.0861295	.0026003	33.12	0.000	.0810331	.0912258
Age_40_49	.15382	.0025944	59.29	0.000	.1487351	.1589049
Age_50_59	.1247297	.0028061	44.45	0.000	.1192298	.1302296
Age_60_65	.1500446	.0028656	52.36	0.000	.1444282	.155661
basic_edu	.0078224	.0008706	8.99	0.000	.006116	.0095287
vocational_edu	.0608997	.0008551	71.22	0.000	.0592237	.0625757
higher_edu	.1743607	.0011185	155.89	0.000	.1721685	.1765528
Married	.0610172	.0004326	141.03	0.000	.0601692	.0618651
children_0_6_yrs	.0328922	.0003936	83.58	0.000	.0321208	.0336635
Copenhagen	.0579626	.0022235	26.07	0.000	.0536046	.0623206
large_city	-.0013454	.0022312	-0.60	0.547	-.0057184	.0030276
small_city	-.0012054	.0023623	-0.51	0.610	-.0058355	.0034246
lnemployee	.0224169	.0000823	272.33	0.000	.0222555	.0225782
share_female	-.2500856	.0006363	-393.03	0.000	-.2513327	-.2488385
share_40_65	-.0583845	.0008979	-65.02	0.000	-.0601443	-.0566246
share_highskill	.2578079	.0019932	129.34	0.000	.2539013	.2617145
share_mskill	.0660427	.0005306	124.48	0.000	.0650028	.0670826
share_UM	.1767504	.0013347	132.43	0.000	.1741345	.1793663
_cons	5.021913	.003357	1495.95	0.000	5.015334	5.028493
q90						
share_of_lowskill	-.1467701	.0196576	-7.47	0.000	-.1852983	-.1082419
share_of_mskill	-.2345469	.0073622	-31.86	0.000	-.2489765	-.2201173
share_of_highskill	.539768	.024221	22.29	0.000	.4922957	.5872403
exp	.0212045	.0001407	150.67	0.000	.0209287	.0214804

exp2	-.0006171	3.99e-06	-154.73	0.000	-.0006249	-.0006093
Tenure	-.0087332	.0001225	-71.29	0.000	-.0089733	-.0084931
ten2	.0002313	4.71e-06	49.12	0.000	.0002221	.0002405
Age_18_24	-.2848154	.0033162	-85.89	0.000	-.2913151	-.2783158
Age_25_29	-.163971	.0033917	-48.34	0.000	-.1706186	-.1573233
Age_30_39	-.0236489	.0031835	-7.43	0.000	-.0298885	-.0174093
Age_40_49	.0863606	.0035673	24.21	0.000	.0793687	.0933524
Age_50_59	.0493791	.0036882	13.39	0.000	.0421503	.0566078
Age_60_65	.0917998	.0036621	25.07	0.000	.0846223	.0989773
basic_edu	.0163117	.0015591	10.46	0.000	.0132559	.0193675
vocational_edu	.0601777	.0014567	41.31	0.000	.0573226	.0630327
higher_edu	.2133431	.0016847	126.64	0.000	.2100412	.216645
Married	.0878236	.0007397	118.72	0.000	.0863738	.0892735
children_0_6_yrs	.0432508	.0008927	48.45	0.000	.0415012	.0450004
Copenhagen	.0498544	.0046574	10.70	0.000	.040726	.0589828
large_city	-.0013606	.005035	-0.27	0.787	-.0112291	.0085079
small_city	-.0012803	.0050367	-0.25	0.799	-.0111521	.0085915
lnemployee	.0230675	.000161	143.26	0.000	.0227519	.023383
share_female	-.2843541	.0010218	-278.29	0.000	-.2863568	-.2823515
share_40_65	-.1741492	.0016313	-106.75	0.000	-.1773466	-.1709518
share_highskill	.2732949	.0033776	80.91	0.000	.2666749	.279915
share_mskill	.0775266	.0015765	49.18	0.000	.0744367	.0806164
share_UM	.2387495	.0021031	113.52	0.000	.2346275	.2428715
_cons	5.359812	.0058353	918.52	0.000	5.348375	5.371249

(est61 stored)

```
.
. test[q10=q25=q50=q75=q90]: share_of_lowskill share_of_mskill share_of_highskill

( 1)  [q10]share_of_lowskill - [q25]share_of_lowskill = 0
( 2)  [q10]share_of_mskill - [q25]share_of_mskill = 0
( 3)  [q10]share_of_highskill - [q25]share_of_highskill = 0
( 4)  [q10]share_of_lowskill - [q50]share_of_lowskill = 0
( 5)  [q10]share_of_mskill - [q50]share_of_mskill = 0
```

```

( 6) [q10]share_of_highskill - [q50]share_of_highskill = 0
( 7) [q10]share_of_lowskill - [q75]share_of_lowskill = 0
( 8) [q10]share_of_mskill - [q75]share_of_mskill = 0
( 9) [q10]share_of_highskill - [q75]share_of_highskill = 0
(10) [q10]share_of_lowskill - [q90]share_of_lowskill = 0
(11) [q10]share_of_mskill - [q90]share_of_mskill = 0
(12) [q10]share_of_highskill - [q90]share_of_highskill = 0

```

```

F( 12,8216468) = 797.53
Prob > F = 0.0000

```

Chapter 3

ANALYSIS ON NATIVE-IMMIGRANT WAGE GAP IN DENMARK

Author

Shihan Du¹

¹Copenhagen Business School, Department of Economics,

Porcelaenshaven 16A, 2000 Frederiksberg, Denmark, E-mail:

sd.eco@cbs.dk

Abstract: This paper analyses native-immigrant wage gap as well as potential discrimination against immigrants for male workers in the labour market in Denmark. I perform an Oaxaca-Blinder and a Melly (2005) decomposition using Danish register data covering the period between 2004 to 2013. With particular interest on the impact of EU enlargement and free movement in the labour market on the wage gap, I focus my empirical analysis on the years 2004, 2007, 2009, 2010 and 2013. I find that low-skilled immigrants experience smaller wage gap and potential discrimination than medium and high-skilled groups. Moreover, after controlling for education, a substantial part of this gap can be explained by coefficient effects which I entirely regard as potential discrimination in this paper. By comparing across different ethnic groups within immigrants, I also find strong empirical evidence showing that the more distant² the home country of the immigrant is from Denmark, the more wage penalty is present. Besides, decomposition methods using quantile regressions indicate larger discrimination at the upper quantiles of the wage distribution in most cases.

Keywords: Native-Immigrant wage gap; Oaxaca-Blinder; Decomposition on conditional quantiles; Discrimination

² In this paper, I measure the distant according to cultural and economic backgrounds.

3.1 Introduction

Immigrants usually earn lower wages than natives in most developed countries. One could come to a quick conclusion that this may be the result of discrimination. In fact, in the wage inequality literature, the wage gap is explained by various factors, including personal and productivity-related characteristics, macroeconomic effects, and policy effects; of course, the influence of discrimination cannot be ignored. Studies examining discrimination in relation to the wage gap have long been an essential part of inequality research. Numerous theoretical and empirical studies have investigated the native-immigrant wage gap for decades. Furthermore, immigrants, in most cases, are willing to enter a country where they can receive equal treatment; hence, the inequality issue is also essential for the policymaking process to some degree. Not surprisingly, although they only account for a minor share of the population in most countries, immigrants have attracted increasing attention within both academia and politics. Moreover, the issue of discrimination has always been hotly debated and frequently studied.

With rich administrative data and a rapidly growing immigrant population, Denmark is a compelling case to investigate. As depicted in Figure 3.1, the number of immigrants entering the labour market in Denmark began to increase dramatically in 2004, when EU enlargement took place. Then, in 2009, Denmark opened its labour market to Eastern European countries. Denmark has a small population but a relatively fast increasing population of immigrants. Concerns regarding wage inequality for immigrants have long been a focal point since immigrants have a relatively significant impact on the Danish labour market and since Denmark is actively working to attract high-skilled immigrants. With a large inflow of immigrants, Denmark is a case worthy of further analysis in terms of how the native-immigrant wage gap differs depending on skill level and nationality and whether potential discrimination plays a role in the wage gap.

In this paper, I focus on the following research questions: First, how native-immigrant wage gap in the labour market in Denmark changes from 2004 to 2013, especially after EU enlargement in 2004 and 2007, and free movement took place in Denmark in 2009. Namely, the year 2004, 2007, 2009, 2010 (one year after free movement), and 2013. Second, whether potential wage discrimination³ against immigrants exists during the study period. Finally, how the native-immigrant wage gap and potential discrimination described in the first and second question vary within different skill levels and under each group of the countries of origins.

This paper provides an overall summary of changes in the population distributions and wage distributions of the native and immigrant groups, between 2004 and 2013. The findings from the empirical analysis contribute to the wage inequality studies on the labour market in Denmark for the period after EU enlargement. Moreover, comprehensive insights into native-immigrant wage gap and potential discrimination against immigrants, at different skill levels as well as within different ethnic groups are obtained. This paper is organised as follows: Section 3.2 discusses studies making a broad sweep across Europe and other regions and provides background information on the immigrant-native wage gap. Section 3.3 gives the motivation and proposes the hypothesis for this paper. Section 3.4 outlines the econometric methodology, and Section 3.5 describes the Danish register data used. Then, Section 3.6 presents and discusses the results, and finally, Section 3.7 concludes the paper.

³When I use the term "discrimination" in this paper, I refer to the remaining component after I have controlled for all possible factors in the wage model I can obtain. This term is not the same as the general political definition in practical.

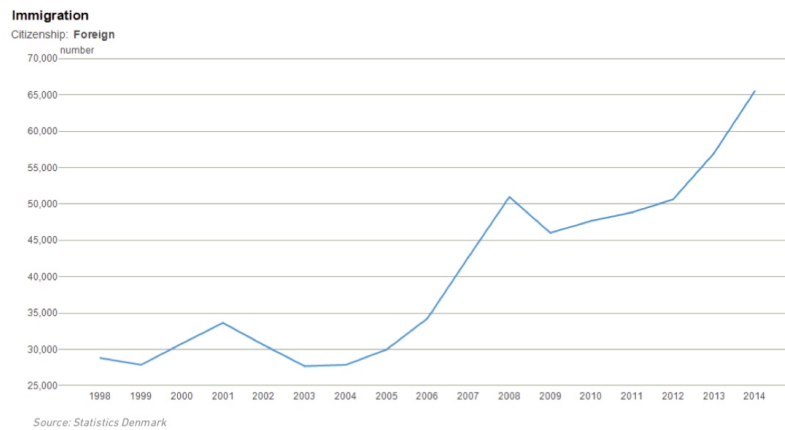


Figure 3.1: Increasing immigration in Denmark

3.2 Literature and background

In this section, I start by looking into the relevant literature to acquire a basic understanding of the wage inequality issue in different countries. I then present background information on changes in the population, skill structure, and immigrant wage distribution in the Danish labour market to gain an overview and to shed light on the motivation for the subsequent empirical analysis.

3.2.1 Literature

Various discrimination models, such as the taste-for-discrimination and statistical discrimination models, have been frequently discussed in the inequality literature. The taste-based discrimination model was proposed by Gary Becker in 1957. Becker modelled discrimination as a personal prejudice, or taste, against associating with a particular group. When wages and employment are determined by employers' prejudices regarding, for example, nationality, race, gender, and religion, discrimination issues arise in the labour market. As a result of the "taste" of employers, wages and employment rates could vary for the same job position.

In the statistical discrimination model, employers' decisions on whom to employ, promote, and reward are based on imperfect information on individuals' ability and productivity, and this can generate inequality. Kenneth Arrow has defined discrimination as "the valuation in the marketplace of personal characteristics of the worker that is unrelated to worker productivity." (Arrow, 1973). Later, Altonji and Blank (1999) provided an overview of this mechanism. For example, according to the theory, a candidate's foreign education background may be undervalued by the employer. In this case, an employer could treat candidates with similar education differently (e.g., divergent decisions on wages or recruitment for native and immigrant workers).

Compared with employment differentials, wage differentials are a more fundamental measure of labour market discrimination (Cain, 1986). Not surprisingly, income inequality has long been one of the critical topics in discrimination studies, in both theoretical and empirical research. Wage discrimination against immigrants occurs when there is unequal treatment in terms of wages which penalises immigrants in comparison to natives. The literature on income inequality is vast. However, in this paper, my focus is limited to empirical studies considering immigrants. In recent years, extensive empirical studies have examined income inequality among immigrants in many countries. There is overwhelming evidence of the existence of an immigrant-native wage gap, and various empirical studies have analysed that differential.

Empirical analyses of immigrant-native wage differentials started early. One classical work on this topic is the seminal 1978 paper by Chiswick. His paper described a wage penalty of 3% for immigrants in the U.S. as compared with comparable native-born individuals, even after controlling for socio-economic characteristics. Other studies based on U.S. data have concluded that immigration has a minor effect on the native-immigrant wage gap at the national level (e.g., Ottaviano and Peri, 2008 and Card, 2009).

When it comes to empirical research on European countries, Kee (1995) conducted an early empirical study to gain evidence on the native-immigrant wage gap in the Netherlands. The

microdata for natives and immigrants came from the 1985 Labour Supply Panel of the Organisation for Strategic Labour Market Research and the 1984 Quality of Life Surveys, respectively. The results illustrated that discrimination existed against Antilleans and Turkish male immigrant groups, but no sign of discrimination against Surinamese and Moroccans was found. The Oaxaca decomposition method was applied to decompose the wage gap. Education and experience were found to be main contributors to the gap.

Studies in the United Kingdom have also reported that immigrants suffer a wage disadvantage compared to their native-born counterparts. Earlier studies reported that the wages of white immigrants and white natives were almost the same, while non-white immigrants suffered from a wage penalty (e.g. Chiswick, 1980 and Bell, 1997). Later, empirical evidence from Miranda and Zhu (2013) suggested a causal effect of English as additional language (EAL) on the native-immigrant wage gap for male employees. A composition-adjusted male native-immigrant wage gap of 12% was found, and it was slightly below the raw wage differential. However, after controlling for the EAL indicator, the gap virtually disappeared. They also applied the IV strategy and demonstrated that a 23% causal negative effect on wages was attributable to the EAL variable. This study potentially indicates that instead of discrimination explaining all differentials, the language barrier could be an important factor contributing to the native-immigrant wage gap.

Having experienced a large inflow of immigrants, and benefiting from a rich large-scale administrative dataset, Germany has provided relatively more fruitful empirical evidence on income inequality for immigrants. Veiling (1995) analysed the native-immigrant wage gap for male workers of different nationality groups using a 1% sample of IAB data for 1989. He first applied a tobit estimation procedure to the wage equation to measure wage discrimination itself and then carried out the Oaxaca-Blinder decomposition. The results illustrated that different endowments of human capital explained most of the wage differentials between natives and foreign nationality groups.

Lang (2000) used empirical data from the German Socio-Economic Panel for 1997 to analyse the income disparity among ethnic groups in West Germany. Lang found that more than 75% of the wage differential between natives and foreign nationalities could be explained by the human capital gap and that statistically significant discrimination could only be measured in the groups of ethnic Germans from Eastern Europe. Other studies adopting a similar approach have also reported that immigrants suffer from a wage disadvantage in Germany (Constant and Massey, 2005 and Büchel and Frick, 2005).

In a later empirical work by (Lehmer and Ludsteck, 2011), the wages of foreign and native male employees were compared based on employment register data. The foreign workers came from various Eastern and Western European states and entered the German labour market from 1995 to 2000. Through Oaxaca/Blinder type decompositions, the researchers found that Poles had the highest wage differentials, but in general, the wage gaps between the targeted groups were relatively small. Then, via quantile decompositions, they found that coefficient effects tended to be larger at the bottom of the wage distribution.

Hirsch and Jahn (2012) stated that observable productivity characteristics (including occupation) could explain approximately 14 to 17 percentage points of the wage differential, after controlling for occupation. Then, from the view of unobservable, Nanos and Schluter (2014) attempted to identify factors responsible for wage differentials between migrants and natives based on differences in search friction, reservation wages, productivity, and other similar factors. They examined the native-migrant wage gap by applying a structural empirical general equilibrium search model with on-the-job search, using a large-scale German administrative dataset covering the period from 1975 to 2004. Counterfactual decomposition was then performed to analyse the native-migrant wage differentials. Migrant effects of different occupation groups were a particular focus. On average, the migrant effect accounts for 19.6% of the wage gap, while the migrant effect for skilled workers contributed 12-15%, and for clerks and service workers, the effect was 23-39% of the average.

As richer datasets become available, more comparative empirical evidence is emerging, as are data on other European countries previously lacking empirical findings on native-immigrant wage differentials. Kahanec and Zaiceva (2009) were the first to comparatively examine the roles of foreign origin and citizenship in the EU8 and EU15 labour markets; they used EU survey data from 2005. Ordinary least squares and Oaxaca-Blinder decomposition techniques using the Neumark (1988) method were applied to measure the native/non-native labour market divide. The researchers found that foreign origin could have a negative impact on employment for both genders in the EU15, and males of foreign origin also suffered from an earnings penalty in the EU15. In the EU8, being both of foreign origin and foreign-born negatively affected employment and wages for both genders.

Berman and Aste (2016) quantitatively examined the effect of immigration on wealth and income inequality through a pure statistical framework. By using empirical data from the U.S. (1986-2012), the UK (1990-2004), and Australia (1995-2005), which all witnessed a large inflow of immigrants into the labour market, they concluded that there is a relatively small direct effect of immigration on inequality in those destination countries where immigrants only account for a small proportion of the total population.

Chletsos and Roupakias (2017) applied the Blinder-Oaxaca methodology and found that the wage gap between natives and immigrants from non-EU countries is greater than that between natives and EU foreigners in Greece. Data from the 2009 Greek Labour Force Survey were used for the analysis. They also found evidence that post-migration human capital is a crucial wage determinant, particularly for non-EU immigrants. Moreover, post-migration experience is more important and offsets the effect of experience obtained in the country of origin. In addition, for those who terminate education in Greece, labour market experience is what makes the greatest difference.

Hofer et al. (2017) provided empirical evidence of the existence of a native-immigrant wage

gap in Austria. They used combined information from labour force surveys and administrative social security data and found a wage penalty of 15 percentage points for immigrants relative to natives. Although the Oaxaca-Blinder decomposition indicated that much of the wage gap could be explained by differences in human capital endowments and job positions, decomposition methods using quantile regression suggest greater discrimination against immigrants in the upper quantile of the wage distribution.

Taking advantage of the rich administrative dataset for Luxembourg, Sologon and Kerm (2018) examined trends in income inequality from 1988 to 2009. Only a small overall increase in income inequality was found, and the study listed three factors contributing to that growth: decreased persistent inequality in native groups, a lower employment share of natives, and a rapidly growing employment rate for cross-border workers.

In Denmark, Nielsen et al. (2004) analysed wage differentials through immigrant wage gaps and gender wage gaps within immigrant groups in Denmark for the period from 1984 to 1995, and they found a higher wage penalty for male foreigners. Later, based on Danish register data for the period from 1995 to 2002, empirical evidence from Nielsen (2011) indicated the probable existence of statistical discrimination against immigrants in Denmark. He found that foreign-educated immigrants were more prone to over-education than either native Danes or immigrants educated in Denmark. Over-educated workers earned slightly more than their co-workers but less than they would have earned given their level of education, and foreign-educated immigrants benefit the least from their over-education. This evidence all points to a classic case of statistical discrimination; as mentioned earlier in this section, an employer could devalue a candidate's foreign education due to imperfect information and thus offer a lower wage than the worker could have earned for this competency.

In summary, native-immigrant wage gaps have been observed in many studies. In empirical works, wage discrimination can be expressed as the unexplained part of wage differentials.

The explainable portion is often attributable to education, training, and other observed productivity characteristics. In this paper, I adopt a narrower definition of discrimination: Holding job characteristics such as industry constant, only unexplainable wage differentials between natives and immigrants working in the same industry are defined as discrimination. The question of how workers obtain their job offers is not considered by this definition.

3.2.2 Background

Despite much empirical evidence on native-immigrant wage differentials, the question of how discrimination against immigrants contributes to that wage gap does not have a clear answer, as seen from the literature mentioned above. There is no doubt that different countries could have unique immigration environments, and this fact could lead to a wide range of empirical evidence. Therefore, it is reasonable to zoom my study in on a single country, given that comparable cross-country datasets are inaccessible or unavailable.

With rich administrative data and rapid immigration growth, Denmark is an interesting case to investigate. As demonstrated in the introduction, the number of immigrants entering the labour market in Denmark increased dramatically starting in 2004, when EU enlargement took place. Immigration fell after 2008, probably due to the financial crisis. The number of immigrants started to grow again after 2009. In particular, 2009 was also the year when Denmark opened its labour market to Eastern European countries.

Hence it is interesting to look into the increase of immigrant population within different groups of countries of origins. In this paper, I followed the classification used by Malchow-Møller et al.(2012) and group immigrants into four categories by country of origin: (1) classical EU-15 countries, plus Norway and Iceland; (2) new EU-12 countries (those that joined the EU in 2004 and 2007); (3) remaining developed countries (according to UN classification); and (4) developing countries, including unknown countries of origin. As illustrated in Figure 3.2, since 2007, immigration from all ethnic groups has grown, with immigration from less developed

countries increasing the most. Meanwhile, immigration from new EU-12 countries has grown at a relatively rapid rate, which indicates that free movement has attracted more immigrants to the Danish labour market.

Moreover, there was also a spike in the number of immigrants from EU15 and less developed countries from 2007 to 2008. As OECD reports on International Migration-Denmark (2015) confirm, the significant increase in immigrant numbers during this period was mostly a consequence of new legislation regarding residence permits for work or study in Denmark (introduction of green card scheme in 2007). According to the report, the substantial increase in long-term immigration in 2008 was due to a rise in the number of foreign workers, and particularly workers from EU countries. I further divide the EU15 group into two groups: the southern EU15, which contains Italy, Spain, Greece, and Portugal, and the remaining EU15 countries. The right-hand part of Figure 3.2 clarifies that the increased immigration from the EU15 group was mostly attributable to arrivals from Germany, the UK, and France coming for working purposes.

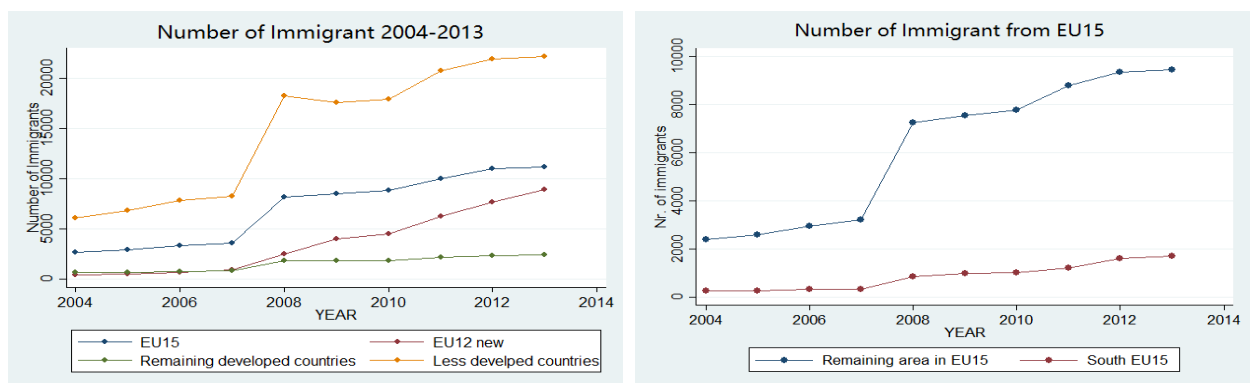


Figure 3.2

Changes in population distribution

The next step is to check further how the immigrant population was distributed within each skill level. I classify skill levels according to educational background, and the classification

scheme is described in detail in Section 3.5. In Figure 3.3, the share of immigrants over total immigration under each skill levels and within different ethnic groups during 2004 to 2013 is presented. As I can conclude from the comparisons, before 2007, immigrants from new EU12 countries were the smallest group of origin. The share of new EU12 immigrants increased the fastest, especially in the low-skilled group. New EU12 immigration even exceeded EU15 immigration, with the former becoming the second largest group of origin in the low-skilled immigrant category. The share of EU12 immigrants did not increase obviously in the medium-skilled and high-skilled groups, but it did exceed the share of immigrants from remaining developed countries after 2010.

Other groups of origin experienced less significant changes within the different skill levels, except for immigrants from less developed countries. This group represents a major immigrant population, but its share of the low-skilled group has been decreasing rapidly. This could be the result of the rapid growth of new EU12 immigrants in the low-skilled group. Moreover, the share of immigrants from less developed countries within the high-skilled group started increasing after 2005. Together with EU15 immigrants, they form a major portion of the high-skilled population. Since 2007, immigrants from less developed countries become the largest group of high-skilled workers. This could be an indication of the success of immigration policies aiming to attract high-skilled workers. Therefore, I further consider the skill structure within each group of origin to paint a more detailed picture.

To compare the skill structure of these groups, I move to Figure 3.4. Overall, EU15 immigrants and arrivals from remaining developed countries share a similar skill structure and have experienced comparable trends. Furthermore, unlike the relatively steady share of native workers and total immigrant workers in the medium-skilled group, the share of medium-skilled immigrant workers from EU15, new EU12 and remaining developed countries decreased, while it increased within immigrants from the group of less developed countries. Specifically focusing on the new EU12 group, the share of medium-skilled workers decreased significantly. More-

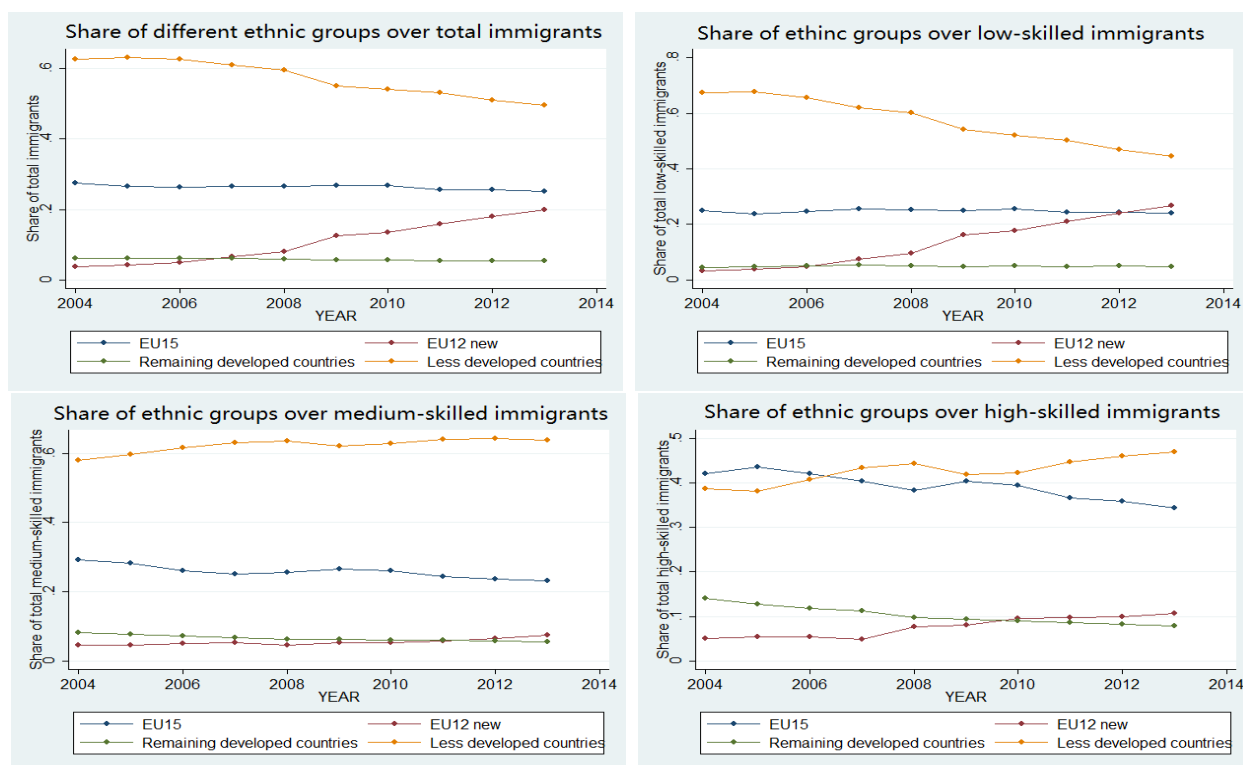


Figure 3.3

over, for all groups, low-skilled individuals comprised the largest population segment since 2004. Regarding the low-skilled population, although accounted for the largest population segment, all immigrant groups saw difference increases/decreases, and certain cases stand out. Again, in the new EU12 group, the share of low-skilled workers increased dramatically among the immigrant group of new EU12 countries. In EU15 and remaining developed countries groups, there were sharp increases from 2004 to 2007. In addition, different from all other groups of origin, there was a drop in the share of low-skilled immigrant workers from less developed countries. High-skilled workers were the smallest segment for all immigrant groups, but it showed increasing trends over the period, except the new EU12 immigrants' group.

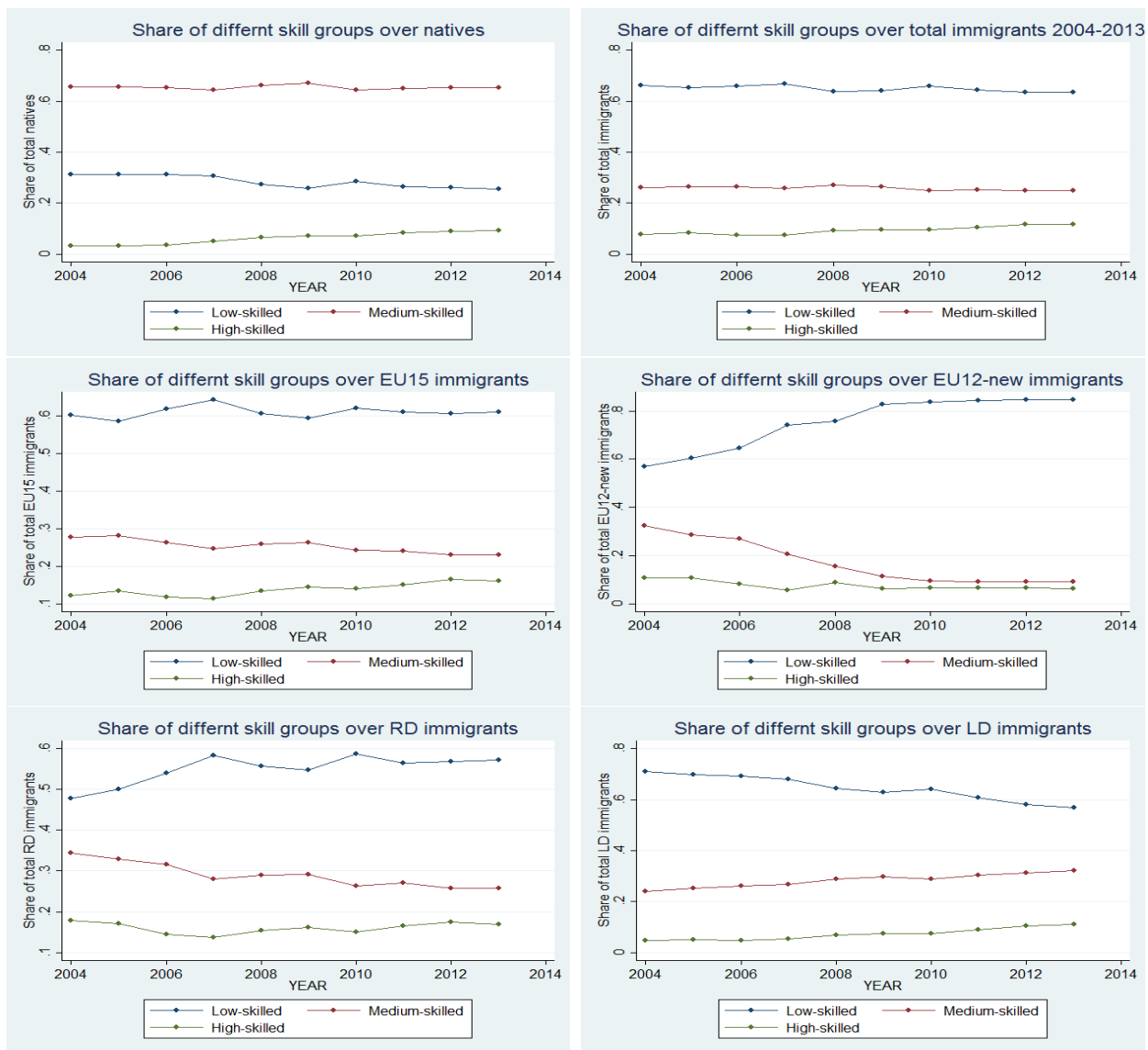


Figure 3.4

Change in wage distribution

As Figures 3.3 and 3.4 illustrate, the distribution of the population across different ethnic groups and different skill levels changed over the period of study. Therefore, I now examine whether the wage distribution shifted correspondingly. From Figure 3.5, I observe that the wage distribution for both natives and immigrants was flatter in 2013 than in 2004. Moreover, the native-immigrant wage gap decreased dramatically before 2007 and then started to increase after that point. The gap emerged in 2009 and reached a peak in 2010, which could be due to the change in the population distribution of the groups of origin or/and skill structures among immigrants. As depicted in the lower part of Figure 3.5, the wage gap for different wage quantiles appears significantly different. The gap between the two tails of the wage distribution is larger than the medium-low and high-medium quantile wage difference for both native and immigrant groups. Therefore, the native-immigrant wage gap can vary depending on the wage quantile and is worthy of further examination.

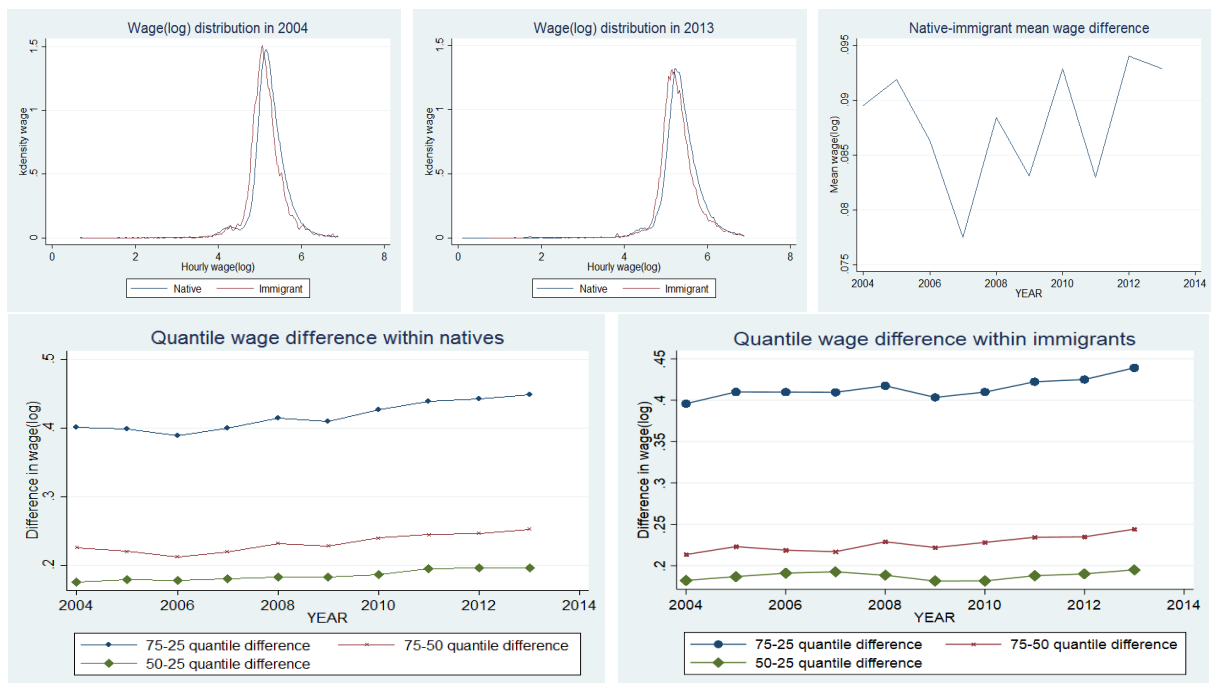


Figure 3.5

3.3 Motivation and hypothesis

3.3.1 Motivation

With the large inflow of immigrants representing different ethnic groups and skill levels, it is worthwhile to analyse the native-immigrant gap for different groups of origin and different skill levels for the Danish labour market from 2004 to 2013. Years that saw major changes, such as 2009 and 2010, when the free movement was introduced, are of particular interest.

Several studies have empirically analysed income inequality in Denmark (e.g. Nielsen et al., 2004 and Nielsen, 2011); However, except for Nielsen et al. (2004), empirical evidence on wage differentials by migration status is very scarce, not to mention empirical evidence for the period after EU enlargement. This research gap provided an incentive to conduct such an empirical study for the period from 2004 to 2013.

Furthermore, the substantial increase in immigration could have had an impact on the skill structure in the local labour market in Denmark. As discussed in the literature (e.g., Nanos and Schluter, 2014), the effect of immigration on native-immigrant wage differentials can vary depending on the skill group. This has also been empirically confirmed by Liu et al. (2004). They performed an empirical analysis on the intra- and inter-occupational wage gap between immigrants and the natives based on Hong Kong 1996 census data, and they found different occupation distributions between immigrants and the native-born. These arguments strongly motivate a more inside look into the native-immigrant wage gap within different skilled groups in Denmark.

Therefore, in this paper, I focus on the following questions on native-immigrant wage gap and whether potential wage discrimination against immigrants existed:

1. How native-immigrant wage gap in the labour market in Denmark changes from 2004 to 2013, especially after EU enlargement in 2004 and 2007, and free movement took place in Denmark in 2009 (namely, the year of 2004, 2007, 2009, 2010 (one year after free movement), and 2013).
2. Whether potential wage discrimination⁴ against immigrants exists during the study period, if it exists, how it changes.
3. How the native-immigrant wage gap and potential discrimination described in the first and second question vary within different skill levels and under each group of the countries of origins.

Following the empirical strategies adopted by many similar studies on native-immigrant wage gaps, I apply decomposition techniques in this paper to analyse the native-immigrant wage gap for various skill levels. Moreover, in order to examine the degree of discrimination against different nationalities, I also focus on particular groups of origin, such as new EU12 countries and less developed countries. Given that the wage gap differs by wage quantile, as demonstrated in Figure 3.5, a decomposition on the quantile regressions is also performed to gain a fuller picture of the whole wage distribution. To be more specific, an Oaxaca-Blinder decomposition is used for the analysis of the mean wage gap; then, decomposition based on conditional quantiles is performed. The decomposition framework is presented in the later methodology section (3.4).

⁴When I use the term "discrimination" in this paper, I refer to the remaining component after I have controlled for all possible factors in the wage model I can obtain. This term is not the same as the general political definition in practical.

The findings from this study contribute to the wage inequality literature on the labour market in Denmark, mainly it addresses the research gap of the relevant studies on the period after EU enlargement. Further, empirical evidence obtained provides comprehensive insights into native-immigrant wage gap and potential discrimination against immigrants, at different skill levels as well as within different ethnic groups are obtained.

3.3.2 Hypothesis

According to the research questions of how native-immigrant wage gap and potential wage discrimination against immigrants varies under different skill levels and within different ethnic groups, I proposed the two hypothesis (H1 and H2), for the situation under different skill levels and different ethnic groups separately. Both hypotheses are based on theoretical models of statistical discrimination and taste-based discrimination (discussed in Section 3.2)

Before a detailed discussion on the hypothesis, I firstly start with all immigrants of all skill levels, the theory of demand and supply suggests that the increased number of immigrants and their relatively low bargaining power should result in immigrants earning lower wages than natives. While the conditions for the native group remain stable, the native-immigrant wage gap could grow. Regarding the discrimination issue, both taste-based and statistical discrimination could occur in this case. With more immigrants arriving, employers could grow more prejudiced against the immigrant population or have imperfect information on immigrant workers' qualifications; these outcomes would result in discrimination positively contributing to the wage gap. Therefore, the hypotheses based on my research questions are as follows:

H1: The higher the skill level is, the larger the wage gap and potential statistical discrimination could present.

In this paper, firstly, I consider the skill levels individually. For low-skilled immigrants, qualifications and productivity-related ability are not a strong requirement for low-skilled work. Hence

the native-immigrant wage gap could be small and mainly due to discrimination. With more low-skilled immigrants arriving in the country, competition might deepen among low-skilled immigrants, which could result in lower wages. However, natives' wages could also be affected, so the direction of change for the wage gap between low-skilled natives and immigrants is unclear. However, as taste-based discrimination increases, the wage gap will increase.

Then, for the medium-skilled group, which is the largest category for the native working population and accounts for a vital proportion of immigrant, statistical discrimination can play a role in terms of the native-immigrant wage gap. Productivity-related characteristics play a more significant role in wage setting for medium-skilled jobs than for low-skill jobs. When an employer devalues the overseas education background of an immigrant worker, his or her wage could be lower than that of a native with similar capabilities and comparable education background. Given that the share of medium-skilled immigrants remained relatively steady though slightly from 2004 to 2013, the wage gap might also be stable due to the unchanged discrimination level. However, due to assimilation effects, the wage gap might decrease. Moreover, a demand-supply-type impact could also play a role in the wage gap.

For the high-skilled group, statistical discrimination might be greater; hence, the high-skilled native-immigrant wage gap might be affected by this factor to a greater degree than for the other two skill groups. Wages are mainly determined based on productivity for high-skilled work; thus, the wage gap might be larger due to employers' imperfect information on foreign candidates' education or working background. With a large inflow of high-skilled immigrants, holding the discrimination level unchanged, the wage gap might increase according to the supply-and-demand theory. While Denmark has implemented various immigration policies to attract high-skilled immigrants, the bargaining power of high-skilled immigrants is relatively stronger than that of immigrants from the other two skill level categories; hence, the policy can play an important role in decreasing the wage gap, especially when the share of high-skilled immigrants falls.

H2: The more distant⁵ the immigrant's home country from Denmark, the more positive the discrimination.

In addition to my focus on different skill levels, I also compare the wage gap between natives and different immigrant groups of origin to examine the discrimination level faced by different nationalities. After 2009, the large inflow of low-skilled immigrants from new EU12 countries could have caused an increased wage gap and more discrimination. As compared in Figures 3.3 and 3.4, the EU15 and remaining developed countries groups shared similar skill structures, with these populations consisting of a higher proportion of high-skilled immigrant than seen in the native group. Therefore, the mean wage of immigrants from EU15 countries and remaining developed countries could be higher than those of native workers; in this scenario, the native-immigrant group would be negative. Lastly, the wage gap between natives and immigrants from less developed countries could fall due to the increase in the high-skilled population share. For all groups of origin, statistical discrimination could exist and contribute to the wage gap. An employer could undervalue qualifications acquired in less developed countries while overvaluing the education background of those from other areas with, for example, more well-known universities. Differences in terms of cultural background also affect the level of statistical discrimination.

To conclude, the hypotheses on the size of the wage gap between natives and different ethnic groups are based mainly on my principal investigation of wage differentials for various skill levels. The unique skill structure of each group of origin could explain the corresponding native-immigrant wage gap. According to the theoretical definition of statistical discrimination, larger cultural differences could cause more statistical discrimination. I tested these hypotheses by applying decomposition techniques to my empirical procedures using the Danish dataset. I present the decomposition methodologies in the following section; then I describe my dataset for the empirical analysis in Section 3.5.

⁵In this paper, "distant" is measured in the aspect of cultural and economic levels.

3.4 Decomposition methodologies

In general, discrimination in labour economics can be identified by comparing the estimated coefficients of wage equations. A likelihood-ratio test for the restriction of equal parameters for the native group and foreigners group was carried out for this purpose. However, if positive and negative discrimination were coinciding, the rejection of equal parameters would still not indicate that discrimination was at work. The typical approach used for studying earnings differentials between immigrants and natives is to estimate earnings equations for the two groups separately using OLS and to then decompose the mean differences between them (Blinder, 1973; Oaxaca, 1973; e.g. Miller and Chiswick, 1985). I employed such typical procedures and then decomposed the whole wage distribution to further investigate the wage gap at the top and bottom of the distribution.

The problem of lack of causal interpretability in conventional decompositions is discussed in relevant studies, such as Huber (2015). He argues that in both linear and nonparametric decompositions, this concern usually arises. The problem comes from the choice of group variables, such as gender or ethnicity which is determined at or even before birth. If the target component is explained by the observed characteristics that are determined before the group variable, the decompositions would bear a causal interpretation in some cases. In this study, I chose immigrant status as the group variable. Unlike gender and ethnicity, immigrant status is not determined before birth and can be decided. For this reason, the argument from Huber (2015) on causal interpretability might not apply in this paper.

Therefore, I carried out a decomposition of wage differentials into an endowment component and a discrimination component, following the approach in Oaxaca (1973) and Blinder (1973). Then I applied Melly (2005) decomposition approach based on quantile regressions to investigate the native-immigrant wage gap for the whole wage distribution.

3.4.1 Oaxaca-Blinder decomposition

I first employed the decomposition method developed by Oaxaca (1973) and Blinder (1973). This approach can be used to obtain the average wage differential between natives and immigrants in a productivity-related difference (Δ_x) and a so-called discrimination component (Δ_u).

Assume two disjoint groups, N and I . Formally, let $g \in N, I$, where g is a group indicator (e.g., 0 for natives and 1 for immigrants). Furthermore, let W_N be the wage of a native ($g = N$), and W_I be the wage of an immigrant ($g = I$). The starting point is the so-called Mincerian wage equation (Equation 3.1), which is separately estimated for each of the two groups (N and I).

$$W_g = X_g\beta_g + u_g \quad (3.1)$$

Where the log hourly wage (W_g) for group g is a linear function of a set of variables denoted by X . In turn, X is a vector of productivity-determining characteristics, while β is a vector of the OLS wage regression coefficients. Then, the average wage gap can be decomposed in the following way:

$$\overline{W_N} - \overline{W_I} = \overline{X_N}\hat{\beta}_N - \overline{X_I}\hat{\beta}_I \quad (3.2)$$

$$= (\overline{X_N} - \overline{X_I})\hat{\beta}_N + \overline{X_I}(\hat{\beta}_N - \hat{\beta}_I) \quad (3.3)$$

$$= (\overline{X_N} - \overline{X_I})\hat{\beta}_I + \overline{X_N}(\hat{\beta}_N - \hat{\beta}_I) \quad (3.4)$$

According to the Oaxaca-Blinder decomposition method, the income gap between labour groups is due to two factors: First, the different characteristics of human capital and employment units. Second, the labour market's discrimination against specific groups. The bar notation indicates that the log wage differential is evaluated for persons with the characteristics of the average wage earner. Equations 3.2-3.4 decompose the native-immigrant observed wage differential into a component resulting from differences in personal characteristics and into a component due to discrimination. The decomposition requires an estimate of the non-discriminatory wage structure. In this study, the native group comprised the non-discriminatory group. Thus Equation 3.3 was applied for the Oaxaca-Blinder decomposition

in the empirical applications. Therefore, I rewrote Equations 3.3 and obtained Equation 3.5:

$$\begin{aligned}\Delta_x &= (\overline{X_N} - \overline{X_I})\hat{\beta}_N \\ \Delta_u &= \overline{X_I}(\hat{\beta}_N - \hat{\beta}_I)\end{aligned}\tag{3.5}$$

Where Δ_x represents the share of the average wage gap due to the different endowments of the two groups with productivity-related characteristics, and Δ_u represents the unexplained residual. It covers all unobservable differences between natives and immigrants. The discrimination component is included in these unobserved effects. Δ_u may overstate actual discrimination if there are important unobserved qualification variables (for instance language proficiency), see, e.g. Altonji and Blank (1999). Due to the data availability, I use variables which captures part of the information as proxies for those unobserved factors, e.g. language. In this paper, the term "potential discrimination" refers to the unexplained component Δ_u . Besides, since the decomposition approach is based on the mean differences between native and immigrant groups which are calculated from the separate earning equations, the explanatory variables X are the same for both groups. The choice of explanatory variables X is further discussed in Section 3.5. Positive potential discrimination indicates that the immigrant should have received more with their current characteristics, while a negative sign of Δ_u shows an indication that immigrants are paid too much compared to their characteristics.

3.4.2 Counterfactual wage distributions

Linear regression models are only based on the regression on the mean of wages, and so as to the decomposition models based on OLS regressions, e.g. the Oaxaca-Blinder model. Rather than relying on the mean, empirical researchers have increasingly been focusing on the entire wage distribution and the associated decomposition procedures (see Fortin et al., 2011). The quantile regression methodology, proposed by Koenker and Bassett (1978), allows the characteristics of individuals to have different impacts at various points of the wage distribution. Consequently, these factors can have effects on the implied decompositions at each point, thus allowing an examination of impacts at the top and bottom of the distribution.

Methods of decomposition based on conditional quantiles have been proposed in studies such as Melly (2005), and Machado and Mata (2005). The approach by Melly (2005) overcomes the problem of crossing quantile curves, and discusses more efficient ways in computing distributions than the approach by Machado and Mata (2005) (Fortin et al., 2011). The method in Melly (2005) is also less restrictive than the JMP (Juhn et al., 1993). In Melly (2005), procedures are developed to convert a conditional distribution to an unconditional one, and this allows the characteristics to influence the whole conditional distribution of W . Therefore, in this paper, I follow the approach by Melly (2005).

Even though that it is much more difficult to decompose quantiles than the mean, the overall decomposition strategy for conditional quantiles for this study is similar to the idea used in Oaxaca-Blender approach. It is briefly presented as follows: Step 1: Obtain estimation results from conditional quantile regressions for both native and immigrant groups.

Step 2: Obtain estimation results from the counterfactual distributions. Procedures based on transforming a wage observation into a counterfactual observation is the object of interest in this step.

Step 3: Calculate and decompose wage differences by using results from step 1 and 2.

To start, the conditional quantiles, which later decomposition is based on, is defined in Eq. 3.6 :

$$F_{W_g|X_g}^{-1}(\tau|X_g) = X_g \cdot \beta_g(\tau) \quad (3.6)$$

where in my application, W represents the log hourly wage, X is a row vector of wage covariates (e.g. education level, experience, etc.), and, g is a group indicator (e.g. 0 for natives and 1 for immigrants). $F_{W_g|X_g}^{-1}$ is the τ th quantile of W_g which is conditional on X_g . Linearity is assumed between the quantiles of W_g and X_g , and this assumption can be relaxed by the use of dummy variables, polynomial expansions and interaction terms.

It is difficult to estimate an entire conditional distribution function for each value of $(W_g|X_g)$, because one needs to know the entire conditional distribution of $F_{W_g|X_g}$. Thus the conditional distribution cannot be directly used to decomposing the changes in wage distribution, and the unconditional quantiles of W_g are needed for decomposition procedures. Follow the approach in Melly (2005), the unconditional distribution is obtained by integrating the conditional distribution over the whole range of the distribution of the regressors. Let q_0 be the θ th quantile of W_g , $q_0 = F_{W_g}^{-1}(\theta)$, then

$$\theta = \int 1 \cdot (W_g \leq q_0) dF_{W_g}(W_g) \quad (3.7)$$

After changing the variable integrations in Eq. 3.7, one can obtain the following Eq. 3.8

$$\theta = \int \left(\int_0^1 1 \cdot (F_{W_g|X_g}^{-1}(\tau|X_g) \leq q_0) d\tau \right) dF_{X_g} \quad (3.8)$$

Define that $\{W_{gi}, X_{gi}\}_{i=1}^n$ is an independent sample, $\beta_g = \{\beta_g(\tau_1), \dots, \beta_g(\tau_J)\}$, and that τ_j is the equidistant grid point. Let the linear quantile regression be estimated at J equidistant grid points. Next, replace $F_{W_g|X_g}^{-1}(\tau_j|X_g)$ by its consistent estimate $X_{gi}\hat{\beta}_g(\tau_j)$, and the sample analogue of q_0 is then given by Eq. 3.9

$$\hat{q}(\hat{\beta}_g, X_g) = \inf \left\{ q : \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^J (\tau_j - \tau_{j-1}) \cdot 1 \cdot (X_{gi}\hat{\beta}_g(\tau_j) \leq q) \geq \theta \right\} \quad (3.9)$$

Therefore, \hat{q} is the θ th quantile of W_g . Then, step 2 can be performed to obtain counterfactual distribution for the decomposition of wage differences. Same as in mean-based decomposition procedures, here I also use native group as a reference group for quantile decomposition. Hence, the counterfactual distribution in my application is the distribution of W for the native group N if its distribution of characteristics would be as in the immigrant group I . I construct the counterfactual distribution based on Eq. 3.9

$$\hat{q}(\hat{\beta}_I, X_N) = \inf \left\{ q : \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^J (\tau_j - \tau_{j-1}) \cdot 1 \cdot (X_{Ni}\hat{\beta}_I(\tau_j) \leq q) \geq \theta \right\} \quad (3.10)$$

Then the counterfactual distribution can be obtained by computing $\hat{q}(\hat{\beta}_I, X_N)$ in Eq. 3.10 for all θ , and using the estimated coefficients for group I and observed X from group N . $\hat{q}(\hat{\beta}_N, X_I)$ is the θ th quantile of this counterfactual distribution of wages.

The total estimated differential can be finally decomposed as :

$$\begin{aligned} & \hat{q}(\hat{\beta}_N, X_N) - \hat{q}(\hat{\beta}_I, X_I) \\ &= \left[\hat{q}(\hat{\beta}_N, X_N) - \hat{q}(\hat{\beta}_I, X_N) \right] \\ & \quad + \left[\hat{q}(\hat{\beta}_I, X_N) - \hat{q}(\hat{\beta}_I, X_I) \right] \end{aligned} \quad (3.11)$$

Where X_N and X_I are vectors with the characteristics of native and immigrant separately, β_N and β_I are the estimated coefficients on characteristics. $\hat{q}(\hat{\beta}_N, X_I)$ is the counterfactual earning distribution of individuals with characteristics in the native group. In other words, what would have native earned if they were in the immigrant group. Eq. 3.11 allows us to decompose the difference between native and immigrant wages (the left-hand side of Eq. 3.11) into the two factors that appear in the two brackets of the right-hand side of Eq. 3.11. The first bracket captures the effect of coefficients, and the second captures the effect of characteristics (the explained component). The explained components explain the part of wage gap which is due to the different endowments of the two groups with productivity-related characteristics, and the effect of coefficients represents the unobserved effects which cannot be explained by the characteristics in X . The characteristics effect and the coefficient effect in Eq. 3.11 have similar indication as the terms Δ_x , and Δ_u discussed in Eq. 3.5. Moreover, the coefficient effect is regarded as "potential discrimination" in the later empirical analysis. Again, a positive coefficient effect indicates that the immigrant is paid less compared to their current characteristics, while a negative coefficient effect shows that immigrants received too much than they would have.

3.5 Data

My empirical analysis was based on administrative data from Statistics Denmark(DST). As the data were register based, measurement problems were negligible. I used individual data as well as linked employer-employee data from the IDA to perform analyses on the native-immigrant wage gap using Oaxaca-Blinder and Melly (2005) decomposition methods. As

already mentioned in Section 3.2, given the differences in female labour market participation in Europe, I only concentrated on full-time male workers aged between 18 and 64 years. I used the full population of both natives and immigrants for my empirical analysis. Benefiting from the rich DST data, my sample included 5,189,960 observations for male individuals of working age from 2004 to 2013. The sample contained 1,184,092 individuals in total, including 1,102,490 natives and 81,732 immigrant workers. The immigrant workers accounted for 6.9% of the total working population.

Regarding the immigrant-related identifiers, I used three primary ones in this paper: immigrant indicator, country of origin, and indicators for each ethnic group in the immigrant population I adopted for the empirical analysis. Firstly, I used the definition of the term immigrant given by DST: Immigrants are born abroad, and neither parent is either a Danish citizen or born in Denmark. If there was no information about either parent and the person had been born abroad, he or she was also regarded as an immigrant. Moreover, the country of origin of an immigrant was based on his parents' countries of birth: When none of the parents is known, the country of origin is defined based on the person's own information. If the person is immigrant, it is assumed that the country of origin is equal to the birth country. When only one of parents' information is known, the country of origin is defined based on the known parent. When both parents are known, the country of origin is defined based on the mother's country of birth or the country of citizenship respectively. Then, I followed the classification used by Malchow-Møller et al.(2012) and grouped the immigrants into four categories by country of origin: (1) classical EU-15 countries, plus Norway and Iceland; (2) new EU-12 countries (those that joined the EU in 2004 and 2007); (3) remaining developed countries (according to classification by the UN); and (4) developing countries, including unknown countries of origin. As illustrated in Figure 3.3 in Section 3.2, immigrants from less developed countries accounted for a major immigrant population.

The dependent variable of the framework is the hourly wage (log). I calculated the hourly

wage using the observation's total contribution divided by the total number of hours worked during the year; the variables used for the computation came from the linked IDA employer-employee data. Wages for all years were deflated by the average CPI of the year. The choice of independent variables for the wage decomposition was essential. In this study, I included the major variables chosen by similar studies. Productivity-based and qualification-related characteristics were selected to explain part of the wage gap. To be more specific, dummies of different age groups (e.g., 18-24 years, 25-29 years, and 30-39 years), education levels, marital statuses, city sizes, experience levels, union statuses, firm sizes, and industry sectors were chosen as explanatory variables. The summary statistics and sensitivity analysis on the selected variables are presented in the Appendix.

Education level was a key variable in my analysis. Education is an essential identifier to classify an employee's skill level. I classified education according to three levels: basic education (low skilled), vocational education (medium skilled), and higher education (high skilled) according to International Standard Classification of Education codes. However, the measurement error in the immigrant sample was relatively high compared to it was the native sample. Data on immigrants' education backgrounds prior to entering Denmark were obtained mainly from DST surveys. There are more missing data on the educational variable in the immigrant sample than they are in the native sample, but the problem is not severe in general, only 3.4% missing for immigrant sample over the whole period while the native sample contains 0.71% missing value on education variables. However, large missing data within the education variable were found in the immigrant sample in the year 2010, 27.45%, due to a change of the coding of the variable. Therefore, I performed correction procedures on the education variable in the year 2010 for the immigrant sample. I replaced the missing education data in 2010 with basic education if the person was basic education in 2009 or 2011, because all observations are full-time working population. As long as one is continuously observed in the dataset, the education status cannot be changed. After the correction approach done on the education variable, the missing value in the immigrant sample for all years reduced to 1.33%, and for

the year 2010, it reduced significantly, from 27.45% to 10.59%. The comparison table of the education variables before and after correction is included in the Appendix.

3.6 Empirical results and discussions

Based on the Danish data, I performed decomposition procedures to analyse the native-immigrant wage gap. In this section, I present the empirical results obtained using the decomposition approaches I discussed in Section 3.4: firstly, I performed an Oaxaca-Blinder decomposition to decompose the native-immigrant wage gap based on the mean for a baseline check, and then I focused on Melly's (2005) approach to examine the gap for all wage quantiles. The key analysis was based on the Melly approach. This section concludes by further discussing the results obtained from the above two approaches. To examine and compare the discrimination level for different skill levels and ethnic groups, for both decomposition approaches, I perform the decomposition on the whole sample, then on subsamples for the various skill levels, and finally, on subsamples combining natives and each group of origin.

Results are reported for selected years: 2004, 2007, 2009, 2010, and 2013. The choice of years is based on the discussion in Section 3.2. The EU enlargement took place in 2004 and 2007; then, in 2009, the labour market freely opened to Eastern EU workers. That year saw an enormous inflow of immigrants, especially immigrants from the low-skilled group. In 2010, one year after the free movement, the share of low-skilled immigrants reached a relatively high point. Finally, I was also interested in the results for the most recent year in my sample, 2013.

3.6.1 Oaxaca-Blinder decomposition

Before I performed the Oaxaca-Blinder procedures, according to the framework in Section 3.3, I firstly set natives as the reference group for the decomposition. Then, based on the Mincer

wage equation, I predicted the mean wages for different skill levels and ethnic groups; I then performed the decomposition for the corresponding samples. Full regression results as well as additional detailed results from the mean predictions for native and immigrants of different skill levels and each group of origin are in the tables in the appendix; the results are also organised by skill level for each group of origin for purposes of comparison. I concluded that the high-skilled group earned the most compared to all other skill groups. Moreover, to compare across groups of origin, immigrants from remaining developed countries had the highest wages. I then decomposed the native-immigrant wage gap under different skill groups for all groups of origin and the wage gap within each group of origin for all skill levels individually. In this paper, I focus on comparing the results from these two samples to check how the discrimination level differs according to the skill group and group of origin.

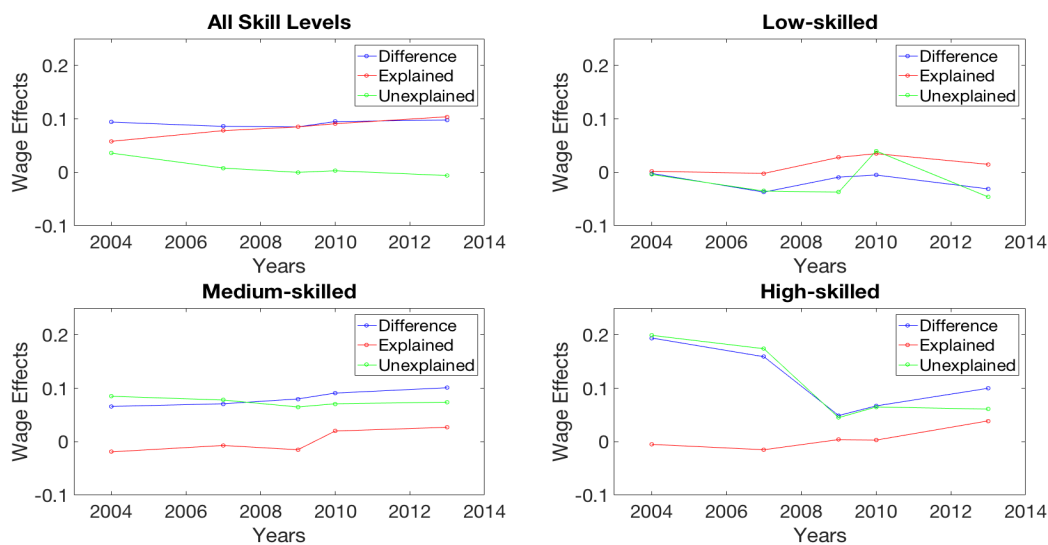


Figure 3.6: Wage gap decomposition under different skill levels

Figure 3.6 provides the decomposition results for all immigrant groups of different skill levels. Additional decomposition results for the various skill levels within each group of origin are presented in the appendix. As Figure 3.6 makes clear, for the sample of natives and all immigrants of all skill levels, the native-immigrant wage gap slightly increased beginning in

2004, while the unexplained effects, which I regarded as discrimination, decreased over the study period and even became negative (-0.006 log pp, and the estimates were significant) in 2013. The effects of characteristics were stronger than those of discrimination during the period.

Then I move to the separate skill groups. As seen in Figure 3.6, the low-skill group had a less clear trend in terms of wage differences over time, and the unexplained coefficient was insignificant in 2004. The large inflow of low-skilled immigrants could have significantly influenced the wage distribution for this immigrant group, and the Oaxaca approach was only the mean decomposition. However, I also investigated whole quantiles for this group later in the decomposition based on conditional quantiles. While one can observe from the comparison of the results for 2004 and 2007 that as the share of low-skilled immigrants from new EU12 countries increased rapidly, the native-immigrant wage gap became more negative, and the explanatory power of discrimination component was likewise increasing. The changes in the wage gap and discrimination effects during the period for the medium-skilled groups were clear. The gap increased, as did the discrimination effect meanwhile the share of immigrants from EU12 countries with that skill level declined significantly. Regarding the high-skilled, the size of both native-immigrant wage gap and unexplained component showed vibrated changes over the period, particularly from 2007 to 2009, decreased fast until 2009 and then started increasing slowly. If I compare across the skill groups, I see that the low-skilled group had the smallest wage gap in terms of absolute value. The high-skilled group had the largest wage gap and the highest discrimination level except in 2009; In 2009, the medium-skilled group had the greatest wage gap and highest discrimination level. Moreover, after the skill level was controlled, the explained component became very small. Therefore, the results also indicated that education, which determined the skill level in this paper, explained most of the characteristics effect.

Given that none of the total shares of natives and immigrants at each skill level changed dra-

matically since 2004, only the share of groups from different countries of origin within each skill level fluctuated during the whole period instead, I assume that country of origin plays a key factor on the changes; Hence this gives the reason to examine the native-immigrant wage gap further from the aspect of countries of origin. The results of the native-immigrant wage gap decomposition for different groups of origin for all skill levels are outlined in Figure 3.7. The EU15 group and the remaining developed countries had negative wage gaps and discrimination effects, and the effects of discrimination outweighed the effects of characteristics in all cases. Moreover, the wage gap showed smooth changes over the study period. In terms of the rapid growth in the new EU12 group, the wage gap was positive and increasing; it was larger than that for the EU15 group and the immigrants from remaining developed countries. The estimations for unexplained differences in all study years were significant, unlike the EU15 group and the remaining developed countries, only minor negative discrimination was found in some cases. Regarding the group from less developed countries, both a positive wage gap and positive discrimination were identified by the decomposition, although these values slightly decreased during the study period. Moreover, the wage gap and the discrimination level were the highest seen for any group of origin in most years except in the year 2009.

From the above results, I can conclude that as regards the different skill levels, the mean wage differentials for the low-skilled groups were the smallest differentials in most cases. Potential discrimination dominants in the medium- and high-skilled groups. This finding is consistent with my hypothesis presented in Section 3.3. In terms of groups of origin, I found that EU15 and remaining developed countries had negative mean wage gaps and discrimination, while new EU12 immigrants and the group from less developed countries presented a positive wage gap and discrimination in most cases. Except for the new EU12 group, discrimination effects were no less than the effects of characteristics in most cases. Instead of ignoring extreme cases within the distribution, which were not surprising given the substantial inflow of immigrants, I was also interested in examining the native-immigrant wage gap for the whole wage distribution. Therefore, I then performed decomposition based on conditional quantiles to gain

a complete picture of the wage gap and discrimination for different wage quantiles.

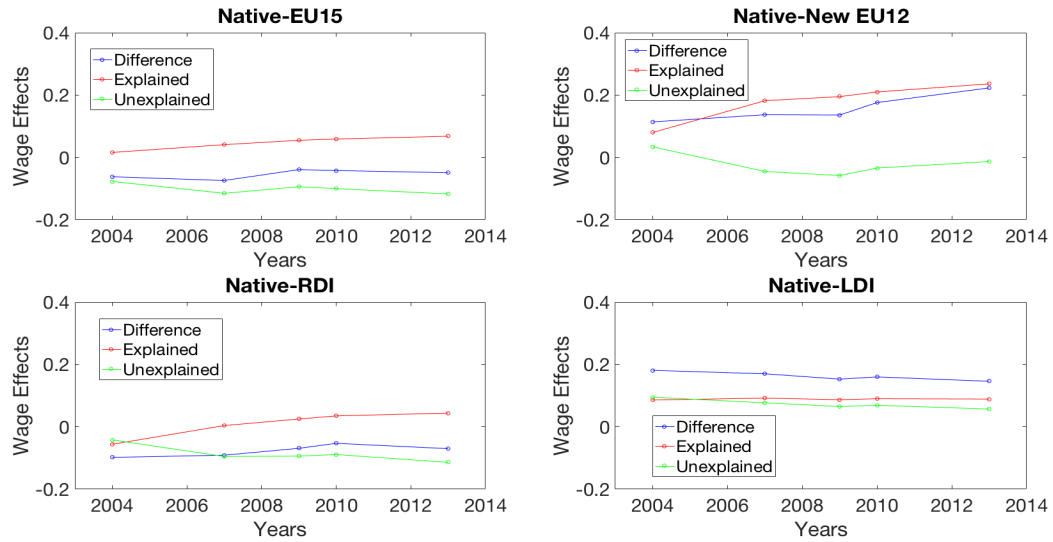


Figure 3.7: Wage gap decomposition under different ethnic groups

3.6.2 Melly decomposition

Moving beyond the analysis based on the mean, I now focus on the entire wage distribution. The decomposition approach proposed by Melly (2005) was applied for my empirical analysis. Similar to my approach to the Oaxaca-Blinder procedures, again, I first decomposed the native-immigrant wage gap for all immigrants to obtain an overall picture. Then, I decomposed the gap for different skill levels and groups of origin to gain a detailed picture of the wage gap. With decomposition based on conditional quantiles, I acquired further information on the wage gap at different wage quantiles. As a departure point, I checked the cumulative density function of wages. As illustrated in Figure 3.20 in the appendix, natives earned higher wages than immigrants in most wage quantiles from 2004 to 2013. I, therefore, defined natives as the non-discriminatory group for the decomposition, just as I did for the earlier Oaxaca-Blinder decomposition. I then approximated the conditional wage distribution with 100 quantile regressions (standard STATA procedures); statistical inferences were based

on bootstrapped standard errors. Fifty bootstrap replications (standard STATA procedures) were used to obtain an estimate of the variance of the estimators. Most estimations were significant. The bootstrapped results and figures presenting the discrimination effects with confidence intervals are presented in the appendix.

All groups of origin and all skill levels

I started the decomposition with the sample of all skill levels and all groups of origin. Figure 3.8 provides an overview of the native-immigrant wage gap and the extent to which the discrimination component contributed to the gap for all immigrants of all skill levels and groups of origin. Firstly, I focused on the middle quantile (the q20-q80 range, the middle section of the graphs). A clear trend was evident: The coefficient effects, which were regarded as potential discrimination in this study, increasingly contributed to the wage gap from 2004 to 2013. After 2007, the effects of discrimination even exceeded the effects of characteristics and dominated the wage effect thereafter. Unlike minor negative discrimination in some cases I observed from the Oaxaca-Blinder decomposition, the discrimination level was positive for the middle range of the quantile. Since 2009, the effects of discrimination even exceeded the effects of characteristics and dominated the wage effect thereafter. I next examined each tail of the distribution. Less discrimination and a smaller wage gap were found at lower quantile, and the case is different for the higher quantile, where coefficient effects explained almost the entire wage gap and potential discrimination had a stronger impact on the wage gap than characteristics. The wage gap for those earning lower wages was mainly explained by productivity-related characteristics, and these individuals experienced less potential discrimination than foreign workers within the lower wage quantiles. The negative discrimination for the entire immigrant group in 2013 that I identified via the Oaxaca decomposition was observed for the lower quantiles in that year.

As discussed in Section 3.2, during the study period, the share of low-skilled immigrants from

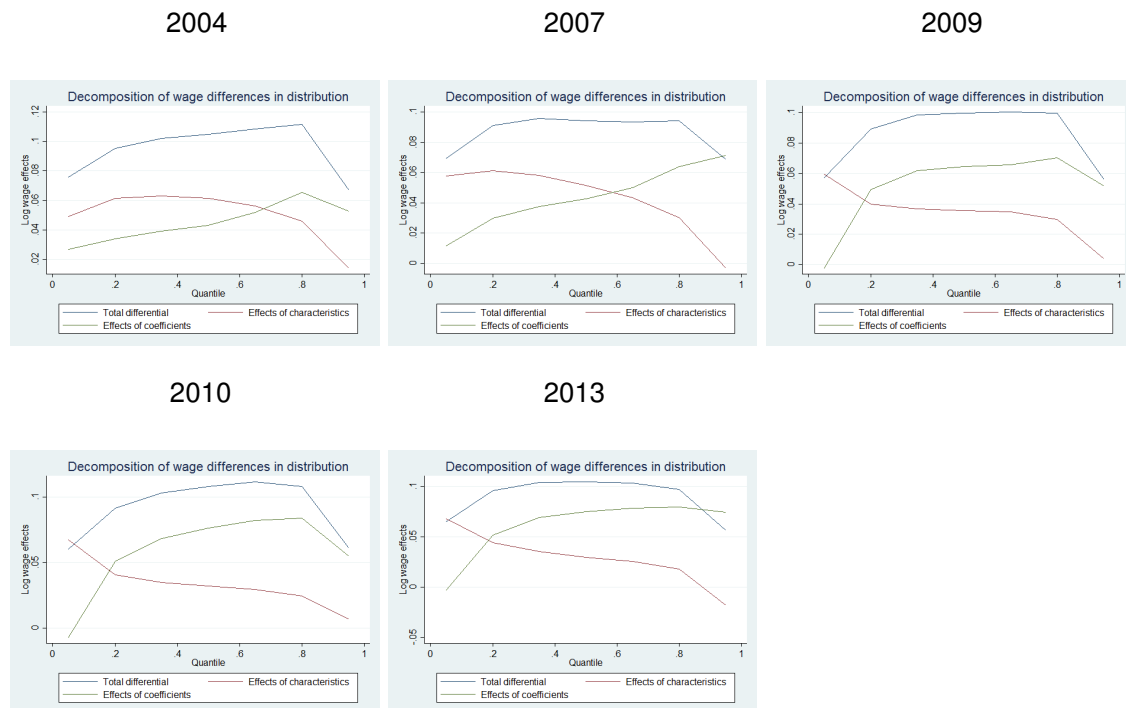


Figure 3.8: Native-total immigrants wage gap decomposition in all skill levels 2004-2013

new EU12 countries increased dramatically. In fact, low-skilled immigrants were the chief immigrant segment in all years, while the share of low-skilled immigrants from less developed countries decreased over the same period. Again as mentioned in empirical analysis from the Oaxaca-Blinder approach, could the potential increased discrimination effects derive from the change of population distributions on country of origin? Some empirical evidence was found through significant estimates obtained from the Oaxaca decomposition for comparisons of different years, but they were only based on the mean. Therefore, I performed the Melly decomposition for different skill groups for all immigrants to check how the discrimination level changed according to skill level.

Individual skill levels

Figures 3.9, 3.10 and Figure 3.11 offer the results for the native-immigrant wage gap decomposition for each skill level for all immigrants for 2004-2013. If I compare skill groups, the medium-skilled and high-skilled groups had larger wage gaps and faced more discrimination than low-skilled workers. This finding is consistent with my previous hypothesis and the results of the Oaxaca-Blinder decomposition. For a job with stricter requirements in terms of qualifications and educational background, stronger statistical discrimination could exist towards immigrants with relevant foreign experience. Another commonality across the three skill levels was that the effects of discrimination were larger than the effects of characteristics. In other words, potential discrimination played an important role for each skill group from 2004 to 2013 for the middle wage quantiles.

Unlike the clear increase in positive discrimination after the year 2007, after I controlled for skill (education), the native-immigrant wage gap in the low-skilled group became negative in particular range of quantiles in all years. This outcome was also in accordance with the Oaxaca results for the low-skilled group. For the middle wage quantiles, before the year 2009, potential discrimination explained more proportion of the wage gap, and it was negative. Then, since 2009, the coefficient effect for this quantile became positive. Regarding the small decrease in the share of low-skilled immigrants after the year 2007, the wage gap also turned positive. However, the wage gap became negative again in 2013. In most years except the year 2004 and 2013, the effects of characteristics explained little of the wage gap for the low-skilled group; this indicated that education played an important role in explaining the wage gap and was again consistent with the findings of the Oaxaca decomposition. The native-immigrant wage gap for middle-wage quantiles in the medium-skilled group was flatter across the quantiles than for the low-skilled group. In addition, as an important contributor to the wage gap, discrimination exhibited decreasing effects until 2010; the wage gap decreased as well. However, in 2013, the coefficient effect increased again. The effects of characteristics

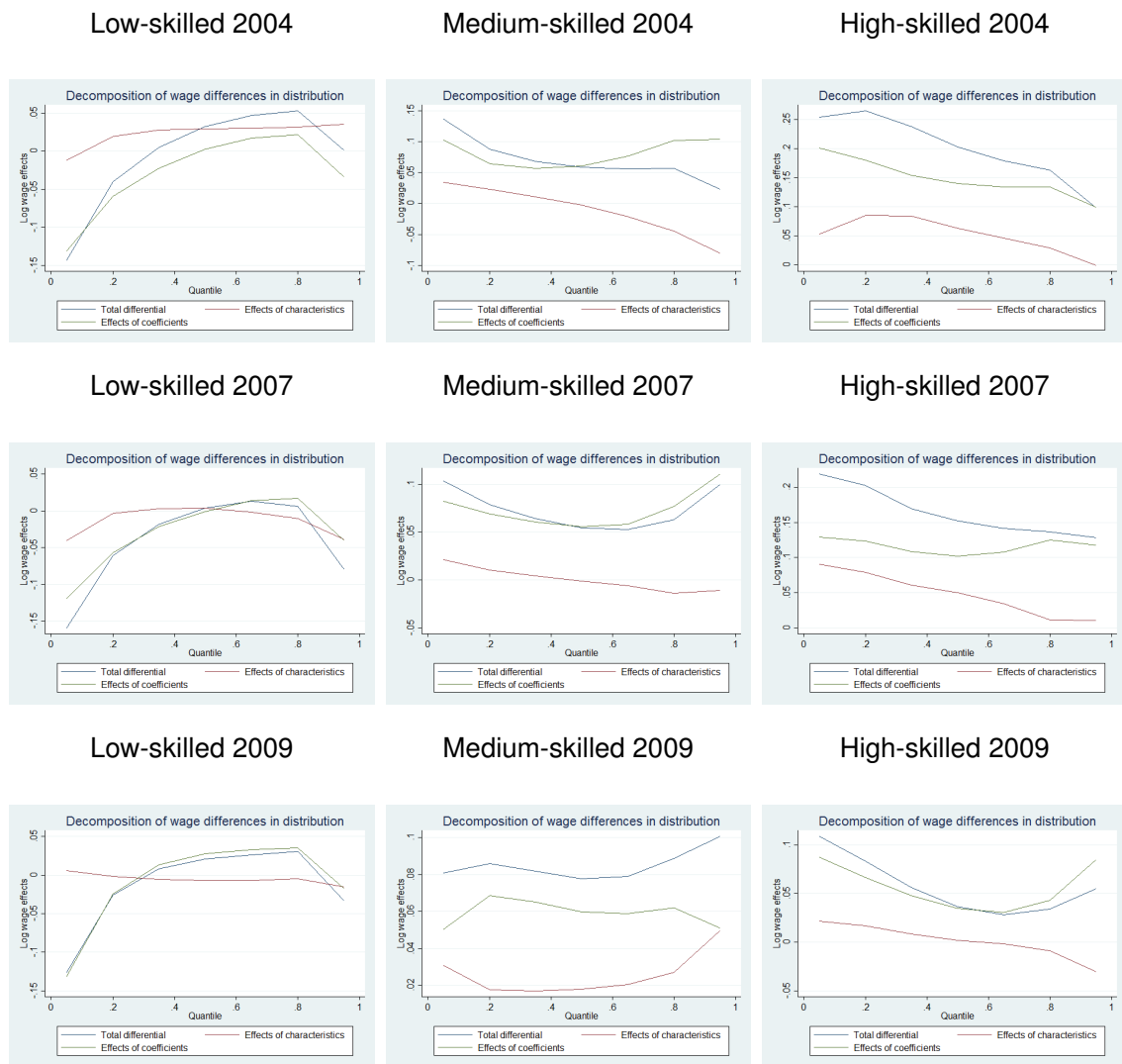
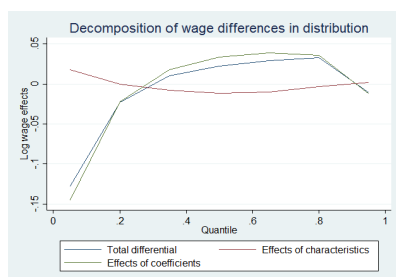
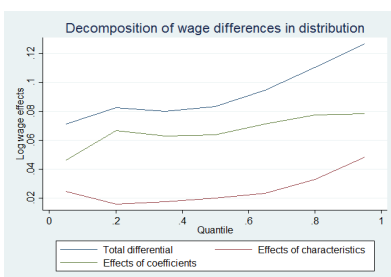


Figure 3.9: Native-total immigrants wage gap decomposition in each skill level 2004-2009

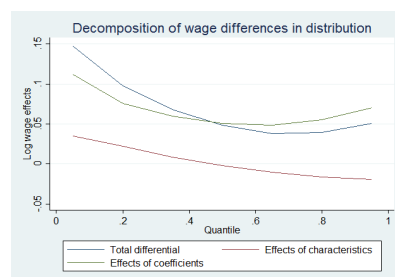
Low-skilled 2010



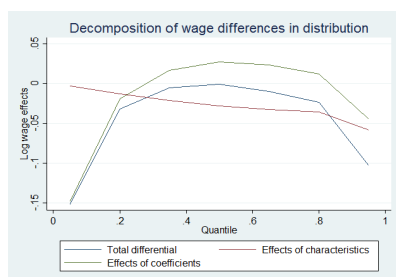
Medium-skilled 2010



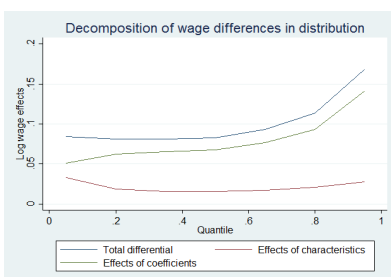
High-skilled 2010



Low-skilled 2013



Medium-skilled 2013



High-skilled 2013

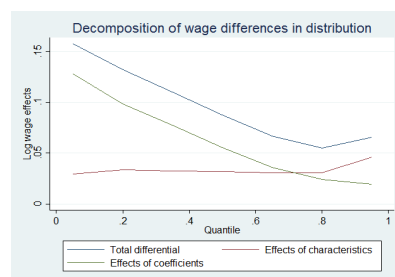


Figure 3.10: Native-total immigrants wage gap decomposition in each skill level 2010-2013

were relatively steady over the years and generally explained a minor portion of the wage gap.

I then investigated the high-skilled group. Similar to the medium-wage group, the level of discrimination decreased until 2009; it then grew again in 2010 and 2013. Since 2007, with the slight increase in the share of high-skilled immigrants, the effects of characteristics became negative after approximately the 60th quantile, and in the year 2013, the effects of characteristics explain more of the wage gap than potential discrimination does. This might explain why in the Oaxaca decomposition, the gap for the medium-skilled group was much larger than that for the high-skilled group after 2009. In the quantile decomposition, the negative wage gap for the upper quantile within the high-skilled group lowered the mean wage gap.

If I consider the results at the tails of the wage distribution in Figure 3.11, I find that in the lower tail of the low-skilled group, on average, the discrimination level was negative, and it was relatively large in absolute terms as compared with that seen for the middle and upper quantiles within the group. The same was true for the high-skilled group, except for the direction of discrimination. The medium-skilled group also faced high positive discrimination at the tails, but it is at both low and high quantiles. For the low-skilled groups, the discrimination level did not change over the years.

In summary, on average, the wage gap and discrimination level for low-skilled immigrants in the low-skilled group was lower than for medium- and high-skilled immigrants within the medium- and high-skilled groups, respectively. The increased power of discrimination observed in Figure 3.8 for all immigrants can be explained partly by the increase in the share of low-skilled immigrants. However, as the discrimination effects were greater than the characteristics effects in most cases for all three skill levels, and as the discrimination level in the low-skilled group was increasing, it was relatively small compared to the medium- and high-skilled groups, which both were decreasing at the same time. Another point regarding the low-skilled immigrant group is that the sign of the direction of discrimination changed with the

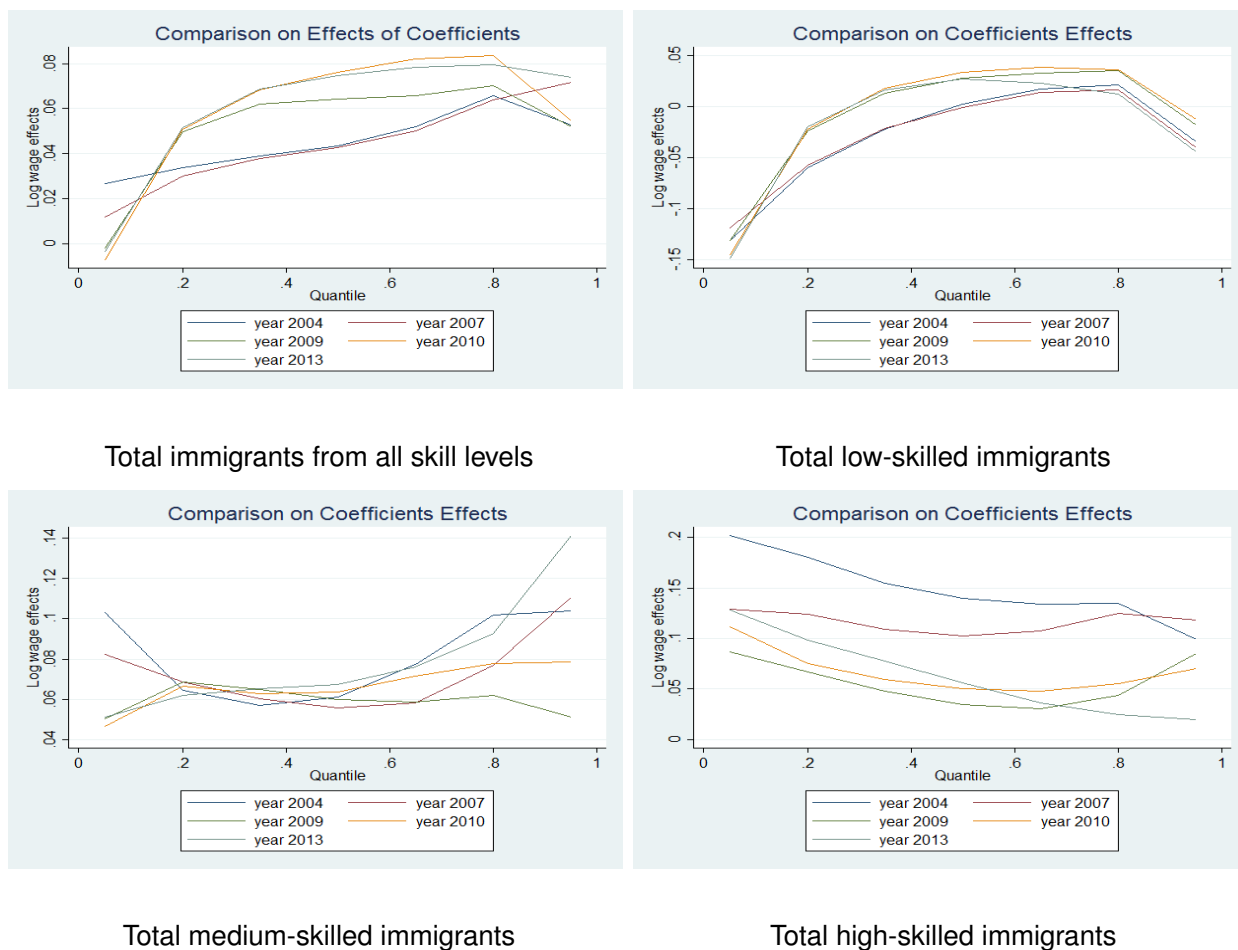


Figure 3.11

share of low-skilled immigrants within each group of origin. Therefore, it is difficult to judge at this stage whether the increase was purely due to low-skilled immigrants. Discrimination against certain nationalities could play a role as well, as the skill structure of the population varied within each group of origin.

Individual groups of origin

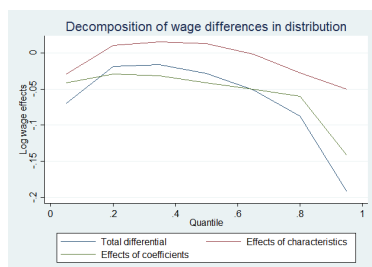
According to the above arguments, I performed decompositions within the individual groups of origin for all skill levels to check whether discrimination levels varied by region of origin. Since

the EU15 group and the remaining developed countries shared similar trends, population structures, and Oaxaca-Blinder decomposition results, and since the EU15 population was much larger than the other, I only performed the decomposition on the EU15 group to compare it with the new EU12 group and the less developed countries. The wage gap resulting from skill differences was captured by education in the explainable part of the decomposition. Therefore, the unexplained coefficient effect can be interpreted as potential discrimination against different ethnic groups. The results of the decomposition are in Figure 3.12 and 3.13, and comparisons of coefficient effect over the years for all wage quantiles in each group of origin are depicted in Figure 3.14.

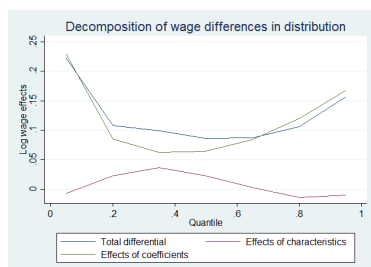
The results indicate a clear trend across the various ethnic groups. For the EU15, EU12, and less developed countries groups, a higher positive discrimination level is evident. To start, I first focus on middle-wage quantiles (q20-q80) in each group of origin. In the EU 15 group, discrimination and productivity-related characteristics almost equally contributed to the wage gap in absolute value. Discrimination and the wage gap were both negative in this group for all years. Positive effects of characteristics were present, but reversed from 2010. The EU15 group's share of all immigrants held relatively steady over the years, but the EU15 group's share of medium-skilled immigrants fell in 2010, while its share of low-skilled immigrants slightly increased in the same year. As Figure 3.11 clarifies, the effects of possible discrimination became more towards positive with the increase in low-skilled immigrants. Hence, the decreasing effects of coefficient in absolute value in 2010 can be explained by the increasing inflow of low-skilled EU15 immigrants. In general, being from an EU15 country can result in a wage advantage since, for this group, potential discrimination is negative in most cases.

The new EU12 group faced positive discrimination effects, which decreased from 2004 to 2007 as the share of new EU12 immigrants increased. Thereafter, productivity-based characteristics explained more of the wage gap. With the large inflow of low-skilled immigrants

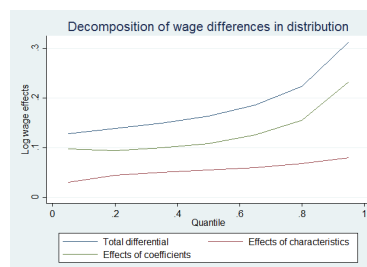
EU15 2004



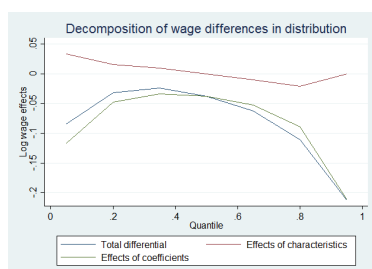
New EU12 2004



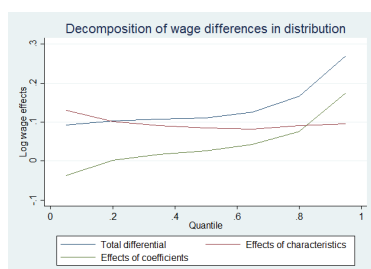
LD 2004



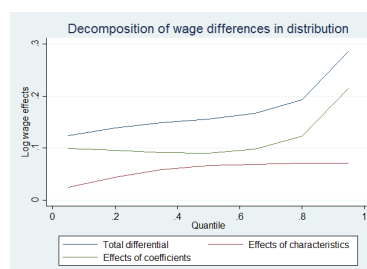
EU15 2007



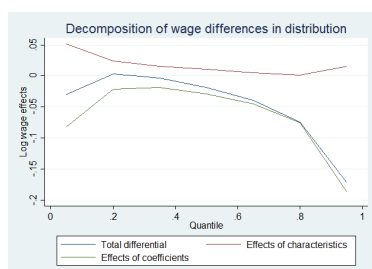
New EU12 2007



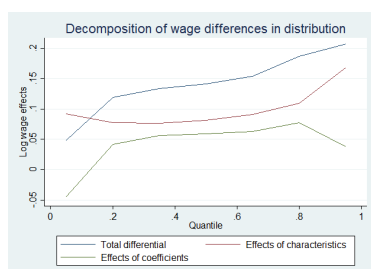
LD 2007



EU15 2009



New EU12 2009



LD 2009

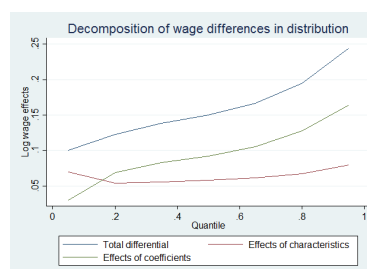


Figure 3.12: Native-immigrants wage gap decomposition in all skill levels under each immigrant group(EU12 New EU12 and Less Developed countries(LD))2004-2009

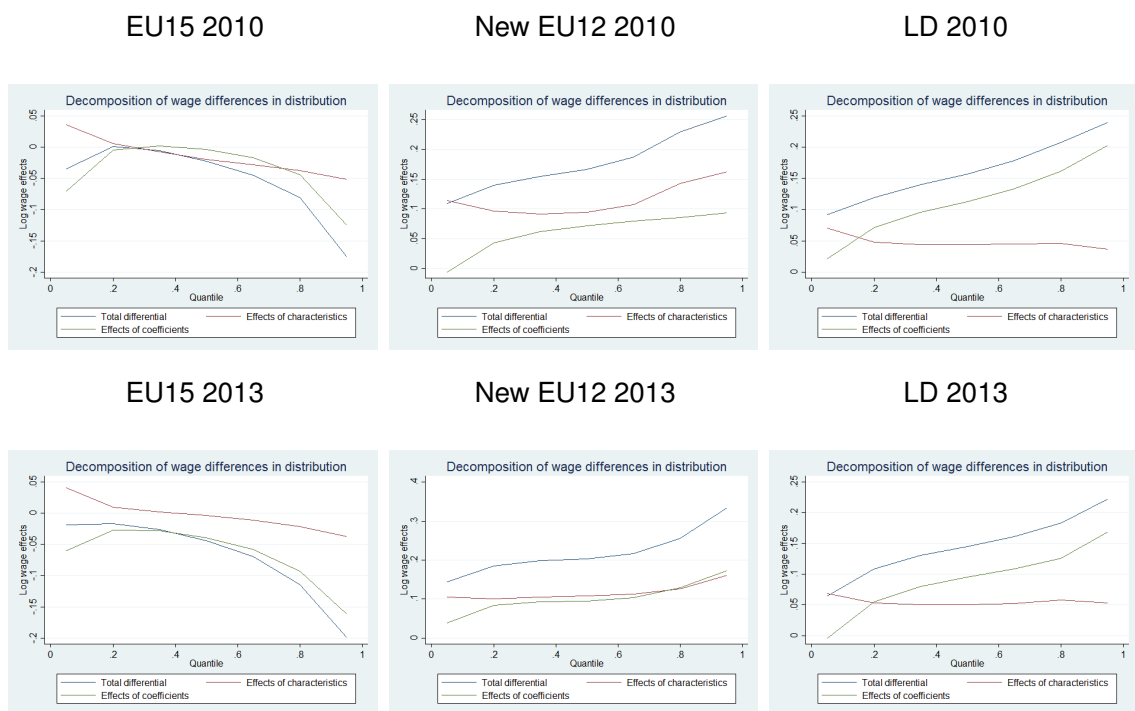


Figure 3.13: Native-immigrants wage gap decomposition in all skill levels under each immigrant group (EU12 New EU12 and Less Developed countries (LD)) 2010-2013

from the new EU12 group, skill explained more of the wage gap since in general, low-skilled workers earn less than those in higher-skill jobs, and low-skilled immigrants account for major part of the new EU12 group, while medium-skilled workers were dominant in the native group. This fact is also the reason for the increasing wage gap from 2004 to 2013.

Thirdly, in the group from less developed countries, discrimination was positive in all years, and this factor's effect was stronger than that of characteristics. Furthermore, the wage gap and discrimination effect increased with the wage quantile. Compared to other two ethnic groups, this group had a lower wage gap than new EU12 immigrants but a higher wage gap than EU15 immigrants. This outcome could be due to the combination of the individual skill structure within each group of origin and the level of discrimination. Additionally, the level of discrimination did not change much for the group from less developed countries over the years for the middle wage quantile.

I then investigated each tail of the wage quantiles. All three groups present a larger wage gap for the upper quantile. Furthermore, for the EU15 and the group from less developed countries, a larger potential discrimination effect emerged, while in the new-EU12 group, the coefficient effects were much lower in 2009 and 2010 than in other years. For the lower quantiles, all groups had a relatively low level of discrimination in absolute terms. For immigrants from less developed countries, for the lower quantile wage, the potential discrimination level decreased.

To summarise, discrimination contributed more to the wage gap in the group from less developed countries than in the other two groups of origin. This finding suggests an increase in the extent of discrimination the more culturally and economically distant the immigrant's home country is from Denmark. I reached the same conclusion via the Oaxaca-Blinder decompositions. I discuss this finding further in the next section. Moreover, the EU15 group had a negative discrimination level, which means these immigrants had advantages over natives

with similar qualifications. Finally, the upper quantiles faced higher levels of discrimination in all three groups of origin.

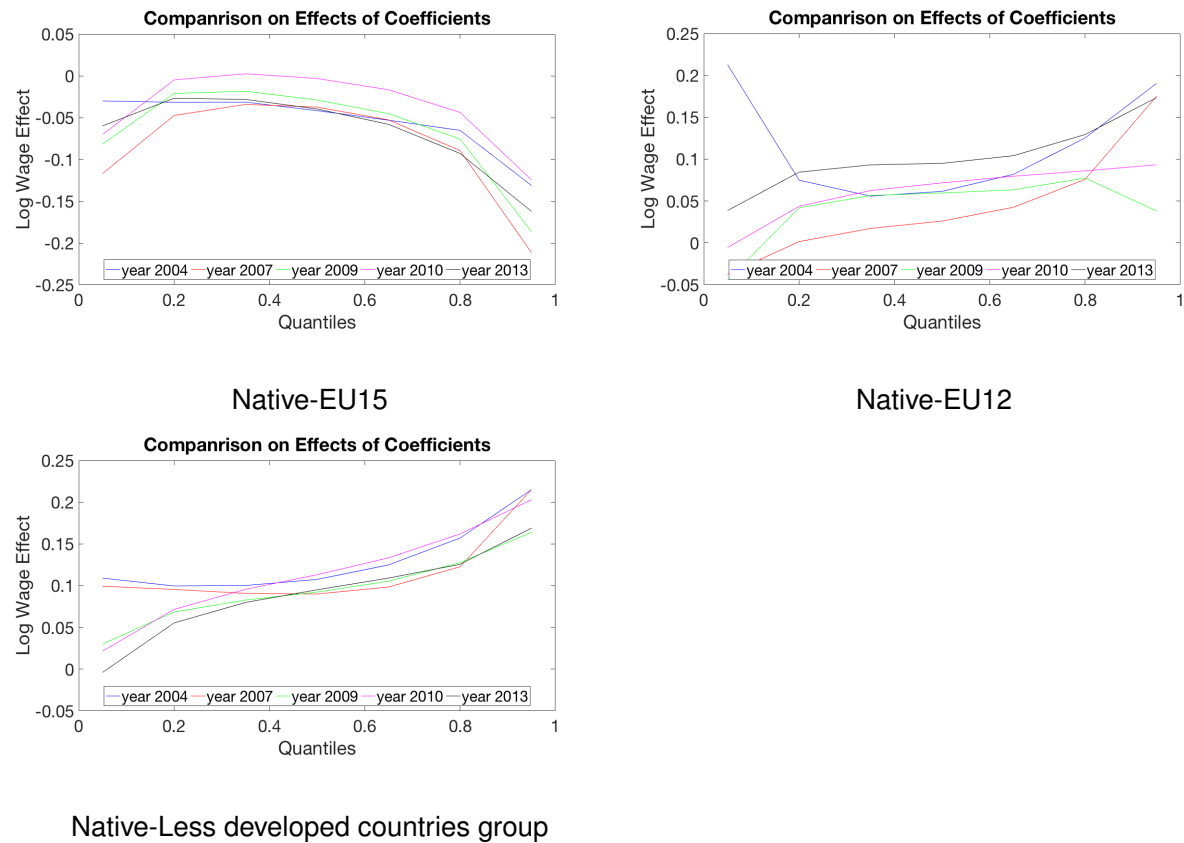


Figure 3.14

3.6.3 Further discussion on empirical results

Through the above analysis using Oaxaca-Blinder and Melly decomposition procedures, I obtained consistent conclusions. According to the empirical results, the size of the wage gap largely depends on the skill level, and whether the wage gap is positive is associated with an immigrant's country of origin. This finding confirms my hypothesis in Section 3.3 that the higher the skill level is, the larger wage gap and potential statistical discrimination could

present; and the more distant⁶ the immigrant's home country from Denmark, the more positive the discrimination. Again, when I use the term "discrimination" in this paper, I refer to the remaining component after I have controlled for all possible factors in the wage model I can obtain. This term is not the same as the general political definition in practical. Moreover, "potential discrimination" in this paper refers to the unexplained components in the decomposition models.

When analysing the native-immigrant wage gap, I first investigated the skill level aspect. After I controlled for education in each skill group, the explained portion of the wage gap became relatively small. This confirmed the hypothesis that education (or skill level) is an important factor in wage setting. In general, the native sample had a relatively high share of medium-skilled workers, while the immigrant sample contained a high share of low-skilled workers. This difference can explain why, on average, before controlling for education level, the effects of characteristics were positive in most cases. With more low-skilled workers in the immigrant group, the coefficient effect played a greater role in the wage gap. The increase in the share of low-skilled immigrants enlarged the wage gap because low-skilled workers often earn less than those at other skill levels. Experience and tenure can also play a role in the wage gap, but the effects were small in my empirical study. The wage for the low-skilled group was the smallest among all skill levels, and the size of the gap was more explained by the "discrimination component". Regarding the wage gaps for the different groups of origin, I found the largest gap within the group from less developed countries, followed by the new EU12 group. For the EU15 and the group from remaining developed countries, the wage gap became negative. Also, the wage gap for high-wage earners was larger for all groups of origin since high-skilled workers were most likely to be at the upper wage quantiles within each group. This finding is also consistent with Hofer et al. (2017), who used Austrian data.

The unexplained portion of the gap may be due to occupational differences or discrimination

⁶In this paper, "distant" is measured in the aspect of cultural and economic levels.

against different ethnic groups. For the various skill groups, the unexplained part constituted the dominant portion of the wage gap and also confirmed my earlier hypothesis, the size of the coefficient effect is smaller in the low-skilled group in general. In terms of the various wage quantiles, higher-wage earners experienced more potential discrimination, except for those in the high-skilled group. The relatively high wages in the high-skilled group meant that many of these individuals held managerial or key positions in their firm. Candidates for such positions often already have a strong background. Statistical discrimination, of course, still exists and could be positive due to different cultural backgrounds. The medium-skilled group is more heterogeneous than the other two skill levels, for the reason that both low-skilled and high-skilled immigrant workers could fall into medium-skilled due to overvalue or undervalue of their qualifications.

For the groups of origin, the direction of the potential discrimination was more explicit. The more distant (in terms of cultural and economic background) the immigrant's home country from Denmark, the more positive the discrimination. This finding was even consistent over time. Immigrants from EU15 and other developed countries faced negative discrimination, which meant that their foreign education or experience was overvalued by employers, giving them the wage advantages in Denmark. While the new EU12 group can have larger cultural differences than the EU15 group and immigrants from the U.S., Australia, Canada, and similar countries, the new EU12 group had a more positive discrimination level. Moreover, for the group from less developed countries, discrimination became positive. The share of high-skilled workers within this group grew, but the wage gap and discrimination level exhibited decreasing trends. This could be a result of the policies aiming to attract high-skilled immigrants. It could also be explained by a situation where there are more international students who have finished their master's or PhD studies in Denmark and they then stay to work. In this case, education and qualifications of those international students would not have been devalued, which would have reduced the discrimination. However, the discrimination against one's cultural background is different from the discrimination against one's educational background;

an assimilation effect would also contribute to the discrimination level⁷. Further research should be done within each group of origin to distinguish whether discrimination is due to policy, assimilation effects, or increases in the number of highly educated international students remaining in the country.

⁷the variable "Years since migration" is constructed to capture the assimilation effect. Since the information on working experience for immigrants are only available after the immigrant entering Denmark, the variable "experience" used in the models in this study has already capture part of the assimilation effect. Therefore, I did not use "Years since migration" for the empirical analysis in this paper. However, additional robustness checks on the inclusion of "years since migration" instead of "experience" have been performed, and they are shown in the Appendix.

3.7 Conclusion

In this paper, I have analysed the native-immigrant wage gap for male immigrants in different wage quantiles for the labour market in Denmark. I also explored the existence of wage discrimination against immigrants in the labour market in Denmark from 2004 to 2013. I particularly focused on changes in the discrimination level over the period within skill levels and groups of origin following EU enlargement in 2004 and 2007, and the introduction of free movement in Denmark in 2009. I started the analysis by assessing changes in the population distributions and wage distributions of the native and immigrant groups. I then applied the Oaxaca-Blinder and Melly (2005) approaches to decomposing the native-immigrant wage gap. To simplify the analysis, I only focused on males.

Results from both decomposition approaches confirm that the size of the wage gap is mostly dependent on skill level and whether the wage gap positive is more associated with an immigrant's country of origin. Wage differentials were generally the smallest within the low-skilled group. For low-wage earners in the low-skilled and high-skilled groups, the wage gap and discrimination level were larger than for other wage quantiles within the same skill category. The medium-skilled group, however, presented a higher wage gap and effect of discrimination among the high-wage population. The absolute value of discrimination increased with the skill level. Furthermore, education (or skill level) was an important factor in wage setting; after conditioning on skill level, the explained effects of characteristics became rather small. The discrimination component was the dominant factor explaining the wage gap for the various skill levels. Moreover, stronger discrimination also occurred for the upper wage quantiles in each group of origin. In terms of whether the wage gap and discrimination were positive, the country of origin played a role. Measured discrimination was the strongest and most positive for immigrants from less developed countries, most of which were non-EU countries. Moreover, regarding the EU and developed countries such as the U.S. and Australia, negative discrimination (i.e., advantages) was observed in the context of the wage gap. In other words,

the more distant the home country of the immigrant from Denmark, the more positive was the discrimination.

This paper has investigated the native-immigrant wage gap for males, but further research should be conducted. At this stage, I cannot fully claim that the conclusions I obtained fit the whole immigrant population in Denmark since results for females might present a different picture. Therefore, comparative research should be performed as the next step, comparing the wage differentials within different groups of origins at various skill levels. Moreover, the assimilation effect has not been distinguished from the effects of discrimination. It is viewed as one kind of discrimination in my paper since I had no information on language skills, which could be a key factor contributing to success in the labour market. However, I could still perform certain checks using the available data. Nielsen et al. (2004) investigated assimilation in Denmark and used years since migration as a tool for measuring the assimilation effect. I could also compare immigrants who received education in Denmark and stayed to work with immigrants possessing a foreign education or work background. Policy effects cannot be ignored, as various immigration policies aiming to attract high-skilled immigrants have been implemented during recent decades. As discussed earlier, further research should be done to analyse the influence of policy and assimilation within each group of immigrant origin in Denmark.

Bibliography

- Altonji, J. G. & Blank, R. M. (1999), 'Race and gender in the labor market', *Handbook of labor economics* **3**, 3143–3259.
- Arrow, K. et al. (1973), 'The theory of discrimination', *Discrimination in labor markets* **3**(10), 3–33.
- Becker, G. S. (2010), *The economics of discrimination*, University of Chicago press.
- Bell, B. D. (1997), 'The performance of immigrants in the united kingdom: evidence from the ghos', *The Economic Journal* **107**(441), 333–344.
- Berman, Y. & Aste, T. (2016), 'To what extent does immigration affect inequality?', *Physica A: Statistical Mechanics and its Applications* **462**, 1029–1039.
- Blinder, A. S. (1973), 'Wage discrimination: reduced form and structural estimates', *Journal of Human resources* pp. 436–455.
- Büchel, F. & Frick, J. R. (2005), 'Immigrants' economic performance across europe—does

immigration policy matter?', *population Research and policy Review* **24**(2), 175–212.

Cain, G. G. (1986), 'The economic analysis of labor market discrimination: A survey', *Handbook of labor economics* **1**, 693–785.

Card, D. (2009), 'Immigration and inequality', *American Economic Review* **99**(2), 1–21.

Chiswick, B. R. (1978), 'The effect of americanization on the earnings of foreign-born men', *Journal of political Economy* **86**(5), 897–921.

Chiswick, B. R. (1980), 'The earnings of white and coloured male immigrants in britain', *Economica* **47**(185), 81–87.

Chletsos, M. & Roupakias, S. (2017), 'Native-immigrant wage differentials in greece: discrimination and assimilation', *Applied Economics* **49**(17), 1732–1736.

Constant, A. & Massey, D. S. (2005), 'Labor market segmentation and the earnings of german guestworkers', *Population Research and Policy Review* **24**(5), 489–512.

Fortin, N., Lemieux, T. & Firpo, S. (2011), Decomposition methods in economics, in 'Handbook of labor economics', Vol. 4, Elsevier, pp. 1–102.

Hirsch, B. & Jahn, E. (2012), Is there monopsonistic discrimination against immigrants? first evidence from linked employer-employee data, Technical report, Institute for the Study of Labor (IZA).

Hofer, H., Titelbach, G., Winter-Ebmer, R. & Ahammer, A. (2017), 'Wage discrimination against immigrants in austria?', *Labour* **31**(2), 105–126.

- Huber, M. (2015), 'Causal pitfalls in the decomposition of wage gaps', *Journal of Business & Economic Statistics* **33**(2), 179–191.
- International Migration Denmark* (2015), Technical report, Immigration, Integration and Housing Ministry, Copenhagen.
- Juhn, C., Murphy, K. M. & Pierce, B. (1993), 'Wage inequality and the rise in returns to skill', *Journal of political Economy* **101**(3), 410–442.
- Kahanec, M. & Zaiceva, A. (2009), 'Labor market outcomes of immigrants and non-citizens in the eu: An east-west comparison', *International Journal of Manpower* **30**(1/2), 97–115.
- Kee, P. (1995), 'Native-immigrant wage differentials in the netherlands: discrimination?', *Oxford Economic Papers* pp. 302–317.
- Koenker, R. & Bassett Jr, G. (1978), 'Regression quantiles', *Econometrica: journal of the Econometric Society* pp. 33–50.
- Lang, G. (2000), Native-immigrant wage differentials in germany: Assimilation, discrimination, or human capital?, Technical report, Volkswirtschaftliche Diskussionsreihe, Institut für Volkswirtschaftslehre der Universität Augsburg.
- Lehmer, F. & Ludsteck, J. (2011), 'The immigrant wage gap in germany: Are east europeans worse off?', *International migration review* **45**(4), 872–906.
- Liu, P.-W., Zhang, J. & Chong, S.-C. (2004), 'Occupational segregation and wage differentials between natives and immigrants: evidence from hong kong', *Journal of Development Economics* **73**(1), 395–413.

- Machado, J. A. & Mata, J. (2005), 'Counterfactual decomposition of changes in wage distributions using quantile regression', *Journal of applied Econometrics* **20**(4), 445–465.
- Malchow-Møller, N., Munch, J. R. & Skaksen, J. R. (2012), 'Do immigrants affect firm-specific wages?', *The Scandinavian Journal of Economics* **114**(4), 1267–1295.
- Melly, B. (2005), 'Decomposition of differences in distribution using quantile regression', *Labour economics* **12**(4), 577–590.
- MILLER, P. W. & CHISWICK, B. R. (1985), 'Immigrant generation and income in australia', *Economic Record* **61**(2), 540–553.
- Miranda, A. & Zhu, Y. (2013), 'English deficiency and the native–immigrant wage gap', *Economics Letters* **118**(1), 38–41.
- Nanos, P. & Schluter, C. (2014), 'The composition of wage differentials between migrants and natives', *European Economic Review* **65**, 23–44.
- Neumark, D. (1988), 'Employers' discriminatory behavior and the estimation of wa', *The Journal of Human Resources* **23**(3), 279.
- Nielsen, C. P. (2011), 'Immigrant over-education: evidence from denmark', *Journal of Population Economics* **24**(2), 499–520.
- Nielsen, H. S., Rosholm, M., Smith, N. & Husted, L. (2004), 'Qualifications, discrimination, or assimilation? an extended framework for analysing immigrant wage gaps', *Empirical Economics* **29**(4), 855–883.

- Oaxaca, R. (1973), 'Male-female wage differentials in urban labor markets', *International economic review* pp. 693–709.
- Ottaviano, G. I. & Peri, G. (2008), Immigration and national wages: Clarifying the theory and the empirics, Technical report, National Bureau of Economic Research.
- Sologon, D. M. & Van Kerm, P. (2018), 'Modelling earnings dynamics and inequality: foreign workers and inequality trends in luxembourg, 1988–2009', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **181**(2), 409–440.
- Velling, J. (1995), Wage discrimination and occupational segregation of foreign male workers in germany, Technical report, ZEW Discussion Papers.

Appendices

Appendix A: Additional tables for Section 3.5

Table 3.1: Correction procedures on education variables

	Native sample	Immigrant sample before correction	Immigrant sample after correction
Missing	0.71%	3.40 %	1.33%
in 2010	0.57 %	27.45%	10.59%
Low-skilled	26.55%	62.39 %	63.10%
in 2010	24.71%	49.42%	55.20%
Medium-skilled	65.78%	24.81%	25.57%
in 2010	67.2%	18.55%	24.78%
High-skilled	6.96%	9.4%	10.00%
in 2010	7.51%	4.57 %	9.43%

Table 3.2: Descriptive statistics for male workers

Variable Names	Natives		Immigrants	
	mean	sd	mean	sd
hourly wage(log)	5.3201549	.41912002	5.236966	.42251175
age 18-24	.08013683	.27150494	.08083912	.27258839
age 25-29	.091306	.28804379	.16123506	.36774831
age 30-39	.26002099	.43864577	.35926799	.47978678
age 40-49	.27369448	.44585407	.28537805	.45159515
age 50-59	.22497798	.41756787	.09880047	.29839448
age 60-65	.0572066	.23223698	.01126541	.10553931
low-skilled	.26229875	.43988425	.631027	.48252751
medium skilled	.65426989	.47560577	.25572674	.43626974
high skilled	.0692822	.25393344	.09998846	.29998517
married	.53412286	.49883432	.56853281	.49528191

Continued on next page...

... table 3.2 continued

Variable Names	Natives		Immigrants	
	mean	sd	mean	sd
children 0-6 yrs	.20817414	.4060021	.29960413	.45808545
Copenhagen	.09877812	.2983639	.21435195	.41037278
large city	.4698621	.49909093	.44932054	.49742589
small city	.43514969	.49577665	.34427847	.47513329
experience	9.8409508	7.4715813	5.7931955	5.7569399
Union	.67953396	.46665576	.85589954	.35119222
number of employees	301.89397	695.42575	396.36167	858.65595
argricuture	.04303977	.20294668	.07366277	.26122178
manufactoring	.13940032	.34636382	.14873697	.35582967
energy	.04234201	.20136826	.03862268	.19269433
construction	.2233234	.4164734	.1703107	.37590623
trade	.08003097	.27134116	.10034225	.30045634
transport	.09176256	.28869051	.06923481	.25385348
finance etc	.06085895	.23907143	.06983439	.25486818
personal service	.16314128	.36949456	.15584628	.36271023
N	4921439		268521	

Addition notes to Table 3.2

1. Experience: In years. Sum of days employed in the labour market, divided by 365.

2.Tenure: In years. Time employed within the workplace. This variable is not included in the empirical tests but used for sensitivity analysis; the results are presented in Appendix F. In this paper, I chose "Experience" in the regression models.

3. The variable "Years since migration(YSM)" has also been constructed at the initial stage. For the native group, it is equal to the age of an individual; For the immigrant group, it is the number of years since one has entered in Denmark. However, since the working history variable "Experience" captured part of features of YSM has been used in this study, I did not report summary statistics for YSM in Table 3.2.

4. Hourly wage: It is calculated through dividing total annual wages including ATP by total hours worked in the year.

5. Age dummies: Age 18-24, Age 25-29, Age 30-39, Age 40-49, Age 50-59, Age 60-65.

6. Children 0-6 years: A dummy indicator of whether there is a child in the household aged 0-6 years.

Appendix B: Full results from Oaxaca-Blinder decomposition for Figure 3.6 and Figure 3.7

Appendix B.1: Full results for Figure 3.6

Table 3.3: Oaxaca-Blinder decomposition results for males of different skilled groups: Native-Total immigrants

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
All Skilled groups					
difference	0.094***	0.086***	0.085***	0.095***	0.098***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
explained	0.058***	0.078***	0.085***	0.091***	0.104***

Continued on next page...

... table 3.3 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
unexplained	0.036***	0.008**	0.000	0.003	-0.006**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Low-skilled					
difference	-0.002	-0.037***	-0.009***	-0.005	-0.031***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
explained	0.002	-0.002	0.028***	0.035***	0.015***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
unexplained	-0.004	-0.035***	-0.037***	-0.040***	-0.046***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Medium-skilled					
difference	0.066***	0.071***	0.080***	0.091***	0.101***
	(0.008)	(0.007)	(0.004)	(0.004)	(0.004)
explained	-0.019***	-0.007***	0.015***	0.020***	0.027***
	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)
unexplained	0.085***	0.078***	0.065***	0.071***	0.074***
	(0.008)	(0.006)	(0.004)	(0.004)	(0.004)
High-skilled					
difference	0.194***	0.159***	0.049***	0.067***	0.100***
	(0.018)	(0.015)	(0.008)	(0.008)	(0.007)
explained	-0.005	-0.015**	0.004	0.003	0.039***
	(0.008)	(0.007)	(0.003)	(0.004)	(0.003)
unexplained	0.199***	0.174***	0.045***	0.065***	0.061***
	(0.018)	(0.014)	(0.008)	(0.008)	(0.007)

Table 3.4: Full results from Oaxaca-Blinder decomposition for Native-Total immigrants wage gap

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	0.058*** (0.002)	0.078*** (0.002)	0.085*** (0.001)	0.091*** (0.001)	0.104*** (0.001)
unexplained	0.036*** (0.004)	0.008** (0.004)	0.000 (0.002)	0.003 (0.002)	-0.006** (0.002)
Explained)					
lgage	0.028*** (0.001)	0.022*** (0.001)	0.028*** (0.000)	0.030*** (0.000)	0.028*** (0.000)
basic edu	0.019*** (0.003)	0.007** (0.003)	-0.007** (0.003)	0.024*** (0.001)	-0.023*** (0.003)
vocational edu	0.048*** (0.003)	0.058*** (0.003)	0.060*** (0.003)	0.025*** (0.001)	0.083*** (0.004)
higher edu	-0.014*** (0.001)	-0.007*** (0.001)	-0.008*** (0.001)	-0.006*** (0.000)	-0.010*** (0.001)
Married	-0.005*** (0.000)	-0.003*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
children 0-6 yrs	-0.017*** (0.001)	-0.013*** (0.001)	-0.008*** (0.000)	-0.007*** (0.000)	-0.006*** (0.000)
Copenhagen	-0.005*** (0.000)	-0.005*** (0.000)	-0.003*** (0.000)	-0.004*** (0.000)	-0.004*** (0.000)
lgexp	0.019*** (0.001)	0.035*** (0.001)	0.026*** (0.001)	0.032*** (0.001)	0.043*** (0.001)
Union	0.001** (0.000)	-0.001*** (0.000)	0.001*** (0.000)	0.002*** (0.000)	-0.001*** (0.000)
Inemployee	-0.005*** (0.000)	-0.006*** (0.000)	-0.006*** (0.000)	-0.007*** (0.000)	-0.006*** (0.000)

Continued on next page...

... table 3.4 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
agriculture	-0.000*** (0.000)	-0.000 (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.000*** (0.000)
manufacturing	0.000*** (0.000)	0.001*** (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000*** (0.000)
energy	0.001*** (0.000)	0.001*** (0.000)	-0.000 (0.000)	-0.000*** (0.000)	0.000*** (0.000)
construction	-0.005*** (0.000)	-0.004*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000 (0.000)
trade	-0.000 (0.000)	0.000* (0.000)	0.001*** (0.000)	0.000*** (0.000)	0.000 (0.000)
transport	-0.001** (0.000)	-0.001* (0.000)	0.004*** (0.000)	0.005*** (0.000)	0.005*** (0.000)
finance etc	-0.007*** (0.001)	-0.007*** (0.001)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
personal service	0.000 (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001** (0.000)	-0.002*** (0.000)
Unexplained)					
lgage	-0.006 (0.065)	0.047 (0.057)	0.043 (0.039)	0.046 (0.041)	-0.056 (0.042)
basic edu	0.124** (0.057)	0.226*** (0.034)	0.218*** (0.030)	-0.035*** (0.005)	0.249*** (0.041)
vocational edu	0.067*** (0.023)	0.109*** (0.013)	0.113*** (0.013)	0.005** (0.003)	0.128*** (0.017)
higher edu	0.028*** (0.007)	0.038*** (0.004)	0.041*** (0.005)	0.004*** (0.001)	0.061*** (0.008)
Married	0.041*** (0.006)	0.031*** (0.005)	0.028*** (0.003)	0.035*** (0.003)	0.037*** (0.003)

Continued on next page...

... table 3.4 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
children 0-6 yrs	0.024*** (0.003)	0.021*** (0.002)	0.014*** (0.002)	0.011*** (0.002)	0.002 (0.001)
Copenhagen	-0.001 (0.002)	-0.005*** (0.002)	0.003** (0.001)	0.003*** (0.001)	0.006*** (0.001)
lgexp	-0.025*** (0.004)	0.002 (0.005)	0.042*** (0.004)	0.058*** (0.004)	0.082*** (0.005)
Union	-0.047*** (0.010)	-0.047*** (0.008)	-0.041*** (0.005)	-0.038*** (0.005)	-0.050*** (0.005)
Inemployee	-0.071*** (0.013)	-0.089*** (0.011)	-0.060*** (0.007)	-0.076*** (0.007)	-0.137*** (0.007)
agriculture	-0.001 (0.001)	-0.000 (0.001)	-0.001 (0.001)	0.001 (0.001)	-0.005*** (0.001)
manufacturing	0.007 (0.005)	-0.002 (0.003)	-0.006*** (0.001)	-0.005*** (0.001)	-0.012*** (0.001)
energy	0.002 (0.002)	-0.000 (0.002)	-0.000 (0.000)	0.001** (0.000)	-0.002*** (0.000)
construction	0.001 (0.002)	-0.000 (0.002)	0.009*** (0.002)	0.007*** (0.002)	-0.004*** (0.001)
trade	0.006* (0.003)	0.003 (0.002)	0.004*** (0.001)	0.005*** (0.001)	-0.001 (0.001)
transport	0.002 (0.002)	0.000 (0.002)	-0.006*** (0.001)	-0.004*** (0.001)	-0.008*** (0.001)
finance etc	0.003 (0.003)	-0.004 (0.002)	-0.004*** (0.001)	-0.002*** (0.001)	-0.005*** (0.001)
personal service	0.001 (0.001)	-0.001 (0.001)	-0.007*** (0.002)	-0.007*** (0.002)	-0.016*** (0.001)
Constant	-0.121	-0.321***	-0.389***	-0.005	-0.276***

Continued on next page...

... table 3.4 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.112)	(0.082)	(0.063)	(0.040)	(0.079)
N	246910	254158	587218	578504	617376

Table 3.5: Full results from Oaxaca-Blinder decomposition for Native-immigrant wage gap in Low-skilled group

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	0.002	-0.002	0.028***	0.035***	0.015***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
unexplained	-0.004	-0.035***	-0.037***	-0.040***	-0.046***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Explained)					
lgage	0.038***	0.024***	0.032***	0.034***	0.024***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Married	-0.007***	-0.007***	-0.004***	-0.006***	-0.008***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
children 0-6 yrs	-0.029***	-0.023***	-0.015***	-0.016***	-0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Copenhagen	-0.006***	-0.006***	-0.004***	-0.005***	-0.006***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
lgexp	0.016***	0.029***	0.024***	0.031***	0.032***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Union	0.000	-0.001*	0.003***	0.003***	-0.001***

Continued on next page...

... table 3.5 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Inemployee	-0.005***	-0.007***	-0.007***	-0.007***	-0.006***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
argricuture	0.000**	0.000***	0.000	0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
manufacturing	0.003***	0.003***	0.000**	0.000***	-0.000***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
energy	0.003***	0.003***	-0.000***	-0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
construction	-0.010***	-0.010***	-0.002***	-0.003***	-0.000**
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
trade	0.001***	0.001***	0.000**	0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
transport	0.003***	0.003***	0.003***	0.004***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
finance etc	-0.008***	-0.009***	-0.001***	-0.001***	-0.002***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
personal service	0.000	-0.000	-0.001***	-0.001***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained)					
lgage	0.416***	0.364***	0.290***	0.323***	0.110**
	(0.077)	(0.069)	(0.050)	(0.055)	(0.056)
Married	0.035***	0.032***	0.027***	0.032***	0.040***
	(0.007)	(0.006)	(0.004)	(0.004)	(0.004)
children 0-6 yrs	0.037***	0.032***	0.023***	0.022***	0.006***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
Copenhagen	0.000	-0.006***	0.001	0.001	0.005***

Continued on next page...

... table 3.5 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
lgexp	-0.024***	0.015**	0.059***	0.082***	0.096***
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)
Union	-0.027**	-0.064***	-0.061***	-0.054***	-0.061***
	(0.012)	(0.010)	(0.007)	(0.007)	(0.007)
lnemployee	-0.057***	-0.070***	-0.053***	-0.064***	-0.185***
	(0.016)	(0.015)	(0.009)	(0.010)	(0.009)
argricuture	-0.003*	-0.003***	-0.002	0.001	-0.008***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
manufactoring	-0.001	-0.014***	-0.012***	-0.011***	-0.019***
	(0.007)	(0.005)	(0.002)	(0.002)	(0.001)
energy	-0.002	-0.007***	-0.001	-0.000	-0.002***
	(0.003)	(0.002)	(0.001)	(0.001)	(0.000)
construction	0.001	-0.003	0.003	0.006*	-0.013***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
trade	0.004	-0.006*	0.005***	0.007***	-0.002
	(0.005)	(0.003)	(0.002)	(0.002)	(0.001)
transport	0.001	-0.003	-0.007***	-0.005***	-0.009***
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
finance etc	0.007**	-0.007***	-0.006***	-0.006***	-0.009***
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
personal service	0.002*	-0.000	-0.004*	-0.002	-0.010***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Constant	-0.394***	-0.293***	-0.299***	-0.372***	0.016
	(0.082)	(0.072)	(0.051)	(0.055)	(0.054)
N	76103	77425	155737	136180	161311

Table 3.6: Full results from Oaxaca-Blinder decomposition for Native-immigrant wage gap in

Medium-skilled group

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	-0.019*** (0.003)	-0.007*** (0.003)	0.015*** (0.001)	0.020*** (0.002)	0.027*** (0.002)
unexplained	0.085*** (0.008)	0.078*** (0.006)	0.065*** (0.004)	0.071*** (0.004)	0.074*** (0.004)
Explained)					
lgage	0.009*** (0.001)	0.007*** (0.000)	0.016*** (0.001)	0.020*** (0.001)	0.019*** (0.001)
Married	-0.006*** (0.001)	-0.002*** (0.001)	-0.000 (0.000)	-0.001** (0.000)	-0.001* (0.000)
children 0-6 yrs	-0.012*** (0.001)	-0.009*** (0.001)	-0.008*** (0.000)	-0.007*** (0.000)	-0.007*** (0.000)
Copenhagen	-0.006*** (0.001)	-0.006*** (0.001)	-0.004*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)
lgexp	0.018*** (0.001)	0.026*** (0.001)	0.017*** (0.001)	0.020*** (0.001)	0.028*** (0.001)
Union	0.001* (0.000)	-0.001*** (0.000)	-0.000 (0.000)	0.001*** (0.000)	-0.001*** (0.000)
Inemployee	-0.006*** (0.001)	-0.008*** (0.001)	-0.006*** (0.000)	-0.008*** (0.000)	-0.007*** (0.000)
agriculture	-0.000* (0.000)	-0.000* (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
manufacturing	0.000 (0.000)	0.000 (0.000)	-0.000** (0.000)	-0.000 (0.000)	-0.000** (0.000)
energy	-0.000 (0.000)	-0.000 (0.000)	0.000** (0.000)	0.000** (0.000)	0.000*** (0.000)

Continued on next page...

... table 3.6 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
construction	-0.001	-0.001**	-0.000	-0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
trade	0.001**	0.002***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
transport	-0.002***	-0.001*	0.001	0.001**	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
finance etc	-0.014***	-0.014***	-0.002***	-0.002***	-0.002***
	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)
personal service	-0.000**	0.001***	0.002***	0.003***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained)					
lgage	-0.111	0.070	0.059	0.120*	0.209***
	(0.134)	(0.104)	(0.068)	(0.068)	(0.067)
Married	0.042***	0.022**	0.028***	0.040***	0.034***
	(0.012)	(0.009)	(0.006)	(0.006)	(0.005)
children 0 6 yrs	0.017***	0.012**	0.009***	0.002	0.006**
	(0.006)	(0.005)	(0.003)	(0.003)	(0.003)
Copenhagen	-0.002	-0.001	0.001	0.005**	0.006***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
lgexp	-0.021**	-0.034***	-0.035***	-0.029***	-0.017*
	(0.009)	(0.011)	(0.009)	(0.010)	(0.010)
Union	-0.061***	-0.012	-0.013	-0.006	-0.036***
	(0.017)	(0.013)	(0.008)	(0.008)	(0.007)
Inemployee	-0.044*	-0.071***	0.006	-0.001	-0.018
	(0.025)	(0.019)	(0.012)	(0.012)	(0.012)
argriculture	0.001	0.003*	-0.000	0.001	-0.000

Continued on next page...

... table 3.6 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
manufacturing	0.008	0.003	-0.004*	-0.004*	-0.004**
	(0.008)	(0.006)	(0.002)	(0.002)	(0.002)
energy	0.002	0.003	-0.001*	0.000	-0.002***
	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)
construction	-0.002	-0.001	0.006**	0.006**	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
trade	0.003	0.008*	0.002	0.003**	0.003**
	(0.006)	(0.004)	(0.002)	(0.001)	(0.001)
transport	0.001	0.002	-0.006***	-0.002	-0.001
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
finance etc	0.000	0.004	-0.003**	-0.001	0.000
	(0.006)	(0.006)	(0.002)	(0.002)	(0.001)
personal service	0.001	0.000	-0.003	-0.005	-0.008***
	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)
Constant	0.252*	0.070	0.020	-0.058	-0.100
	(0.139)	(0.107)	(0.067)	(0.067)	(0.065)
N	160233	161781	385936	365009	397490

Table 3.7: Full results from Oaxaca-Blinder decomposition for Native-immigrant wage gap in

High-skilled group

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	-0.005	-0.015**	0.004	0.003	0.039***
	(0.008)	(0.007)	(0.003)	(0.004)	(0.003)
unexplained	0.199***	0.174***	0.045***	0.065***	0.061***
Explained)					
	(0.018)	(0.014)	(0.008)	(0.008)	(0.007)
lgage	-0.004	-0.001	0.003**	0.004**	0.021***
	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
Married	-0.014***	-0.009***	-0.006***	-0.005***	-0.002**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
children 0-6 yrs	-0.003**	-0.003**	-0.001***	-0.001***	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Copenhagen	-0.002**	-0.000	-0.000	-0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
lgexp	0.018***	0.031***	0.015***	0.016***	0.028***
	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
Union	-0.000	-0.002**	0.003***	0.002***	0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Inemployee	0.001	0.003***	-0.009***	-0.011***	-0.012***
	(0.000)	(0.001)	(0.002)	(0.002)	(0.001)
argricuture	-0.002	0.000	0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
manufacturing	-0.001	-0.004**	-0.003***	-0.003***	-0.003***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)
energy	-0.007***	-0.004**	-0.000	-0.000**	-0.000
	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)

Continued on next page...

... table 3.7 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
construction	0.005** (0.002)	0.001 (0.002)	0.001 (0.001)	-0.000 (0.001)	-0.001*** (0.000)
trade	-0.009*** (0.002)	-0.004* (0.002)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
transport	-0.005* (0.003)	-0.004* (0.002)	0.006*** (0.001)	0.007*** (0.001)	0.008*** (0.001)
finance etc	0.023*** (0.003)	0.009*** (0.002)	0.000 (0.000)	-0.000 (0.000)	-0.001*** (0.000)
personal service	-0.005*** (0.001)	-0.027*** (0.003)	-0.006*** (0.001)	-0.007*** (0.001)	-0.002*** (0.001)
Unexplained)					
lgage	0.994*** (0.330)	0.883*** (0.295)	0.410** (0.161)	0.390** (0.167)	0.249* (0.138)
Married	0.085*** (0.026)	0.050** (0.021)	0.030** (0.012)	0.038*** (0.012)	0.033*** (0.010)
children 0-6 yrs	0.019 (0.013)	0.012 (0.009)	0.000 (0.006)	-0.004 (0.006)	0.004 (0.005)
Copenhagen	-0.007 (0.009)	-0.011 (0.008)	0.014*** (0.004)	0.013*** (0.004)	0.018*** (0.004)
lgexp	-0.034* (0.020)	-0.032 (0.026)	0.018 (0.017)	0.048*** (0.018)	-0.004 (0.017)
Union	-0.155*** (0.046)	-0.077*** (0.029)	-0.067*** (0.018)	-0.048** (0.019)	-0.037** (0.016)
Inemployee	-0.259*** (0.053)	-0.214*** (0.043)	-0.196*** (0.025)	-0.185*** (0.025)	-0.169*** (0.021)
argricuture	0.006* (0.003)	0.003 (0.002)	0.003*** (0.001)	0.003*** (0.001)	0.002*** (0.001)

Continued on next page...

... table 3.7 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
manufacturing	0.008 (0.009)	0.020*** (0.007)	0.006 (0.004)	0.011*** (0.003)	0.003 (0.002)
energy	0.012* (0.007)	0.011** (0.005)	0.002** (0.001)	0.004*** (0.001)	0.001 (0.001)
construction	0.000 (0.004)	0.004 (0.003)	0.015*** (0.003)	0.014*** (0.003)	0.008*** (0.002)
trade	0.004 (0.008)	0.016*** (0.006)	0.006*** (0.002)	0.007*** (0.002)	0.002* (0.001)
transport	0.014* (0.008)	0.015*** (0.005)	0.005 (0.003)	0.009** (0.003)	-0.002 (0.002)
finance etc	-0.049** (0.020)	-0.016 (0.013)	-0.004 (0.003)	0.002 (0.003)	-0.006*** (0.002)
personal service	-0.017** (0.008)	-0.008 (0.007)	-0.021** (0.010)	-0.030*** (0.009)	-0.038*** (0.007)
Constant	-0.422 (0.347)	-0.484 (0.297)	-0.176 (0.155)	-0.208 (0.163)	-0.002 (0.131)
N	8118	12763	41764	42018	55622

Appendix B.2: Full results for Figure 3.7

Table 3.8: Wage differentials and decomposition results across different ethnic groups: all skilled

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
Native-total immigrant)					
difference	0.094*** (0.004)	0.086*** (0.004)	0.085*** (0.002)	0.095*** (0.002)	0.098*** (0.002)
explained	0.058*** (0.002)	0.078*** (0.002)	0.085*** (0.001)	0.091*** (0.001)	0.104*** (0.001)
unexplained	0.036*** (0.004)	0.008** (0.004)	0.000 (0.002)	0.003 (0.002)	-0.006** (0.002)
EU15					
difference	-0.062*** (0.009)	-0.074*** (0.008)	-0.039*** (0.005)	-0.042*** (0.005)	-0.049*** (0.005)
explained	0.016*** (0.004)	0.041*** (0.004)	0.055*** (0.002)	0.059*** (0.002)	0.068*** (0.002)
unexplained	-0.077*** (0.008)	-0.115*** (0.007)	-0.094*** (0.005)	-0.100*** (0.005)	-0.117*** (0.005)
EU12					
difference	0.114*** (0.022)	0.137*** (0.012)	0.136*** (0.006)	0.176*** (0.006)	0.223*** (0.005)
explained	0.080*** (0.012)	0.182*** (0.007)	0.195*** (0.003)	0.210*** (0.003)	0.236*** (0.003)
unexplained	0.034* (0.020)	-0.045*** (0.012)	-0.058*** (0.006)	-0.034*** (0.005)	-0.013*** (0.004)
RDI					
difference	-0.098*** (0.019)	-0.091*** (0.016)	-0.069*** (0.011)	-0.053*** (0.011)	-0.070*** (0.012)

Continued on next page...

... table 3.8 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	-0.056*** (0.010)	0.004 (0.008)	0.025*** (0.005)	0.035*** (0.005)	0.044*** (0.005)
unexplained	-0.042** (0.017)	-0.095*** (0.016)	-0.094*** (0.011)	-0.089*** (0.011)	-0.114*** (0.012)
LDI					
difference	0.181*** (0.005)	0.170*** (0.004)	0.153*** (0.003)	0.160*** (0.003)	0.146*** (0.003)
explained	0.086*** (0.003)	0.092*** (0.002)	0.087*** (0.001)	0.090*** (0.002)	0.089*** (0.002)
unexplained	0.095*** (0.005)	0.077*** (0.004)	0.065*** (0.003)	0.069*** (0.003)	0.057*** (0.003)

Table 3.9: Full results from Oaxaca-Blinder decomposition for Native-EU15 wage gap

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	0.016*** (0.004)	0.041*** (0.004)	0.055*** (0.002)	0.059*** (0.002)	0.068*** (0.002)
unexplained	-0.077*** (0.008)	-0.115*** (0.007)	-0.094*** (0.005)	-0.100*** (0.005)	-0.117*** (0.005)
Explained)					
lgage	0.014*** (0.001)	0.010*** (0.001)	0.017*** (0.001)	0.016*** (0.001)	0.013*** (0.001)
basic_edu	0.016*** (0.003)	0.008** (0.003)	-0.005** (0.002)	0.024*** (0.001)	-0.020*** (0.003)
vocational_edu	0.047*** (0.003)	0.061*** (0.004)	0.061*** (0.003)	0.024*** (0.001)	0.087*** (0.004)
higher_edu	-0.027*** (0.002)	-0.017*** (0.002)	-0.024*** (0.001)	-0.017*** (0.001)	-0.028*** (0.002)
Married	0.001** (0.001)	0.001** (0.001)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)
children_0_6_yrs	-0.015*** (0.001)	-0.011*** (0.001)	-0.008*** (0.000)	-0.007*** (0.000)	-0.005*** (0.000)
Copenhagen	-0.006*** (0.001)	-0.006*** (0.001)	-0.004*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)
lgexp	0.014*** (0.001)	0.026*** (0.001)	0.024*** (0.001)	0.029*** (0.001)	0.039*** (0.001)
Union	0.001** (0.000)	-0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	-0.001*** (0.000)
Inemployee	-0.006*** (0.001)	-0.006*** (0.001)	-0.006*** (0.000)	-0.008*** (0.000)	-0.009*** (0.000)
agricuture	-0.000	-0.000	-0.000***	-0.000***	-0.000***

Continued on next page...

... table 3.9 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
manufacturing	-0.002***	-0.001***	0.000	0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
energy	0.001***	0.000**	-0.000	-0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
construction	-0.004***	-0.005***	-0.001***	-0.002***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
trade	-0.000**	-0.000	0.001***	0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
transport	-0.003***	-0.001	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
finance_etc	-0.016***	-0.018***	-0.003***	-0.004***	-0.004***
	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)
personal_service	0.001***	0.001***	-0.001*	0.001**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained)					
lgage	-0.188	0.192	-0.097	-0.173*	-0.447***
	(0.144)	(0.139)	(0.095)	(0.100)	(0.126)
basic_edu	0.050	0.149***	0.151***	-0.011	0.148***
	(0.088)	(0.022)	(0.029)	(0.012)	(0.052)
vocational_edu	0.043	0.083***	0.090***	0.021***	0.087***
	(0.040)	(0.009)	(0.013)	(0.006)	(0.021)
higher_edu	0.031	0.050***	0.059***	0.020***	0.067***
	(0.019)	(0.005)	(0.008)	(0.003)	(0.014)
Married	-0.008	-0.013	-0.003	0.002	0.011*
	(0.009)	(0.008)	(0.006)	(0.006)	(0.007)
children_0_6_yrs	0.037***	0.018***	0.012***	0.004	-0.011***

Continued on next page...

... table 3.9 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.006)	(0.005)	(0.003)	(0.003)	(0.003)
Copenhagen	0.002	-0.002	0.004	0.003	0.009**
	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
lgexp	0.023**	0.053***	0.072***	0.092***	0.149***
	(0.010)	(0.014)	(0.009)	(0.010)	(0.013)
Union	-0.100***	-0.078***	-0.087***	-0.051***	-0.057***
	(0.023)	(0.017)	(0.013)	(0.012)	(0.012)
Inemployee	-0.065**	-0.127***	-0.086***	-0.127***	-0.151***
	(0.026)	(0.022)	(0.014)	(0.014)	(0.014)
agriculture	-0.002	0.001	0.002**	0.002**	-0.002***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
manufacturing	0.003	0.011**	-0.003	-0.002	-0.021***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
energy	-0.000	0.004	0.003***	0.003***	-0.003***
	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)
construction	0.002	0.007***	0.008***	0.001	-0.027***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
trade	-0.006	0.003	0.014***	0.012***	-0.007***
	(0.007)	(0.005)	(0.003)	(0.003)	(0.002)
transport	0.005	0.007**	-0.005**	-0.002	-0.015***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
finance_etc	0.000	0.002	-0.001	0.000	-0.014***
	(0.007)	(0.006)	(0.002)	(0.002)	(0.002)
personal_service	0.004	0.003	0.007*	0.011***	-0.022***
	(0.003)	(0.002)	(0.004)	(0.004)	(0.003)
Constant	0.094	-0.478***	-0.235**	0.094	0.189
	(0.207)	(0.151)	(0.114)	(0.100)	(0.154)

Continued on next page...

... table 3.9 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
N	240537	245712	566100	556698	588927

Table 3.10: Full results from Oaxaca-Blinder decomposition for Native-EU12 wage gap

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	0.080*** (0.012)	0.182*** (0.007)	0.195*** (0.003)	0.210*** (0.003)	0.236*** (0.003)
unexplained	0.034* (0.020)	-0.045*** (0.012)	-0.058*** (0.006)	-0.034*** (0.005)	-0.013*** (0.004)
Explained)					
lgage	0.037*** (0.004)	0.039*** (0.002)	0.052*** (0.001)	0.057*** (0.001)	0.053*** (0.001)
basic_edu	0.014*** (0.003)	0.010*** (0.004)	-0.008** (0.004)	0.038*** (0.001)	-0.031*** (0.005)
vocational_edu	0.041*** (0.004)	0.068*** (0.004)	0.082*** (0.004)	0.034*** (0.001)	0.116*** (0.005)
higher_edu	-0.023*** (0.005)	-0.002 (0.002)	0.001 (0.001)	0.001 (0.001)	0.010*** (0.001)
Married	-0.001 (0.002)	0.006*** (0.001)	0.006*** (0.001)	0.005*** (0.001)	0.004*** (0.000)
children_0_6_yrs	0.006** (0.002)	0.009*** (0.001)	0.005*** (0.001)	0.003*** (0.001)	0.001*** (0.000)
Copenhagen	-0.002*** (0.001)	-0.003*** (0.001)	-0.000*** (0.000)	-0.000 (0.000)	-0.000*** (0.000)
lgexp	0.016*** (0.002)	0.054*** (0.002)	0.051*** (0.001)	0.062*** (0.001)	0.078*** (0.001)
Union	0.001*** (0.000)	-0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	-0.001*** (0.000)
Inemployee	-0.004* (0.002)	0.002* (0.001)	-0.005*** (0.001)	-0.004*** (0.001)	0.003*** (0.000)
agriculture	0.000	-0.000	0.004***	0.006***	0.001***

Continued on next page...

... table 3.10 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
manufacturing	0.000	0.001***	-0.000	0.000*	-0.000**
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
energy	0.000	-0.001**	-0.000	-0.000	0.000***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
construction	-0.003***	0.001	-0.001***	-0.002***	-0.000**
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
trade	-0.000	0.000**	0.000***	0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
transport	0.005***	0.005***	0.009***	0.011***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
finance_etc	-0.007*	-0.003	0.001***	0.002***	0.002***
	(0.004)	(0.002)	(0.000)	(0.000)	(0.000)
personal_service	-0.000	-0.003***	-0.005***	-0.007***	-0.010***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Unexplained)					
lgage	0.292	0.187	0.310***	0.558***	0.185***
	(0.299)	(0.174)	(0.085)	(0.088)	(0.070)
basic_edu	-0.065	0.036	0.303***	-0.083***	0.334***
	(0.048)	(0.025)	(0.026)	(0.014)	(0.021)
vocational_edu	-0.010	0.029***	0.052***	-0.005**	0.043***
	(0.031)	(0.003)	(0.005)	(0.002)	(0.003)
higher_edu	0.012	0.019***	0.021***	-0.009***	0.026***
	(0.011)	(0.004)	(0.001)	(0.002)	(0.001)
Married	0.035	-0.003	0.008	0.008	0.027***
	(0.024)	(0.010)	(0.006)	(0.006)	(0.005)
children_0_6_yrs	0.003	0.005	0.007**	0.005*	0.001

Continued on next page...

... table 3.10 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.008)	(0.004)	(0.003)	(0.003)	(0.002)
Copenhagen	-0.009	0.004	0.003	0.001	0.006**
	(0.009)	(0.005)	(0.002)	(0.002)	(0.003)
Igexp	-0.037	0.012	0.017***	0.014**	0.056***
	(0.023)	(0.012)	(0.005)	(0.006)	(0.005)
Union	-0.099**	0.007	-0.002	0.004	0.011
	(0.044)	(0.035)	(0.026)	(0.024)	(0.015)
Inemployee	-0.100*	-0.156***	-0.107***	-0.130***	-0.129***
	(0.060)	(0.047)	(0.019)	(0.019)	(0.015)
agriculture	0.000	-0.011***	-0.021***	-0.009*	-0.035***
	(0.004)	(0.003)	(0.006)	(0.005)	(0.003)
manufacturing	0.009	-0.012	-0.009***	-0.012***	-0.020***
	(0.021)	(0.011)	(0.003)	(0.003)	(0.002)
energy	0.005	-0.001	-0.001	0.000	-0.003***
	(0.008)	(0.003)	(0.001)	(0.001)	(0.001)
construction	0.004	-0.027***	-0.004	-0.002	-0.011***
	(0.011)	(0.009)	(0.004)	(0.003)	(0.003)
trade	0.017	-0.003	0.002	0.002	-0.005***
	(0.015)	(0.006)	(0.002)	(0.002)	(0.002)
transport	0.009	0.003	-0.007***	-0.002	-0.003***
	(0.006)	(0.003)	(0.002)	(0.002)	(0.001)
finance_etc	0.001	-0.008	-0.000	0.001	0.001
	(0.012)	(0.007)	(0.001)	(0.001)	(0.001)
personal_service	0.004	-0.002	-0.022***	-0.020***	-0.017***
	(0.005)	(0.002)	(0.004)	(0.003)	(0.002)
Constant	-0.038	-0.125	-0.607***	-0.353***	-0.482***
	(0.311)	(0.182)	(0.100)	(0.092)	(0.077)

Continued on next page...

... table 3.10 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
N	238442	243413	561615	552407	586254

Table 3.11: Full results from Oaxaca-Blinder decomposition for Native-RDI wage gap

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	-0.056*** (0.010)	0.004 (0.008)	0.025*** (0.005)	0.035*** (0.005)	0.044*** (0.005)
unexplained	-0.042** (0.017)	-0.095*** (0.016)	-0.094*** (0.011)	-0.089*** (0.011)	-0.114*** (0.012)
Explained)					
lgage	0.014*** (0.002)	0.011*** (0.002)	0.014*** (0.001)	0.013*** (0.001)	0.009*** (0.001)
basic_edu	0.009*** (0.002)	0.007*** (0.003)	-0.004** (0.002)	0.022*** (0.001)	-0.017*** (0.003)
vocational_edu	0.039*** (0.004)	0.056*** (0.004)	0.057*** (0.003)	0.024*** (0.001)	0.082*** (0.004)
higher_edu	-0.046*** (0.005)	-0.024*** (0.004)	-0.031*** (0.003)	-0.021*** (0.002)	-0.033*** (0.003)
Married	-0.010*** (0.001)	-0.007*** (0.001)	-0.004*** (0.001)	-0.005*** (0.001)	-0.006*** (0.001)
children_0_6_yrs	-0.022*** (0.002)	-0.020*** (0.002)	-0.009*** (0.001)	-0.009*** (0.001)	-0.007*** (0.001)
Copenhagen	-0.007*** (0.001)	-0.008*** (0.001)	-0.005*** (0.000)	-0.007*** (0.001)	-0.008*** (0.001)
lgexp	0.015*** (0.001)	0.034*** (0.002)	0.025*** (0.001)	0.031*** (0.001)	0.040*** (0.001)
Union	0.001*** (0.000)	-0.001*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	-0.001*** (0.000)
Inemployee	-0.008*** (0.002)	-0.009*** (0.002)	-0.010*** (0.001)	-0.009*** (0.001)	-0.013*** (0.001)
agricuture	0.000	0.000	-0.000***	-0.001***	-0.000***

Continued on next page...

... table 3.11 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
manufacturing	-0.003***	-0.002***	0.000	0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
energy	0.000	0.000	-0.000	-0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
construction	-0.005***	-0.005***	-0.001***	-0.003***	-0.001***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
trade	0.000	0.000	0.001***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
transport	-0.002	0.001	-0.002**	-0.000	0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
finance_etc	-0.034***	-0.031***	-0.005***	-0.005***	-0.005***
	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)
personal_service	0.001**	0.003***	-0.000	0.002**	0.002**
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Unexplained)					
lgage	-0.059	-0.479	-0.325	-0.501**	-0.237
	(0.349)	(0.304)	(0.223)	(0.241)	(0.271)
basic_edu	0.059	0.144***	0.169***	-0.014	-0.010
	(0.050)	(0.038)	(0.025)	(0.024)	(0.015)
vocational_edu	0.079**	0.108***	0.134***	0.027**	0.053***
	(0.037)	(0.018)	(0.013)	(0.013)	(0.003)
higher_edu	0.065***	0.064***	0.089***	0.035***	0.052***
	(0.021)	(0.011)	(0.009)	(0.008)	(0.006)
Married	0.066**	0.041*	0.020	0.019	0.020
	(0.029)	(0.022)	(0.017)	(0.016)	(0.018)
children_0_6_yrs	0.031**	0.029**	0.005	0.013*	0.003

Continued on next page...

... table 3.11 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.014)	(0.012)	(0.007)	(0.007)	(0.007)
Copenhagen	0.006	-0.019*	-0.002	0.004	0.012
	(0.011)	(0.010)	(0.007)	(0.008)	(0.010)
Igexp	-0.033	0.059***	0.072***	0.112***	0.134***
	(0.021)	(0.022)	(0.020)	(0.021)	(0.028)
Union	-0.089***	-0.127***	-0.088***	-0.106***	-0.079***
	(0.033)	(0.034)	(0.023)	(0.023)	(0.024)
Inemployee	-0.069	-0.074	-0.109***	-0.115***	-0.157***
	(0.056)	(0.047)	(0.030)	(0.030)	(0.030)
agriculture	0.001	-0.002	0.001	0.003	-0.001
	(0.004)	(0.003)	(0.002)	(0.002)	(0.001)
manufacturing	0.015	-0.012	-0.008	-0.009*	-0.026***
	(0.010)	(0.010)	(0.008)	(0.005)	(0.005)
energy	0.006	-0.009	0.002	0.000	-0.005**
	(0.007)	(0.007)	(0.002)	(0.002)	(0.002)
construction	0.006	-0.005	0.005	0.000	-0.010*
	(0.004)	(0.005)	(0.007)	(0.005)	(0.005)
trade	0.004	-0.018	0.007	0.002	-0.013**
	(0.013)	(0.011)	(0.007)	(0.006)	(0.005)
transport	0.004	-0.003	-0.012*	-0.012**	-0.019***
	(0.009)	(0.006)	(0.007)	(0.005)	(0.005)
finance_etc	-0.012	-0.032*	-0.004	-0.010**	-0.016***
	(0.020)	(0.018)	(0.007)	(0.005)	(0.005)
personal_service	0.001	-0.003	0.015	0.006	-0.023***
	(0.007)	(0.006)	(0.011)	(0.009)	(0.008)
Constant	-0.124	0.241	-0.067	0.458*	0.206
	(0.387)	(0.372)	(0.244)	(0.237)	(0.265)

Continued on next page...

... table 3.11 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
N	238666	243336	559786	550222	581104

Table 3.12: Full results from Oaxaca-Blinder decomposition for Native-LDI wage gap

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
explained	0.086*** (0.003)	0.092*** (0.002)	0.087*** (0.001)	0.090*** (0.002)	0.089*** (0.002)
unexplained	0.095*** (0.005)	0.077*** (0.004)	0.065*** (0.003)	0.069*** (0.003)	0.057*** (0.003)
Explained)					
lgage	0.034*** (0.001)	0.027*** (0.001)	0.029*** (0.001)	0.031*** (0.001)	0.027*** (0.001)
basic_edu	0.022*** (0.003)	0.008** (0.003)	-0.006** (0.003)	0.024*** (0.001)	-0.018*** (0.003)
vocational_edu	0.050*** (0.003)	0.057*** (0.003)	0.056*** (0.003)	0.023*** (0.001)	0.069*** (0.003)
higher_edu	-0.005*** (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	-0.007*** (0.001)
Married	-0.008*** (0.000)	-0.006*** (0.000)	-0.004*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)
children_0_6_yrs	-0.019*** (0.001)	-0.016*** (0.001)	-0.010*** (0.000)	-0.010*** (0.000)	-0.009*** (0.000)
Copenhagen	-0.004*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.004*** (0.000)	-0.004*** (0.000)
lgexp	0.022*** (0.001)	0.038*** (0.001)	0.025*** (0.001)	0.029*** (0.001)	0.039*** (0.001)
Union	0.001*** (0.000)	-0.001*** (0.000)	0.001*** (0.000)	0.002*** (0.000)	-0.001*** (0.000)
Inemployee	-0.004*** (0.000)	-0.006*** (0.000)	-0.007*** (0.000)	-0.008*** (0.000)	-0.007*** (0.000)
agriculture	-0.000***	-0.000	0.001***	0.001***	0.000***

Continued on next page...

... table 3.12 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
manufacturing	0.002***	0.002***	0.000	0.000*	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
energy	0.001***	0.001***	-0.000	-0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
construction	-0.005***	-0.005***	-0.000***	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
trade	0.000	0.000*	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
transport	-0.000	-0.001***	0.006***	0.007***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
finance_etc	-0.001	0.000	0.001***	0.001***	0.000*
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
personal_service	-0.000***	-0.002***	-0.001**	-0.000	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained)					
lgage	0.399***	0.372***	0.350***	0.410***	0.479***
	(0.074)	(0.064)	(0.045)	(0.046)	(0.045)
basic_edu	0.271***	0.062***	0.210***	-0.037***	0.155***
	(0.016)	(0.008)	(0.011)	(0.006)	(0.008)
vocational_edu	0.110***	0.040***	0.118***	-0.001	0.116***
	(0.006)	(0.002)	(0.005)	(0.003)	(0.005)
higher_edu	0.030***	0.015***	0.030***	0.002*	0.043***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Married	0.023***	0.029***	0.030***	0.036***	0.030***
	(0.007)	(0.006)	(0.004)	(0.004)	(0.004)
children_0_6_yrs	0.026***	0.026***	0.020***	0.018***	0.012***

Continued on next page...

... table 3.12 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
Copenhagen	-0.002	-0.003	0.003*	0.004***	0.008***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Igexp	-0.025***	-0.009	0.025***	0.045***	0.054***
	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)
Union	-0.003	-0.004	0.000	-0.004	-0.027***
	(0.011)	(0.009)	(0.006)	(0.005)	(0.005)
Inemployee	-0.045***	-0.054***	-0.028***	-0.030***	-0.096***
	(0.015)	(0.013)	(0.009)	(0.009)	(0.009)
argricuture	-0.003**	-0.001	-0.004***	-0.004***	-0.003***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
manufactoring	-0.013*	-0.017***	-0.007***	-0.007***	-0.004***
	(0.008)	(0.005)	(0.002)	(0.002)	(0.001)
energy	-0.003	-0.004*	-0.002***	-0.000	-0.000
	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)
construction	-0.001	-0.000	0.006**	0.006**	0.009***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
trade	0.008*	0.007**	0.001	0.002**	0.004***
	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)
transport	-0.001	-0.004*	-0.002**	-0.001	-0.002**
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
finance_etc	0.007**	0.003	-0.002***	0.000	-0.001
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
personal_service	0.000	-0.001	-0.015***	-0.015***	-0.010***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Constant	-0.680***	-0.379***	-0.669***	-0.355***	-0.710***
	(0.091)	(0.066)	(0.054)	(0.047)	(0.046)

Continued on next page...

... table 3.12 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
N	243607	249575	574104	564698	598052

Appendix C: Statistical inference for estimations in Melly(2005) approach

in Section 3.6.2

Appendix C.1: Statistical inference for Figure 3.8 (native-total immigrants wage gap)

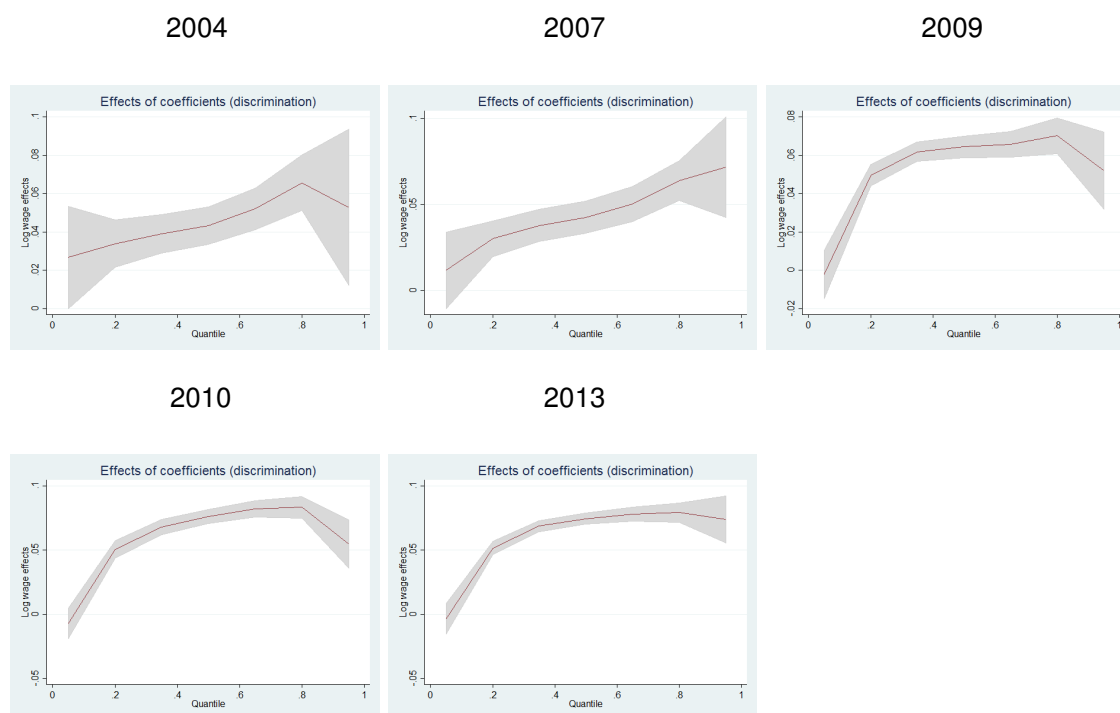


Figure 3.15: Statistical inference for native-total immigrants wage gap decomposition in all skill levels 2004-2013

Appendix C.2: Statistical inference for Figure 3.9 and Figure 3.10 (native-immigrants wage gap under different skill levels)

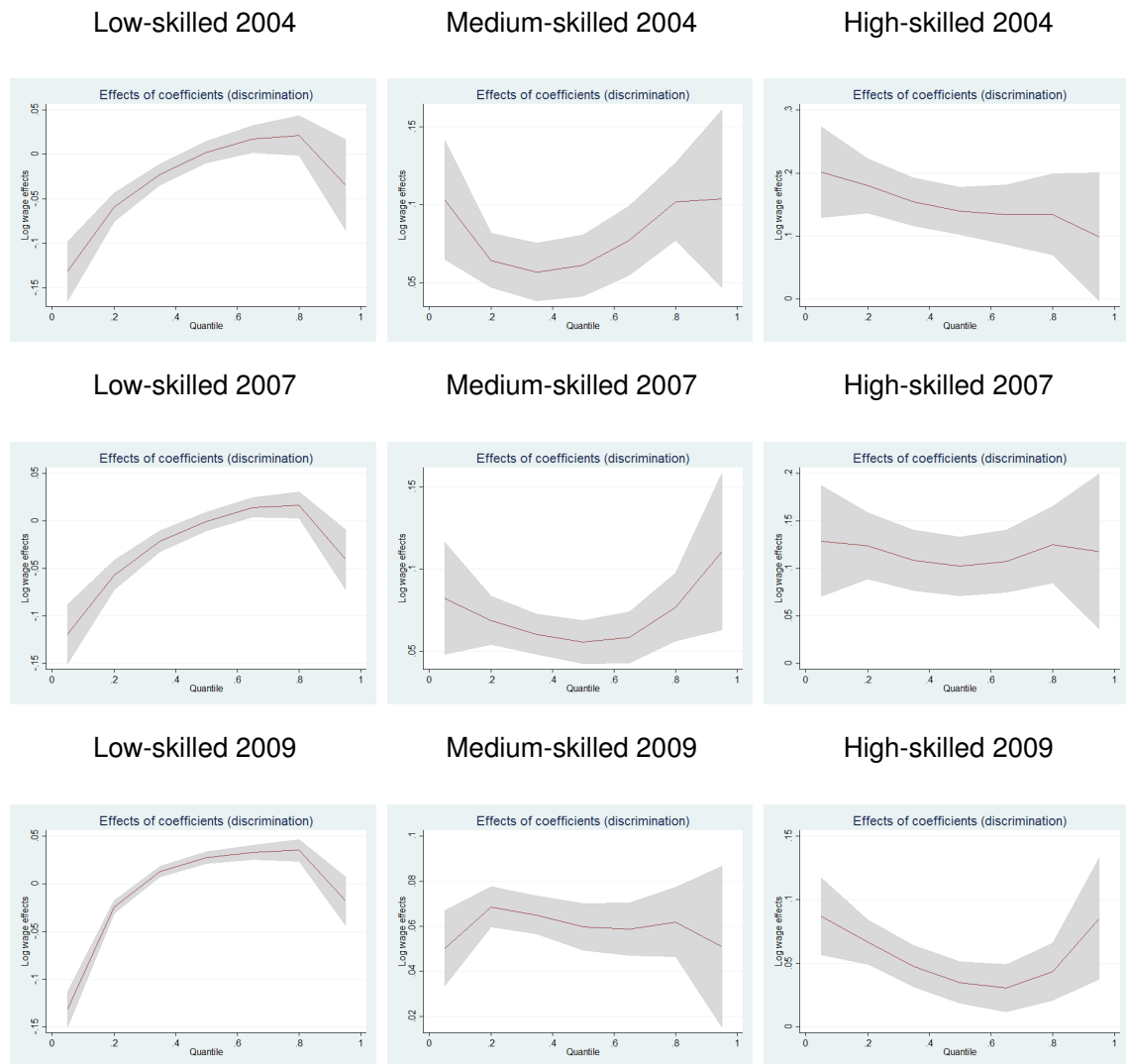


Figure 3.16: Statistical inference of native-immigrants wage gap decomposition in different skill levels 2004-2009

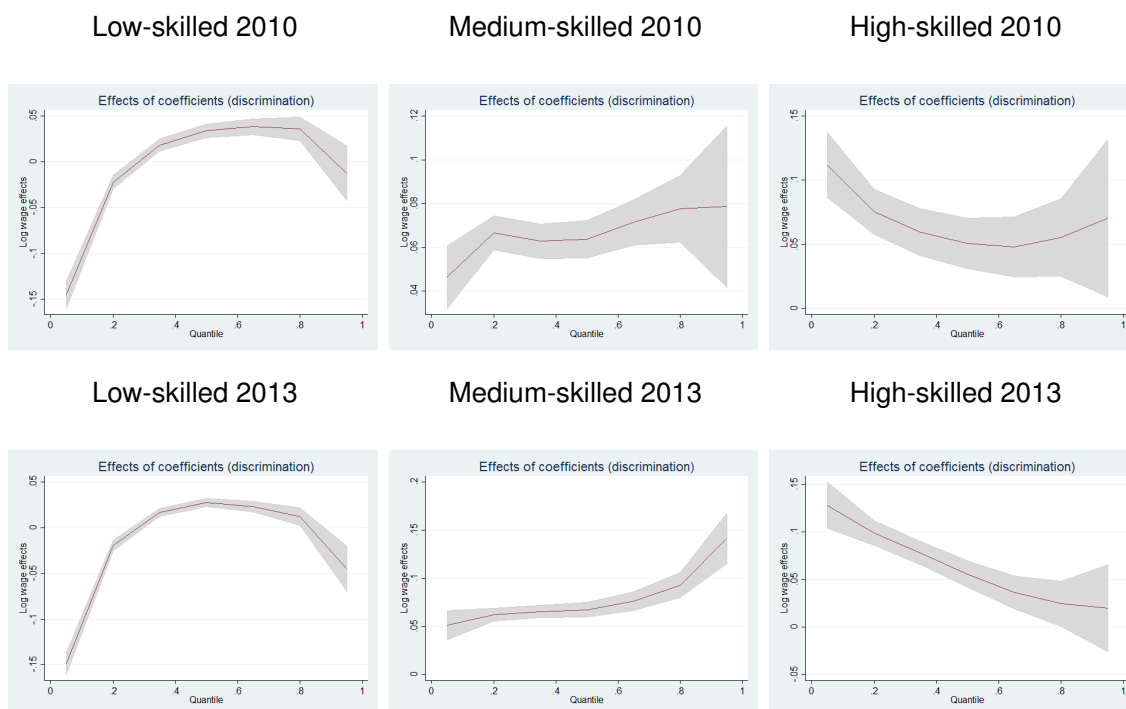


Figure 3.17: Statistical inference of native-immigrants wage gap decomposition in different skill levels 2010-2013

Appendix C.3: Statistical inference for Figure 3.12 and Figure 3.13(native-immigrants wage gap under different groups of countries of origin)



Figure 3.18: Statistical inference of native-immigrants wage gap decomposition in all skill levels under each immigrant group (EU12 New EU12 and Less Developed countries (LD)) 2004-2009

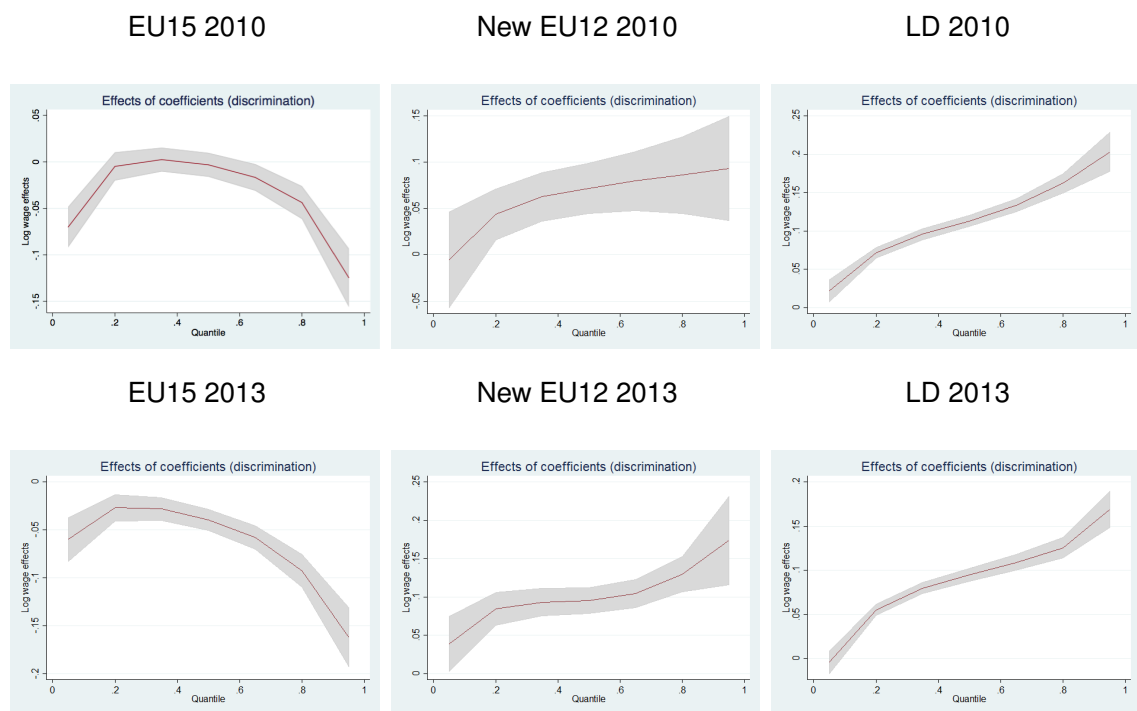


Figure 3.19: Statistical inference of native-immigrants wage gap decomposition in all skill levels under each immigrant group (EU12 New EU12 and Less Developed countries (LD)) 2010-2013

Appendix D: Full results (beta-estimates and bootstrapped standard errors) for decomposition results from Melly (2005) approach-From STATA log files

Appendix D.0: Cumulative density function of wage(log) for the empirical analysis in Section 3.6.

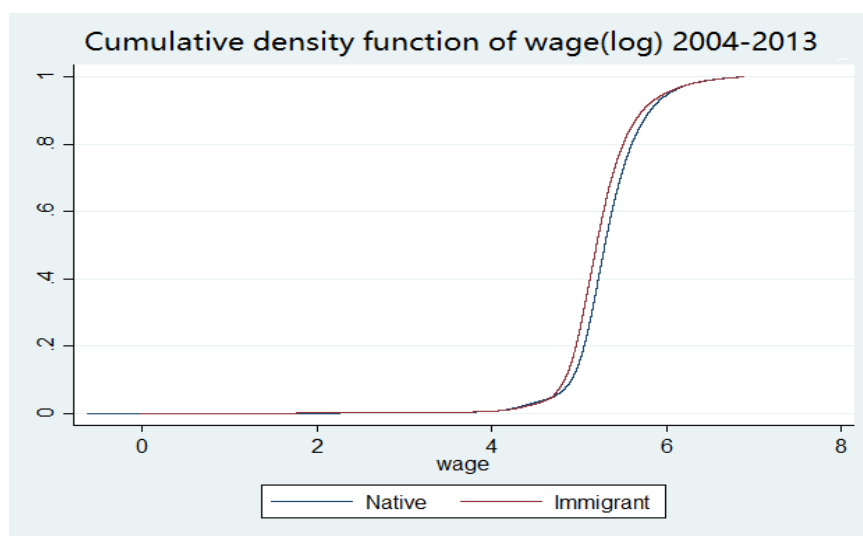


Figure 3.20: Cumulative density function of wage (log) 2004-2013

Appendix D.1: Full results for Figure 3.8(native-total immigrants wage gap)

***Quantile regression estimation results and bootstrapped standard errors

**Native-total immigrants for all skilled levels

**Year 2004

. rqdeco wage `X' if YEAR==2004, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)

Fitting base model

(bootstrapping)

Decomposition of differences in distribution using quantile regression

Total number of observations 246910
 Number of observations in group 0 8796
 Number of observations in group 1 238114

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.075853	.009276	8.18	0.000	.057673	.094033
Characteristics	.049143	.009447	5.20	0.000	.030628	.067658
Coefficients	.02671	.0136	1.96	0.050	.000054	.053366
Quantile .2						
Raw difference	.095188	.004279	22.25	0.000	.086802	.103575
Characteristics	.06132	.005204	11.78	0.000	.05112	.07152
Coefficients	.033869	.006306	5.37	0.000	.02151	.046227
Quantile .35						
Raw difference	.10195	.00324	31.46	0.000	.095599	.108301

Characteristics	.062947	.004399	14.31	0.000	.054325	.071569
Coefficients	.039003	.005072	7.69	0.000	.029063	.048944
<hr/>						
Quantile .5						
Raw difference	.104849	.003052	34.35	0.000	.098867	.110831
Characteristics	.061533	.004274	14.40	0.000	.053155	.06991
Coefficients	.043316	.004962	8.73	0.000	.033592	.053041
<hr/>						
Quantile .65						
Raw difference	.108445	.003698	29.32	0.000	.101197	.115694
Characteristics	.056448	.004952	11.40	0.000	.046743	.066154
Coefficients	.051997	.005483	9.48	0.000	.04125	.062744
<hr/>						
Quantile .8						
Raw difference	.111653	.005631	19.83	0.000	.100617	.122688
Characteristics	.046055	.006984	6.59	0.000	.032367	.059743
Coefficients	.065598	.007382	8.89	0.000	.051129	.080066
<hr/>						
Quantile .95						
Raw difference	.066998	.013697	4.89	0.000	.040153	.093843
Characteristics	.014168	.017202	0.82	0.496	-.019547	.047883
Coefficients	.05283	.020821	2.54	0.011	.012021	.093639
<hr/>						

```
. *** Year 2007
. rqdeco wage `X' if YEAR==2007, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

```
Decomposition of differences in distribution using quantile regression
Total number of observations      254158
Number of observations in group 0    11532
Number of observations in group 1    242626
```

Number of quantile regressions estimated

100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.069246	.008636	8.02	0.000	.05232	.086173
Characteristics	.05765	.008599	6.70	0.000	.040796	.074505
Coefficients	.011596	.011333	1.02	0.306	-.010617	.033809
Quantile .2						
Raw difference	.091203	.00411	22.19	0.000	.083148	.099258
Characteristics	.061237	.004428	13.83	0.000	.052558	.069916
Coefficients	.029966	.005314	5.64	0.000	.019551	.040381
Quantile .35						
Raw difference	.095754	.003755	25.50	0.000	.088394	.103114
Characteristics	.058057	.003653	15.89	0.000	.050897	.065217
Coefficients	.037697	.004723	7.98	0.000	.028441	.046953
Quantile .5						
Raw difference	.094099	.00389	24.19	0.000	.086475	.101723
Characteristics	.051591	.003646	14.15	0.000	.044446	.058737
Coefficients	.042508	.004737	8.97	0.000	.033224	.051792
Quantile .65						
Raw difference	.093455	.004345	21.51	0.000	.084939	.101971
Characteristics	.043368	.004212	10.30	0.000	.035112	.051624
Coefficients	.050087	.005176	9.68	0.000	.039943	.060231
Quantile .8						
Raw difference	.094153	.005321	17.70	0.000	.083725	.104581
Characteristics	.030319	.005236	5.79	0.000	.020055	.040582

Coefficients	.063835	.005859	10.90	0.000	.052352	.075317
Quantile .95						
Raw difference	.068535	.011184	6.13	0.000	.046614	.090456
Characteristics	-.003009	.010616	-0.28	0.840	-.023817	.017798
Coefficients	.071544	.014933	4.79	0.000	.042275	.100813

. **2009

.
.
.

. rqdeco wage `X' if YEAR==2009, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping)

Decomposition of differences in distribution using quantile regression

Total number of observations 587218
Number of observations in group 0 29089
Number of observations in group 1 558129

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.057348	.00464	12.36	0.000	.048253	.066444
Characteristics	.059522	.005504	10.81	0.000	.048735	.07031
Coefficients	-.002174	.006444	-0.34	0.736	-.014804	.010456

Quantile .2						
Raw difference	.08942	.001965	45.50	0.000	.085568	.093272
Characteristics	.039814	.002588	15.39	0.000	.034742	.044886
Coefficients	.049606	.002856	17.37	0.000	.044008	.055203
Quantile .35						
Raw difference	.098446	.001781	55.27	0.000	.094955	.101937
Characteristics	.036567	.00222	16.47	0.000	.032216	.040918
Coefficients	.061879	.002601	23.79	0.000	.056782	.066976
Quantile .5						
Raw difference	.099908	.001935	51.62	0.000	.096115	.103702
Characteristics	.035517	.002278	15.59	0.000	.031053	.039981
Coefficients	.064391	.002865	22.47	0.000	.058776	.070007
Quantile .65						
Raw difference	.100427	.002455	40.91	0.000	.095616	.105238
Characteristics	.034697	.002607	13.31	0.000	.029587	.039807
Coefficients	.06573	.003448	19.06	0.000	.058972	.072489
Quantile .8						
Raw difference	.099802	.003737	26.71	0.000	.092478	.107125
Characteristics	.029605	.003813	7.76	0.000	.022132	.037078
Coefficients	.070197	.004729	14.84	0.000	.060929	.079465
Quantile .95						
Raw difference	.055948	.009049	6.18	0.000	.038212	.073684
Characteristics	.004066	.008284	0.49	0.693	-.012171	.020302
Coefficients	.051883	.01029	5.04	0.000	.031715	.072051

. **2010

.

.

```

.   rqdeco wage `X' if YEAR==2010, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      578504
Number of observations in group 0    29997
Number of observations in group 1   548507
```

```

Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.06022	.004983	12.09	0.000	.050453	.069986
Characteristics	.067323	.005582	12.06	0.000	.056382	.078263
Coefficients	-.007103	.006118	-1.16	0.246	-.019095	.004889
Quantile .2						
Raw difference	.091282	.001983	46.02	0.000	.087394	.095169
Characteristics	.04055	.002587	15.67	0.000	.035479	.045621
Coefficients	.050732	.003492	14.53	0.000	.043888	.057575
Quantile .35						
Raw difference	.102899	.001997	51.53	0.000	.098986	.106812
Characteristics	.034701	.002155	16.10	0.000	.030478	.038925
Coefficients	.068198	.003018	22.60	0.000	.062283	.074112
Quantile .5						
Raw difference	.108024	.00212	50.94	0.000	.103868	.11218

Characteristics	.031863	.002087	15.26	0.000	.027772	.035954
Coefficients	.076161	.002863	26.61	0.000	.070551	.081772
<hr/>						
Quantile .65						
Raw difference	.111317	.002328	47.82	0.000	.106754	.115879
Characteristics	.029222	.002344	12.47	0.000	.024628	.033816
Coefficients	.082094	.003231	25.41	0.000	.075762	.088427
<hr/>						
Quantile .8						
Raw difference	.107955	.003404	31.72	0.000	.101284	.114626
Characteristics	.024414	.003146	7.76	0.000	.018248	.03058
Coefficients	.083541	.00434	19.25	0.000	.075034	.092048
<hr/>						
Quantile .95						
Raw difference	.061123	.007968	7.67	0.000	.045506	.07674
Characteristics	.006459	.007291	0.89	0.497	-.007832	.02075
Coefficients	.054664	.009504	5.75	0.000	.036036	.073292

. **2013

```
.
.   rqdeco wage `X' if YEAR==2013, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

```
Decomposition of differences in distribution using quantile regression
Total number of observations      617376
Number of observations in group 0    38389
Number of observations in group 1   578987

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.064687	.00476	13.59	0.000	.055358	.074016
Characteristics	.068029	.004908	13.86	0.000	.05841	.077647
Coefficients	-.003342	.005999	-0.56	0.578	-.0151	.008417
Quantile .2						
Raw difference	.09552	.001878	50.87	0.000	.091839	.0992
Characteristics	.043787	.002092	20.93	0.000	.039688	.047887
Coefficients	.051732	.002607	19.84	0.000	.046622	.056843
Quantile .35						
Raw difference	.103882	.001659	62.63	0.000	.100631	.107133
Characteristics	.035134	.001592	22.07	0.000	.032014	.038254
Coefficients	.068748	.002211	31.10	0.000	.064415	.07308
Quantile .5						
Raw difference	.104277	.001886	55.29	0.000	.10058	.107974
Characteristics	.029652	.001573	18.86	0.000	.02657	.032734
Coefficients	.074625	.002286	32.64	0.000	.070144	.079106
Quantile .65						
Raw difference	.103328	.002382	43.37	0.000	.098658	.107997
Characteristics	.025103	.001883	13.33	0.000	.021413	.028794
Coefficients	.078225	.002778	28.16	0.000	.07278	.083669
Quantile .8						
Raw difference	.096989	.00365	26.57	0.000	.089835	.104143
Characteristics	.017648	.002829	6.24	0.000	.012103	.023194
Coefficients	.079341	.00389	20.40	0.000	.071716	.086965
Quantile .95						
Raw difference	.055899	.008466	6.60	0.000	.039306	.072493

Characteristics	-.018097	.006365	-2.84	0.051	-.030572	-.005623
Coefficients	.073997	.009272	7.98	0.000	.055824	.092169

.

Appendix D.2: Full results for Figure 3.9(native-immigrants wage gap under different skill levels)

***Quantile regression estimation results and bootstrapped standard errors

**Native-total immigrants in low-skilled group

. *2004

```

.   rqdeco wage `X' if YEAR==2004&basic_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      76103
Number of observations in group 0   5787
Number of observations in group 1  70316
```

```

Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	

Quantile .05						
Raw difference	-.143252	.014457	-9.91	0.000	-.171588	-.114916
Characteristics	-.011964	.007858	-1.52	0.481	-.027365	.003437
Coefficients	-.131288	.016972	-7.74	0.000	-.164552	-.098023

Quantile .2						
Raw difference	-.039921	.006365	-6.27	0.000	-.052397	-.027445
Characteristics	.01964	.005892	3.33	0.016	.008092	.031187
Coefficients	-.05956	.008137	-7.32	0.000	-.075508	-.043613

Quantile .35						
Raw difference	.004824	.004152	1.16	0.245	-.003315	.012962
Characteristics	.027344	.004789	5.71	0.000	.017959	.03673
Coefficients	-.022521	.005997	-3.76	0.000	-.034275	-.010766
Quantile .5						
Raw difference	.03207	.003773	8.50	0.000	.024675	.039466
Characteristics	.029544	.0047	6.29	0.000	.020333	.038756
Coefficients	.002526	.006076	0.42	0.678	-.009382	.014435
Quantile .65						
Raw difference	.046821	.004736	9.89	0.000	.037538	.056105
Characteristics	.029705	.005134	5.79	0.000	.019643	.039768
Coefficients	.017116	.007688	2.23	0.026	.002048	.032184
Quantile .8						
Raw difference	.052776	.007459	7.08	0.000	.038157	.067394
Characteristics	.031503	.006576	4.79	0.005	.018615	.044392
Coefficients	.021273	.01129	1.88	0.060	-.000856	.043401
Quantile .95						
Raw difference	.000989	.018152	0.05	0.957	-.034588	.036566
Characteristics	.035324	.016956	2.08	0.173	.00209	.068558
Coefficients	-.034335	.025912	-1.33	0.185	-.085123	.016452

. *2009

```
.
. rqdeco wage `X' if YEAR==2009&basic_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

Total number of observations 155737
Number of observations in group 0 18511
Number of observations in group 1 137226

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	-.125832	.008501	-14.80	0.000	-.142492	-.109171
Characteristics	.0057	.004534	1.26	0.545	-.003186	.014586
Coefficients	-.131532	.009424	-13.96	0.000	-.150003	-.11306
Quantile .2						
Raw difference	-.026231	.002446	-10.72	0.000	-.031026	-.021437
Characteristics	-.002065	.002196	-0.94	0.542	-.006369	.002238
Coefficients	-.024166	.003387	-7.14	0.000	-.030804	-.017528
Quantile .35						
Raw difference	.007571	.002067	3.66	0.000	.003521	.011621
Characteristics	-.005525	.001983	-2.79	0.046	-.009412	-.001638
Coefficients	.013096	.002775	4.72	0.000	.007657	.018535
Quantile .5						
Raw difference	.020542	.002389	8.60	0.000	.015859	.025225
Characteristics	-.00717	.002115	-3.39	0.020	-.011315	-.003024
Coefficients	.027712	.003072	9.02	0.000	.021691	.033732
Quantile .65						
Raw difference	.025713	.002729	9.42	0.000	.020364	.031062
Characteristics	-.007433	.002612	-2.85	0.051	-.012552	-.002313

Coefficients	.033146	.003816	8.68	0.000	.025665	.040626
Quantile .8						
Raw difference	.030239	.004137	7.31	0.000	.022131	.038348
Characteristics	-.004915	.003567	-1.38	0.401	-.011906	.002076
Coefficients	.035154	.00585	6.01	0.000	.023688	.046621
Quantile .95						
Raw difference	-.033182	.011537	-2.88	0.004	-.055795	-.010569
Characteristics	-.015465	.007268	-2.13	0.231	-.02971	-.001219
Coefficients	-.017718	.012915	-1.37	0.170	-.043031	.007596

. *2010

```
.
. rqdeco wage `X' if YEAR==2010&basic_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      136180
Number of observations in group 0    16680
Number of observations in group 1   119500
```

```
Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]
Quantile .05					
Raw difference	-.127849	.006981	-18.31	0.000	-.141531 -.114166

Characteristics	.017461	.00476	3.67	0.023	.008132	.026789
Coefficients	-.14531	.007668	-18.95	0.000	-.160338	-.130281
Quantile .2						
Raw difference	-.022979	.003059	-7.51	0.000	-.028975	-.016983
Characteristics	-.00071	.002853	-0.25	0.849	-.006302	.004883
Coefficients	-.022269	.00373	-5.97	0.000	-.02958	-.014959
Quantile .35						
Raw difference	.009868	.002748	3.59	0.000	.004483	.015254
Characteristics	-.008167	.002513	-3.25	0.018	-.013093	-.003241
Coefficients	.018035	.00344	5.24	0.000	.011292	.024778
Quantile .5						
Raw difference	.022611	.002633	8.59	0.000	.017451	.027771
Characteristics	-.011319	.002461	-4.60	0.002	-.016142	-.006496
Coefficients	.03393	.003711	9.14	0.000	.026657	.041203
Quantile .65						
Raw difference	.028702	.002852	10.06	0.000	.023113	.034292
Characteristics	-.009756	.003078	-3.17	0.028	-.015789	-.003723
Coefficients	.038458	.004444	8.65	0.000	.029747	.04717
Quantile .8						
Raw difference	.032522	.004401	7.39	0.000	.023897	.041147
Characteristics	-.003358	.004535	-0.74	0.603	-.012246	.00553
Coefficients	.03588	.006453	5.56	0.000	.023233	.048527
Quantile .95						
Raw difference	-.010643	.011751	-0.91	0.365	-.033675	.01239
Characteristics	.001961	.00998	0.20	0.897	-.017599	.02152
Coefficients	-.012603	.015175	-0.83	0.406	-.042346	.017139

. *2013

```

. rqdeco wage `X' if YEAR==2013&basic_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      161311
Number of observations in group 0    24133
Number of observations in group 1   137178
```

```

Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	-.151615	.00624	-24.30	0.000	-.163844	-.139385
Characteristics	-.00336	.004229	-0.79	0.590	-.01165	.004929
Coefficients	-.148255	.006243	-23.75	0.000	-.16049	-.136019
Quantile .2						
Raw difference	-.031869	.002634	-12.10	0.000	-.037032	-.026706
Characteristics	-.012579	.001985	-6.34	0.000	-.016469	-.00869
Coefficients	-.01929	.003057	-6.31	0.000	-.025281	-.013298
Quantile .35						
Raw difference	-.005064	.002262	-2.24	0.025	-.009497	-.000631
Characteristics	-.021558	.001713	-12.58	0.000	-.024916	-.0182
Coefficients	.016494	.002294	7.19	0.000	.011999	.02099
Quantile .5						

Raw difference	-.000946	.002326	-0.41	0.684	-.005506	.003613
Characteristics	-.028417	.001807	-15.73	0.000	-.031957	-.024876
Coefficients	.02747	.002216	12.40	0.000	.023128	.031813
<hr/>						
Quantile .65						
Raw difference	-.009529	.002756	-3.46	0.001	-.01493	-.004127
Characteristics	-.032893	.002132	-15.43	0.000	-.037071	-.028715
Coefficients	.023364	.00294	7.95	0.000	.017601	.029127
<hr/>						
Quantile .8						
Raw difference	-.023436	.00425	-5.51	0.000	-.031765	-.015107
Characteristics	-.035733	.003106	-11.50	0.000	-.04182	-.029645
Coefficients	.012297	.004853	2.53	0.011	.002784	.02181
<hr/>						
Quantile .95						
Raw difference	-.103544	.010667	-9.71	0.000	-.12445	-.082637
Characteristics	-.058827	.007597	-7.74	0.000	-.073717	-.043938
Coefficients	-.044717	.012513	-3.57	0.000	-.069242	-.020191

***Quantile regression estimation results and bootstrapped standard errors

**Native-total immigrants in medium-skilled

. **medium

. *2004

```
. rqdeco wage `X' if YEAR==2004&vocational_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      160233
Number of observations in group 0    2287
Number of observations in group 1   157946
```

```
Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.136908	.019801	6.91	0.000	.098098	.175717
Characteristics	.033959	.012902	2.63	0.081	.008671	.059247
Coefficients	.102949	.019454	5.29	0.000	.064819	.141078
Quantile .2						
Raw difference	.087611	.007556	11.59	0.000	.072802	.102421
Characteristics	.023221	.006809	3.41	0.009	.009876	.036567

Coefficients	.06439	.008892	7.24	0.000	.046963	.081817
Quantile .35						
Raw difference	.067861	.007318	9.27	0.000	.053519	.082203
Characteristics	.010923	.006985	1.56	0.242	-.002767	.024614
Coefficients	.056938	.009342	6.09	0.000	.038628	.075248
Quantile .5						
Raw difference	.058674	.007712	7.61	0.000	.043558	.07379
Characteristics	-.00241	.007574	-0.32	0.810	-.017256	.012435
Coefficients	.061084	.01001	6.10	0.000	.041465	.080704
Quantile .65						
Raw difference	.055617	.007957	6.99	0.000	.040022	.071212
Characteristics	-.021535	.00867	-2.48	0.057	-.038527	-.004542
Coefficients	.077151	.011321	6.81	0.000	.054962	.09934
Quantile .8						
Raw difference	.057314	.009496	6.04	0.000	.038702	.075927
Characteristics	-.044637	.010258	-4.35	0.000	-.064741	-.024533
Coefficients	.101952	.012681	8.04	0.000	.077097	.126806
Quantile .95						
Raw difference	.022943	.024361	0.94	0.346	-.024803	.07069
Characteristics	-.080882	.019841	-4.08	0.005	-.119768	-.041995
Coefficients	.103825	.028944	3.59	0.000	.047096	.160554

. *2007

```
. rqdeco wage `X' if YEAR==2007&vocational_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```


Decomposition of differences in distribution using quantile regression

Total number of observations 161781
 Number of observations in group 0 2950
 Number of observations in group 1 158831

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.103491	.01515	6.83	0.000	.073797	.133185
Characteristics	.021323	.01046	2.04	0.220	.000821	.041825
Coefficients	.082168	.017367	4.73	0.000	.04813	.116207
Quantile .2						
Raw difference	.078865	.006065	13.00	0.000	.066977	.090753
Characteristics	.010239	.005633	1.82	0.171	-.000801	.021278
Coefficients	.068626	.007474	9.18	0.000	.053977	.083276
Quantile .35						
Raw difference	.064216	.005663	11.34	0.000	.053117	.075316
Characteristics	.003976	.004891	0.81	0.522	-.005611	.013562
Coefficients	.060241	.006208	9.70	0.000	.048072	.072409
Quantile .5						
Raw difference	.054548	.005843	9.34	0.000	.043096	.066
Characteristics	-.001104	.005207	-0.21	0.869	-.01131	.009102
Coefficients	.055652	.006677	8.33	0.000	.042565	.068738
Quantile .65						
Raw difference	.052555	.006022	8.73	0.000	.040752	.064357

Characteristics	-.005772	.006378	-0.91	0.467	-.018272	.006728
Coefficients	.058327	.007933	7.35	0.000	.042778	.073875
<hr/>						
Quantile .8						
Raw difference	.062886	.007674	8.19	0.000	.047845	.077926
Characteristics	-.013913	.00806	-1.73	0.188	-.02971	.001883
Coefficients	.076799	.010557	7.27	0.000	.056108	.09749
<hr/>						
Quantile .95						
Raw difference	.099481	.017823	5.58	0.000	.064548	.134413
Characteristics	-.011074	.014965	-0.74	0.648	-.040405	.018258
Coefficients	.110554	.02429	4.55	0.000	.062947	.158161

. *2009

```

.
. rqrdeco wage `X' if YEAR==2009&vocational_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      385936
Number of observations in group 0    7767
Number of observations in group 1   378169
```

```

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]
Quantile .05					

Raw difference	.080948	.008444	9.59	0.000	.064399	.097498
Characteristics	.030704	.005728	5.36	0.000	.019477	.04193
Coefficients	.050245	.008511	5.90	0.000	.033563	.066926
<hr/>						
Quantile .2						
Raw difference	.086079	.003839	22.42	0.000	.078554	.093604
Characteristics	.017461	.003195	5.46	0.000	.011198	.023723
Coefficients	.068618	.004558	15.05	0.000	.059684	.077552
<hr/>						
Quantile .35						
Raw difference	.081956	.003302	24.82	0.000	.075484	.088429
Characteristics	.016953	.002873	5.90	0.000	.011323	.022584
Coefficients	.065003	.004278	15.20	0.000	.056619	.073387
<hr/>						
Quantile .5						
Raw difference	.077604	.004068	19.08	0.000	.069631	.085577
Characteristics	.017826	.003268	5.46	0.001	.011421	.024231
Coefficients	.059778	.00524	11.41	0.000	.049507	.070049
<hr/>						
Quantile .65						
Raw difference	.078993	.004896	16.14	0.000	.069397	.088588
Characteristics	.020346	.003759	5.41	0.001	.012978	.027714
Coefficients	.058647	.005951	9.86	0.000	.046984	.07031
<hr/>						
Quantile .8						
Raw difference	.088755	.006323	14.04	0.000	.076362	.101147
Characteristics	.026885	.004852	5.54	0.001	.017376	.036394
Coefficients	.06187	.007813	7.92	0.000	.046556	.077184
<hr/>						
Quantile .95						
Raw difference	.100691	.012877	7.82	0.000	.075453	.125929
Characteristics	.049715	.010965	4.53	0.006	.028224	.071207
Coefficients	.050976	.01817	2.81	0.005	.015363	.086589

. *2010

```
. : rqdeco wage `X' if YEAR==2010&vocational_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      365009
Number of observations in group 0    7509
Number of observations in group 1   357500
```

```
Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
<hr/>						
Quantile .05						
Raw difference	.071134	.006997	10.17	0.000	.05742	.084848
Characteristics	.024733	.004121	6.00	0.001	.016656	.032811
Coefficients	.0464	.007342	6.32	0.000	.03201	.060791
<hr/>						
Quantile .2						
Raw difference	.082496	.003335	24.73	0.000	.075959	.089033
Characteristics	.015908	.002803	5.68	0.000	.010414	.021402
Coefficients	.066588	.00397	16.77	0.000	.058807	.074369
<hr/>						
Quantile .35						
Raw difference	.080171	.003463	23.15	0.000	.073383	.086958
Characteristics	.017428	.00258	6.75	0.000	.012371	.022486
Coefficients	.062742	.004085	15.36	0.000	.054736	.070749
<hr/>						

Quantile .5						
Raw difference		.083473	.003598	23.20	0.000	.076421 .090526
Characteristics		.019811	.002813	7.04	0.000	.014297 .025325
Coefficients		.063663	.004432	14.36	0.000	.054976 .072349
<hr/>						
Quantile .65						
Raw difference		.094785	.004312	21.98	0.000	.086332 .103237
Characteristics		.023442	.00308	7.61	0.000	.017406 .029478
Coefficients		.071343	.005273	13.53	0.000	.061008 .081677
<hr/>						
Quantile .8						
Raw difference		.110462	.005979	18.48	0.000	.098744 .122181
Characteristics		.032869	.004244	7.75	0.000	.024551 .041186
Coefficients		.077594	.007833	9.91	0.000	.062242 .092945
<hr/>						
Quantile .95						
Raw difference		.126699	.015046	8.42	0.000	.097209 .156189
Characteristics		.048148	.008977	5.36	0.011	.030554 .065742
Coefficients		.078551	.018868	4.16	0.000	.04157 .115532
<hr/>						

. *2013

```
.
. rqdeco wage `X' if YEAR==2013&vocational_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

```
Decomposition of differences in distribution using quantile regression
Total number of observations      397490
Number of observations in group 0    9860
Number of observations in group 1   387630

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.084162	.007203	11.68	0.000	.070044	.09828
Characteristics	.033167	.004447	7.46	0.000	.024451	.041882
Coefficients	.050995	.007739	6.59	0.000	.035827	.066164
Quantile .2						
Raw difference	.080817	.003384	23.88	0.000	.074185	.087449
Characteristics	.018791	.002486	7.56	0.000	.013918	.023663
Coefficients	.062026	.003366	18.43	0.000	.05543	.068623
Quantile .35						
Raw difference	.080939	.003092	26.17	0.000	.074878	.087
Characteristics	.015507	.002408	6.44	0.000	.010788	.020227
Coefficients	.065432	.003282	19.94	0.000	.059	.071864
Quantile .5						
Raw difference	.082891	.003376	24.56	0.000	.076276	.089507
Characteristics	.015612	.00258	6.05	0.000	.010556	.020668
Coefficients	.067279	.003928	17.13	0.000	.059581	.074978
Quantile .65						
Raw difference	.093352	.004485	20.81	0.000	.084561	.102143
Characteristics	.01709	.002814	6.07	0.001	.011575	.022606
Coefficients	.076262	.004924	15.49	0.000	.066612	.085912
Quantile .8						
Raw difference	.113666	.006459	17.60	0.000	.101006	.126325
Characteristics	.020797	.00324	6.42	0.002	.014446	.027147
Coefficients	.092869	.006598	14.08	0.000	.079938	.1058

Quantile .95						
Raw difference	.168675	.011625	14.51	0.000	.14589	.191459
Characteristics	.027349	.005989	4.57	0.041	.01561	.039089
Coefficients	.141325	.013364	10.57	0.000	.115131	.167519

```

***Quantile regression estimation results and bootstrapped standard errors
**Native-total immigrants in high-skilled group
. **High-Skilled*****
.
. *2004
.
. rqdeco wage `X' if YEAR==2004&higher_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      8118
Number of observations in group 0    712
Number of observations in group 1   7406
```

```

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
<hr/>						
Quantile .05						
Raw difference	.254181	.02614	9.72	0.000	.202948	.305414
Characteristics	.052548	.027706	1.90	0.153	-.001754	.10685
Coefficients	.201633	.036813	5.48	0.000	.129481	.273785
<hr/>						
Quantile .2						
Raw difference	.265256	.020703	12.81	0.000	.224678	.305834
Characteristics	.085167	.01641	5.19	0.000	.053003	.117331
Coefficients	.180089	.021941	8.21	0.000	.137086	.223092

Quantile .35						
Raw difference	.237855	.023642	10.06	0.000	.191518	.284192
Characteristics	.083419	.017601	4.74	0.000	.048922	.117915
Coefficients	.154436	.019172	8.06	0.000	.11686	.192013
Quantile .5						
Raw difference	.203002	.024417	8.31	0.000	.155147	.250858
Characteristics	.063008	.019827	3.18	0.001	.024148	.101867
Coefficients	.139995	.01889	7.41	0.000	.102972	.177018
Quantile .65						
Raw difference	.179575	.02335	7.69	0.000	.133809	.22534
Characteristics	.045803	.02162	2.12	0.057	.003429	.088178
Coefficients	.133771	.0241	5.55	0.000	.086537	.181006
Quantile .8						
Raw difference	.163273	.027694	5.90	0.000	.108994	.217552
Characteristics	.028885	.024422	1.18	0.378	-.018982	.076751
Coefficients	.134388	.03279	4.10	0.000	.07012	.198656
Quantile .95						
Raw difference	.098252	.048669	2.02	0.044	.002862	.193642
Characteristics	-.001026	.034187	-0.03	0.984	-.068032	.065979
Coefficients	.099278	.051815	1.92	0.055	-.002278	.200835

. *2007

```
.
. rqdeco wage `X' if YEAR==2007&higher_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

Total number of observations 12763
Number of observations in group 0 857
Number of observations in group 1 11906

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.218691	.022033	9.93	0.000	.175507	.261876
Characteristics	.089964	.027624	3.26	0.003	.035822	.144106
Coefficients	.128728	.029782	4.32	0.000	.070356	.1871
Quantile .2						
Raw difference	.202193	.013683	14.78	0.000	.175375	.22901
Characteristics	.078574	.014808	5.31	0.000	.049551	.107598
Coefficients	.123618	.017735	6.97	0.000	.088859	.158378
Quantile .35						
Raw difference	.169444	.014046	12.06	0.000	.141915	.196973
Characteristics	.060718	.012931	4.70	0.000	.035374	.086061
Coefficients	.108727	.016231	6.70	0.000	.076914	.140539
Quantile .5						
Raw difference	.152434	.013947	10.93	0.000	.125098	.179769
Characteristics	.050255	.013279	3.78	0.001	.024228	.076282
Coefficients	.102179	.015828	6.46	0.000	.071157	.1332
Quantile .65						
Raw difference	.141584	.015211	9.31	0.000	.111772	.171396
Characteristics	.034103	.0138	2.47	0.041	.007054	.061151

Coefficients	.107481	.016657	6.45	0.000	.074834	.140128
Quantile .8						
Raw difference	.136467	.019686	6.93	0.000	.097884	.17505
Characteristics	.011528	.015314	0.75	0.573	-.018486	.041542
Coefficients	.124939	.020477	6.10	0.000	.084804	.165074
Quantile .95						
Raw difference	.128429	.034265	3.75	0.000	.061271	.195586
Characteristics	.01052	.029143	0.36	0.801	-.046599	.06764
Coefficients	.117908	.041643	2.83	0.005	.036289	.199527

. *2009

```
.
. rqdeco wage `X' if YEAR==2009&higher_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      41764
Number of observations in group 0   2805
Number of observations in group 1  38959
```

```
Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]
Quantile .05					
Raw difference	.108804	.012969	8.39	0.000	.083386 .134222

Characteristics	.021715	.009342	2.32	0.159	.003406	.040025
Coefficients	.087089	.015416	5.65	0.000	.056873	.117304
Quantile .2						
Raw difference	.083372	.008553	9.75	0.000	.06661	.100135
Characteristics	.016729	.006221	2.69	0.059	.004536	.028923
Coefficients	.066643	.008857	7.52	0.000	.049283	.084002
Quantile .35						
Raw difference	.05628	.00732	7.69	0.000	.041933	.070627
Characteristics	.008586	.005937	1.45	0.299	-.003051	.020223
Coefficients	.047694	.008264	5.77	0.000	.031498	.063891
Quantile .5						
Raw difference	.036419	.007066	5.15	0.000	.022569	.050269
Characteristics	.001579	.005844	0.27	0.849	-.009874	.013033
Coefficients	.03484	.008307	4.19	0.000	.018558	.051121
Quantile .65						
Raw difference	.028417	.008154	3.48	0.000	.012435	.0444
Characteristics	-.001974	.006228	-0.32	0.835	-.014181	.010234
Coefficients	.030391	.009446	3.22	0.001	.011876	.048905
Quantile .8						
Raw difference	.034182	.010406	3.28	0.001	.013787	.054578
Characteristics	-.009053	.007945	-1.14	0.431	-.024625	.006518
Coefficients	.043236	.011492	3.76	0.000	.020712	.065759
Quantile .95						
Raw difference	.054857	.020271	2.71	0.007	.015126	.094588
Characteristics	-.030234	.015495	-1.95	0.213	-.060603	.000135
Coefficients	.085091	.0243	3.50	0.000	.037465	.132718

. *2010

```

. rqdeco wage `X' if YEAR==2010&higher_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      42018
Number of observations in group 0    2823
Number of observations in group 1   39195
```

```

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.146798	.012568	11.68	0.000	.122165	.17143
Characteristics	.034963	.008038	4.35	0.007	.019209	.050718
Coefficients	.111834	.012856	8.70	0.000	.086636	.137032
Quantile .2						
Raw difference	.097377	.007971	12.22	0.000	.081754	.112999
Characteristics	.022164	.006738	3.29	0.013	.008958	.035371
Coefficients	.075213	.008924	8.43	0.000	.057722	.092703
Quantile .35						
Raw difference	.068234	.006942	9.83	0.000	.054627	.081841
Characteristics	.008792	.005748	1.53	0.350	-.002474	.020057
Coefficients	.059442	.009403	6.32	0.000	.041012	.077873
Quantile .5						

Raw difference	.04855	.007503	6.47	0.000	.033844	.063256
Characteristics	-.002017	.006049	-0.33	0.842	-.013874	.009839
Coefficients	.050568	.010092	5.01	0.000	.030788	.070348
<hr/>						
Quantile .65						
Raw difference	.037944	.009658	3.93	0.000	.019014	.056874
Characteristics	-.010136	.007828	-1.29	0.396	-.025478	.005207
Coefficients	.04808	.011943	4.03	0.000	.024672	.071488
<hr/>						
Quantile .8						
Raw difference	.039179	.013705	2.86	0.004	.012317	.066041
Characteristics	-.016208	.010277	-1.58	0.293	-.03635	.003935
Coefficients	.055387	.015427	3.59	0.000	.02515	.085624
<hr/>						
Quantile .95						
Raw difference	.05076	.027485	1.85	0.065	-.003111	.10463
Characteristics	-.01964	.016447	-1.19	0.530	-.051874	.012595
Coefficients	.070399	.031265	2.25	0.024	.009121	.131678

. *2013

```
. rqdeco wage `X' if YEAR==2013&higher_edu==1, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

```
Decomposition of differences in distribution using quantile regression
Total number of observations      55622
Number of observations in group 0    4393
Number of observations in group 1    51229

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.157499	.011972	13.16	0.000	.134035	.180963
Characteristics	.029514	.008425	3.50	0.017	.013002	.046026
Coefficients	.127985	.012417	10.31	0.000	.103649	.152322
Quantile .2						
Raw difference	.131923	.006492	20.32	0.000	.1192	.144647
Characteristics	.033374	.005535	6.03	0.000	.022526	.044222
Coefficients	.098549	.006653	14.81	0.000	.085509	.111589
Quantile .35						
Raw difference	.110065	.006129	17.96	0.000	.098052	.122077
Characteristics	.032262	.00489	6.60	0.000	.022677	.041847
Coefficients	.077803	.006273	12.40	0.000	.065508	.090097
Quantile .5						
Raw difference	.087452	.006748	12.96	0.000	.074227	.100677
Characteristics	.031732	.005363	5.92	0.000	.02122	.042243
Coefficients	.05572	.007124	7.82	0.000	.041757	.069684
Quantile .65						
Raw difference	.066844	.008214	8.14	0.000	.050744	.082943
Characteristics	.030684	.006383	4.81	0.001	.018175	.043194
Coefficients	.036159	.008868	4.08	0.000	.018778	.05354
Quantile .8						
Raw difference	.054782	.010024	5.47	0.000	.035135	.074428
Characteristics	.030381	.008402	3.62	0.012	.013912	.046849
Coefficients	.024401	.01206	2.02	0.043	.000763	.048039

Quantile .95							
Raw difference		.065824	.016343	4.03	0.000	.033792	.097855
Characteristics		.046157	.014887	3.10	0.047	.016979	.075335
Coefficients		.019667	.023289	0.84	0.398	-.025979	.065312

Appendix D.3: Full results for Figure 3.12(native-immigrants wage gap under different groups of countries of origin)

***Quantile regression estimation results and bootstrapped standard errors for native-immigrant wage gap decomposition under different groups of countries of origin.

**Native-EU15

```
.   rqdeco wage `X' if YEAR==2004, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      240517
Number of observations in group 0   2420
Number of observations in group 1  238097
```

```
Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
<hr/>						
Quantile .05						
Raw difference	-.064525	.014266	-4.52	0.000	-.092487	-.036564
Characteristics	-.034482	.031192	-1.11	0.342	-.095618	.026654
Coefficients	-.030043	.036303	-0.83	0.408	-.101196	.041109
<hr/>						
Quantile .2						
Raw difference	-.01918	.007996	-2.40	0.016	-.034852	-.003507
Characteristics	.012456	.011546	1.08	0.399	-.010174	.035085
Coefficients	-.031636	.014784	-2.14	0.032	-.060611	-.00266
<hr/>						
Quantile .35						
Raw difference	-.014194	.007999	-1.77	0.076	-.029872	.001484
Characteristics	.017293	.011292	1.53	0.226	-.004839	.039425

Coefficients	-.031487	.014281	-2.20	0.027	-.059477	-.003497
Quantile .5						
Raw difference	-.028198	.00839	-3.36	0.001	-.044643	-.011754
Characteristics	.013435	.011839	1.13	0.385	-.009769	.036639
Coefficients	-.041633	.015451	-2.69	0.007	-.071918	-.011349
Quantile .65						
Raw difference	-.053459	.010008	-5.34	0.000	-.073074	-.033844
Characteristics	-.000361	.012267	-0.03	0.983	-.024403	.023681
Coefficients	-.053098	.01703	-3.12	0.002	-.086476	-.019721
Quantile .8						
Raw difference	-.088898	.013624	-6.52	0.000	-.115601	-.062195
Characteristics	-.023686	.01474	-1.61	0.256	-.052575	.005204
Coefficients	-.065213	.020865	-3.13	0.002	-.106107	-.024319
Quantile .95						
Raw difference	-.184986	.026564	-6.96	0.000	-.237051	-.132921
Characteristics	-.053709	.029904	-1.80	0.181	-.11232	.004902
Coefficients	-.131276	.040163	-3.27	0.001	-.209994	-.052559

***2007

```
.
.
.
.   rqdeco wage `X' if YEAR==2007, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression
Total number of observations 245710

Number of observations in group 0 3086
Number of observations in group 1 242624

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	-.083901	.017414	-4.82	0.000	-.118033	-.04977
Characteristics	.032971	.01542	2.14	0.112	.002748	.063194
Coefficients	-.116873	.020732	-5.64	0.000	-.157507	-.076238
Quantile .2						
Raw difference	-.032244	.007865	-4.10	0.000	-.047658	-.016829
Characteristics	.015233	.007518	2.03	0.120	.000497	.029968
Coefficients	-.047476	.009796	-4.85	0.000	-.066675	-.028277
Quantile .35						
Raw difference	-.024217	.006118	-3.96	0.000	-.036208	-.012226
Characteristics	.009978	.007073	1.41	0.182	-.003885	.02384
Coefficients	-.034195	.007484	-4.57	0.000	-.048863	-.019527
Quantile .5						
Raw difference	-.037415	.00672	-5.57	0.000	-.050585	-.024244
Characteristics	.000036	.007901	0.00	0.996	-.01545	.015523
Coefficients	-.037451	.007898	-4.74	0.000	-.052931	-.02197
Quantile .65						
Raw difference	-.062568	.009829	-6.37	0.000	-.081833	-.043303
Characteristics	-.0098	.009797	-1.00	0.352	-.029002	.009403
Coefficients	-.052768	.010524	-5.01	0.000	-.073395	-.032141

Quantile .8						
Raw difference	-.110472	.015466	-7.14	0.000	-.140784	-.08016
Characteristics	-.02104	.013234	-1.59	0.203	-.046979	.004898
Coefficients	-.089432	.016533	-5.41	0.000	-.121835	-.057028
Quantile .95						
Raw difference	-.212099	.022365	-9.48	0.000	-.255934	-.168264
Characteristics	-.000317	.023457	-0.01	0.992	-.046292	.045657
Coefficients	-.211781	.030303	-6.99	0.000	-.271174	-.152388

```
. **2009
.
.
.
. rqdeco wage `X' if YEAR==2009, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      566100
Number of observations in group 0    7971
Number of observations in group 1   558129
```

```
Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]
Quantile .05					
Raw difference	-.030218	.009845	-3.07	0.002	-.049514 - .010922

Characteristics	.051437	.008104	6.35	0.000	.035554	.067319
Coefficients	-.081655	.01324	-6.17	0.000	-.107604	-.055705
Quantile .2						
Raw difference	.003226	.004251	0.76	0.448	-.005106	.011557
Characteristics	.02428	.005181	4.69	0.001	.014125	.034435
Coefficients	-.021054	.007009	-3.00	0.003	-.034793	-.007316
Quantile .35						
Raw difference	-.002875	.003728	-0.77	0.441	-.010181	.004431
Characteristics	.015601	.004625	3.37	0.015	.006536	.024666
Coefficients	-.018476	.006382	-2.90	0.004	-.030984	-.005968
Quantile .5						
Raw difference	-.018458	.004193	-4.40	0.000	-.026677	-.01024
Characteristics	.010424	.00513	2.03	0.129	.000369	.02048
Coefficients	-.028883	.006876	-4.20	0.000	-.042358	-.015407
Quantile .65						
Raw difference	-.03964	.00513	-7.73	0.000	-.049695	-.029584
Characteristics	.005663	.006096	0.93	0.467	-.006286	.017611
Coefficients	-.045302	.007794	-5.81	0.000	-.060577	-.030027
Quantile .8						
Raw difference	-.074417	.007753	-9.60	0.000	-.089614	-.059221
Characteristics	.001194	.007874	0.15	0.909	-.014239	.016627
Coefficients	-.075611	.010396	-7.27	0.000	-.095987	-.055236
Quantile .95						
Raw difference	-.171591	.013189	-13.01	0.000	-.197441	-.145741
Characteristics	.015159	.017753	0.85	0.532	-.019636	.049954
Coefficients	-.18675	.024229	-7.71	0.000	-.234238	-.139262

.

```

*2010
rqdeco wage `X' if YEAR==2010, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      556698
Number of observations in group 0    8191
Number of observations in group 1   548507
```

```

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
<hr/>						
Quantile .05						
Raw difference	-.034217	.006521	-5.25	0.000	-.046997	-.021437
Characteristics	.035837	.008873	4.04	0.001	.018446	.053228
Coefficients	-.070055	.011157	-6.28	0.000	-.091921	-.048188
<hr/>						
Quantile .2						
Raw difference	.001489	.003919	0.38	0.704	-.006193	.00917
Characteristics	.006375	.006508	0.98	0.411	-.00638	.019131
Coefficients	-.004887	.007762	-0.63	0.529	-.0201	.010326
<hr/>						
Quantile .35						
Raw difference	-.00503	.004047	-1.24	0.214	-.012963	.002903
Characteristics	-.007468	.005562	-1.34	0.255	-.01837	.003435
Coefficients	.002438	.006566	0.37	0.710	-.010432	.015307
<hr/>						

Quantile .5						
Raw difference	-.022318	.004671	-4.78	0.000	-.031472	-.013164
Characteristics	-.019142	.005651	-3.39	0.004	-.030217	-.008067
Coefficients	-.003177	.006572	-0.48	0.629	-.016058	.009705
Quantile .65						
Raw difference	-.044868	.005512	-8.14	0.000	-.055671	-.034065
Characteristics	-.028152	.005764	-4.88	0.000	-.039449	-.016854
Coefficients	-.016716	.007311	-2.29	0.022	-.031045	-.002387
Quantile .8						
Raw difference	-.08096	.006961	-11.63	0.000	-.094604	-.067316
Characteristics	-.037179	.006471	-5.75	0.000	-.049862	-.024497
Coefficients	-.04378	.009049	-4.84	0.000	-.061517	-.026044
Quantile .95						
Raw difference	-.175758	.012135	-14.48	0.000	-.199542	-.151973
Characteristics	-.05117	.012965	-3.95	0.001	-.07658	-.025759
Coefficients	-.124588	.01608	-7.75	0.000	-.156104	-.093072

. **2013

```
.
.   rqdeco wage `X' if YEAR==2013, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

Total number of observations	588927
Number of observations in group 0	9940
Number of observations in group 1	578987

Number of quantile regressions estimated	100
--	-----

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	-.019001	.010261	-1.85	0.064	-.039112	.001109
Characteristics	.040895	.010691	3.83	0.000	.019942	.061848
Coefficients	-.059896	.011719	-5.11	0.000	-.082864	-.036928
Quantile .2						
Raw difference	-.01691	.003897	-4.34	0.000	-.024548	-.009271
Characteristics	.00993	.006318	1.57	0.158	-.002454	.022313
Coefficients	-.026839	.007027	-3.82	0.000	-.040613	-.013066
Quantile .35						
Raw difference	-.025696	.003238	-7.93	0.000	-.032043	-.019348
Characteristics	.0025	.005265	0.47	0.680	-.00782	.012819
Coefficients	-.028195	.006054	-4.66	0.000	-.04006	-.01633
Quantile .5						
Raw difference	-.043661	.003605	-12.11	0.000	-.050726	-.036596
Characteristics	-.003984	.00504	-0.79	0.484	-.013863	.005894
Coefficients	-.039677	.005688	-6.98	0.000	-.050825	-.028528
Quantile .65						
Raw difference	-.069542	.004873	-14.27	0.000	-.079092	-.059991
Characteristics	-.011508	.005626	-2.05	0.065	-.022535	-.00048
Coefficients	-.058034	.006243	-9.30	0.000	-.070269	-.045799
Quantile .8						
Raw difference	-.11469	.007187	-15.96	0.000	-.128777	-.100603
Characteristics	-.021784	.00719	-3.03	0.013	-.035876	-.007693
Coefficients	-.092906	.008748	-10.62	0.000	-.110051	-.07576

Quantile .95						
Raw difference	-.199738	.011076	-18.03	0.000	-.221448	-.178029
Characteristics	-.037558	.012243	-3.07	0.017	-.061553	-.013562
Coefficients	-.162181	.015694	-10.33	0.000	-.192941	-.131421

***Quantile regression estimation results and bootstrapped standard errors

**Native-EU12

****2004

. rqdeco wage `X' if YEAR==2004, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
 Fitting base model
 (bootstrapping)

Decomposition of differences in distribution using quantile regression

Total number of observations 238424
 Number of observations in group 0 327
 Number of observations in group 1 238097

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]
Quantile .05					
Raw difference	.201716	.07715	2.61	0.009	.050505 .352927
Characteristics	-.011094	.062407	-0.18	0.893	-.133409 .111221
Coefficients	.21281	.082442	2.58	0.010	.051226 .374394

Quantile .2						
Raw difference	.109189	.027655	3.95	0.000	.054987	.163391
Characteristics	.034298	.041281	0.83	0.441	-.046611	.115207
Coefficients	.074891	.044551	1.68	0.093	-.012429	.16221
Quantile .35						
Raw difference	.096056	.023554	4.08	0.000	.049892	.142221
Characteristics	.040152	.038184	1.05	0.269	-.034688	.114991
Coefficients	.055905	.036305	1.54	0.124	-.015251	.127061
Quantile .5						
Raw difference	.084628	.02213	3.82	0.000	.041254	.128003
Characteristics	.023024	.03493	0.66	0.495	-.045437	.091485
Coefficients	.061604	.033717	1.83	0.068	-.00448	.127688
Quantile .65						
Raw difference	.089546	.024551	3.65	0.000	.041427	.137665
Characteristics	.007598	.034683	0.22	0.830	-.06038	.075575
Coefficients	.081949	.035389	2.32	0.021	.012588	.151309
Quantile .8						
Raw difference	.116516	.029455	3.96	0.000	.058786	.174246
Characteristics	-.008785	.039364	-0.22	0.832	-.085936	.068367
Coefficients	.1253	.041528	3.02	0.003	.043907	.206694
Quantile .95						
Raw difference	.135809	.063613	2.13	0.033	.011129	.260489
Characteristics	-.05468	.05208	-1.05	0.465	-.156755	.047394
Coefficients	.190489	.07491	2.54	0.011	.043669	.33731

. ***2007

```

.
.
.
. rqdeco wage `X' if YEAR==2007, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      243411
Number of observations in group 0    787
Number of observations in group 1   242624
```

```

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
<hr/>						
Quantile .05						
Raw difference	.09232	.033133	2.79	0.005	.027381	.157259
Characteristics	.129648	.033941	3.82	0.001	.063124	.196171
Coefficients	-.037328	.03848	-0.97	0.332	-.112748	.038092
<hr/>						
Quantile .2						
Raw difference	.103298	.013184	7.84	0.000	.077457	.129138
Characteristics	.101732	.025629	3.97	0.000	.051501	.151964
Coefficients	.001565	.024756	0.06	0.950	-.046955	.050085
<hr/>						
Quantile .35						
Raw difference	.107989	.010907	9.90	0.000	.086612	.129366
Characteristics	.090666	.020947	4.33	0.000	.049611	.131722
Coefficients	.017323	.021199	0.82	0.414	-.024227	.058872
<hr/>						

Quantile .5						
Raw difference	.110755	.009802	11.30	0.000	.091543	.129968
Characteristics	.084646	.019611	4.32	0.000	.046209	.123082
Coefficients	.02611	.019662	1.33	0.184	-.012427	.064646
Quantile .65						
Raw difference	.124976	.008623	14.49	0.000	.108075	.141876
Characteristics	.08237	.023294	3.54	0.001	.036715	.128025
Coefficients	.042606	.023834	1.79	0.074	-.004109	.08932
Quantile .8						
Raw difference	.166029	.011918	13.93	0.000	.14267	.189389
Characteristics	.090137	.029461	3.06	0.007	.032393	.14788
Coefficients	.075892	.033643	2.26	0.024	.009954	.141831
Quantile .95						
Raw difference	.27061	.029393	9.21	0.000	.213001	.32822
Characteristics	.095511	.039961	2.39	0.072	.017188	.173833
Coefficients	.1751	.053033	3.30	0.001	.071156	.279043

```

. **2009
.
.
.
. rqdeco wage `X' if YEAR==2009, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

```

Decomposition of differences in distribution using quantile regression
Total number of observations      561615
Number of observations in group 0    3486
```

Number of observations in group 1 558129

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.048199	.012546	3.84	0.000	.02361	.072788
Characteristics	.092328	.024029	3.84	0.000	.045231	.139424
Coefficients	-.044129	.02521	-1.75	0.080	-.09354	.005283
Quantile .2						
Raw difference	.11923	.0048	24.84	0.000	.109822	.128638
Characteristics	.077308	.01272	6.08	0.000	.052377	.102239
Coefficients	.041922	.012641	3.32	0.001	.017146	.066699
Quantile .35						
Raw difference	.133777	.004907	27.26	0.000	.12416	.143394
Characteristics	.077129	.009887	7.80	0.000	.057752	.096507
Coefficients	.056647	.009917	5.71	0.000	.03721	.076084
Quantile .5						
Raw difference	.141618	.004756	29.77	0.000	.132296	.15094
Characteristics	.081959	.009704	8.45	0.000	.062939	.100979
Coefficients	.059659	.009546	6.25	0.000	.04095	.078368
Quantile .65						
Raw difference	.15432	.005127	30.10	0.000	.144272	.164368
Characteristics	.090831	.010596	8.57	0.000	.070063	.1116
Coefficients	.063488	.010504	6.04	0.000	.042901	.084076
Quantile .8						

Raw difference	.187171	.006436	29.08	0.000	.174556	.199786
Characteristics	.109723	.013541	8.10	0.000	.083183	.136263
Coefficients	.077448	.014521	5.33	0.000	.048988	.105908

Quantile .95						
Raw difference	.207373	.023521	8.82	0.000	.161273	.253473
Characteristics	.169178	.036334	4.66	0.000	.097965	.240391
Coefficients	.038194	.041475	0.92	0.357	-.043095	.119484

```
***2010
.   rqdeco wage `X' if YEAR==2010, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      552407
Number of observations in group 0    3900
Number of observations in group 1   548507
```

```
Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	

Quantile .05						
Raw difference	.109186	.013457	8.11	0.000	.082811	.135561
Characteristics	.114612	.023888	4.80	0.000	.067793	.161431
Coefficients	-.005426	.02633	-0.21	0.837	-.057032	.046181

Quantile .2						
Raw difference	.140143	.005993	23.39	0.000	.128397	.151888
Characteristics	.0964	.012525	7.70	0.000	.071853	.120948
Coefficients	.043742	.013962	3.13	0.002	.016378	.071107
Quantile .35						
Raw difference	.154298	.005152	29.95	0.000	.1442	.164395
Characteristics	.091709	.012127	7.56	0.000	.06794	.115479
Coefficients	.062588	.013243	4.73	0.000	.036632	.088545
Quantile .5						
Raw difference	.165872	.005474	30.30	0.000	.155143	.176601
Characteristics	.094155	.012754	7.38	0.000	.069157	.119153
Coefficients	.071717	.013928	5.15	0.000	.044419	.099014
Quantile .65						
Raw difference	.187345	.006124	30.59	0.000	.175343	.199347
Characteristics	.107684	.014357	7.50	0.000	.079545	.135823
Coefficients	.07966	.016161	4.93	0.000	.047985	.111336
Quantile .8						
Raw difference	.229012	.007867	29.11	0.000	.213594	.244431
Characteristics	.142956	.018386	7.78	0.000	.10692	.178992
Coefficients	.086056	.021137	4.07	0.000	.044628	.127484
Quantile .95						
Raw difference	.255652	.023367	10.94	0.000	.209853	.30145
Characteristics	.162406	.02917	5.57	0.000	.105234	.219579
Coefficients	.093245	.028809	3.24	0.001	.036781	.149709

. **2013


```

.      rqdeco wage `X' if YEAR==2013, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      586254
Number of observations in group 0    7267
Number of observations in group 1   578987
```

```

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.144269	.010104	14.28	0.000	.124465	.164072
Characteristics	.105567	.015474	6.82	0.000	.075238	.135896
Coefficients	.038702	.018116	2.14	0.033	.003194	.074209
Quantile .2						
Raw difference	.184959	.00438	42.23	0.000	.176375	.193544
Characteristics	.100466	.010007	10.04	0.000	.080853	.120079
Coefficients	.084493	.011061	7.64	0.000	.062814	.106173
Quantile .35						
Raw difference	.197907	.004228	46.81	0.000	.189621	.206193
Characteristics	.10469	.00823	12.72	0.000	.08856	.120821
Coefficients	.093217	.009162	10.17	0.000	.075259	.111175
Quantile .5						
Raw difference	.203472	.004237	48.02	0.000	.195166	.211777

Characteristics	.108431	.007736	14.02	0.000	.093269	.123593
Coefficients	.09504	.008572	11.09	0.000	.078239	.111842
<hr/>						
Quantile .65						
Raw difference	.216517	.004385	49.38	0.000	.207923	.22511
Characteristics	.112336	.008285	13.56	0.000	.096098	.128575
Coefficients	.10418	.009301	11.20	0.000	.085951	.122409
<hr/>						
Quantile .8						
Raw difference	.256171	.004737	54.08	0.000	.246886	.265456
Characteristics	.126757	.01037	12.22	0.000	.106432	.147081
Coefficients	.129414	.011776	10.99	0.000	.106334	.152495
<hr/>						
Quantile .95						
Raw difference	.3338	.013817	24.16	0.000	.306719	.360882
Characteristics	.160374	.024596	6.52	0.000	.112167	.20858
Coefficients	.173427	.029546	5.87	0.000	.115518	.231336

***Quantile regression estimation results and bootstrapped standard errors

**Native-LDI:

```
.   rqdeco wage `X' if YEAR==2004, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```
Total number of observations      243590
  Number of observations in group 0    5493
  Number of observations in group 1   238097

Number of quantile regressions estimated    100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.126441	.013457	9.40	0.000	.100066	.152816
Characteristics	.017398	.015282	1.14	0.318	-.012555	.04735
Coefficients	.109043	.017438	6.25	0.000	.074865	.143222
Quantile .2						
Raw difference	.136652	.005289	25.84	0.000	.126287	.147018
Characteristics	.03694	.00833	4.43	0.000	.020614	.053267
Coefficients	.099712	.008634	11.55	0.000	.08279	.116634
Quantile .35						
Raw difference	.149994	.003494	42.93	0.000	.143146	.156841
Characteristics	.049732	.006876	7.23	0.000	.036255	.063209
Coefficients	.100262	.006773	14.80	0.000	.086987	.113537
Quantile .5						
Raw difference	.165382	.003063	53.99	0.000	.159377	.171386
Characteristics	.057878	.006672	8.67	0.000	.0448	.070955
Coefficients	.107504	.006734	15.96	0.000	.094306	.120702
Quantile .65						
Raw difference	.187857	.003478	54.02	0.000	.181041	.194674
Characteristics	.062908	.007366	8.54	0.000	.048471	.077344
Coefficients	.12495	.007804	16.01	0.000	.109654	.140246
Quantile .8						
Raw difference	.225515	.004669	48.30	0.000	.216364	.234666
Characteristics	.06844	.009324	7.34	0.000	.050166	.086714
Coefficients	.157075	.009459	16.61	0.000	.138535	.175615

Quantile .95						
Raw difference		.304798	.011971	25.46	0.000	.281335 .328261
Characteristics		.089695	.020643	4.34	0.000	.049235 .130156
Coefficients		.215102	.024281	8.86	0.000	.167513 .262692

```

rqdeco wage `X' if YEAR==2007, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

Total number of observations	249573
Number of observations in group 0	6949
Number of observations in group 1	242624

Number of quantile regressions estimated	100
--	-----

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]
Quantile .05					
Raw difference	.12394	.011456	10.82	0.000	.101487 .146393
Characteristics	.0246	.013388	1.84	0.109	-.001641 .05084
Coefficients	.099341	.015364	6.47	0.000	.069228 .129453
Quantile .2					
Raw difference	.138719	.004732	29.32	0.000	.129444 .147993
Characteristics	.043156	.007252	5.95	0.000	.028943 .05737
Coefficients	.095562	.008208	11.64	0.000	.079475 .111649
Quantile .35					

Raw difference	.149176	.004031	37.01	0.000	.141275	.157077
Characteristics	.058377	.005996	9.74	0.000	.046625	.070129
Coefficients	.090799	.007203	12.61	0.000	.076682	.104916

Quantile .5						
Raw difference	.15584	.004175	37.33	0.000	.147658	.164022
Characteristics	.065761	.005572	11.80	0.000	.054841	.076682
Coefficients	.090079	.007172	12.56	0.000	.076023	.104135

Quantile .65						
Raw difference	.167017	.004777	34.96	0.000	.157654	.176379
Characteristics	.068572	.005922	11.58	0.000	.056965	.080178
Coefficients	.098445	.007621	12.92	0.000	.083507	.113383

Quantile .8						
Raw difference	.193115	.006255	30.87	0.000	.180855	.205375
Characteristics	.07043	.006478	10.87	0.000	.057734	.083127
Coefficients	.122685	.008567	14.32	0.000	.105894	.139476

Quantile .95						
Raw difference	.286476	.010372	27.62	0.000	.266146	.306805
Characteristics	.071315	.01153	6.19	0.000	.048717	.093913
Coefficients	.215161	.01651	13.03	0.000	.182802	.24752

**2009

```

.
.
.
. rqdeco wage `X' if YEAR==2009, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

Total number of observations 574104
 Number of observations in group 0 15975
 Number of observations in group 1 558129

Number of quantile regressions estimated 100

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.100041	.007592	13.18	0.000	.085161	.11492
Characteristics	.069723	.006863	10.16	0.000	.056272	.083173
Coefficients	.030318	.009475	3.20	0.001	.011748	.048888
Quantile .2						
Raw difference	.122412	.00276	44.35	0.000	.117003	.127821
Characteristics	.053896	.003406	15.82	0.000	.047221	.060572
Coefficients	.068516	.004391	15.60	0.000	.059909	.077122
Quantile .35						
Raw difference	.138555	.002297	60.33	0.000	.134053	.143056
Characteristics	.055734	.002559	21.78	0.000	.050718	.06075
Coefficients	.082821	.003603	22.99	0.000	.075759	.089883
Quantile .5						
Raw difference	.150295	.002721	55.23	0.000	.144961	.155628
Characteristics	.057868	.002534	22.84	0.000	.052901	.062834
Coefficients	.092427	.003655	25.29	0.000	.085264	.09959
Quantile .65						
Raw difference	.166892	.003348	49.85	0.000	.160331	.173454

Characteristics	.061512	.002978	20.66	0.000	.055676	.067348
Coefficients	.10538	.004537	23.23	0.000	.096489	.114272
<hr/>						
Quantile .8						
Raw difference	.194952	.004533	43.01	0.000	.186067	.203837
Characteristics	.067351	.004291	15.70	0.000	.058941	.075761
Coefficients	.127601	.00654	19.51	0.000	.114782	.14042
<hr/>						
Quantile .95						
Raw difference	.244092	.011037	22.12	0.000	.22246	.265724
Characteristics	.079995	.007947	10.07	0.000	.06442	.09557
Coefficients	.164097	.01464	11.21	0.000	.135402	.192791

```

rqdeco wage `X' if YEAR==2010, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      564698
Number of observations in group 0    16191
Number of observations in group 1    548507
```

```

Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]
<hr/>					
Quantile .05					
Raw difference	.092173	.00644	14.31	0.000	.079551 .104795

Characteristics	.070452	.005044	13.97	0.000	.060566	.080337
Coefficients	.021721	.007467	2.91	0.004	.007086	.036357
Quantile .2						
Raw difference	.119552	.002733	43.74	0.000	.114195	.124909
Characteristics	.047882	.003191	15.01	0.000	.041628	.054136
Coefficients	.07167	.003506	20.44	0.000	.064798	.078542
Quantile .35						
Raw difference	.139643	.002309	60.47	0.000	.135117	.144169
Characteristics	.043936	.003465	12.68	0.000	.037144	.050727
Coefficients	.095708	.003807	25.14	0.000	.088246	.103169
Quantile .5						
Raw difference	.157193	.002249	69.91	0.000	.152786	.1616
Characteristics	.044071	.003471	12.70	0.000	.037267	.050875
Coefficients	.113122	.003844	29.43	0.000	.105588	.120657
Quantile .65						
Raw difference	.178534	.002401	74.36	0.000	.173828	.18324
Characteristics	.04513	.003683	12.25	0.000	.03791	.052349
Coefficients	.133404	.004304	30.99	0.000	.124968	.141841
Quantile .8						
Raw difference	.207797	.00386	53.84	0.000	.200232	.215362
Characteristics	.0458	.004759	9.62	0.000	.036473	.055127
Coefficients	.161997	.006304	25.70	0.000	.149642	.174352
Quantile .95						
Raw difference	.239761	.009223	26.00	0.000	.221685	.257837
Characteristics	.036684	.009441	3.89	0.005	.01818	.055187
Coefficients	.203077	.012986	15.64	0.000	.177625	.22853

**2013


```

.      rqdeco wage `X' if YEAR==2013, by(native) qlow(.05) qhigh(.95) qste(.15) vce(boot)
Fitting base model
(bootstrapping .....)
```

Decomposition of differences in distribution using quantile regression

```

Total number of observations      598052
Number of observations in group 0    19065
Number of observations in group 1    578987
```

```

Number of quantile regressions estimated      100
```

The variance has been estimated by bootstrapping the results 50 times

Component	Effects	Std. Err.	t	P> t	[95% Conf. Interval]	
Quantile .05						
Raw difference	.064602	.005734	11.27	0.000	.053363	.075841
Characteristics	.068405	.0052	13.16	0.000	.058214	.078597
Coefficients	-.003803	.006598	-0.58	0.564	-.016735	.009128
Quantile .2						
Raw difference	.108598	.002555	42.51	0.000	.10359	.113605
Characteristics	.053167	.002601	20.44	0.000	.048069	.058265
Coefficients	.055431	.003238	17.12	0.000	.049085	.061777
Quantile .35						
Raw difference	.130367	.002408	54.14	0.000	.125648	.135086
Characteristics	.050423	.002335	21.59	0.000	.045846	.055
Coefficients	.079944	.003208	24.92	0.000	.073656	.086231
Quantile .5						

Raw difference	.145255	.002772	52.40	0.000	.139822	.150688
Characteristics	.050131	.002437	20.57	0.000	.045355	.054908
Coefficients	.095123	.003737	25.46	0.000	.0878	.102447
<hr/>						
Quantile .65						
Raw difference	.161484	.003186	50.68	0.000	.155239	.16773
Characteristics	.052366	.002958	17.71	0.000	.046569	.058162
Coefficients	.109119	.004516	24.16	0.000	.100268	.11797
<hr/>						
Quantile .8						
Raw difference	.183406	.004084	44.91	0.000	.175402	.191409
Characteristics	.057764	.004195	13.77	0.000	.049542	.065985
Coefficients	.125642	.005942	21.15	0.000	.113997	.137287
<hr/>						
Quantile .95						
Raw difference	.221972	.007687	28.88	0.000	.206905	.237038
Characteristics	.05292	.00824	6.42	0.000	.036769	.069071
Coefficients	.169052	.010518	16.07	0.000	.148437	.189666

Appendix E: Additional results from Oaxaca-Blinder decompositions

Appendix E.1: Mean predictions by native and immigrant males workers of different skilled groups

Table 3.13: Mean predictions by native and immigrant males workers of different skilled groups

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
All skilled group)					
Native	5.244*** (0.001)	5.364*** (0.001)	5.324*** (0.001)	5.352*** (0.001)	5.395*** (0.001)
Immigrant Total	5.150*** (0.004)	5.278*** (0.004)	5.239*** (0.002)	5.257*** (0.002)	5.297*** (0.002)
EU15	5.306*** (0.009)	5.438*** (0.008)	5.364*** (0.005)	5.393*** (0.005)	5.445*** (0.005)
EU12	5.130*** (0.022)	5.228*** (0.012)	5.188*** (0.006)	5.175*** (0.006)	5.172*** (0.005)
RD	5.343*** (0.018)	5.456*** (0.016)	5.393*** (0.011)	5.405*** (0.011)	5.465*** (0.011)
LD	5.063*** (0.005)	5.195*** (0.004)	5.172*** (0.003)	5.192*** (0.003)	5.249*** (0.003)
Low-skilled					
Native	5.086*** (0.002)	5.195*** (0.002)	5.175*** (0.001)	5.205*** (0.001)	5.210*** (0.001)
Immigrant Total	5.088*** (0.005)	5.232*** (0.005)	5.184*** (0.003)	5.210*** (0.003)	5.240*** (0.003)
EU15	5.237*** (0.010)	5.400*** (0.010)	5.308*** (0.007)	5.345*** (0.007)	5.404*** (0.007)
EU12	5.077*** (0.030)	5.193*** (0.014)	5.142*** (0.006)	5.140*** (0.006)	5.131*** (0.005)

Continued on next page...

... table 3.13 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
RD	5.254*** (0.028)	5.416*** (0.023)	5.372*** (0.016)	5.379*** (0.016)	5.465*** (0.017)
LD	5.022*** (0.005)	5.149*** (0.005)	5.121*** (0.003)	5.148*** (0.004)	5.186*** (0.004)
Medium-skilled					
Native	5.300*** (0.001)	5.422*** (0.001)	5.357*** (0.001)	5.386*** (0.001)	5.431*** (0.001)
Immigrant Total	5.234*** (0.008)	5.350*** (0.007)	5.277*** (0.004)	5.295*** (0.004)	5.329*** (0.004)
EU15	5.365*** (0.016)	5.478*** (0.015)	5.387*** (0.009)	5.408*** (0.009)	5.445*** (0.009)
EU12	5.161*** (0.034)	5.338*** (0.028)	5.257*** (0.015)	5.264*** (0.016)	5.251*** (0.013)
RD	5.366*** (0.029)	5.474*** (0.027)	5.365*** (0.018)	5.394*** (0.018)	5.412*** (0.018)
LD	5.157*** (0.009)	5.287*** (0.008)	5.222*** (0.005)	5.240*** (0.005)	5.288*** (0.004)
High-skilled					
Native	5.573*** (0.005)	5.609*** (0.004)	5.544*** (0.002)	5.587*** (0.002)	5.637*** (0.002)
Immigrant Total	5.379*** (0.017)	5.450*** (0.014)	5.494*** (0.008)	5.520*** (0.008)	5.537*** (0.007)
EU15	5.508*** (0.026)	5.573*** (0.022)	5.547*** (0.013)	5.575*** (0.012)	5.601*** (0.011)
EU12	5.315*** (0.059)	5.305*** (0.053)	5.641*** (0.032)	5.633*** (0.031)	5.593*** (0.024)
RD	5.520***	5.582***	5.511***	5.510***	5.548***

Continued on next page...

... table 3.13 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.040)	(0.038)	(0.022)	(0.021)	(0.023)
LD	5.185***	5.317***	5.407***	5.441***	5.471***
	(0.025)	(0.020)	(0.012)	(0.013)	(0.010)

Appendix E.2: Additional Oaxaca-Blinder decompositions on sub-samples

Table 3.14: Oaxaca-Blinder decomposition results for males of different skilled groups: Native-EU15 Immigrants

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
All Skilled groups					
difference	-0.062***	-0.074***	-0.039***	-0.042***	-0.049***
	(0.009)	(0.008)	(0.005)	(0.005)	(0.005)
explained	0.016***	0.041***	0.055***	0.059***	0.068***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
unexplained	-0.077***	-0.115***	-0.094***	-0.100***	-0.117***
	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)
Low-skilled					
	-0.151***	-0.205***	-0.133***	-0.140***	-0.194***
	(0.011)	(0.010)	(0.007)	(0.007)	(0.007)
explained	-0.034***	-0.039***	0.011***	0.015***	-0.014***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)

Continued on next page...

... table 3.14 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
unexplained	-0.117*** (0.010)	-0.166*** (0.010)	-0.144*** (0.007)	-0.155*** (0.007)	-0.180*** (0.007)
Medium-skilled					
difference	-0.065*** (0.016)	-0.057*** (0.015)	-0.030*** (0.009)	-0.021** (0.009)	-0.014 (0.009)
explained	-0.036*** (0.005)	-0.023*** (0.005)	0.001 (0.003)	0.002 (0.003)	0.004 (0.003)
unexplained	-0.029* (0.015)	-0.034** (0.013)	-0.031*** (0.009)	-0.024*** (0.008)	-0.018** (0.009)
High-skilled					
difference	0.065** (0.026)	0.036 (0.022)	-0.003 (0.013)	0.012 (0.012)	0.035*** (0.011)
explained	-0.013 (0.011)	-0.046*** (0.010)	0.001 (0.005)	-0.003 (0.005)	0.029*** (0.005)
unexplained	0.078*** (0.025)	0.082*** (0.021)	-0.004 (0.012)	0.015 (0.011)	0.007 (0.010)

Table 3.15: Oaxaca-Blinder decomposition results for males of different skilled groups: Native-
New EU12 Immigrants

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
All Skilled groups					
difference	0.114***	0.137***	0.136***	0.176***	0.223***

Continued on next page...

... table 3.15 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.022)	(0.012)	(0.006)	(0.006)	(0.005)
explained	0.080***	0.182***	0.195***	0.210***	0.236***
	(0.012)	(0.007)	(0.003)	(0.003)	(0.003)
unexplained	0.034*	-0.045***	-0.058***	-0.034***	-0.013***
Low-skilled					
difference	0.009	0.001	0.033***	0.065***	0.078***
	(0.030)	(0.014)	(0.006)	(0.006)	(0.005)
explained	0.058***	0.110***	0.119***	0.136***	0.099***
	(0.015)	(0.007)	(0.003)	(0.003)	(0.003)
unexplained	-0.049*	-0.109***	-0.086***	-0.071***	-0.021***
	(0.027)	(0.013)	(0.006)	(0.006)	(0.005)
Medium-skilled					
difference	0.139***	0.084***	0.100***	0.123***	0.180***
	(0.034)	(0.028)	(0.015)	(0.016)	(0.013)
explained	0.031**	0.020**	0.050***	0.066***	0.088***
	(0.012)	(0.010)	(0.006)	(0.007)	(0.006)
unexplained	0.107***	0.063**	0.050***	0.057***	0.092***
	(0.032)	(0.026)	(0.014)	(0.014)	(0.012)
High-skilled					
	0.258***	0.304***	-0.097***	-0.046	0.044*
	(0.059)	(0.053)	(0.032)	(0.031)	(0.024)
explained	-0.043	-0.010	0.016	0.046***	0.070***
	(0.031)	(0.029)	(0.010)	(0.011)	(0.009)
unexplained	0.301***	0.313***	-0.113***	-0.092***	-0.026
	(0.074)	(0.054)	(0.030)	(0.029)	(0.021)

Table 3.16: Oaxaca-Blinder decomposition results for males of different skilled groups:Native-Immigrants from remaining developed countries

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
All Skilled groups					
difference	-0.098*** (0.019)	-0.091*** (0.016)	-0.069*** (0.011)	-0.053*** (0.011)	-0.070*** (0.012)
explained	-0.056*** (0.010)	0.004 (0.008)	0.025*** (0.005)	0.035*** (0.005)	0.044*** (0.005)
unexplained	-0.042** (0.017)	-0.095*** (0.016)	-0.094*** (0.011)	-0.089*** (0.011)	-0.114*** (0.012)
Low-skilled					
difference	-0.168*** (0.028)	-0.222*** (0.023)	-0.197*** (0.016)	-0.174*** (0.016)	-0.256*** (0.017)
explained	-0.067*** (0.012)	-0.048*** (0.010)	-0.008 (0.005)	0.002 (0.006)	-0.017*** (0.005)
unexplained	-0.101*** (0.026)	-0.173*** (0.022)	-0.189*** (0.015)	-0.176*** (0.016)	-0.239*** (0.017)
Medium-skilled					
difference	-0.066** (0.029)	-0.053* (0.027)	-0.008 (0.018)	-0.008 (0.018)	0.019 (0.018)
explained	-0.067*** (0.010)	-0.069*** (0.011)	-0.014*** (0.005)	-0.012* (0.006)	-0.015*** (0.006)
unexplained	0.001 (0.027)	0.017 (0.025)	0.006 (0.017)	0.004 (0.017)	0.033** (0.017)
High-skilled					
difference	0.052 (0.041)	0.026 (0.039)	0.032 (0.022)	0.077*** (0.021)	0.088*** (0.023)
explained	-0.032* (0.017)	-0.030* (0.016)	-0.018* (0.011)	-0.019* (0.011)	-0.018** (0.012)

Continued on next page...

... table 3.16 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
	(0.017)	(0.018)	(0.009)	(0.011)	(0.009)
unexplained	0.085**	0.057*	0.051**	0.096***	0.107***
	(0.039)	(0.034)	(0.021)	(0.020)	(0.022)

Table 3.17: Oaxaca-Blinder decomposition results for males of different skilled groups: Native-Immigrants from less developed countries

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
All Skilled groups					
difference	0.181***	0.170***	0.153***	0.160***	0.146***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)
explained	0.086***	0.092***	0.087***	0.090***	0.089***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)
unexplained	0.095***	0.077***	0.065***	0.069***	0.057***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)
Low-skilled					
difference	0.064***	0.046***	0.054***	0.057***	0.023***
	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)
explained	0.013***	0.004	0.020***	0.024***	0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
unexplained	0.051***	0.042***	0.034***	0.033***	0.023***
	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)
Medium-skilled					

Continued on next page...

... table 3.17 continued

	(2004)	(2007)	(2009)	(2010)	(2013)
	log wage	log wage	log wage	log wage	log wage
	b/se	b/se	b/se	b/se	b/se
difference	0.142*** (0.009)	0.135*** (0.008)	0.135*** (0.005)	0.146*** (0.005)	0.143*** (0.004)
explained	-0.009*** (0.003)	0.004 (0.003)	0.021*** (0.002)	0.027*** (0.002)	0.033*** (0.002)
unexplained	0.151*** (0.009)	0.131*** (0.007)	0.114*** (0.005)	0.120*** (0.005)	0.110*** (0.004)
High-skilled					
difference	0.388*** (0.026)	0.292*** (0.020)	0.136*** (0.012)	0.146*** (0.013)	0.166*** (0.010)
explained	-0.011 (0.012)	-0.001 (0.010)	0.008 (0.005)	0.001 (0.005)	0.050*** (0.005)
unexplained	0.399*** (0.028)	0.293*** (0.021)	0.129*** (0.012)	0.145*** (0.013)	0.116*** (0.009)

Appendix F: Additional sensitivity analysis on the choice of explanatory variables (Based on Oaxaca-Blender decompositions)

Appendix F.1: Use the variables "Tenure and " and "Experience"

**Regression results before education correction for total immigrants in all skill groups.
 **Use the variables "experience and tenure" (Shown as "exp" and "Tenure" in the table) instead of "Years since migration".

*Note:

*Group 1: immigrant = 0

*Group 2: immigrant = 1

. esttab male_2004 male_2005 male_2006 male_2007 male_2008,se star(* 0.10 ** 0.05 *** 0.01)

	2004 wage	2005 wage	2006 wage	2007 wage	2008 wage
overall					
group_1	5.232*** (0.000827)	5.269*** (0.000849)	5.303*** (0.000807)	5.350*** (0.000837)	5.281*** (0.000524)
group_2	5.174*** (0.00339)	5.203*** (0.00347)	5.240*** (0.00326)	5.289*** (0.00317)	5.205*** (0.00204)
difference	0.0580*** (0.00349)	0.0655*** (0.00357)	0.0634*** (0.00336)	0.0614*** (0.00328)	0.0760*** (0.00210)
explained	-0.0621*** (0.00190)	-0.0588*** (0.00197)	-0.0514*** (0.00190)	-0.0428*** (0.00201)	-0.0239*** (0.00126)
unexplained	0.120*** (0.00336)	0.124*** (0.00345)	0.115*** (0.00323)	0.104*** (0.00322)	0.0999*** (0.00218)
explained ind	-0.000971** (0.000463)	-0.00143*** (0.000408)	-0.000224 (0.000339)	0.000565* (0.000323)	-0.000608*** (0.000607)

exp	0.0298*** (0.00103)	0.0455*** (0.00240)	0.0502*** (0.00183)	0.0586*** (0.00184)	0.0532*** (0.00110)
Tenure	-0.129*** (0.00308)	-0.00231 (0.00170)	0.00229** (0.00110)	0.00276*** (0.000926)	0.000156 (0.000423)
Union	-0.000195 (0.000170)	-0.000731*** (0.000176)	-0.00118*** (0.000169)	-0.00115*** (0.000175)	-0.000200* (0.000112)
Age_18_24	0.00559*** (0.00105)	0.00570*** (0.000937)	0.00360*** (0.000978)	0.00280*** (0.000936)	0.000385 (0.000308)
Age_25_29	-0.00855*** (0.00211)	-0.00664*** (0.00188)	-0.00997*** (0.00182)	-0.0113*** (0.00172)	-0.0119*** (0.000750)
Age_30_39	-0.0417*** (0.00351)	-0.0364*** (0.00321)	-0.0347*** (0.00308)	-0.0338*** (0.00288)	-0.0238*** (0.00129)
Age_40_49	-0.0233*** (0.00338)	-0.0316*** (0.00322)	-0.0298*** (0.00312)	-0.0289*** (0.00296)	-0.0134*** (0.00135)
Age_50_59	0.0488*** (0.00261)	0.0472*** (0.00241)	0.0461*** (0.00234)	0.0432*** (0.00219)	0.0336*** (0.00117)
Age_60_65	0.00585*** (0.00112)	0.00747*** (0.00103)	0.0115*** (0.00101)	0.0143*** (0.000955)	0.0125*** (0.000532)
basic_edu	-0.0197*** (0.00134)	-0.0176*** (0.00143)	-0.0218*** (0.00130)	-0.0258*** (0.00124)	-0.0219*** (0.000740)
vocational~u	-0.00321*** (0.000479)	-0.00331*** (0.000481)	-0.00111*** (0.000267)	0.0000326 (0.0000494)	-0.000594** (0.000275)
higher_edu	-0.00848*** (0.000602)	-0.00922*** (0.000625)	-0.00592*** (0.000465)	-0.00277*** (0.000315)	-0.00405*** (0.000271)

Married	-0.0105*** (0.000391)	-0.0100*** (0.000395)	-0.00837*** (0.000355)	-0.00745*** (0.000364)	-0.00472*** (0.000206)
children_0~s	-0.00379*** (0.000248)	-0.00358*** (0.000242)	-0.00365*** (0.000231)	-0.00298*** (0.000209)	-0.00223*** (0.000119)
Copenhagen	-0.00513*** (0.00133)	-0.00565*** (0.00148)	-0.00789*** (0.00137)	-0.00556*** (0.00126)	-0.00151* (0.000781)
large_city	-0.00217*** (0.000812)	-0.00144** (0.000734)	-0.00236*** (0.000693)	0.000445 (0.000451)	-0.000431* (0.000252)
small_city	0.00133 (0.00209)	-0.000162 (0.00216)	0.00279 (0.00198)	0.000802 (0.000760)	-0.000455 (0.000504)
lnemployee	-0.00716*** (0.000402)	-0.00687*** (0.000392)	-0.00603*** (0.000380)	-0.00711*** (0.000383)	-0.00503*** (0.000188)
<hr/>					
unexplained ind	-0.0273*** (0.00947)	-0.0379*** (0.00969)	-0.0268*** (0.00886)	-0.0275*** (0.00837)	-0.0338*** (0.00452)
exp	-0.0524*** (0.00931)	-0.0671*** (0.0248)	-0.0571*** (0.0170)	-0.0557*** (0.0157)	-0.0384*** (0.00898)
Tenure	-0.145*** (0.00301)	-0.0170* (0.0102)	-0.0126** (0.00612)	-0.00978** (0.00482)	-0.00674*** (0.00212)
Union	-0.0229*** (0.00645)	-0.0179*** (0.00642)	-0.0231*** (0.00592)	-0.0235*** (0.00587)	-0.00886** (0.00377)
Age_18_24	0.0152* (0.00818)	0.0168*** (0.00629)	0.00521 (0.00635)	0.0113** (0.00503)	0.0216*** (0.00344)

Age_25_29	0.0343** (0.0139)	0.0361*** (0.00984)	0.0129 (0.00926)	0.0244*** (0.00703)	0.0463*** (0.00589)
Age_30_39	0.112*** (0.0411)	0.121*** (0.0297)	0.0489* (0.0273)	0.0772*** (0.0204)	0.129*** (0.0159)
Age_40_49	0.0992*** (0.0337)	0.119*** (0.0268)	0.0583** (0.0260)	0.0909*** (0.0202)	0.148*** (0.0161)
Age_50_59	0.0425** (0.0185)	0.0504*** (0.0138)	0.0220 (0.0134)	0.0415*** (0.0105)	0.0786*** (0.00926)
Age_60_65	0.00628* (0.00329)	0.00873*** (0.00258)	0.00241 (0.00238)	0.00612*** (0.00200)	0.0135*** (0.00192)
basic_edu	0.00409 (0.00267)	0.00953*** (0.00282)	0.0192*** (0.00271)	0.0258*** (0.00270)	0.0283*** (0.00181)
vocational~u	0.0556*** (0.00970)	0.0756*** (0.0101)	0.0935*** (0.00890)	0.108*** (0.00809)	0.114*** (0.00575)
higher_edu	0.00662*** (0.00122)	0.00796*** (0.00126)	0.00980*** (0.00115)	0.0102*** (0.00111)	0.0108*** (0.000928)
Married	0.0437*** (0.00507)	0.0446*** (0.00504)	0.0372*** (0.00445)	0.0401*** (0.00435)	0.0379*** (0.00292)
children_0~s	0.00905*** (0.00239)	0.0113*** (0.00229)	0.0126*** (0.00213)	0.00695*** (0.00206)	0.00654*** (0.00138)
Copenhagen	0.00346 (0.00703)	0.00197 (0.00659)	0.00556 (0.00739)	-0.00464 (0.00643)	0.00642 (0.00444)
large_city	0.00567 (0.0133)	0.00430 (0.0123)	0.0103 (0.0136)	-0.00540 (0.0183)	0.00527 (0.0121)

small_city	0.00311 (0.0222)	-0.000325 (0.0214)	0.0136 (0.0243)	-0.00729 (0.0136)	0.00895 (0.00818)
lnemployee	-0.0513*** (0.00982)	-0.0763*** (0.0103)	-0.0847*** (0.00962)	-0.0722*** (0.00936)	-0.0531*** (0.00645)
_cons	-0.182 (0.127)	-0.208** (0.0989)	-0.0683 (0.0982)	-0.170** (0.0776)	-0.445*** (0.0592)
N	254278	256008	271069	261812	631497

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

```
.
.
.
. esttab male_2009 male_2010 male_2011 male_2012 male_2013,se star(* 0.10 ** 0.05 *** 0.01)
```

	2009 wage	2010 wage	2011 wage	2012 wage	2013 wage
overall					
group_1	5.319*** (0.000544)	5.347*** (0.000540)	5.338*** (0.000542)	5.358*** (0.000544)	5.390*** (0.000546)
group_2	5.245*** (0.00201)	5.264*** (0.00203)	5.265*** (0.00196)	5.274*** (0.00205)	5.304*** (0.00206)
difference	0.0736*** (0.00209)	0.0823*** (0.00210)	0.0728*** (0.00203)	0.0847*** (0.00212)	0.0858*** (0.00213)
explained	-0.0194***	0.0617***	-0.0179***	-0.0104***	-0.00744***

	(0.00127)	(0.00114)	(0.00145)	(0.00154)	(0.00159)
unexplained	0.0930*** (0.00215)	0.0206*** (0.00209)	0.0907*** (0.00211)	0.0951*** (0.00221)	0.0932*** (0.00224)
explained					
ind	-0.000314*** (0.0000437)	-0.000489*** (0.0000987)	-0.000676*** (0.0000733)	-0.000959*** (0.0000796)	-0.00101*** (0.0000781)
exp	0.0539*** (0.00117)	0.0735*** (0.00127)	0.0887*** (0.00133)	0.0997*** (0.00139)	0.107*** (0.00145)
Tenure	0.000964** (0.000483)	0.00432*** (0.000420)	0.00284*** (0.000324)	0.00178*** (0.000309)	0.00288*** (0.000299)
Union	-0.000332** (0.000132)	0.000809*** (0.000134)	-0.0000918 (0.000138)	0.000372*** (0.000143)	-0.000224 (0.000149)
Age_18_24	0.000777 (0.000579)	0.00145*** (0.000195)	0.000234** (0.000107)	-0.000650*** (0.000117)	-0.00108*** (0.000126)
Age_25_29	-0.0304*** (0.00121)	-0.0174*** (0.000737)	-0.0107*** (0.000475)	-0.00798*** (0.000414)	-0.00543*** (0.000356)
Age_30_39	-0.0349*** (0.00184)	-0.0203*** (0.00111)	-0.0143*** (0.000699)	-0.0147*** (0.000625)	-0.0148*** (0.000616)
Age_40_49	-0.0100*** (0.00193)	-0.00685*** (0.00119)	-0.00453*** (0.000742)	-0.000368 (0.000602)	0.00202*** (0.000520)
Age_50_59	0.0549*** (0.00161)	0.0317*** (0.00108)	0.0191*** (0.000673)	0.0165*** (0.000590)	0.0138*** (0.000535)
Age_60_65	0.0192*** (0.000782)	0.0115*** (0.000498)	0.00865*** (0.000337)	0.00688*** (0.000284)	0.00576*** (0.000254)

basic_edu	-0.0206*** (0.000727)	0.0267*** (0.000505)	-0.0278*** (0.000725)	-0.0275*** (0.000745)	-0.0288*** (0.000718)
vocational~u	-0.000646** (0.000275)	0.0112*** (0.000290)	-0.00437*** (0.000502)	-0.00443*** (0.000633)	-0.00627*** (0.000703)
higher_edu	-0.00386*** (0.000276)	0.00694*** (0.000192)	-0.00342*** (0.000266)	-0.00444*** (0.000287)	-0.00371*** (0.000271)
Married	-0.00342*** (0.000190)	-0.00427*** (0.000219)	-0.00460*** (0.000232)	-0.00413*** (0.000218)	-0.00422*** (0.000220)
children_0~s	-0.00192*** (0.000111)	-0.00125*** (0.0000904)	-0.00152*** (0.0000946)	-0.00164*** (0.0000957)	-0.00159*** (0.0000946)
Copenhagen	-0.00159* (0.000814)	-0.000742 (0.000810)	-0.00374*** (0.000797)	-0.00310*** (0.000878)	-0.00260*** (0.000891)
large_city	-0.000481 (0.000318)	-0.000669** (0.000311)	-0.000216 (0.000330)	-0.000174 (0.000381)	-0.000821* (0.000427)
small_city	-0.000337 (0.000481)	-0.000615 (0.000487)	-0.0000968 (0.000451)	-0.0000563 (0.000478)	-0.000673 (0.000451)
lnemployee	-0.00629*** (0.000213)	-0.00723*** (0.000232)	-0.00698*** (0.000232)	-0.00703*** (0.000228)	-0.00613*** (0.000227)

unexplained ind	-0.0288*** (0.00452)	-0.0362*** (0.00460)	-0.0236*** (0.00376)	-0.0148*** (0.00403)	-0.00880** (0.00406)
exp	-0.0416*** (0.00908)	0.0638*** (0.00915)	0.0160* (0.00863)	-0.00621 (0.00890)	0.00892 (0.00921)

Tenure	0.0122** (0.00478)	0.00219 (0.00455)	0.00562 (0.00376)	0.0315*** (0.00375)	0.0199*** (0.00372)
Union	-0.0101** (0.00391)	-0.0347*** (0.00386)	-0.0193*** (0.00376)	-0.0180*** (0.00388)	-0.0235*** (0.00400)
Age_18_24	-0.00314 (0.00596)	0.0125*** (0.00328)	0.00894*** (0.00195)	0.00990*** (0.00192)	0.0102*** (0.00188)
Age_25_29	0.00577 (0.0123)	0.0362*** (0.00795)	0.0270*** (0.00447)	0.0284*** (0.00469)	0.0279*** (0.00461)
Age_30_39	0.0163 (0.0295)	0.0799*** (0.0187)	0.0553*** (0.00992)	0.0588*** (0.0103)	0.0580*** (0.0106)
Age_40_49	0.0341 (0.0292)	0.0983*** (0.0187)	0.0685*** (0.00963)	0.0660*** (0.00946)	0.0620*** (0.00940)
Age_50_59	0.0128 (0.0168)	0.0511*** (0.0109)	0.0346*** (0.00550)	0.0364*** (0.00559)	0.0380*** (0.00585)
Age_60_65	0.00112 (0.00373)	0.00914*** (0.00234)	0.00627*** (0.00133)	0.00605*** (0.00138)	0.00620*** (0.00141)
basic_edu	0.0355*** (0.00178)	0.01000*** (0.000783)	0.0377*** (0.00199)	0.0412*** (0.00211)	0.0427*** (0.00206)
vocational~u	0.140*** (0.00603)	0.00358*** (0.000548)	0.136*** (0.00612)	0.150*** (0.00650)	0.153*** (0.00636)
higher_edu	0.0110*** (0.000968)	0.00297*** (0.000371)	0.0139*** (0.00106)	0.0198*** (0.00116)	0.0200*** (0.00115)
Married	0.0360*** (0.00282)	0.0425*** (0.00286)	0.0464*** (0.00259)	0.0488*** (0.00276)	0.0479*** (0.00277)

children_0~s	0.00413*** (0.00131)	0.00141 (0.00131)	0.00281** (0.00126)	-0.00253** (0.00125)	-0.00283** (0.00125)
Copenhagen	0.00663 (0.00437)	0.00245 (0.00547)	0.00367 (0.00450)	0.0160*** (0.00598)	0.0131* (0.00692)
large_city	0.00558 (0.0111)	0.000165 (0.0137)	-0.00133 (0.0110)	0.0128 (0.0139)	0.00808 (0.0158)
small_city	0.00515 (0.00758)	0.00226 (0.00939)	0.00314 (0.00759)	0.0113 (0.00967)	0.0118 (0.0112)
lnemployee	-0.0546*** (0.00605)	-0.0702*** (0.00619)	-0.105*** (0.00612)	-0.0990*** (0.00611)	-0.123*** (0.00600)
_cons	-0.111 (0.100)	-0.237*** (0.0684)	-0.224*** (0.0414)	-0.299*** (0.0457)	-0.275*** (0.0495)
N	599315	591553	647327	645573	635822

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Appendix F.2: Use the variables "Years since migration" only

**Regression results before education correction for total immigrants in high-skilled group.
 **Use the variable "Year since migration" (Shown as "YSM" in the table) instead of working history variables (experience and tenure).

*Note:

*Group 1: immigrant = 0

*Group 2: immigrant = 1

. esttab male_2004 male_2005 male_2006 male_2007 male_2008,se star(* 0.10 ** 0.05 *** 0.01)

	2004 wage	2005 wage	2006 wage	2007 wage	2008 wage
overall					
group_1	5.232*** (0.000827)	5.269*** (0.000849)	5.303*** (0.000807)	5.350*** (0.000837)	5.281*** (0.000524)
group_2	5.145*** (0.00423)	5.178*** (0.00423)	5.218*** (0.00391)	5.273*** (0.00372)	5.194*** (0.00239)
difference	0.0867*** (0.00431)	0.0907*** (0.00431)	0.0855*** (0.00399)	0.0771*** (0.00381)	0.0876*** (0.00245)
explained	-0.0181** (0.00753)	-0.0121 (0.00752)	-0.00232 (0.00689)	0.0334*** (0.00680)	0.0597*** (0.00465)
unexplained	0.105*** (0.00885)	0.103*** (0.00880)	0.0879*** (0.00803)	0.0438*** (0.00771)	0.0278*** (0.00527)
explained					
Age_18_24	-0.00551*** (0.00146)	-0.00380*** (0.00128)	-0.00702*** (0.00133)	-0.00704*** (0.00126)	-0.00454*** (0.000452)
Age_25_29	-0.0365***	-0.0285***	-0.0304***	-0.0305***	-0.0259***

	(0.00292)	(0.00256)	(0.00243)	(0.00228)	(0.00108)
Age_30_39	-0.103*** (0.00458)	-0.0919*** (0.00420)	-0.0872*** (0.00398)	-0.0811*** (0.00373)	-0.0584*** (0.00186)
Age_40_49	-0.00397 (0.00401)	-0.0162*** (0.00383)	-0.0168*** (0.00372)	-0.0207*** (0.00356)	-0.0139*** (0.00170)
Age_50_59	0.121*** (0.00268)	0.110*** (0.00257)	0.108*** (0.00245)	0.100*** (0.00237)	0.0810*** (0.00166)
Age_60_65	0.0221*** (0.000758)	0.0227*** (0.000750)	0.0268*** (0.000748)	0.0302*** (0.000758)	0.0270*** (0.000583)
basic_edu	-0.0224*** (0.00166)	-0.0197*** (0.00174)	-0.0231*** (0.00155)	-0.0235*** (0.00145)	-0.0175*** (0.000911)
vocational~u	-0.00474*** (0.000732)	-0.00464*** (0.000697)	-0.00194*** (0.000399)	0.000148 (0.000182)	0.00257*** (0.000403)
higher_edu	-0.0116*** (0.000811)	-0.0120*** (0.000820)	-0.00807*** (0.000613)	-0.00388*** (0.000426)	-0.00568*** (0.000378)
Married	-0.00779*** (0.000429)	-0.00756*** (0.000439)	-0.00572*** (0.000396)	-0.00433*** (0.000413)	-0.00206*** (0.000240)
children_0~s	-0.00595*** (0.000354)	-0.00571*** (0.000342)	-0.00588*** (0.000326)	-0.00480*** (0.000295)	-0.00385*** (0.000177)
YSM	0.0384*** (0.00713)	0.0446*** (0.00711)	0.0487*** (0.00638)	0.0751*** (0.00604)	0.0824*** (0.00393)
Copenhagen	-0.00525*** (0.00146)	-0.00535*** (0.00160)	-0.00833*** (0.00147)	-0.00558*** (0.00140)	-0.00137* (0.000830)

large_city	-0.00179** (0.000714)	-0.00109* (0.000649)	-0.00216*** (0.000628)	0.000457 (0.000456)	-0.000384 (0.000238)
small_city	0.00113 (0.00211)	-0.000713 (0.00218)	0.00320 (0.00200)	0.000961 (0.000890)	-0.000420 (0.000566)
ind	-0.000334 (0.000549)	-0.00127*** (0.000477)	0.000297 (0.000390)	0.00117*** (0.000365)	-0.000959*** (0.000776)
Union	-0.000668** (0.000285)	-0.00181*** (0.000291)	-0.00264*** (0.000271)	-0.00260*** (0.000272)	-0.00106*** (0.000176)
lnemployee	-0.00515*** (0.000450)	-0.00585*** (0.000443)	-0.00525*** (0.000429)	-0.00643*** (0.000429)	-0.00490*** (0.000215)
<hr/>					
unexplained					
Age_18_24	-0.00700 (0.0109)	0.00269 (0.00816)	-0.00755 (0.00786)	-0.000979 (0.00608)	0.00613 (0.00520)
Age_25_29	0.00317 (0.0183)	0.0197 (0.0125)	-0.000317 (0.0110)	0.0116 (0.00828)	0.0226*** (0.00873)
Age_30_39	0.0211 (0.0499)	0.0684* (0.0355)	0.0128 (0.0309)	0.0422* (0.0229)	0.0638*** (0.0224)
Age_40_49	0.0176 (0.0312)	0.0535** (0.0250)	0.0175 (0.0233)	0.0443** (0.0185)	0.0653*** (0.0191)
Age_50_59	0.00164 (0.00790)	0.00910 (0.00660)	0.00115 (0.00653)	0.0100* (0.00552)	0.0166*** (0.00610)
Age_60_65	-0.000682 (0.000820)	0.000908 (0.000723)	-0.000401 (0.000666)	0.000477 (0.000574)	0.00116 (0.000709)
basic_edu	0.00948***	0.0129***	0.0215***	0.0213***	0.0186***

	(0.00288)	(0.00301)	(0.00280)	(0.00271)	(0.00175)
vocational~u	0.0752*** (0.0129)	0.105*** (0.0129)	0.111*** (0.0105)	0.104*** (0.00904)	0.102*** (0.00616)
higher_edu	0.00907*** (0.00178)	0.0111*** (0.00174)	0.0111*** (0.00145)	0.00982*** (0.00133)	0.0104*** (0.00107)
Married	0.0525*** (0.00583)	0.0472*** (0.00579)	0.0403*** (0.00494)	0.0437*** (0.00466)	0.0371*** (0.00323)
children_0~s	0.00531* (0.00318)	0.00868*** (0.00297)	0.0118*** (0.00274)	0.00610** (0.00259)	0.00538*** (0.00174)
YSM	-0.000536 (0.00407)	0.00558 (0.00437)	-0.0122*** (0.00411)	-0.0284*** (0.00436)	-0.0346*** (0.00294)
Copenhagen	0.00111 (0.00850)	0.00310 (0.00760)	0.000586 (0.00822)	-0.00927 (0.00804)	0.00459 (0.00532)
large_city	0.00378 (0.0149)	0.00782 (0.0133)	0.000959 (0.0142)	-0.0142 (0.0216)	0.00848 (0.0140)
small_city	0.00962 (0.0252)	0.0166 (0.0234)	0.00516 (0.0256)	-0.00911 (0.0157)	0.0122 (0.00923)
ind	-0.0242** (0.0114)	-0.0317*** (0.0117)	-0.0238** (0.0103)	-0.0336*** (0.00943)	-0.0256*** (0.00516)
Union	-0.0315*** (0.00969)	-0.0218** (0.00958)	-0.0274*** (0.00842)	-0.0167** (0.00814)	0.000381 (0.00530)
lnemployee	-0.0973*** (0.0125)	-0.112*** (0.0124)	-0.105*** (0.0114)	-0.103*** (0.0108)	-0.0736*** (0.00746)

_cons	0.101 (0.131)	-0.0498 (0.102)	0.0753 (0.0964)	0.00595 (0.0794)	-0.187*** (0.0693)
N	250379	252282	267213	258022	622279

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

```
.
.
.
. esttab male_2009 male_2010 male_2011 male_2012 male_2013,se star(* 0.10 ** 0.05 *** 0.01)
```

	2009 wage	2010 wage	2011 wage	2012 wage	2013 wage
overall					
group_1	5.319*** (0.000544)	5.347*** (0.000540)	5.338*** (0.000542)	5.358*** (0.000544)	5.390*** (0.000546)
group_2	5.237*** (0.00232)	5.255*** (0.00231)	5.255*** (0.00221)	5.263*** (0.00230)	5.294*** (0.00229)
difference	0.0823*** (0.00239)	0.0922*** (0.00237)	0.0829*** (0.00227)	0.0958*** (0.00237)	0.0958*** (0.00235)
explained	0.0643*** (0.00425)	0.150*** (0.00338)	0.0732*** (0.00461)	0.0629*** (0.00480)	0.0733*** (0.00479)
unexplained	0.0180*** (0.00489)	-0.0577*** (0.00423)	0.00970* (0.00511)	0.0329*** (0.00538)	0.0225*** (0.00535)
explained Age_18_24	-0.00686***	-0.000476**	-0.0000458	-0.000543***	-0.000932***

	(0.000761)	(0.000216)	(0.0000327)	(0.000104)	(0.000148)
Age_25_29	-0.0525*** (0.00160)	-0.0284*** (0.00103)	-0.0160*** (0.000670)	-0.0111*** (0.000586)	-0.00677*** (0.000505)
Age_30_39	-0.0791*** (0.00238)	-0.0457*** (0.00157)	-0.0297*** (0.00102)	-0.0270*** (0.000915)	-0.0251*** (0.000897)
Age_40_49	-0.00966*** (0.00233)	-0.00841*** (0.00149)	-0.00632*** (0.000956)	-0.00176** (0.000787)	0.00117* (0.000682)
Age_50_59	0.118*** (0.00186)	0.0693*** (0.00153)	0.0416*** (0.000998)	0.0346*** (0.000889)	0.0287*** (0.000821)
Age_60_65	0.0402*** (0.000704)	0.0245*** (0.000569)	0.0181*** (0.000436)	0.0149*** (0.000397)	0.0124*** (0.000376)
basic_edu	-0.0127*** (0.000865)	0.0303*** (0.000543)	-0.0171*** (0.000889)	-0.0174*** (0.000934)	-0.0167*** (0.000898)
vocational~u	0.00486*** (0.000491)	0.0145*** (0.000319)	0.00752*** (0.000789)	0.00951*** (0.000967)	0.0121*** (0.00105)
higher_edu	-0.00562*** (0.000402)	0.00850*** (0.000205)	-0.00500*** (0.000407)	-0.00690*** (0.000427)	-0.00606*** (0.000424)
Married	-0.000884*** (0.000219)	-0.00173*** (0.000256)	-0.00200*** (0.000273)	-0.00176*** (0.000255)	-0.00207*** (0.000257)
children_0~s	-0.00331*** (0.000160)	-0.00232*** (0.000137)	-0.00260*** (0.000137)	-0.00270*** (0.000132)	-0.00257*** (0.000130)
YSM	0.0769*** (0.00345)	0.0879*** (0.00313)	0.0835*** (0.00345)	0.0690*** (0.00349)	0.0739*** (0.00342)

Copenhagen	-0.00144* (0.000872)	-0.000185 (0.000867)	-0.00329*** (0.000856)	-0.00258*** (0.000942)	-0.00231** (0.000934)
large_city	-0.000469 (0.000309)	-0.000729** (0.000311)	-0.000309 (0.000355)	-0.000218 (0.000394)	-0.000755* (0.000426)
small_city	-0.000364 (0.000551)	-0.000764 (0.000544)	-0.000150 (0.000483)	-0.0000488 (0.000531)	-0.000575 (0.000493)
ind	-0.000521*** (0.000572)	-0.000927*** (0.000109)	-0.000887*** (0.0000794)	-0.00117*** (0.0000845)	-0.00108*** (0.0000796)
Union	-0.00114*** (0.000192)	0.000549*** (0.000190)	-0.000799*** (0.000195)	-0.0000841 (0.000195)	-0.000913*** (0.000198)
lnemployee	-0.00634*** (0.000240)	-0.00753*** (0.000263)	-0.00708*** (0.000262)	-0.00727*** (0.000257)	-0.00634*** (0.000256)
<hr/>					
unexplained					
Age_18_24	-0.0126* (0.00730)	-0.00419 (0.00683)	-0.00246 (0.00434)	-0.00799* (0.00450)	-0.00336 (0.00469)
Age_25_29	-0.00888 (0.0147)	0.00231 (0.0165)	0.00369 (0.0102)	-0.0141 (0.0115)	-0.00299 (0.0120)
Age_30_39	-0.0159 (0.0336)	0.00438 (0.0370)	0.00893 (0.0218)	-0.0275 (0.0245)	-0.00702 (0.0270)
Age_40_49	0.000144 (0.0282)	0.0116 (0.0321)	0.0175 (0.0187)	-0.0143 (0.0203)	0.00228 (0.0216)
Age_50_59	-0.00476 (0.00945)	-0.00172 (0.0117)	0.00237 (0.00698)	-0.00871 (0.00822)	-0.000975 (0.00965)
Age_60_65	-0.00114	-0.000683	-0.0000676	-0.00177	-0.000543

	(0.00129)	(0.00147)	(0.000902)	(0.00108)	(0.00124)
basic_edu	0.0216*** (0.00171)	0.00885*** (0.000817)	0.0199*** (0.00188)	0.0262*** (0.00201)	0.0258*** (0.00197)
vocational~u	0.110*** (0.00614)	-0.000672 (0.000550)	0.0989*** (0.00611)	0.119*** (0.00653)	0.116*** (0.00639)
higher_edu	0.00830*** (0.00106)	0.00159*** (0.000375)	0.0110*** (0.00113)	0.0184*** (0.00125)	0.0174*** (0.00123)
Married	0.0357*** (0.00307)	0.0421*** (0.00308)	0.0477*** (0.00278)	0.0503*** (0.00297)	0.0496*** (0.00297)
children_0~s	0.00239 (0.00162)	-0.00203 (0.00157)	0.00123 (0.00150)	-0.00491*** (0.00146)	-0.00597*** (0.00145)
YSM	-0.0364*** (0.00264)	-0.00820*** (0.00224)	-0.0397*** (0.00277)	-0.0483*** (0.00289)	-0.0506*** (0.00285)
Copenhagen	0.00289 (0.00500)	-0.000913 (0.00657)	-0.000720 (0.00507)	0.0154** (0.00713)	0.00705 (0.00735)
large_city	0.00308 (0.0122)	0.00117 (0.0159)	-0.00532 (0.0120)	0.0150 (0.0159)	0.000355 (0.0163)
small_city	0.00404 (0.00810)	0.00457 (0.0106)	0.000959 (0.00818)	0.0135 (0.0109)	0.00633 (0.0113)
ind	-0.0272*** (0.00508)	-0.0335*** (0.00505)	-0.0163*** (0.00414)	-0.00475 (0.00447)	0.00325 (0.00449)
Union	-0.00306 (0.00538)	-0.0138*** (0.00521)	-0.0143*** (0.00498)	-0.0192*** (0.00493)	-0.0300*** (0.00502)

lnemployee	-0.0718*** (0.00687)	-0.0765*** (0.00691)	-0.119*** (0.00678)	-0.116*** (0.00673)	-0.137*** (0.00660)
_cons	0.0328 (0.0986)	0.0293 (0.110)	0.0102 (0.0685)	0.0459 (0.0781)	0.0341 (0.0847)
N	591124	583878	639509	638098	628873

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Conclusion

This PhD thesis compared and applied various econometric frameworks on important topics within the field of labour economics. The empirical studies were based on individual-level administrative labour market data from Germany and Denmark. Each of the three chapters in this thesis has made a unique and novel contribution within their perspective fields.

Chapter 1, "On Omitted Variables, Proxies and Unobserved Effects in Analysis of Administrative Labour Market Data", provided a unified framework that nests various approaches aiming at reducing omitted variable bias in linear regression analysis. Two empirical applications—wage regression and a linear probability model on labour market transition—were conducted using the linked German administrative data. Evidence of sizable omitted variable bias for a number of variables was found in the wage regression, while only a small number of coefficients was systematically affected in the transition analysis. Additional survey variables only contribute to the wage model, but the use of work history variables and panel models could result in changes in coefficients in the two models. Empirical evidence obtained suggested that panel analysis was expected to capture more relevant unobservable components than an expanded regressor set at one point of time. Our results can also be used as a guide for both researchers and data providers on the choice of variables in the data application.

Chapter 2, "Impact of Immigration on the Wages of Native Workers in Denmark", have investigated on the case where an increase in the number of immigrants, especially the high-skilled workers, were found in the local labour market in Denmark. Based on the hypothesised mechanisms, I argued that the positive wage effects of low-skilled immigrants could be due to wage efficiency theory and the occupation upgrading behaviour of natives. Then, the increase in medium-skilled natives and immigrants generates even fiercer competition in this market as the share of medium-skilled immigrants increases; the outcome is reduced wages for natives. For high-skilled groups, hiring standards and wage levels both rise due to high-wage firms' motivation to reduce costs associated with labour replacement and to hire more productive workers. OLS, FE, FE-IV and quantile regression models were presented for the empirical application. I tested the impact of immigrants on the wages of natives using Danish register data on the full population in Denmark for the period from 2004 to 2013. The evidence obtained confirmed the prediction of wage efficiency theory that there is a positive impact from high-skilled immigrants on the wages of natives, especially high-wage earners, while a positive impact from low-skilled immigrant was presented in low and medium wage quantiles. Finally, a negative impact from medium-skilled immigrants on the wages of natives was shown by the results. Chapter 2 implemented the existing knowledge on the impact of immigration on the wages of native workers in Denmark, during the period (2004-2013) and provided empirical evidence within both various wage quantiles and different skill levels.

Chapter 3, "Analysis on Native-Immigrant Wage Gap in Denmark", explored the native-immigrant wage gap for male immigrants in different wage quantiles for the labour market in Denmark, using Danish register data covering the same period as in Chapter 2. I started the analysis by assessing changes in the population distributions and wage distributions of the native and immigrant groups. I then applied the Oaxaca-Blinder and Melly (2005) decomposition approaches to gain empirical insights from the native-immigrant wage gap. I also examined the existence of possible wage discrimination⁸ against immigrants in the labour mar-

⁸When I use the term "discrimination" in this chapter, I refer to the remaining component after I have controlled

ket in Denmark from 2004 to 2013. It was found that for low-wage earners in the low-skilled and high-skilled groups, the wage gap and potential discrimination level were larger than for other wage quantiles within the same skill category. The medium-skilled group, however, presented a higher wage gap and potential discrimination among the high-wage population. The absolute value of possible discrimination component increased with the skill level. Moreover, stronger potential discrimination also occurred for the upper wage quantiles in each group of origin. Immigrants from less developed countries, most of which were non-EU countries were empirically found to experience the strongest and most positive wage penalties and potential discrimination.

Overall, this thesis discussed various econometrics frameworks and compared the empirical evidence obtained from them. Each chapter contributed to our understanding of critical issues within the topics of labour economics. Large-scale individual-level administrative labour market data is used for the empirical studies in this paper, and the results obtained provide fruitful insights into widely discussed research questions in the field of labour economics and applied econometrics. In the future, based on such rich and high-quality data, it could be interesting to examine the impact from additional factors, e.g. policy or assimilation effects, on topics such as wage and immigration in labour market, applying or extending existing empirical framework.

for all available possible factors in the wage model I can obtain. This term is not the same as the general political definition in practical.

TITLER I PH.D.SERIEN:

– *a Field Study of the Rise and Fall of a Bottom-Up Process*

2004

1. Martin Grieger
Internet-based Electronic Marketplaces and Supply Chain Management
2. Thomas Basbøll
*LIKENESS
A Philosophical Investigation*
3. Morten Knudsen
*Beslutningens vaklen
En systemteoretisk analyse af moderniseringen af et amtskommunalt sundhedsvæsen 1980-2000*
4. Lars Bo Jeppesen
*Organizing Consumer Innovation
A product development strategy that is based on online communities and allows some firms to benefit from a distributed process of innovation by consumers*
5. Barbara Dragsted
*SEGMENTATION IN TRANSLATION AND TRANSLATION MEMORY SYSTEMS
An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process*
6. Jeanet Hardis
*Sociale partnerskaber
Et socialkonstruktivistisk casestudie af partnerskabsaktørers virkelighedsopfattelse mellem identitet og legitimitet*
7. Henriette Hallberg Thygesen
System Dynamics in Action
8. Carsten Mejer Plath
Strategisk Økonomistyring
9. Annemette Kjærgaard
Knowledge Management as Internal Corporate Venturing
10. Knut Arne Hovdal
*De professionelle i endring
Norsk ph.d., ej til salg gennem Samfundslitteratur*
11. Søren Jeppesen
*Environmental Practices and Greening Strategies in Small Manufacturing Enterprises in South Africa
– A Critical Realist Approach*
12. Lars Frode Frederiksen
*Industriel forskningsledelse
– på sporet af mønstre og samarbejde i danske forskningsintensive virksomheder*
13. Martin Jes Iversen
*The Governance of GN Great Nordic
– in an age of strategic and structural transitions 1939-1988*
14. Lars Pynt Andersen
*The Rhetorical Strategies of Danish TV Advertising
A study of the first fifteen years with special emphasis on genre and irony*
15. Jakob Rasmussen
Business Perspectives on E-learning
16. Sof Thrane
*The Social and Economic Dynamics of Networks
– a Weberian Analysis of Three Formalised Horizontal Networks*
17. Lene Nielsen
Engaging Personas and Narrative Scenarios – a study on how a user-centered approach influenced the perception of the design process in the e-business group at AstraZeneca
18. S.J Valstad
*Organisationsidentitet
Norsk ph.d., ej til salg gennem Samfundslitteratur*

19. Thomas Lyse Hansen
Six Essays on Pricing and Weather risk in Energy Markets
 20. Sabine Madsen
Emerging Methods – An Interpretive Study of ISD Methods in Practice
 21. Evis Sinani
The Impact of Foreign Direct Investment on Efficiency, Productivity Growth and Trade: An Empirical Investigation
 22. Bent Meier Sørensen
Making Events Work Or, How to Multiply Your Crisis
 23. Pernille Schnoor
Brand Ethos
Om troværdige brand- og virksomhedsidentiteter i et retorisk og diskursteoretisk perspektiv
 24. Sidsel Fabech
Von welchem Österreich ist hier die Rede?
Diskursive forhandlinger og magtkampe mellem rivaliserende nationale identitetskonstruktioner i østrigske pressediskurser
 25. Klavs Odgaard Christensen
Sprogpolitik og identitetsdannelse i flersprogede forbundsstater
Et komparativt studie af Schweiz og Canada
 26. Dana B. Minbaeva
Human Resource Practices and Knowledge Transfer in Multinational Corporations
 27. Holger Højlund
Markedets politiske fornuft
Et studie af velfærdens organisering i perioden 1990-2003
 28. Christine Mølgaard Frandsen
A.s erfaring
Om mellemværendets praktik i en transformation af mennesket og subjektiviteten
 29. Sine Nørholm Just
The Constitution of Meaning – A Meaningful Constitution?
Legitimacy, identity, and public opinion in the debate on the future of Europe
- 2005**
1. Claus J. Varnes
Managing product innovation through rules – The role of formal and structured methods in product development
 2. Helle Hedegaard Hein
Mellem konflikt og konsensus – Dialogudvikling på hospitalsklinikker
 3. Axel Rosenø
Customer Value Driven Product Innovation – A Study of Market Learning in New Product Development
 4. Søren Buhl Pedersen
Making space
An outline of place branding
 5. Camilla Funck Ellehave
Differences that Matter
An analysis of practices of gender and organizing in contemporary workplaces
 6. Rigmor Madeleine Lond
Styring af kommunale forvaltninger
 7. Mette Aagaard Andreassen
Supply Chain versus Supply Chain Benchmarking as a Means to Managing Supply Chains
 8. Caroline Aggestam-Pontoppidan
From an idea to a standard
The UN and the global governance of accountants' competence
 9. Norsk ph.d.
 10. Vivienne Heng Ker-ni
An Experimental Field Study on the

- Effectiveness of Grocer Media Advertising*
Measuring Ad Recall and Recognition, Purchase Intentions and Short-Term Sales
11. Allan Mortensen
Essays on the Pricing of Corporate Bonds and Credit Derivatives
 12. Remo Stefano Chiari
Figure che fanno conoscere
Itinerario sull'idea del valore cognitivo e espressivo della metafora e di altri troppi da Aristotele e da Vico fino al cognitivismo contemporaneo
 13. Anders McIlquham-Schmidt
Strategic Planning and Corporate Performance
An integrative research review and a meta-analysis of the strategic planning and corporate performance literature from 1956 to 2003
 14. Jens Geersbro
The TDF – PMI Case
Making Sense of the Dynamics of Business Relationships and Networks
 15. Mette Andersen
Corporate Social Responsibility in Global Supply Chains
Understanding the uniqueness of firm behaviour
 16. Eva Boxenbaum
Institutional Genesis: Micro – Dynamic Foundations of Institutional Change
 17. Peter Lund-Thomsen
Capacity Development, Environmental Justice NGOs, and Governance: The Case of South Africa
 18. Signe Jarlov
Konstruktioner af offentlig ledelse
 19. Lars Stæhr Jensen
Vocabulary Knowledge and Listening Comprehension in English as a Foreign Language
 20. Christian Nielsen
Essays on Business Reporting
Production and consumption of strategic information in the market for information
 21. Marianne Thejls Fischer
Egos and Ethics of Management Consultants
 22. Annie Bekke Kjær
Performance management i Process-innovation
– belyst i et social-konstruktivistisk perspektiv
 23. Suzanne Dee Pedersen
GENTAGELSENS METAMORFOSE
Om organisering af den kreative gøren i den kunstneriske arbejdspraksis
 24. Benedikte Dorte Rosenbrink
Revenue Management
Økonomiske, konkurrencemæssige & organisatoriske konsekvenser
 25. Thomas Riise Johansen
Written Accounts and Verbal Accounts
The Danish Case of Accounting and Accountability to Employees
 26. Ann Fogelgren-Pedersen
The Mobile Internet: Pioneering Users' Adoption Decisions
 27. Birgitte Rasmussen
Ledelse i fællesskab – de tillidsvalgte fornyende rolle
 28. Gitte Thit Nielsen
Remerger
– skabende ledelseskrafter i fusion og opkøb
 29. Carmine Gioia
A MICROECONOMETRIC ANALYSIS OF MERGERS AND ACQUISITIONS

30. Ole Hinz
Den effektive forandringsleder: pilot, pædagog eller politiker?
Et studie i arbejdslederes meningstilskrivninger i forbindelse med vellykket gennemførelse af ledelsesinitierede forandringsprojekter
31. Kjell-Åge Gotvassli
Et praksisbasert perspektiv på dynamiske læringsnettverk i toppidretten
Norsk ph.d., ej til salg gennem Samfundslitteratur
32. Henriette Langstrup Nielsen
Linking Healthcare
An inquiry into the changing performances of web-based technology for asthma monitoring
33. Karin Tweddell Levinsen
Virtuel Uddannelsespraksis
Master i IKT og Læring – et casestudie i hvordan proaktiv proceshåndtering kan forbedre praksis i virtuelle læringsmiljøer
34. Anika Liversage
Finding a Path
Labour Market Life Stories of Immigrant Professionals
35. Kasper Elmquist Jørgensen
Studier i samspillet mellem stat og erhvervsliv i Danmark under 1. verdenskrig
36. Finn Janning
A DIFFERENT STORY
Seduction, Conquest and Discovery
37. Patricia Ann Plackett
Strategic Management of the Radical Innovation Process
Leveraging Social Capital for Market Uncertainty Management
2. Niels Rom-Poulsen
Essays in Computational Finance
3. Tina Brandt Husman
Organisational Capabilities, Competitive Advantage & Project-Based Organisations
The Case of Advertising and Creative Good Production
4. Mette Rosenkrands Johansen
Practice at the top
– how top managers mobilise and use non-financial performance measures
5. Eva Parum
Corporate governance som strategisk kommunikations- og ledelsesværktøj
6. Susan Aagaard Petersen
Culture's Influence on Performance Management: The Case of a Danish Company in China
7. Thomas Nicolai Pedersen
The Discursive Constitution of Organizational Governance – Between unity and differentiation
The Case of the governance of environmental risks by World Bank environmental staff
8. Cynthia Selin
Volatile Visions: Transactions in Anticipatory Knowledge
9. Jesper Banghøj
Financial Accounting Information and Compensation in Danish Companies
10. Mikkel Lucas Overby
Strategic Alliances in Emerging High-Tech Markets: What's the Difference and does it Matter?
11. Tine Aage
External Information Acquisition of Industrial Districts and the Impact of Different Knowledge Creation Dimensions

2006

1. Christian Vintergaard
Early Phases of Corporate Venturing

- A case study of the Fashion and Design Branch of the Industrial District of Montebelluna, NE Italy*
12. Mikkel Flyverbom
Making the Global Information Society Governable
On the Governmentality of Multi-Stakeholder Networks
 13. Anette Grønning
Personen bag
Tilstedevær i e-mail som interaktionsform mellem kunde og medarbejder i dansk forsikringskontekst
 14. Jørn Helder
One Company – One Language?
The NN-case
 15. Lars Bjerregaard Mikkelsen
Differing perceptions of customer value
Development and application of a tool for mapping perceptions of customer value at both ends of customer-supplier dyads in industrial markets
 16. Lise Granerud
Exploring Learning
Technological learning within small manufacturers in South Africa
 17. Esben Rahbek Pedersen
Between Hopes and Realities: Reflections on the Promises and Practices of Corporate Social Responsibility (CSR)
 18. Ramona Samson
The Cultural Integration Model and European Transformation.
The Case of Romania
- 2007**
1. Jakob Vestergaard
Discipline in The Global Economy
Panopticism and the Post-Washington Consensus
 2. Heidi Lund Hansen
Spaces for learning and working
A qualitative study of change of work, management, vehicles of power and social practices in open offices
 3. Sudhanshu Rai
Exploring the internal dynamics of software development teams during user analysis
A tension enabled Institutionalization Model; "Where process becomes the objective"
 4. Norsk ph.d.
Ej til salg gennem Samfundslitteratur
 5. Serden Ozcan
EXPLORING HETEROGENEITY IN ORGANIZATIONAL ACTIONS AND OUTCOMES
A Behavioural Perspective
 6. Kim Sundtoft Hald
Inter-organizational Performance Measurement and Management in Action
– An Ethnography on the Construction of Management, Identity and Relationships
 7. Tobias Lindeberg
Evaluative Technologies
Quality and the Multiplicity of Performance
 8. Merete Wedell-Wedellsborg
Den globale soldat
Identitetsdannelse og identitetsledelse i multinationale militære organisationer
 9. Lars Frederiksen
Open Innovation Business Models
Innovation in firm-hosted online user communities and inter-firm project ventures in the music industry
– A collection of essays
 10. Jonas Gabrielsen
Retorisk toposlære – fra statisk 'sted' til persuasiv aktivitet

11. Christian Moldt-Jørgensen
Fra meningsløs til meningsfuld evaluering.
Anvendelsen af studentertilfredsheds-målinger på de korte og mellemlange videregående uddannelser set fra et psykodynamisk systemperspektiv
12. Ping Gao
Extending the application of actor-network theory
Cases of innovation in the telecommunications industry
13. Peter Mejlby
Frihed og fængsel, en del af den samme drøm?
Et phronetisk baseret casestudie af frigørelsens og kontrollens sam-eksistens i værdibaseret ledelse!
14. Kristina Birch
Statistical Modelling in Marketing
15. Signe Poulsen
Sense and sensibility:
The language of emotional appeals in insurance marketing
16. Anders Bjerre Trolle
Essays on derivatives pricing and dynamic asset allocation
17. Peter Feldhütter
Empirical Studies of Bond and Credit Markets
18. Jens Henrik Eggert Christensen
Default and Recovery Risk Modeling and Estimation
19. Maria Theresa Larsen
Academic Enterprise: A New Mission for Universities or a Contradiction in Terms?
Four papers on the long-term implications of increasing industry involvement and commercialization in academia
20. Morten Wellendorf
Postimplementering af teknologi i den offentlige forvaltning
Analyser af en organisations kontinuerlige arbejde med informations-teknologi
21. Ekaterina Mhaanna
Concept Relations for Terminological Process Analysis
22. Stefan Ring Thorbjørnsen
Forsvaret i forandring
Et studie i officerers kapabiliteter under påvirkning af omverdenens forandringspres mod øget styring og læring
23. Christa Breum Amhøj
Det selvskabte medlemskab om managementstaten, dens styringsteknologier og indbyggere
24. Karoline Bromose
Between Technological Turbulence and Operational Stability
– An empirical case study of corporate venturing in TDC
25. Susanne Justesen
Navigating the Paradoxes of Diversity in Innovation Practice
– A Longitudinal study of six very different innovation processes – in practice
26. Luise Noring Henler
Conceptualising successful supply chain partnerships
– Viewing supply chain partnerships from an organisational culture perspective
27. Mark Mau
Kampen om telefonen
Det danske telefonvæsen under den tyske besættelse 1940-45
28. Jakob Halskov
The semiautomatic expansion of existing terminological ontologies using knowledge patterns discovered

- on the WWW – an implementation and evaluation*
29. Gergana Koleva
European Policy Instruments Beyond Networks and Structure: The Innovative Medicines Initiative
 30. Christian Geisler Asmussen
Global Strategy and International Diversity: A Double-Edged Sword?
 31. Christina Holm-Petersen
*Stolthed og fordom
Kultur- og identitetsarbejde ved skabelsen af en ny sengeafdeling gennem fusion*
 32. Hans Peter Olsen
*Hybrid Governance of Standardized States
Causes and Contours of the Global Regulation of Government Auditing*
 33. Lars Bøge Sørensen
Risk Management in the Supply Chain
 34. Peter Aagaard
*Det unikkes dynamikker
De institutionelle mulighedsbetingelser bag den individuelle udforskning i professionelt og frivilligt arbejde*
 35. Yun Mi Antorini
*Brand Community Innovation
An Intrinsic Case Study of the Adult Fans of LEGO Community*
 36. Joachim Lynggaard Boll
*Labor Related Corporate Social Performance in Denmark
Organizational and Institutional Perspectives*
- 2008**
1. Frederik Christian Vinten
Essays on Private Equity
 2. Jesper Clement
Visual Influence of Packaging Design on In-Store Buying Decisions
 3. Marius Brostrøm Kousgaard
*Tid til kvalitetsmåling?
– Studier af indrulleringsprocesser i forbindelse med introduktionen af kliniske kvalitetsdatabaser i speciallægepraksissektoren*
 4. Irene Skovgaard Smith
*Management Consulting in Action
Value creation and ambiguity in client-consultant relations*
 5. Anders Rom
*Management accounting and integrated information systems
How to exploit the potential for management accounting of information technology*
 6. Marina Candi
Aesthetic Design as an Element of Service Innovation in New Technology-based Firms
 7. Morten Schnack
*Teknologi og tværfaglighed
– en analyse af diskussionen omkring indførelse af EPJ på en hospitalsafdeling*
 8. Helene Balslev Clausen
Juntos pero no revueltos – un estudio sobre emigrantes norteamericanos en un pueblo mexicano
 9. Lise Justesen
*Kunsten at skrive revisionsrapporter.
En beretning om forvaltningsrevisions beretninger*
 10. Michael E. Hansen
The politics of corporate responsibility: CSR and the governance of child labor and core labor rights in the 1990s
 11. Anne Roepstorff
Holdning for handling – en etnologisk undersøgelse af Virksomheders Sociale Ansvar/CSR

12. Claus Bajlum
Essays on Credit Risk and Credit Derivatives
 13. Anders Bojesen
The Performative Power of Competence – an Inquiry into Subjectivity and Social Technologies at Work
 14. Satu Reijonen
*Green and Fragile
A Study on Markets and the Natural Environment*
 15. Ilduara Busta
*Corporate Governance in Banking
A European Study*
 16. Kristian Anders Hvass
*A Boolean Analysis Predicting Industry Change: Innovation, Imitation & Business Models
The Winning Hybrid: A case study of isomorphism in the airline industry*
 17. Trine Paludan
*De uvidende og de udviklingsparate
Identitet som mulighed og restriktion
blandt fabriksarbejdere på det aftayloriserede fabriksgulv*
 18. Kristian Jakobsen
Foreign market entry in transition economies: Entry timing and mode choice
 19. Jakob Elming
Syntactic reordering in statistical machine translation
 20. Lars Brømsøe Termansen
*Regional Computable General Equilibrium Models for Denmark
Three papers laying the foundation for regional CGE models with agglomeration characteristics*
 21. Mia Reinholt
The Motivational Foundations of Knowledge Sharing
 22. Frederikke Krogh-Meibom
*The Co-Evolution of Institutions and Technology
– A Neo-Institutional Understanding of Change Processes within the Business Press – the Case Study of Financial Times*
 23. Peter D. Ørberg Jensen
OFFSHORING OF ADVANCED AND HIGH-VALUE TECHNICAL SERVICES: ANTECEDENTS, PROCESS DYNAMICS AND FIRMLEVEL IMPACTS
 24. Pham Thi Song Hanh
Functional Upgrading, Relational Capability and Export Performance of Vietnamese Wood Furniture Producers
 25. Mads Vangkilde
*Why wait?
An Exploration of first-mover advantages among Danish e-grocers through a resource perspective*
 26. Hubert Buch-Hansen
*Rethinking the History of European Level Merger Control
A Critical Political Economy Perspective*
- 2009**
1. Vivian Lindhardsen
From Independent Ratings to Communal Ratings: A Study of CWA Raters' Decision-Making Behaviours
 2. Guðrið Weihe
Public-Private Partnerships: Meaning and Practice
 3. Chris Nøkkentved
*Enabling Supply Networks with Collaborative Information Infrastructures
An Empirical Investigation of Business Model Innovation in Supplier Relationship Management*
 4. Sara Louise Muhr
Wound, Interrupted – On the Vulnerability of Diversity Management

5. Christine Sestoft
Forbrugeradfærd i et Stats- og Livsformsteoretisk perspektiv
6. Michael Pedersen
Tune in, Breakdown, and Reboot: On the production of the stress-fit self-managing employee
7. Salla Lutz
Position and Reposition in Networks – Exemplified by the Transformation of the Danish Pine Furniture Manufacturers
8. Jens Forssbæk
Essays on market discipline in commercial and central banking
9. Tine Murphy
Sense from Silence – A Basis for Organised Action
How do Sensemaking Processes with Minimal Sharing Relate to the Reproduction of Organised Action?
10. Sara Malou Strandvad
Inspirations for a new sociology of art: A sociomaterial study of development processes in the Danish film industry
11. Nicolaas Mouton
On the evolution of social scientific metaphors: A cognitive-historical enquiry into the divergent trajectories of the idea that collective entities – states and societies, cities and corporations – are biological organisms.
12. Lars Andreas Knutsen
Mobile Data Services: Shaping of user engagements
13. Nikolaos Theodoros Korfiatis
Information Exchange and Behavior
A Multi-method Inquiry on Online Communities
14. Jens Albæk
Forestillinger om kvalitet og tværfaglighed på sygehuse
– skabelse af forestillinger i læge- og plejegrupperne angående relevans af nye idéer om kvalitetsudvikling gennem tolkningsprocesser
15. Maja Lotz
The Business of Co-Creation – and the Co-Creation of Business
16. Gitte P. Jakobsen
Narrative Construction of Leader Identity in a Leader Development Program Context
17. Dorte Hermansen
“Living the brand” som en brandorienteret dialogisk praxis: Om udvikling af medarbejdernes brandorienterede dømmekraft
18. Aseem Kinra
Supply Chain (logistics) Environmental Complexity
19. Michael Nørager
How to manage SMEs through the transformation from non innovative to innovative?
20. Kristin Wallevik
Corporate Governance in Family Firms
The Norwegian Maritime Sector
21. Bo Hansen Hansen
Beyond the Process
Enriching Software Process Improvement with Knowledge Management
22. Annemette Skot-Hansen
Franske adjektivisk afledte adverbier, der tager præpositionssyntagmer indledt med præpositionen à som argumenter
En valensgrammatisk undersøgelse
23. Line Gry Knudsen
Collaborative R&D Capabilities
In Search of Micro-Foundations

- | | |
|--|---|
| <p>24. Christian Scheuer
<i>Employers meet employees
Essays on sorting and globalization</i></p> <p>25. Rasmus Johnsen
<i>The Great Health of Melancholy
A Study of the Pathologies of Performativity</i></p> <p>26. Ha Thi Van Pham
<i>Internationalization, Competitiveness
Enhancement and Export Performance
of Emerging Market Firms:
Evidence from Vietnam</i></p> <p>27. Henriette Balieu
<i>Kontrolbegrebets betydning for kausalalternationen i spansk
En kognitiv-typologisk analyse</i></p> | <p><i>End User Participation between Processes of Organizational and Architectural Design</i></p> <p>7. Rex Degnegaard
<i>Strategic Change Management
Change Management Challenges in the Danish Police Reform</i></p> <p>8. Ulrik Schultz Brix
<i>Værdi i rekruttering – den sikre beslutning
En pragmatisk analyse af perception og synliggørelse af værdi i rekrutterings- og udvælgelsesarbejdet</i></p> <p>9. Jan Ole Similä
<i>Kontraktsledelse
Relasjonen mellom virksomhetsledelse og kontraktshåndtering, belyst via fire norske virksomheter</i></p> |
| 2010 | |
| <p>1. Yen Tran
<i>Organizing Innovation in Turbulent Fashion Market
Four papers on how fashion firms create and appropriate innovation value</i></p> <p>2. Anders Raastrup Kristensen
<i>Metaphysical Labour
Flexibility, Performance and Commitment in Work-Life Management</i></p> <p>3. Margrét Sigrún Sigurdardóttir
<i>Dependently independent
Co-existence of institutional logics in the recorded music industry</i></p> <p>4. Ásta Dis Óladóttir
<i>Internationalization from a small domestic base:
An empirical analysis of Economics and Management</i></p> <p>5. Christine Secher
<i>E-deltagelse i praksis – politikernes og forvaltningens medkonstruktion og konsekvenserne heraf</i></p> <p>6. Marianne Stang Våland
<i>What we talk about when we talk about space:</i></p> | <p>10. Susanne Boch Waldorff
<i>Emerging Organizations: In between local translation, institutional logics and discourse</i></p> <p>11. Brian Kane
<i>Performance Talk
Next Generation Management of Organizational Performance</i></p> <p>12. Lars Ohnemus
<i>Brand Thrust: Strategic Branding and Shareholder Value
An Empirical Reconciliation of two Critical Concepts</i></p> <p>13. Jesper Schlamovitz
<i>Håndtering af usikkerhed i film- og byggeprojekter</i></p> <p>14. Tommy Moesby-Jensen
<i>Det faktiske livs forbindtlighed
Førsokratisk informeret, ny-aristotelisk ἦθος-tænkning hos Martin Heidegger</i></p> <p>15. Christian Fich
<i>Two Nations Divided by Common Values
French National Habitus and the Rejection of American Power</i></p> |

16. Peter Beyer
Processer, sammenhængskraft og fleksibilitet
Et empirisk casestudie af omstillingsforløb i fire virksomheder
17. Adam Buchhorn
Markets of Good Intentions
Constructing and Organizing Biogas Markets Amid Fragility and Controversy
18. Cecilie K. Moesby-Jensen
Social læring og fælles praksis
Et mixed method studie, der belyser læringskonsekvenser af et lederkursus for et praksisfællesskab af offentlige mellemledere
19. Heidi Boye
Fødevarer og sundhed i senmodernismen
– En indsigt i hyggefænomenet og de relaterede fødevarerpraksisser
20. Kristine Munkgård Pedersen
Flygtige forbindelser og midlertidige mobiliseringer
Om kulturel produktion på Roskilde Festival
21. Oliver Jacob Weber
Causes of Intercompany Harmony in Business Markets – An Empirical Investigation from a Dyad Perspective
22. Susanne Ekman
Authority and Autonomy
Paradoxes of Modern Knowledge Work
23. Anette Frey Larsen
Kvalitetsledelse på danske hospitaler
– Ledelsernes indflydelse på introduktion og vedligeholdelse af kvalitetsstrategier i det danske sundhedsvæsen
24. Toyoko Sato
Performativity and Discourse: Japanese Advertisements on the Aesthetic Education of Desire
25. Kenneth Brinch Jensen
Identifying the Last Planner System
Lean management in the construction industry
26. Javier Busquets
Orchestrating Network Behavior for Innovation
27. Luke Patey
The Power of Resistance: India's National Oil Company and International Activism in Sudan
28. Mette Vedel
Value Creation in Triadic Business Relationships. Interaction, Interconnection and Position
29. Kristian Tørning
Knowledge Management Systems in Practice – A Work Place Study
30. Qingxin Shi
An Empirical Study of Thinking Aloud
Usability Testing from a Cultural Perspective
31. Tanja Juul Christiansen
Corporate blogging: Medarbejderes kommunikative handlekraft
32. Malgorzata Ciesielska
Hybrid Organisations.
A study of the Open Source – business setting
33. Jens Dick-Nielsen
Three Essays on Corporate Bond Market Liquidity
34. Sabrina Speiermann
Modstandens Politik
Kampagnestyling i Velfærdsstaten.
En diskussion af trafikcampagners styringspotentiale
35. Julie Uldam
Fickle Commitment. Fostering political engagement in 'the flighty world of online activism'

36. Annegrete Juul Nielsen
Traveling technologies and transformations in health care
37. Athur Mühlen-Schulte
Organising Development Power and Organisational Reform in the United Nations Development Programme
38. Louise Rygaard Jonas
Branding på butiksgulvet Et case-studie af kultur- og identitets-arbejdet i Kvickly
8. Ole Helby Petersen
Public-Private Partnerships: Policy and Regulation – With Comparative and Multi-level Case Studies from Denmark and Ireland
9. Morten Krogh Petersen
'Good' Outcomes. Handling Multiplicity in Government Communication
10. Kristian Tangsgaard Hvelplund
Allocation of cognitive resources in translation - an eye-tracking and key-logging study

2011

1. Stefan Fraenkel
Key Success Factors for Sales Force Readiness during New Product Launch A Study of Product Launches in the Swedish Pharmaceutical Industry
2. Christian Plesner Rossing
International Transfer Pricing in Theory and Practice
3. Tobias Dam Hede
Samtalekunst og ledelsesdisciplin – en analyse af coachingsdiskursens genealogi og governmentality
4. Kim Pettersson
Essays on Audit Quality, Auditor Choice, and Equity Valuation
5. Henrik Merkelsen
The expert-lay controversy in risk research and management. Effects of institutional distances. Studies of risk definitions, perceptions, management and communication
6. Simon S. Torp
Employee Stock Ownership: Effect on Strategic Management and Performance
7. Mie Harder
Internal Antecedents of Management Innovation
11. Moshe Yonatany
The Internationalization Process of Digital Service Providers
12. Anne Vestergaard
Distance and Suffering Humanitarian Discourse in the age of Mediatization
13. Thorsten Mikkelsen
Personlighedens indflydelse på forretningsrelationer
14. Jane Thostrup Jagd
Hvorfor fortsætter fusionsbølgen ud-over "the tipping point"? – en empirisk analyse af information og kognitioner om fusioner
15. Gregory Gimpel
Value-driven Adoption and Consumption of Technology: Understanding Technology Decision Making
16. Thomas Stengade Sønderskov
Den nye mulighed Social innovation i en forretningsmæssig kontekst
17. Jeppe Christoffersen
Donor supported strategic alliances in developing countries
18. Vibeke Vad Baunsgaard
Dominant Ideological Modes of Rationality: Cross functional

- integration in the process of product innovation*
19. Throstur Olaf Sigurjonsson
Governance Failure and Iceland's Financial Collapse
 20. Allan Sall Tang Andersen
Essays on the modeling of risks in interest-rate and inflation markets
 21. Heidi Tscherning
Mobile Devices in Social Contexts
 22. Birgitte Gorm Hansen
*Adapting in the Knowledge Economy
Lateral Strategies for Scientists and Those Who Study Them*
 23. Kristina Vaarst Andersen
*Optimal Levels of Embeddedness
The Contingent Value of Networked Collaboration*
 24. Justine Grønbæk Pors
*Noisy Management
A History of Danish School Governing from 1970-2010*
 25. Stefan Linder
*Micro-foundations of Strategic Entrepreneurship
Essays on Autonomous Strategic Action*
 26. Xin Li
*Toward an Integrative Framework of National Competitiveness
An application to China*
 27. Rune Thorbjørn Clausen
*Værdifuld arkitektur
Et eksplorativt studie af bygningers rolle i virksomheders værdiskabelse*
 28. Monica Viken
Markedsundersøkelser som bevis i varemerke- og markedsføringsrett
 29. Christian Wymann
*Tattooing
The Economic and Artistic Constitution of a Social Phenomenon*
 30. Sanne Frandsen
*Productive Incoherence
A Case Study of Branding and Identity Struggles in a Low-Prestige Organization*
 31. Mads Stenbo Nielsen
Essays on Correlation Modelling
 32. Ivan Häuser
*Følelse og sprog
Etablering af en ekspressiv kategori, eksemplificeret på russisk*
 33. Sebastian Schwenen
Security of Supply in Electricity Markets
- 2012**
1. Peter Holm Andreasen
*The Dynamics of Procurement Management
- A Complexity Approach*
 2. Martin Haulrich
Data-Driven Bitext Dependency Parsing and Alignment
 3. Line Kirkegaard
*Konsulent i den anden nat
En undersøgelse af det intense arbejdsliv*
 4. Tonny Stenheim
Decision usefulness of goodwill under IFRS
 5. Morten Lind Larsen
*Produktivit, vækst og velfærd
Industrirådet og efterkrigstidens Danmark 1945 - 1958*
 6. Petter Berg
Cartel Damages and Cost Asymmetries
 7. Lynn Kahle
*Experiential Discourse in Marketing
A methodical inquiry into practice and theory*
 8. Anne Roelsgaard Obling
*Management of Emotions
in Accelerated Medical Relationships*

9. Thomas Frandsen
Managing Modularity of Service Processes Architecture
10. Carina Christine Skovmøller
*CSR som noget særligt
Et casestudie om styring og menings-
skabelse i relation til CSR ud fra en
intern optik*
11. Michael Tell
*Fradragsbeskæring af selskabers
finansieringsudgifter
En skatteretlig analyse af SEL §§ 11,
11B og 11C*
12. Morten Holm
*Customer Profitability Measurement
Models
Their Merits and Sophistication
across Contexts*
13. Katja Joo Dyppel
*Beskatning af derivater
En analyse af dansk skatteret*
14. Esben Anton Schultz
*Essays in Labor Economics
Evidence from Danish Micro Data*
15. Carina Risvig Hansen
*"Contracts not covered, or not fully
covered, by the Public Sector Directive"*
16. Anja Svejgaard Pors
*Iværksættelse af kommunikation
- patientfigurer i hospitalets strategiske
kommunikation*
17. Frans Bévort
*Making sense of management with
logics
An ethnographic study of accountants
who become managers*
18. René Kallestrup
*The Dynamics of Bank and Sovereign
Credit Risk*
19. Brett Crawford
*Revisiting the Phenomenon of Interests
in Organizational Institutionalism
The Case of U.S. Chambers of
Commerce*
20. Mario Daniele Amore
Essays on Empirical Corporate Finance
21. Arne Stjernholm Madsen
*The evolution of innovation strategy
Studied in the context of medical
device activities at the pharmaceutical
company Novo Nordisk A/S in the
period 1980-2008*
22. Jacob Holm Hansen
*Is Social Integration Necessary for
Corporate Branding?
A study of corporate branding
strategies at Novo Nordisk*
23. Stuart Webber
*Corporate Profit Shifting and the
Multinational Enterprise*
24. Helene Ratner
*Promises of Reflexivity
Managing and Researching
Inclusive Schools*
25. Therese Strand
*The Owners and the Power: Insights
from Annual General Meetings*
26. Robert Gavin Strand
*In Praise of Corporate Social
Responsibility Bureaucracy*
27. Nina Sormunen
*Auditor's going-concern reporting
Reporting decision and content of the
report*
28. John Bang Mathiasen
*Learning within a product development
working practice:
- an understanding anchored
in pragmatism*
29. Philip Holst Riis
*Understanding Role-Oriented Enterprise
Systems: From Vendors to Customers*
30. Marie Lisa Dacanay
*Social Enterprises and the Poor
Enhancing Social Entrepreneurship and
Stakeholder Theory*

- | | |
|---|---|
| <p>31. Fumiko Kano Glückstad
<i>Bridging Remote Cultures: Cross-lingual concept mapping based on the information receiver's prior-knowledge</i></p> <p>32. Henrik Barslund Fosse
<i>Empirical Essays in International Trade</i></p> <p>33. Peter Alexander Albrecht
<i>Foundational hybridity and its reproduction
Security sector reform in Sierra Leone</i></p> <p>34. Maja Rosenstock
<i>CSR - hvor svært kan det være?
Kulturanalytisk casestudie om udfordringer og dilemmaer med at forankre Coops CSR-strategi</i></p> <p>35. Jeanette Rasmussen
<i>Tweens, medier og forbrug
Et studie af 10-12 årige danske børns brug af internettet, opfattelse og forståelse af markedsføring og forbrug</i></p> <p>36. Ib Tunby Gulbrandsen
<i>'This page is not intended for a US Audience'
A five-act spectacle on online communication, collaboration & organization.</i></p> <p>37. Kasper Aalling Teilmann
<i>Interactive Approaches to Rural Development</i></p> <p>38. Mette Mogensen
<i>The Organization(s) of Well-being and Productivity
(Re)assembling work in the Danish Post</i></p> <p>39. Søren Friis Møller
<i>From Disinterestedness to Engagement
Towards Relational Leadership In the Cultural Sector</i></p> <p>40. Nico Peter Berhausen
<i>Management Control, Innovation and Strategic Objectives – Interactions and Convergence in Product Development Networks</i></p> | <p>41. Balder Onarheim
<i>Creativity under Constraints
Creativity as Balancing 'Constrainedness'</i></p> <p>42. Haoyong Zhou
<i>Essays on Family Firms</i></p> <p>43. Elisabeth Naima Mikkelsen
<i>Making sense of organisational conflict
An empirical study of enacted sense-making in everyday conflict at work</i></p> <p>2013</p> <p>1. Jacob Lyngsie
<i>Entrepreneurship in an Organizational Context</i></p> <p>2. Signe Groth-Brodersen
<i>Fra ledelse til selvet
En socialpsykologisk analyse af forholdet imellem selvledelse, ledelse og stress i det moderne arbejdsliv</i></p> <p>3. Nis Høyrup Christensen
<i>Shaping Markets: A Neoinstitutional Analysis of the Emerging Organizational Field of Renewable Energy in China</i></p> <p>4. Christian Edelvold Berg
<i>As a matter of size
THE IMPORTANCE OF CRITICAL MASS AND THE CONSEQUENCES OF SCARCITY FOR TELEVISION MARKETS</i></p> <p>5. Christine D. Isakson
<i>Coworker Influence and Labor Mobility
Essays on Turnover, Entrepreneurship and Location Choice in the Danish Maritime Industry</i></p> <p>6. Niels Joseph Jerne Lennon
<i>Accounting Qualities in Practice
Rhizomatic stories of representational faithfulness, decision making and control</i></p> <p>7. Shannon O'Donnell
<i>Making Ensemble Possible
How special groups organize for collaborative creativity in conditions of spatial variability and distance</i></p> |
|---|---|

8. Robert W. D. Veitch
Access Decisions in a Partly-Digital World
Comparing Digital Piracy and Legal Modes for Film and Music
9. Marie Mathiesen
Making Strategy Work
An Organizational Ethnography
10. Arisa Shollo
The role of business intelligence in organizational decision-making
11. Mia Kaspersen
The construction of social and environmental reporting
12. Marcus Møller Larsen
The organizational design of offshoring
13. Mette Ohm Rørdam
EU Law on Food Naming
The prohibition against misleading names in an internal market context
14. Hans Peter Rasmussen
GIV EN GED!
Kan giver-idealtyper forklare støtte til vælgørenhed og understøtte relationsopbygning?
15. Ruben Schachtenhaufen
Fonetisk reduktion i dansk
16. Peter Koerver Schmidt
Dansk CFC-beskatning
I et internationalt og komparativt perspektiv
17. Morten Froholdt
Strategi i den offentlige sektor
En kortlægning af styringsmæssig kontekst, strategisk tilgang, samt anvendte redskaber og teknologier for udvalgte danske statslige styrelser
18. Annette Camilla Sjørup
Cognitive effort in metaphor translation
An eye-tracking and key-logging study
19. Tamara Stucchi
The Internationalization of Emerging Market Firms: A Context-Specific Study
20. Thomas Lopdrup-Hjorth
"Let's Go Outside": The Value of Co-Creation
21. Ana Alačovska
Genre and Autonomy in Cultural Production
The case of travel guidebook production
22. Marius Gudmand-Høyer
Stemningssindssygdommenes historie i det 19. århundrede
Omtydningen af melankolien og manien som bipolære stemningslidelser i dansk sammenhæng under hensyn til dannelsen af det moderne følelseslivs relative autonomi.
En problematiserings- og erfarings-analytisk undersøgelse
23. Lichen Alex Yu
Fabricating an S&OP Process
Circulating References and Matters of Concern
24. Esben Alfort
The Expression of a Need
Understanding search
25. Trine Pallesen
Assembling Markets for Wind Power
An Inquiry into the Making of Market Devices
26. Anders Koed Madsen
Web-Visions
Repurposing digital traces to organize social attention
27. Lærke Højgaard Christiansen
BREWING ORGANIZATIONAL RESPONSES TO INSTITUTIONAL LOGICS
28. Tommy Kjær Lassen
EGENTLIG SELVLEDELSE
En ledelsesfilosofisk afhandling om selvledelsens paradoksale dynamik og eksistentielle engagement

- | | |
|--|---|
| <p>29. Morten Rossing
<i>Local Adaption and Meaning Creation in Performance Appraisal</i></p> <p>30. Søren Obed Madsen
<i>Lederen som oversætter
Et oversættelsesteoretisk perspektiv på strategisk arbejde</i></p> <p>31. Thomas Høgenhaven
<i>Open Government Communities
Does Design Affect Participation?</i></p> <p>32. Kirstine Zinck Pedersen
<i>Failsafe Organizing?
A Pragmatic Stance on Patient Safety</i></p> <p>33. Anne Petersen
<i>Hverdagslogikker i psykiatrisk arbejde
En institutionsetnografisk undersøgelse af hverdagen i psykiatriske organisationer</i></p> <p>34. Didde Maria Humle
<i>Fortællinger om arbejde</i></p> <p>35. Mark Holst-Mikkelsen
<i>Strategieksekverering i praksis – barrierer og muligheder!</i></p> <p>36. Malek Maalouf
<i>Sustaining lean
Strategies for dealing with organizational paradoxes</i></p> <p>37. Nicolaj Tofte Brenneche
<i>Systemic Innovation In The Making
The Social Productivity of Cartographic Crisis and Transitions in the Case of SEEIT</i></p> <p>38. Morten Gylling
<i>The Structure of Discourse
A Corpus-Based Cross-Linguistic Study</i></p> <p>39. Binzhang YANG
<i>Urban Green Spaces for Quality Life - Case Study: the landscape architecture for people in Copenhagen</i></p> | <p>40. Michael Friis Pedersen
<i>Finance and Organization:
The Implications for Whole Farm Risk Management</i></p> <p>41. Even Fallan
<i>Issues on supply and demand for environmental accounting information</i></p> <p>42. Ather Nawaz
<i>Website user experience
A cross-cultural study of the relation between users' cognitive style, context of use, and information architecture of local websites</i></p> <p>43. Karin Beukel
<i>The Determinants for Creating Valuable Inventions</i></p> <p>44. Arjan Markus
<i>External Knowledge Sourcing and Firm Innovation
Essays on the Micro-Foundations of Firms' Search for Innovation</i></p> <p>2014</p> <p>1. Solon Moreira
<i>Four Essays on Technology Licensing and Firm Innovation</i></p> <p>2. Karin Strzeletz Ivertsen
<i>Partnership Drift in Innovation Processes
A study of the Think City electric car development</i></p> <p>3. Kathrine Hoffmann Pii
<i>Responsibility Flows in Patient-centred Prevention</i></p> <p>4. Jane Bjørn Vedel
<i>Managing Strategic Research
An empirical analysis of science-industry collaboration in a pharmaceutical company</i></p> <p>5. Martin Gylling
<i>Processuel strategi i organisationer
Monografi om dobbeltheden i tænkning af strategi, dels som vidensfelt i organisationsteori, dels som kunstnerisk tilgang til at skabe i erhvervsmæssig innovation</i></p> |
|--|---|

6. Linne Marie Lauesen
Corporate Social Responsibility in the Water Sector: How Material Practices and their Symbolic and Physical Meanings Form a Colonising Logic
7. Maggie Qiuzhu Mei
LEARNING TO INNOVATE: The role of ambidexterity, standard, and decision process
8. Inger Høedt-Rasmussen
Developing Identity for Lawyers Towards Sustainable Lawyering
9. Sebastian Fux
Essays on Return Predictability and Term Structure Modelling
10. Thorbjørn N. M. Lund-Poulsen
Essays on Value Based Management
11. Oana Brindusa Albu
Transparency in Organizing: A Performative Approach
12. Lena Olaison
Entrepreneurship at the limits
13. Hanne Sørum
DRESSED FOR WEB SUCCESS? An Empirical Study of Website Quality in the Public Sector
14. Lasse Folke Henriksen
Knowing networks How experts shape transnational governance
15. Maria Halbinger
Entrepreneurial Individuals Empirical Investigations into Entrepreneurial Activities of Hackers and Makers
16. Robert Spliid
Kapitalfondenes metoder og kompetencer
17. Christiane Stelling
Public-private partnerships & the need, development and management of trusting A processual and embedded exploration
18. Marta Gasparin
Management of design as a translation process
19. Kåre Moberg
Assessing the Impact of Entrepreneurship Education From ABC to PhD
20. Alexander Cole
Distant neighbors Collective learning beyond the cluster
21. Martin Møller Boje Rasmussen
Is Competitiveness a Question of Being Alike? How the United Kingdom, Germany and Denmark Came to Compete through their Knowledge Regimes from 1993 to 2007
22. Anders Ravn Sørensen
Studies in central bank legitimacy, currency and national identity Four cases from Danish monetary history
23. Nina Bellak
Can Language be Managed in International Business? Insights into Language Choice from a Case Study of Danish and Austrian Multinational Corporations (MNCs)
24. Rikke Kristine Nielsen
Global Mindset as Managerial Meta-competence and Organizational Capability: Boundary-crossing Leadership Cooperation in the MNC The Case of 'Group Mindset' in Solar A/S.
25. Rasmus Koss Hartmann
User Innovation inside government Towards a critically performative foundation for inquiry

- | | |
|---|---|
| <p>26. Kristian Gylling Olesen
<i>Flertydig og emergerende ledelse i folkeskolen</i>
<i>Et aktør-netværksteoretisk ledelsesstudie af politiske evalueringsreformers betydning for ledelse i den danske folkeskole</i></p> <p>27. Troels Riis Larsen
<i>Kampen om Danmarks omdømme 1945-2010</i>
<i>Omdømmearbejde og omdømmepolitik</i></p> <p>28. Klaus Majgaard
<i>Jagten på autenticitet i offentlig styring</i></p> <p>29. Ming Hua Li
<i>Institutional Transition and Organizational Diversity: Differentiated internationalization strategies of emerging market state-owned enterprises</i></p> <p>30. Sofie Blinkenberg Federspiel
<i>IT, organisation og digitalisering: Institutionelt arbejde i den kommunale digitaliseringsproces</i></p> <p>31. Elvi Weinreich
<i>Hvilke offentlige ledere er der brug for når velfærdstænkningen flytter sig – er Diplomuddannelsens lederprofil svaret?</i></p> <p>32. Ellen Mølgaard Korsager
<i>Self-conception and image of context in the growth of the firm</i>
<i>– A Penrosian History of Fiberline Composites</i></p> <p>33. Else Skjold
<i>The Daily Selection</i></p> <p>34. Marie Louise Conradsen
<i>The Cancer Centre That Never Was</i>
<i>The Organisation of Danish Cancer Research 1949-1992</i></p> <p>35. Virgilio Failla
<i>Three Essays on the Dynamics of Entrepreneurs in the Labor Market</i></p> | <p>36. Nicky Nedergaard
<i>Brand-Based Innovation</i>
<i>Relational Perspectives on Brand Logics and Design Innovation Strategies and Implementation</i></p> <p>37. Mads Gjedsted Nielsen
<i>Essays in Real Estate Finance</i></p> <p>38. Kristin Martina Brandl
<i>Process Perspectives on Service Offshoring</i></p> <p>39. Mia Rosa Koss Hartmann
<i>In the gray zone</i>
<i>With police in making space for creativity</i></p> <p>40. Karen Ingerslev
<i>Healthcare Innovation under The Microscope</i>
<i>Framing Boundaries of Wicked Problems</i></p> <p>41. Tim Neerup Thomsen
<i>Risk Management in large Danish public capital investment programmes</i></p> <p>2015</p> <p>1. Jakob Ion Wille
<i>Film som design</i>
<i>Design af levende billeder i film og tv-serier</i></p> <p>2. Christiane Mossin
<i>Interzones of Law and Metaphysics</i>
<i>Hierarchies, Logics and Foundations of Social Order seen through the Prism of EU Social Rights</i></p> <p>3. Thomas Tøth
<i>TRUSTWORTHINESS: ENABLING GLOBAL COLLABORATION</i>
<i>An Ethnographic Study of Trust, Distance, Control, Culture and Boundary Spanning within Offshore Outsourcing of IT Services</i></p> <p>4. Steven Højlund
<i>Evaluation Use in Evaluation Systems – The Case of the European Commission</i></p> |
|---|---|

5. Julia Kirch Kirkegaard
AMBIGUOUS WINDS OF CHANGE – OR FIGHTING AGAINST WINDMILLS IN CHINESE WIND POWER
A CONSTRUCTIVIST INQUIRY INTO CHINA'S PRAGMATICS OF GREEN MARKETISATION MAPPING
CONTROVERSIES OVER A POTENTIAL TURN TO QUALITY IN CHINESE WIND POWER
6. Michelle Carol Antero
A Multi-case Analysis of the Development of Enterprise Resource Planning Systems (ERP) Business Practices

Morten Friis-Olivarius
The Associative Nature of Creativity
7. Mathew Abraham
New Cooperativism: A study of emerging producer organisations in India
8. Stine Hedegaard
Sustainability-Focused Identity: Identity work performed to manage, negotiate and resolve barriers and tensions that arise in the process of constructing or ganizational identity in a sustainability context
9. Cecilie Glerup
Organizing Science in Society – the conduct and justification of resposable research
10. Allan Salling Pedersen
Implementering af ITIL® IT-governance - når best practice konflikter med kulturen Løsning af implementerings-problemer gennem anvendelse af kendte CSF i et aktionsforskningsforløb.
11. Nihat Misir
A Real Options Approach to Determining Power Prices
12. Mamdouh Medhat
MEASURING AND PRICING THE RISK OF CORPORATE FAILURES
13. Rina Hansen
Toward a Digital Strategy for Omnichannel Retailing
14. Eva Pallesen
In the rhythm of welfare creation
A relational processual investigation moving beyond the conceptual horizon of welfare management
15. Gouya Harirchi
In Search of Opportunities: Three Essays on Global Linkages for Innovation
16. Lotte Holck
Embedded Diversity: A critical ethnographic study of the structural tensions of organizing diversity
17. Jose Daniel Balarezo
Learning through Scenario Planning
18. Louise Pram Nielsen
Knowledge dissemination based on terminological ontologies. Using eye tracking to further user interface design.
19. Sofie Dam
PUBLIC-PRIVATE PARTNERSHIPS FOR INNOVATION AND SUSTAINABILITY TRANSFORMATION
An embedded, comparative case study of municipal waste management in England and Denmark
20. Ulrik Hartmyer Christiansen
Follwoing the Content of Reported Risk Across the Organization
21. Guro Refsum Sanden
Language strategies in multinational corporations. A cross-sector study of financial service companies and manufacturing companies.
22. Linn Gevoll
Designing performance management for operational level
- A closer look on the role of design choices in framing coordination and motivation

23. Frederik Larsen
*Objects and Social Actions
– on Second-hand Valuation Practices*
24. Thorhildur Hansdottir Jetzek
*The Sustainable Value of Open
Government Data
Uncovering the Generative Mechanisms
of Open Data through a Mixed
Methods Approach*
25. Gustav Toppenberg
*Innovation-based M&A
– Technological-Integration
Challenges – The Case of
Digital-Technology Companies*
26. Mie Plotnikof
*Challenges of Collaborative
Governance
An Organizational Discourse Study
of Public Managers' Struggles
with Collaboration across the
Daycare Area*
27. Christian Garmann Johnsen
*Who Are the Post-Bureaucrats?
A Philosophical Examination of the
Creative Manager, the Authentic Leader
and the Entrepreneur*
28. Jacob Brogaard-Kay
*Constituting Performance Management
A field study of a pharmaceutical
company*
29. Rasmus Ploug Jenle
*Engineering Markets for Control:
Integrating Wind Power into the Danish
Electricity System*
30. Morten Lindholst
*Complex Business Negotiation:
Understanding Preparation and
Planning*
31. Morten Grynings
*TRUST AND TRANSPARENCY FROM AN
ALIGNMENT PERSPECTIVE*
32. Peter Andreas Norn
*Byregimer og styringsevne: Politisk
lederskab af store byudviklingsprojekter*
33. Milan Miric
*Essays on Competition, Innovation and
Firm Strategy in Digital Markets*
34. Sanne K. Hjordrup
*The Value of Talent Management
Rethinking practice, problems and
possibilities*
35. Johanna Sax
*Strategic Risk Management
– Analyzing Antecedents and
Contingencies for Value Creation*
36. Pernille Rydén
Strategic Cognition of Social Media
37. Mimmi Sjöklint
*The Measurable Me
- The Influence of Self-tracking on the
User Experience*
38. Juan Ignacio Staricco
*Towards a Fair Global Economic
Regime? A critical assessment of Fair
Trade through the examination of the
Argentinean wine industry*
39. Marie Henriette Madsen
*Emerging and temporary connections
in Quality work*
40. Yangfeng CAO
*Toward a Process Framework of
Business Model Innovation in the
Global Context
Entrepreneurship-Enabled Dynamic
Capability of Medium-Sized
Multinational Enterprises*
41. Carsten Scheibye
*Enactment of the Organizational Cost
Structure in Value Chain Configuration
A Contribution to Strategic Cost
Management*

2016

1. Signe Sofie Dyrby
Enterprise Social Media at Work
2. Dorte Boesby Dahl
The making of the public parking attendant
Dirt, aesthetics and inclusion in public service work
3. Verena Girschik
Realizing Corporate Responsibility
Positioning and Framing in Nascent Institutional Change
4. Anders Ørding Olsen
IN SEARCH OF SOLUTIONS
Inertia, Knowledge Sources and Diversity in Collaborative Problem-solving
5. Pernille Steen Pedersen
Udkast til et nyt copingbegreb
En kvalifikation af ledelsesmuligheder for at forebygge sygefravær ved psykiske problemer.
6. Kerli Kant Hvass
Weaving a Path from Waste to Value: Exploring fashion industry business models and the circular economy
7. Kasper Lindskow
Exploring Digital News Publishing Business Models – a production network approach
8. Mikkel Mouritz Marfelt
The chameleon workforce: Assembling and negotiating the content of a workforce
9. Marianne Bertelsen
Aesthetic encounters
Rethinking autonomy, space & time in today's world of art
10. Louise Hauberg Wilhelmsen
EU PERSPECTIVES ON INTERNATIONAL COMMERCIAL ARBITRATION
11. Abid Hussain
On the Design, Development and Use of the Social Data Analytics Tool (SODATO): Design Propositions, Patterns, and Principles for Big Social Data Analytics
12. Mark Bruun
Essays on Earnings Predictability
13. Tor Bøe-Lillegraven
BUSINESS PARADOXES, BLACK BOXES, AND BIG DATA: BEYOND ORGANIZATIONAL AMBIDEXTERITY
14. Hadis Khonsary-Atighi
ECONOMIC DETERMINANTS OF DOMESTIC INVESTMENT IN AN OIL-BASED ECONOMY: THE CASE OF IRAN (1965-2010)
15. Maj Lervad Grasten
Rule of Law or Rule by Lawyers?
On the Politics of Translation in Global Governance
16. Lene Granzau Juel-Jacobsen
SUPERMARKEDETS MODUS OPERANDI – en hverdagssociologisk undersøgelse af forholdet mellem rum og handlen og understøtte relationsopbygning?
17. Christine Thalsgård Henriques
In search of entrepreneurial learning – Towards a relational perspective on incubating practices?
18. Patrick Bennett
Essays in Education, Crime, and Job Displacement
19. Søren Korsgaard
Payments and Central Bank Policy
20. Marie Kruse Skibsted
Empirical Essays in Economics of Education and Labor
21. Elizabeth Benedict Christensen
The Constantly Contingent Sense of Belonging of the 1.5 Generation
Undocumented Youth
An Everyday Perspective

22. Lasse J. Jessen
Essays on Discounting Behavior and Gambling Behavior
23. Kalle Johannes Rose
*Når stiftertiljen dør...
Et retsøkonomisk bidrag til 200 års
juridisk konflikt om ejendomsretten*
24. Andreas Søeborg Kirkedal
*Danish Stød and Automatic Speech
Recognition*
25. Ida Lunde Jørgensen
*Institutions and Legitimations in
Finance for the Arts*
26. Olga Rykov Ibsen
*An empirical cross-linguistic study of
directives: A semiotic approach to the
sentence forms chosen by British,
Danish and Russian speakers in native
and ELF contexts*
27. Desi Volker
Understanding Interest Rate Volatility
28. Angeli Elizabeth Weller
*Practice at the Boundaries of Business
Ethics & Corporate Social Responsibility*
29. Ida Danneskiold-Samsøe
*Levende læring i kunstneriske
organisationer
En undersøgelse af læringsprocesser
mellem projekt og organisation på
Aarhus Teater*
30. Leif Christensen
*Quality of information – The role of
internal controls and materiality*
31. Olga Zarzecka
Tie Content in Professional Networks
32. Henrik Mahncke
*De store gaver
- Filantropiens gensidighedsrelationer i
teori og praksis*
33. Carsten Lund Pedersen
*Using the Collective Wisdom of
Frontline Employees in Strategic Issue
Management*
34. Yun Liu
Essays on Market Design
35. Denitsa Hazarbassanova Blagoeva
The Internationalisation of Service Firms
36. Manya Jaura Lind
*Capability development in an off-
shoring context: How, why and by
whom*
37. Luis R. Boscán F.
*Essays on the Design of Contracts and
Markets for Power System Flexibility*
38. Andreas Philipp Distel
*Capabilities for Strategic Adaptation:
Micro-Foundations, Organizational
Conditions, and Performance
Implications*
39. Lavinia Bleoca
*The Usefulness of Innovation and
Intellectual Capital in Business
Performance: The Financial Effects of
Knowledge Management vs. Disclosure*
40. Henrik Jensen
*Economic Organization and Imperfect
Managerial Knowledge: A Study of the
Role of Managerial Meta-Knowledge
in the Management of Distributed
Knowledge*
41. Stine Mosekjær
*The Understanding of English Emotion
Words by Chinese and Japanese
Speakers of English as a Lingua Franca
An Empirical Study*
42. Hallur Tor Sigurdarson
*The Ministry of Desire - Anxiety and
entrepreneurship in a bureaucracy*
43. Kätlin Pulk
*Making Time While Being in Time
A study of the temporality of
organizational processes*
44. Valeria Giacomini
*Contextualizing the cluster Palm oil in
Southeast Asia in global perspective
(1880s–1970s)*

- | | | |
|--|--------------------|--|
| <p>45. Jeanette Willert
<i>Managers' use of multiple Management Control Systems: The role and interplay of management control systems and company performance</i></p> <p>46. Mads Vestergaard Jensen
<i>Financial Frictions: Implications for Early Option Exercise and Realized Volatility</i></p> <p>47. Mikael Reimer Jensen
<i>Interbank Markets and Frictions</i></p> <p>48. Benjamin Faigen
<i>Essays on Employee Ownership</i></p> <p>49. Adela Michea
<i>Enacting Business Models An Ethnographic Study of an Emerging Business Model Innovation within the Frame of a Manufacturing Company.</i></p> <p>50. Iben Sandal Stjerne
<i>Transcending organization in temporary systems Aesthetics' organizing work and employment in Creative Industries</i></p> <p>51. Simon Krogh
<i>Anticipating Organizational Change</i></p> <p>52. Sarah Netter
<i>Exploring the Sharing Economy</i></p> <p>53. Lene Tolstrup Christensen
<i>State-owned enterprises as institutional market actors in the marketization of public service provision: A comparative case study of Danish and Swedish passenger rail 1990–2015</i></p> <p>54. Kyoung(Kay) Sun Park
<i>Three Essays on Financial Economics</i></p> | <p>2017</p> | <p>1. Mari Bjerck
<i>Apparel at work. Work uniforms and women in male-dominated manual occupations.</i></p> <p>2. Christoph H. Flöthmann
<i>Who Manages Our Supply Chains? Backgrounds, Competencies and Contributions of Human Resources in Supply Chain Management</i></p> <p>3. Aleksandra Anna Rzeźnik
<i>Essays in Empirical Asset Pricing</i></p> <p>4. Claes Bäckman
<i>Essays on Housing Markets</i></p> <p>5. Kirsti Reitan Andersen
<i>Stabilizing Sustainability in the Textile and Fashion Industry</i></p> <p>6. Kira Hoffmann
<i>Cost Behavior: An Empirical Analysis of Determinants and Consequences of Asymmetries</i></p> <p>7. Tobin Hanspal
<i>Essays in Household Finance</i></p> <p>8. Nina Lange
<i>Correlation in Energy Markets</i></p> <p>9. Anjum Fayyaz
<i>Donor Interventions and SME Networking in Industrial Clusters in Punjab Province, Pakistan</i></p> <p>10. Magnus Paulsen Hansen
<i>Trying the unemployed. Justification and critique, emancipation and coercion towards the 'active society'. A study of contemporary reforms in France and Denmark</i></p> <p>11. Sameer Azizi
<i>Corporate Social Responsibility in Afghanistan – a critical case study of the mobile telecommunications industry</i></p> |
|--|--------------------|--|

12. Malene Myhre
The internationalization of small and medium-sized enterprises: A qualitative study
13. Thomas Presskorn-Thygesen
The Significance of Normativity – Studies in Post-Kantian Philosophy and Social Theory
14. Federico Clementi
Essays on multinational production and international trade
15. Lara Anne Hale
Experimental Standards in Sustainability Transitions: Insights from the Building Sector
16. Richard Pucci
*Accounting for Financial Instruments in an Uncertain World
Controversies in IFRS in the Aftermath of the 2008 Financial Crisis*
17. Sarah Maria Denta
*Kommunale offentlige private partnerskaber
Regulering i skyggen af Farumsagen*
18. Christian Östlund
Design for e-training
19. Amalie Martinus Hauge
Organizing Valuations – a pragmatic inquiry
20. Tim Holst Celik
Tension-filled Governance? Exploring the Emergence, Consolidation and Reconfiguration of Legitimatory and Fiscal State-crafting
21. Christian Bason
Leading Public Design: How managers engage with design to transform public governance
22. Davide Tomio
Essays on Arbitrage and Market Liquidity
23. Simone Stæhr
*Financial Analysts' Forecasts
Behavioral Aspects and the Impact of Personal Characteristics*
24. Mikkel Godt Gregersen
*Management Control, Intrinsic Motivation and Creativity
– How Can They Coexist*
25. Kristjan Johannes Suse Jespersen
Advancing the Payments for Ecosystem Service Discourse Through Institutional Theory
26. Kristian Bondo Hansen
Crowds and Speculation: A study of crowd phenomena in the U.S. financial markets 1890 to 1940
27. Lars Balslev
Actors and practices – An institutional study on management accounting change in Air Greenland
28. Sven Klingler
Essays on Asset Pricing with Financial Frictions
29. Klement Ahrensbach Rasmussen
*Business Model Innovation
The Role of Organizational Design*
30. Giulio Zichella
*Entrepreneurial Cognition.
Three essays on entrepreneurial behavior and cognition under risk and uncertainty*
31. Richard Ledborg Hansen
En forkærlighed til det eksisterende – mellemlederens oplevelse af forandringsmodstand i organisatoriske forandringer
32. Vilhelm Stefan Holsting
Militært chefvirke: Kritik og retfærdiggørelse mellem politik og profession

- | | | | |
|-----|---|-------------|---|
| 33. | Thomas Jensen
<i>Shipping Information Pipeline: An information infrastructure to improve international containerized shipping</i> | 2018 | |
| 34. | Dzmitry Bartalevich
<i>Do economic theories inform policy? Analysis of the influence of the Chicago School on European Union competition policy</i> | | 1. Vishv Priya Kohli
<i>Combatting Falsification and Counterfeiting of Medicinal Products in the European Union – A Legal Analysis</i> |
| 35. | Kristian Roed Nielsen
<i>Crowdfunding for Sustainability: A study on the potential of reward-based crowdfunding in supporting sustainable entrepreneurship</i> | | 2. Helle Haurum
<i>Customer Engagement Behavior in the context of Continuous Service Relationships</i> |
| 36. | Emil Husted
<i>There is always an alternative: A study of control and commitment in political organization</i> | | 3. Nis Grünberg
<i>The Party-state order: Essays on China's political organization and political economic institutions</i> |
| 37. | Anders Ludvig Sevelsted
<i>Interpreting Bonds and Boundaries of Obligation. A genealogy of the emergence and development of Protestant voluntary social work in Denmark as shown through the cases of the Copenhagen Home Mission and the Blue Cross (1850 – 1950)</i> | | 4. Jesper Christensen
<i>A Behavioral Theory of Human Capital Integration</i> |
| 38. | Niklas Kohl
<i>Essays on Stock Issuance</i> | | 5. Poula Marie Helth
<i>Learning in practice</i> |
| 39. | Maya Christiane Flensburg Jensen
<i>BOUNDARIES OF PROFESSIONALIZATION AT WORK An ethnography-inspired study of care workers' dilemmas at the margin</i> | | 6. Rasmus Vendler Toft-Kehler
<i>Entrepreneurship as a career? An investigation of the relationship between entrepreneurial experience and entrepreneurial outcome</i> |
| 40. | Andreas Kamstrup
<i>Crowdsourcing and the Architectural Competition as Organisational Technologies</i> | | 7. Szymon Furtak
<i>Sensing the Future: Designing sensor-based predictive information systems for forecasting spare part demand for diesel engines</i> |
| 41. | Louise Lyngfeldt Gorm Hansen
<i>Triggering Earthquakes in Science, Politics and Chinese Hydropower - A Controversy Study</i> | | 8. Mette Brehm Johansen
<i>Organizing patient involvement. An ethnographic study</i> |
| | | | 9. Iwona Sulinska
<i>Complexities of Social Capital in Boards of Directors</i> |
| | | | 10. Cecilie Fanøe Petersen
<i>Award of public contracts as a means to conferring State aid: A legal analysis of the interface between public procurement law and State aid law</i> |
| | | | 11. Ahmad Ahmad Barirani
<i>Three Experimental Studies on Entrepreneurship</i> |

12. Carsten Allerslev Olsen
Financial Reporting Enforcement: Impact and Consequences
13. Irene Christensen
New product fumbles – Organizing for the Ramp-up process
14. Jacob Taarup-Esbensen
Managing communities – Mining MNEs' community risk management practices
15. Lester Allan Lasrado
Set-Theoretic approach to maturity models
16. Mia B. Münster
Intention vs. Perception of Designed Atmospheres in Fashion Stores
17. Anne Sluhan
Non-Financial Dimensions of Family Firm Ownership: How Socioemotional Wealth and Familiness Influence Internationalization
18. Henrik Yde Andersen
Essays on Debt and Pensions
19. Fabian Heinrich Müller
Valuation Reversed – When Valuators are Valuated. An Analysis of the Perception of and Reaction to Reviewers in Fine-Dining
20. Martin Jarmatz
Organizing for Pricing
21. Niels Joachim Christfort Gormsen
Essays on Empirical Asset Pricing
22. Diego Zunino
Socio-Cognitive Perspectives in Business Venturing
23. Benjamin Asmussen
Networks and Faces between Copenhagen and Canton, 1730-1840
24. Dalia Bagdziunaite
Brains at Brand Touchpoints A Consumer Neuroscience Study of Information Processing of Brand Advertisements and the Store Environment in Compulsive Buying
25. Erol Kazan
Towards a Disruptive Digital Platform Model
26. Andreas Bang Nielsen
Essays on Foreign Exchange and Credit Risk
27. Anne Krebs
Accountable, Operable Knowledge Toward Value Representations of Individual Knowledge in Accounting
28. Matilde Fogh Kirkegaard
A firm- and demand-side perspective on behavioral strategy for value creation: Insights from the hearing aid industry
29. Agnieszka Nowinska
SHIPS AND RELATION-SHIPS Tie formation in the sector of shipping intermediaries in shipping
30. Stine Evald Bentsen
The Comprehension of English Texts by Native Speakers of English and Japanese, Chinese and Russian Speakers of English as a Lingua Franca. An Empirical Study.
31. Stine Louise Daetz
Essays on Financial Frictions in Lending Markets
32. Christian Skov Jensen
Essays on Asset Pricing
33. Anders Kryger
Aligning future employee action and corporate strategy in a resource-scarce environment

34. Maitane Elorriaga-Rubio
The behavioral foundations of strategic decision-making: A contextual perspective
 35. Roddy Walker
Leadership Development as Organisational Rehabilitation: Shaping Middle-Managers as Double Agents
 36. Jinsun Bae
Producing Garments for Global Markets Corporate social responsibility (CSR) in Myanmar's export garment industry 2011–2015
 37. Queralt Prat-i-Pubill
Axiological knowledge in a knowledge driven world. Considerations for organizations.
 38. Pia Mølgaard
Essays on Corporate Loans and Credit Risk
 39. Marzia Aricò
Service Design as a Transformative Force: Introduction and Adoption in an Organizational Context
 40. Christian Dyrland Wåhlin-Jacobsen
Constructing change initiatives in workplace voice activities Studies from a social interaction perspective
 41. Peter Kalum Schou
Institutional Logics in Entrepreneurial Ventures: How Competing Logics arise and shape organizational processes and outcomes during scale-up
 42. Per Henriksen
Enterprise Risk Management Rationaler og paradokser i en moderne ledelsesteknologi
 43. Maximilian Schellmann
The Politics of Organizing Refugee Camps
 44. Jacob Halvas Bjerre
Excluding the Jews: The Aryanization of Danish-German Trade and German Anti-Jewish Policy in Denmark 1937-1943
 45. Ida Schrøder
Hybridising accounting and caring: A symmetrical study of how costs and needs are connected in Danish child protection work
 46. Katrine Kunst
Electronic Word of Behavior: Transforming digital traces of consumer behaviors into communicative content in product design
 47. Viktor Avlonitis
Essays on the role of modularity in management: Towards a unified perspective of modular and integral design
 48. Anne Sofie Fischer
Negotiating Spaces of Everyday Politics: -An ethnographic study of organizing for social transformation for women in urban poverty, Delhi, India
- 2019**
1. Shihan Du
ESSAYS IN EMPIRICAL STUDIES BASED ON ADMINISTRATIVE LABOUR MARKET DATA

TITLER I ATV PH.D.-SERIEN

1992

1. Niels Kornum
Servicesamkørsel – organisation, økonomi og planlægningsmetode

1995

2. Verner Worm
*Nordiske virksomheder i Kina
Kulturspecifikke interaktionsrelationer
ved nordiske virksomhedsetableringer i Kina*

1999

3. Mogens Bjerre
*Key Account Management of Complex Strategic Relationships
An Empirical Study of the Fast Moving Consumer Goods Industry*

2000

4. Lotte Darsø
*Innovation in the Making
Interaction Research with heterogeneous Groups of Knowledge Workers
creating new Knowledge and new Leads*

2001

5. Peter Hobolt Jensen
*Managing Strategic Design Identities
The case of the Lego Developer Network*

2002

6. Peter Lohmann
The Deleuzian Other of Organizational Change – Moving Perspectives of the Human
7. Anne Marie Jess Hansen
To lead from a distance: The dynamic interplay between strategy and strategizing – A case study of the strategic management process

2003

8. Lotte Henriksen
*Videndeling
– om organisatoriske og ledelsesmæssige udfordringer ved videndeling i praksis*
9. Niels Christian Nickelsen
Arrangements of Knowing: Coordinating Procedures Tools and Bodies in Industrial Production – a case study of the collective making of new products

2005

10. Carsten Ørts Hansen
Konstruktion af ledelsesteknologier og effektivitet

TITLER I DBA PH.D.-SERIEN

2007

1. Peter Kastrup-Misir
Endeavoring to Understand Market Orientation – and the concomitant co-mutation of the researched, the researcher, the research itself and the truth

2009

1. Torkild Leo Thellefsen
*Fundamental Signs and Significance effects
A Semeiotic outline of Fundamental Signs, Significance-effects, Knowledge Profiling and their use in Knowledge Organization and Branding*
2. Daniel Ronzani
When Bits Learn to Walk Don't Make Them Trip. Technological Innovation and the Role of Regulation by Law in Information Systems Research: the Case of Radio Frequency Identification (RFID)

2010

1. Alexander Carnera
*Magten over livet og livet som magt
Studier i den biopolitiske ambivalens*