Utrecht University

School of Economics

Local labour markets, job displacement and agglomeration economies

Jordy Meekes

U.S.E. 046

LOCAL LABOUR MARKETS, JOB DISPLACEMENT AND AGGLOMERATION ECONOMIES

Jordy Meekes

Manuscript committee: Prof. dr. M. Goos
Prof. dr. H.L.F. de Groot
Prof. dr. J.N. van Ommeren
Prof. dr. F.G. van Oort
Prof. dr. C.N. Teulings

ISBN 978-94-91870-31-6
Printed by Ridderprint, Ridderkerk

LOCAL LABOUR MARKETS, JOB DISPLACEMENT AND AGGLOMERATION ECONOMIES

Lokale arbeidsmarkten, baanverlies en agglomeratievoordelen
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus,
prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
maandag 18 februari 2019 des middags te 12.45 uur

door

Jordy Meekes
geboren op 4 maart 1991 te Lelystad

Promotor: Prof. dr. W.H.J. Hassink

Acknowledgements

The writing of my dissertation has been a terrific journey and I am indebted to many people for their valuable contributions.

I am very grateful to my advisor, Wolter Hassink, for his continuous support and guidance. I truly enjoyed our talks about both research and life. You have motivated me to try and reach my full potential as a researcher, which has led to the pursuit of a job at the University of Melbourne. I would also like to thank all members of my manuscript committee for devoting some of their precious time to provide commentary on my dissertation.

A special mention to Bram van Besouw, Zori Kambourova, Vincent Schippers and Joep Steegmans. It was amazing to undertake the Research Master and PhD together. I am also grateful to my colleagues in room 2.08 (Emilie, Jingyang, Margot, Milande, Milena, Oleg, Peter, Rebean, Sarah), everyone from USE Your Peers (yes, all of you!) and U.S.E. (notably Adriaan, Bastian, Carla, Elena, Florian, Jochem and Thomas). Thanks to Maarten Goos and Anna Salomons for your advice and the opportunity to go out for dinners with you and seminar speakers, as this inspired me to go on the economics job market. Thanks also to the U.S.E. administrative staff for their time and energy in helping me make all sorts of arrangements.

I would like to thank Roberto Fernandez for hosting me at the MIT Sloan School of Management (Cambridge, the United States) and Erik Volkers for hosting me at Randstad Holding (Diemen, the Netherlands). It was fantastic and very productive to have the opportunity to work in such dynamic and multidisciplinary environments.

I also truly appreciate all social gatherings with Tim Kraay, Bram, Vincent, IBB133 and, in particular, the Guaps: Aad, Bob, Jelle, Jeroen, Joren, Justin, Maarten, Marcel, Rik, Sebastiaan, Thomas, Tim and Vincent. Without all of you, doing a PhD and life in general would not have been so much fun.

I would especially like to bestow some gratitude on my mom, dad and sister for all of their support and love throughout my life. Naturally, the same goes for all members of the families Rosmalen, van Spanje and Steeman.

A final and very special shout out goes to Ellis van Spanje. You have supported me immensely through everything over the past few years. I will never forget your words down in the lobby of 111 Barry Street, Carlton VIC 3053, Australia, when you wished me good luck right before my job market meetings and seminar at the Melbourne Institute. I dearly appreciate the sacrifices you have made to join me on our journey to the other side of the world – Australia.

Jordy Meekes
Melbourne, November 2018

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Motivation

From an urban and housing economics perspective, the economic value of living in a regional area is determined based on the measurement of the value of urban amenities (Smith, 1978). Urban amenities affect the supply and demand equilibrium of housing, which in turn determines the value of residential properties. From the perspective of urban and labour economics, an alternative view of the economic value of living in a regional area is that the agglomeration spillovers associated with the geographical location could lead to better employment prospects and a more productive job match with higher wages (Glaeser and Maré, 2001). Over the last decades, agglomeration of workers and firms has been increasing in many parts of the world. For example, spatial concentration in the U.S. and other member countries of the OECD has increased, which can be gleaned from the clustering of population, employment and skill level in urban areas (Brezzi et al., 2012).

Agglomeration economies are one of the main reasons for productivity differences on a national scale, which has led to strong variances in labour market statistics across regional areas in terms of employment density, employment opportunities, productivity and wages (Moretti, 2011). The role of agglomeration in the differences across regional areas can be explained by urbanisation economies, which refer to positive spillovers from overall economic activity, and by localisation economies, which refer to positive spillovers from intra-industry spatial concentration of economic activities. The positive spillovers may benefit workers and firms in various ways, including improved matching of labour and firms, increased sharing of suppliers, customers and risk, and improved learning due to the better generation, diffusion and accumulation of knowledge (Duranton and Puga, 2004). Thereby, the worker's geographical location could be important for the urban wage premium, which is explained by the returns to agglomeration in wages. Alternatively, the worker's geographical location could be important for the resilience to a negative employment shock such as job displacement. Moreover, the effects of agglomeration spillovers and of a negative employment shock on workers' labour market outcomes can be subgroup-specific, because of differences in the supply and demand of the labour market. Researchers and politicians identify and explain within-country regional differences by using the concept of local labour markets.

A local labour market can be described as a regional area that is relatively self-contained in terms of residence and work activity. Local labour markets are conceptually different from cities or provinces, as local labour markets are defined based on regional economic processes and activities instead of historically defined administrative borders and requirements. To study differences across local labour markets, a researcher should define the geographic space. That is, a researcher should define local labour markets or rely on a pre-defined regional classification to operate local labour markets. Studies in the field of economic geography use various pre-defined regional classifications to operate self-contained local labour markets (De Groot et al., 2016; Fang and Yu, 2017). The structure of workers' local labour market, however, could be changing over time and could be location- or subgroup-specific. Recently, there has been a renewal in interest from economists in the effective size of a local labour market (e.g., see Manning and Petrongolo (2017)). From a policy perspective, a better understanding of the structure of workers' local labour market is important as it may improve the efficiency of regional policies (OECD, 2007, 2014) and place-based policies (Glaeser and Gottlieb, 2008; Neumark and Simpson, 2015).

1.2 Objectives and Contributions

The main research question of this dissertation is: “What is the structure of workers’ local labour markets, and what are its economic consequences?”. The three core chapters of this dissertation integrate concepts of labour economics, urban economics and regional economics, which include the topics of local labour markets, job displacement and agglomeration economies. The research objectives and contributions are as follows.

First of all, I will analyse workers’ spatial and labour response to job displacement. A type of a negative employment shock is job displacement, which in this dissertation refers to an involuntary and unforeseen event involving job loss due to bankruptcy of the firm. The research question is: “To what extent is the spatial structure of homes and jobs relevant for workers’ response to job displacement?”. The explicit focus on workers’ spatial response to job displacement reveals whether workers use the commuting margin and/or the changing home margin in their labour adjustment following job displacement. In addition, I examine the sources of differences in workers’ response to job displacement, by focusing on the importance of workers’ housing state, demographic characteristics and job characteristics. I define a worker’s housing state as being a tenant or a homeowner, and I distinguish between types of homeowners based on their loan-to-value ratio. The worker’s housing state could be important, as it affects the financial incentive structures to become employed. The main contribution is to show that the spatial margins of adjustment are relevant for workers’ response to job displacement.

Secondly, I will develop a flow-based cluster algorithm that is able to define meaningful local labour markets, using commuting flows across regional areas as main input. The flow-based cluster algorithm is capable of defining local labour markets at both various levels of regional aggregation and for different subgroups of workers. Thus, the algorithm allows me to operate a local labour market so that it depends on the worker’s geographical location, demographic characteristics and level of regional aggregation. As such, I am able to assess the differences in the structure of workers’ local labour markets. The main contribution is methodological, by introducing a cluster algorithm in Stata that is capable of defining meaningful clusters based on the relational data of flows.

Finally, I will analyse the consequences of the differences between workers' local labour markets. The research question is: "What is the role of regional aggregation in the agglomeration benefits for wages and employment?". The level of regional aggregation refers to the spatial scale of local labour markets. The impact of matching, sharing and learning micro-foundations of agglomeration economies could be sensitive to the level of regional aggregation, as each of the mechanisms can be more prevalent at different spatial scales. I will examine the returns to agglomeration by focusing on two economic applications. The first application concerns the estimation of the urban wage premium, which is explained by the agglomeration benefits for wages. The second application concerns the analysis of regional differences in the impact of job displacement, which combines the literature of agglomeration economies with the literature of job displacement. Moreover, I use subgroup-specific local labour markets to examine whether subgroups of workers are characterised by differences in the agglomeration benefits. The main contribution is to identify the importance of the structure of local labour markets and the level of regional aggregation for the estimation of agglomeration benefits.

1.3 Methodology

The main share of the detailed empirical analyses are based on Dutch administrative data sets linking employees to employers and covering the period of 2006 to 2014 (CBS, 2018). The administrative data sets contain data on the entire Dutch population in terms of individuals, households and firms. The advantage of these rich data sets is that it allowed me to study on the individual level: (i) workers' spatial response to job displacement and the importance of workers' housing state, (ii) the structure of workers' local labour markets for different subgroups of workers based on commuting flows, and (iii) the economic consequences of agglomeration benefits for individual workers. The time period of 2006 to 2014 is particularly interesting, as it includes the Great Recession that started in 2008 and has been largely understudied. Below, I will discuss my empirical approach.

In Chapter 2, I use a quasi-experimental empirical design to examine the response to job displacement. I assess the displacement effects on four different outcome variables, which include employment, wages, commuting distance and changing home. The key identifying restriction in each of the models is that there are parallel trends in the outcome variables prior to displacement. I also examine whether the displacement effects differ among workers who vary in observed worker characteristics (Chapter 2) and local labour market density (Chapter 4). The main limitation of my empirical approach, which is based on reduced-form models, is that the four outcome variables are likely to be jointly determined. My results reveal to some extent how workers jointly determine their post-displacement labour market outcomes.

The flow-based cluster algorithm that is introduced in Chapter 3 allows me to define mutually exclusive sets of aggregate and subgroup-specific local labour markets at various levels of regional aggregation. A higher level of regional aggregation represents a lower number of distinct local labour markets. The algorithm is able to define meaningful local labour markets based on relational data of flows, such as commuting flows. The main limitation of Chapter 3 is that the algorithm does not allow me to define continuous local labour markets (e.g., in the spirit of Manning and Petrongolo (2017)). Notably, the algorithm is able to define subgroup-specific local labour markets, which has been largely overlooked in the literature.

In Chapter 4, I estimate the returns to agglomeration in wages, i.e. the urban wage premium. The empirical model corresponds to regressing wages on several covariates, including demographic characteristics, job characteristics and regional characteristics. The returns to agglomeration are represented by the employment density and area size at the local labour market level. Hence, I capture mostly agglomeration benefits that are related to urbanisation economies and do not distinguish between urbanisation economies and localisation economies. Notably, I analyse the role of regional aggregation in the returns to agglomeration by running multiple regressions where only the employment density and area size differ. I also use subgroup-specific local labour markets to examine gender differentials and education differentials in the returns to agglomeration.

1.4 Conceptual Considerations

In this subsection, I will discuss the conceptual considerations concerning the structures and its economic consequences of workers' local labour markets. My conceptual analysis starts with a discussion of a worker's response to a negative employment shock. If a worker has become unemployed following a negative employment shock, he or she is assumed to be in search of a job. Following the matching-function literature (e.g., see Petrongolo and Pissarides (2001) for an overview), the labour market outcomes of the unemployed depend on the intersection of labour supply and labour demand.

From the supply side of the labour market, the exit rate into employment is determined by the worker's search behaviour. A worker could affect his or her exit rate into employment by adjusting the search behaviour to increase the arrival rate of job opportunities. Workers may increase the arrival rate of job opportunities by decreasing their reservation wage. However, workers are limited in the set of potential vacant jobs and work locations because of their home location. The link between a worker's place of work and place of home is the commuting distance.[1] In a micro-economic model that consists of a utility function, a worker will not have an infinite willingness to commute, as – given the cost in leisure time and money – commuting negatively affects utility. In other words, a worker is limited by his or her labour market, as the worker's labour market is local due to travel time constraints and financial constraints of commuting. Besides adjusting labour margins such as the reservation wage, it is important to note that workers could increase the number of job opportunities by adjusting spatial margins. For example, if the workers' willingness to commute would grow, it could be considered as an increase in the size of their local labour market (Brueckner et al., 2002). Alternatively, workers could change home with the goal of improving labour market outcomes, which would happen if they change home to a distant local labour market (Topel, 1986). Consequently, the spatial response to a negative employment shock might lead to differences in the structure of workers' local labour market. In Chapter 2, I will analyse workers' spatial response and labour response to a negative employment shock involving job displacement.

The conceptual analysis becomes more complicated if I recognise that workers are heterogeneous, which affects both the supply and the demand side of the labour market. Workers differ in their search behaviour for various reasons, including differences in their financial incentives to be employed and in their willingness to commute. Workers' financial incentive to be employed can also be affected by their housing state.

[1] In this dissertation, I focus on workers' commuting distance instead of commuting time, as the commuting distance is computable with the Dutch administrative data. In general, workers' commuting time and commuting distance are characterised by a relatively high and positive correlation.

Workers' housing state is important, as it affects the payment obligations. For example, outright homeowners, mortgage homeowners and tenants face differences in housing costs. With respect to workers' willingness to commute there can be differences in compensating differentials, as workers experience differences in the opportunity costs of commuting through financial and time constraints. Women, as compared to men, may put a higher value on working close to home, for various reasons including women's dual role as a mother and worker (Madden, 1981). Moreover, the opportunity costs of commuting can differ among workers who vary in education level. Notably, the wage offer curve is increasing in the level of education, because more educated workers are characterised by a thinner local labour market (Manning, 2003). A relatively thin local labour market leads to a balance of power in favour of the worker, which pushes wages upward. In this regard, it is likely that the spatial response to a negative employment shock differs among subgroups of workers. Chapter 2 studies the role of worker characteristics in the response to job displacement, focusing on workers' housing state, demographic characteristics and job characteristics. Importantly, subgroups of workers can experience differences in the structure of local labour markets, as they have different opportunity costs of commuting. I introduce a flow-based cluster algorithm in Chapter 3, which allows me to study the differences in the structure of workers' local labour market.

From the demand side of the labour market, workers' local labour market could have important consequences for their labour market outcomes. For example, the local labour market could make workers more resilient to a negative employment shock if it weakens the negative effects on employment or wages. The role of local labour markets in labour market outcomes can be approached by using both a labour economics perspective and an urban economics perspective. From that first perspective, an increase in the number of local vacant jobs or a decrease in the number of local unemployed workers could increase the tightness of the labour market. A higher labour-market tightness is associated with a higher rate at which workers meet firms and a higher exit rate into employment. To examine regional differences in matching functions, Petrongolo and Pissarides (2006) study the impact of the size of local labour markets in markets with search. Research on regional matching functions aggregate the number of vacancies and unemployed workers within a region, based on a regional classification that is chosen to operate workers' local labour market. Scale effects are then tested by, e.g., regressing wage offers, reservation wages or unemployment duration on the market size. The market size is approximated by the number of vacancies by a skill group in the local labour market. Petrongolo and Pissarides (2006) conclude that scale effects in larger markets exist and lead to variation in wages through a higher productivity match, but not to variation in employment.

From an urban economics perspective, agglomeration economies lead to benefits in labour market outcomes across local labour markets in two ways (Combes and Gobillon, 2015). First, local labour markets differ in the amount of economic activity that takes place. The benefits for workers and firms from a larger amount of economic activity in a local labour market are referred to as urbanisation economies. The degree of urbanisation in a local region is operated by variables that represent the employment density and area size of a region. The effect of density on wages, holding the area size constant, represents the agglomeration benefits from increasing the number of workers or the density in a local labour market. Second, local labour markets differ in their demand for the amount and type of labour. The benefits from a context where many related firms are located within close proximity of each other are referred to as localisation economies. The degree of localisation in a local region is operated by a specialisation variable that represents the Herfindahl index (Henderson et al., 1995), which is constructed from the shares of industries within the local industry employment of a region. The agglomeration economies, including urbanisation economies and localisation economies, may lead to positive and negative spillovers. Examples of positive spillovers from agglomeration include improved job matching, increased sharing of risk and resources, and increased accumulation of knowledge (Duranton and Puga, 2004). Examples of negative spillovers from agglomeration include increased competition and congestion. The field of urban economics generally finds positive effects of agglomeration on workers' wages (e.g., see Melo et al. (2009) and De Groot et al. (2016) for comprehensive overviews). In Chapter 4, I use the urban economics perspective to study the structures of workers' local labour markets and its economic consequences for labour market outcomes.

1.5 Workers' Spatial Response to Job Displacement

In Chapter 2, I will examine the role of geography in how workers respond to job displacement. So far, the literature on job displacement uses human capital theory to predict workers' labour response to this. Human capital theory predicts severe losses in employment and wage outcomes when workers lose their human capital and wage premiums following job displacement (Jacobson et al., 1993). Firm-specific human capital is lost upon the incidence of job displacement. In addition, general human capital depreciates during the post-displacement unemployment period. The displaced worker's hazard rate into employment depends on the set of potential job opportunities and the reservation wage. A greater search effort or a lower reservation rate will lead to a higher arrival rate of job offers, which will realise a higher hazard rate into employment. It is important to note that the literature on job displacement has mostly neglected the role of geography in how workers respond to job displacement. Focusing on the supply side of the labour market, I examine whether the spatial structure of homes and jobs represents relevant margins of labour adjustment other than employment and wages.

Chapter 2 focuses on the extent to which displaced workers use margins of adjustment related to space, such as commuting patterns and home relocations. A displaced worker may expand the geographic scope of their search. A larger geographic scope of search may increase the set of potential job opportunities, which in turn may positively affect the post-displacement hazard rate into employment and post-displacement wages. A worker could change his or her geographic scope of search in two ways: by allowing for a further commuting distance that will expand the local labour market, and by changing home to a distant local labour market. In Chapter 2, I will also examine to what extent the response to job displacement differs by worker characteristics, including demographic characteristics, job characteristics and housing characteristics. My research contributes to the field of labour economics by emphasising the importance of geography for the response to job displacement. However, from a policy perspective my research is relevant also, as I analyse whether there are societal costs involved with displacement other than losses in employment and wages.

1.6 A Flow-Based Cluster Algorithm to Define Local Labour Markets

Chapter 3 provides the implementation of flowbca, which is a so-called agglomerative hierarchical clustering algorithm that is based on relational data of flows. The main motivation to write Chapter 3 was to complement and build on the research documented in Chapter 2 by examining regional differences in the response to job displacement. Flowbca allows me to define local labour markets endogenous to worker characteristics and commuting outcomes. The algorithm is able to define meaningful clusters that can be of use in various research fields, including economic geography, industrial input-output analysis and social network analysis. So far, most cluster algorithms focus on distance-based clustering, which uses as main input the distance between a pair of units as a measure of similarity (e.g., see Duranton and Overman (2005); Delgado et al. (2016)). I contribute to the scientific community by introducing a flow-based cluster algorithm in Stata based on the Mata programming language, which uses relational data of flows as main input. I argue that the implementation of a flow-based cluster algorithm is relevant, because of the increasing availability of relational data on various types of flows.

For decades, geographers paid a lot of attention to the delineation of local labour market areas (Masser and Brown, 1975; Van der Laan and Schalke, 2001). In most studies, local labour markets are defined based on commuting flows across regional units. Alternatively, a researcher could use other types of worker flows, e.g., flows between employment locations following job-to-job turnover. The delineation of local labour markets is an important research topic, as the spatial scale of economic processes depends on the type of processes and type of workers. For example, job matching between workers and firms takes place at a higher spatial scale for more educated workers, as the commuting distance generally increases when the worker's education level does. Moreover, the delineation of self-contained areas is important, as the use of an unrepresentative regional classification drives a measurement error that leads to statistical bias (Openshaw, 1983; Fotheringham and Wong, 1991; Briant et al., 2010).

A main limitation of current economic applications in the field of economic geography is that they are based on pre-defined regional classifications to operate regional areas. The pre-defined regional classifications are generally defined in line with administrative borders and according to administrative needs, instead of being based on economic relevance. A Dutch example of a pre-defined regional classification is the set of forty Nomenclature of Territorial Units for Statistics (NUTS) 3 areas, which were drawn up in 1971. The NUTS 3 areas were defined in line with provincial borders that were not crossed, according to commuting statistics of the typical employed worker. However, since 1971, the regional labour force composition in the Netherlands experienced dramatic changes (Poorthuis et al., 2015). Also in other countries, changes were caused by global technological, demographic and economic shifts that decreased the demand for low-educated or average-educated workers and increased wage dispersion (Fernandez, 2001; Goos et al., 2009, 2014), increased the women's labour force participation (Costa, 2000), and increased the average commute of workers (Van der Laan and Schalke, 2001). In this dissertation, I apply flowbca to the delineation of local labour markets. The algorithm allows me to operate a local labour market so that it depends on the worker's geographical location, demographic characteristics and level of regional aggregation.

1.7 Local Labour Markets, Regional Aggregation and Agglomeration Economies

Chapter 4 builds on the groundwork set out in Chapters 2 and 3 in the following ways. First, in Chapter 4, I use the flow-based algorithm that is introduced in Chapter 3 to define local labour markets at different levels of regional aggregation. The level of regional aggregation represents the spatial scale that is used to operate local labour markets. I define local labour markets at different levels of regional aggregation to study the role of regional aggregation in the estimation of agglomeration economies. This exercise is important, as the spatial scale at which agglomeration economies should be estimated is not clear (Rosenthal and Strange, 2001; Combes and Gobillon, 2015). The estimate of the urban wage premium may depend on the level of regional aggregation that is used to operate local labour markets, as the micro-foundations of agglomeration economies could be more prevalent at different spatial scales (Rosenthal and Strange, 2001, 2003, 2008). For example, while increased accumulation of knowledge from more urbanised areas could be higher at a relatively low spatial scale, improved job matching and sharing of resources and risk could be more prevalent at a relatively high spatial scale. The empirical analysis in Chapter 4 documents the consequences of differences in local labour markets and the level of regional aggregation for workers' returns to agglomeration in wages and employment.

Then, I use the algorithm to define subgroup-specific local labour markets. Subgroup-specific local labour markets are relevant, as subgroups of workers are characterised by differences in the structure of their local labour market (Farmer and Fotheringham, 2011). For example, compared to men, women are more likely to have a relatively small local labour market, as a result of differences in the opportunity costs of commuting due to time and financial constraints. This could be because, in general, women have a dual role as a worker and mother, as they are responsible for the household and children. Therefore, women generally prefer working closer to home than men (Madden, 1981; Crane, 2007). Another example concerns workers with different levels of education. High-educated workers are likely to commute further away, as their local labour market is relatively thin (Manning, 2003). Consequently, the wage offer curve for high-educated workers is relatively steep due to higher benefits in wages from longer commuting. In this regard, I argue that there is a need to redefine local labour markets and make them subgroup-specific, as for policy purposes a local labour market should represent a generalised area of job opportunities.

The limited data availability of commuting flows at a relatively low level of regional aggregation for different subgroups of workers could be an important reason why the consequences of subgroup-specific local labour markets for labour market outcomes have received little attention in the literature. I make use of the subgroup-specific local labour markets to analyse whether the agglomeration benefits vary in the subgroups of workers. So far, the literature focuses on differentials in the magnitude of returns to agglomeration. However, the literature neglects the fact that subgroups differ in local labour market structure. Moreover, the literature does not pay attention to the concept that the role of regional aggregation in agglomeration economies could be more prevalent for specific subgroups. The empirical analysis in Chapter 4 examines whether the structures and its economic consequences of local labour markets vary among subgroups of workers that differ in gender and education level.

Finally, Chapter 4 complements Chapter 2 by studying whether the response to job displacement differs across local labour markets. This economic application focuses specifically on the matching micro-foundation of agglomeration economies. The field of economic geography argues that dense local labour markets may realise positive spillovers that could lead to better labour market outcomes. A priori, the role of agglomeration in matching of firms to workers is ambiguous (Helsley and Strange, 1990). Agglomeration economies in local labour markets may lead to positive spillovers from improved matching of workers to firms through lower search costs. The lower costs of searching could lead to more modest losses in employment and wages for workers who are located in larger and denser labour markets. In contrast, the complexity of search is increasing in the size and density of a local labour market, which could lead to worse post-displacement labour market outcomes (Wheeler, 2001). My analysis documents the role of agglomeration economies in the consequences of job displacement.

CHAPTER 2

Workers' Spatial Response to Job Displacement[1]

2.1 Introduction

As in many other countries, the Dutch owner-occupied housing market and labour market suffered from strong negative developments during the Great Recession that started in 2008 (OECD, 2010).[2] The large scale at which the transaction prices and home property values fell in the Dutch housing market is very rare – it previously occurred in the period 1978 to 1982. The central question we address is how workers adjust after job displacement, by focusing on margins of adjustment that are related to space and the importance of workers' housing state.[3] A better

[1] A paper based on Chapter 2 is published in the Journal of Urban Economics entitled "The role of the housing market in workers' resilience to job displacement after firm bankruptcy" (Meekes and Hassink, 2019).

[2] In the Netherlands, the owner-occupied housing sector experienced a decline of almost 50 per cent in the number of transactions. Moreover, transaction prices decreased with over 20 per cent. The number of households with negative home equity, relative to the total number of owner-occupied households, increased from 13 per cent in 2008 to 34 per cent in 2014. Unemployment rose from 3.4 per cent in the third quarter of 2008 to its peak of 8.1 per cent in the first quarter of 2014. The number of bankruptcies of firms increased from 3,589 in 2007 to 8,376 in 2013 (CBS, 2018).

[3] We define workers' housing state as being a tenant or owner-occupier, where owner-occupiers are classified in five different groups based on the loan-to-value ratio

understanding of the use of margins of adjustment by displaced workers is relevant for policies that aim to limit the impact of negative employment shocks (Crépon and Van den Berg, 2016) and policies that affect housing state choice through subsidising homeownership or stimulating mortgage debt (DiPasquale and Glaeser, 1999).

Our first aim is to examine whether the spatial structure of homes and jobs represents relevant margins of adjustment for displaced workers. The literature on job displacement argues, based on human capital theory, that displaced workers lose human capital and wage premiums, and consequently experience substantial losses in post-displacement employment and wage outcomes (Hamermesh, 1987; Topel, 1990; Jacobson et al., 1993). The focus on losses in human capital, however, ignores the way displaced workers could use margins of adjustment related to space such as commuting patterns and household relocations to different homes, which are key to employment outcomes and wage dispersion.[4] Indeed, Huttunen et al. (2018) show that job displacement increases household relocations to different homes, and argue that the losses in employment and income are highest for displaced workers who change home. Hence, Huttunen et al. (2018) suggest that displaced workers change home for non-economic reasons such as family ties. We would argue that the commuting distance, compared to changing home, is a more relevant margin of labour adjustment for displaced workers, as displacement generates a negative income shock that lowers individuals' ability to change home. To better understand the use of margins of adjustment by displaced workers, we assess the displacement effects on employment, wages, commuting distance and changing home.

Our second aim is to examine the importance of workers' housing state for the displacement effects on employment, wages, commuting and changing home. The impact of workers' housing state on employment is theoretically ambiguous. On the one hand, homeowners relative to tenants, and mortgage owners relative to outright owners (i.e. owners who have paid off their entire mortgage), are less able to change home because

on their home. Note that in the housing economics literature, the definition of housing tenure is used to distinguish only between homeowners and tenants (e.g., see Henley et al. (1994)).

[4] See, e.g., Zax (1991); Simpson and Van der Veen (1992); Smith and Zenou (1997); Van Ommeren et al. (2000); Brueckner et al. (2002); Manning (2003); Smith and Zenou (2003); Fernandez and Su (2004); Van Ommeren and Fosgerau (2009); Mulalic et al. (2014).

of higher transaction costs and a more severe equity constraint (Stein, 1995; Chan, 2001; Ferreira et al., 2010; Andersson and Mayock, 2014).[5] This geographic immobility could hinder employment prospects. On the other, homeowners, relative to tenants, may search more efficiently for jobs due to the use of more effective search methods (Goss and Phillips, 1997; Morescalchi, 2016). Moreover, mortgage owners, relative to outright owners, search more intensively for jobs due to differences in housing costs and payment obligations. In turn, the higher search efficiency or intensity could increase employment prospects. The difference in the exit rate into employment among displaced workers who differ in homeownership and home equity can thus be explained by differences in geographic mobility, search efficiency and search intensity. An alternative hypothesis is that unemployed workers who are highly leveraged and face a home equity constraint are more willing to accept a larger wage loss or a job that is further away. For example, Brown and Matsa (2017) show that unemployed workers who experience a home equity constraint are more willing to apply for positions that are closer in distance but pay lower wages. Ultimately, it is an empirical question how the worker's housing state affects the use of the margins of adjustment.

For our empirical analysis, we created a monthly panel of employees based on rich administrative data sets that contain Dutch data on firms, employees and households in the period from January 2006 to December 2014. This time period is particularly suited as it includes the Great Recession and thereby allows us to incorporate data on the declining property values of homes and increasing number of bankruptcies of firms. We used data on job displacement due to firm bankruptcies (hereafter: job displacement) as an exogenous negative shock to the employment status of workers. The data on job displacement set the stage for a quasi-experimental design. This empirical design is important, since we examine incentive effects of workers' housing state in which endogenous selection into labour turnover should play no role. However, there could exist confounding factors that affect the probability of job displacement. The potential of selection into job displacement based on observables

[5] In contrast, several studies find that negative home equity positively affects geographic mobility by increasing the foreclosure and default rate (e.g., Coulson and Grieco, 2013; Bricker and Bucks, 2016; Demyanyk et al., 2017). In the Dutch institutional context, however, this mechanism is less relevant, as Dutch borrowers have full recourse loans and are characterised by a relatively low default and foreclosure rate (Steegmans and Hassink, 2018).

was minimised by exact matching on coarsened observables of treated (displaced) to similar control (non-displaced) workers (Iacus et al., 2011). We included individual-specific fixed effects and applied the double-differences (DD) and triple-differences (DDD) estimator to control for various sources of unobserved heterogeneity.

The displaced and matched non-displaced workers are followed for eighteen months prior until thirty-six months after job displacement. The framework of the DD estimator was applied to the coarsened exact matched sample to estimate the displacement effects. The displacement effects are inferred from reduced-form models on four margins of adjustment, i.e. changes in employment, hourly wage, commuting distance and changing home. The DDD estimator was applied to assess the role of workers' housing state in the effects of job displacement. Workers' housing state was categorised by tenancy and homeownership, where owners were classified in five groups based on their loan-to-value ratio (hereafter: LTV) that was expressed as a percentage.

Our analysis provides two sets of novel results. First, the estimates show that commuting and household relocations to a different home are significant margins of adjustment in response to job displacement. Specifically, the average displacement effects we estimated show that displaced workers during the post-displacement period of thirty-six months, on average, (i) are about 25 percentage points less employed, (ii) experience a loss in wage of 6 per cent, (iii) experience an increase in the commuting distance of 3 kilometres and (iv) have a 0.06 percentage points lower probability of changing home. The displacement effects on employment and wage are consistent with those reported in the job displacement literature for European countries (e.g., see Schwerdt (2011); Huttunen et al. (2011); Ichino et al. (2017)). To the best of our knowledge, a novel finding of Chapter 2 is the increase in the commuting distance and decrease in changing home for workers who have been displaced, which represent compared to the mean value about a 20 per cent and 14 per cent change, respectively. Also, the results reveal a remarkable pattern over time since job displacement: the negative displacement effect on wages becomes more pronounced, whereas the positive effect on commuting distance becomes smaller. Our results show that for displaced workers the commuting distance is a more relevant margin of adjustment for economic reasons, while changing home is more relevant for family reasons (Huttunen et al., 2018). The first contribution of Chapter 2 is to document the spatial response to job displacement, focusing on the displacement effects on workers' commutes and home changes.

Second, we find that workers' housing state plays a significant role in the displacement effects on employment, wages and the commuting distance, but not on the probability of changing home. We are not the first to examine the role of workers' housing state in employment outcomes.[6] However, the literature ignores various selection problems that arise due to the correlations among workers' exit rate into unemployment, cause of unemployment and housing state choice.[7] We contribute to this literature by exploiting a quasi-experimental design involving job displacement that eliminates the potential of endogenous selection into labour turnover. We find that displaced underwater owners (i.e. owners who face negative home equity), relative to displaced tenants, experience a 7 percentage points lower loss in employment. From the group of displaced homeowners, outright owners experience the highest loss in employment. Also, the results suggest that displaced tenants are more selective in post-displacement wages, and highly leveraged owners become employed relatively close to home. Displaced outright owners experience the highest increase in the commuting distance. The second contribution is to show that the effect of workers' housing state on the post-displacement outcomes in employment, wages and commute is substantial, and comparable to that of other relevant worker characteristics such as gender and sector.

2.2 Conceptual Framework

In this section, we discuss the conceptual considerations concerning the labour response and spatial response to job displacement and the effect of the worker's housing state. The post-displacement labour market

[6] So far, there is abundant evidence of positive homeownership and mortgage effects on the exit rate into employment (e.g., Goss and Phillips, 1997; Coulson and Fisher, 2002; Flatau et al., 2003; Munch et al., 2006; Battu et al., 2008; Baert et al., 2014; Caliendo et al., 2015; Morescalchi, 2016). However, there is little to no evidence that negative home equity of homeowners affects the labour market. At the micro level, Valletta (2013) finds that being an underwater homeowner (i.e. an owner who faces negative home equity) does not affect the exit rate into employment. At the aggregate level, Kothari et al. (2013) and Modestino and Dennett (2013) find a very small impact of the lower mobility caused by the housing bust on the U.S. employment rate.

[7] For a discussion of these correlations, see, e.g., Van Leuvensteijn and Koning (2004); Munch et al. (2008); Moriizumi and Naoi (2011); Head and Lloyd-Ellis (2012); Bricker and Bucks (2016).

outcomes depend on the intersection of the supply and demand of the labour market.

From the supply side of the labour market, the key choice behind the length of a displaced worker's unemployment spell is whether employment is preferred to the alternative of remaining unemployed while searching for better job offers. Taken from Rogerson et al. (2005), (2.1) shows the exit rate into employment H to be equal to the product of the job offer arrival rate α and the probability of accepting the job offer $1 - F(w_r)$

$$H = \alpha[1 - F(w_r)] \tag{2.1}$$

where F denotes the wage-offer distribution and w_r represents the reservation wage. Observe from (2.1) that H can be increased by either accepting a lower w_r or by accepting a higher search effort. A higher search effort enables the unemployed worker to attract a higher α. The optimal choice of each worker depends on the continuous time reservation wage equation

$$w_r \geq b - g(\alpha) + \alpha/r \int_{w_r}^{\infty} (w - w_r) dF(w) \tag{2.2}$$

where b represents the unemployment benefits and $g(\alpha)$ represents the costs of a job match, which includes the cost of searching, the cost of commuting and the cost of changing home. The first-order condition for an interior solution equals

$$rg'(\alpha) = \int_{w_r}^{\infty} (w - w_r) dF(w) \tag{2.3}$$

Holding α fixed, (2.2) and (2.3) show that a decrease in b would decrease w_r.

2.2.1 A Spatial Response to Job Displacement

Besides adjusting the search effort to increase α, an alternative option for the displaced worker would be to expand his or her geographical search area. The geographical scope could be expanded by accepting a higher commuting distance while keeping the same home. For a relatively large increase in commute, workers are more likely to relocate to a different home. This mechanism suggests that displaced workers who change home have a relatively small increase in commute. However, a priori, it is likely that the changing home margin of adjustment is less relevant than the commute margin, as the ability to change home is reduced following displacement. A displaced worker might be less able to relocate due to the loss in income that makes it difficult to get a new mortgage or a rental agreement. Moreover, changing home for employment reasons is less likely in the Dutch-specific context, as the Netherlands has a relatively high employment density and a small area size.

Note that there is an ambiguous effect of α. On the one hand, a higher α increases H (see (2.1)). On the other, a higher α leads to a higher w_r (see (2.2) and (2.3)), which decreases H. Similarly, a higher w_r might lead to a lower willingness to commute or to change home, as the costs of commuting and changing home increase when the wage does (Van den Berg, 1992; Glaeser et al., 2008). In this regard, there exists a trade-off between the reservation wage, the willingness to commute and the willingness to change home, as these are jointly determined by the worker (Mulalic et al., 2014; Mayock, 2015). We predict that the post-displacement loss in wages and the post-displacement increase in commutes are negatively correlated. If we assume that b would become less generous over time and are finite in time, the length of the post-displacement unemployment period becomes important. Based on decreasing b over time, we predict that the wage of displaced workers is likely to decrease over the post-displacement period. To compensate an unemployed worker for lower wages, the displaced workers' commute and probability of changing home are also expected to decrease over the post-displacement period.

2.2.2 Housing Status and the Response to Job Displacement

From the demand side of the labour market, we should recognise that firms have different preferences for types of workers. Consequently, the intensity and pay-offs of the search could be endogenous to the worker's individual attributes, including their gender, education, age and household state. However, the individual attributes also affect the supply side of the labour market, through differences in the reservation wage, willingness to commute and willingness to move. The worker's housing state is relevant, because differences in payment obligations affect the financial incentive to become re-employed. We start our conceptual analysis from Munch et al. (2006), who focus on the home change margin of adjustment and compare unemployed homeowners to unemployed tenants. They argue that homeowners experience a higher $g(\alpha)$, because the transaction costs of changing home are higher for homeowners than tenants. Consequently, unemployed homeowners will set relatively low reservation wages in order to avoid having to change home to a distant local labour market (see (2.2)). The study distinguishes between local labour markets and distant labour markets. Munch et al. (2006) predict that homeowners have a relatively high exit rate into employment in their local labour market, but a relatively low exit rate in a distant labour market. Importantly, they recognise that the commute margin is relevant, and predict that homeowners are more likely to use the commute margin if the costs of changing home are higher for homeowners than for tenants.

One of the contributions to the literature on housing and unemployment duration of Chapter 2 is that we incorporate the commute response to job displacement and recognise that homeowners vary in the degree of leverage. Consider a displaced worker who is a mortgage owner. The unemployed mortgage owner faces a financial constraint and is obliged to amortise and pay off their mortgage. The payment obligations are likely to induce relatively high opportunity costs of continued unemployment and forgone wage. Conversely, displaced owners who have paid off their full mortgage have relatively weak financial incentives to become employed. A key question is how displaced workers who are relatively leveraged can increase their exit rate into employment. Morescalchi (2016) argues that unemployed leveraged owners search more intensively than other unemployed workers. If this holds, an unemployed mortgage owner is expected to have a relatively high α, keeping w_r and the willingness to commute and change home constant. Alternatively, leveraged workers could increase H by accepting a lower w_r, a higher commuting distance, or a relocation to a different home (Brown and Matsa, 2017).

A priori, it is an empirical question whether and to what extent unemployed leveraged workers prefer to use the wage margin over the commute margin to become employed. Interestingly, the role of the unemployed worker's housing status in the use of the changing home margin is ambiguous. Unemployed negative home equity owners, compared to unemployed outright owners, could be more willing to change home for employment reasons. However, they are less able to change home for financial reasons, as they face a liquidity constraint. Note that in the

Netherlands, negative home equity owners are unable to strategically default on their home, as all mortgages are full recourse loans. Together with the credit constraint that is caused by negative home equity and the job displacement, we assume that relatively leveraged owners lack the ability to get a new mortgage or rental agreements. In the spirit of Huttunen et al. (2018), our reasoning suggests that displaced underwater owners are able and willing to change home mostly for family reasons, for example by moving in with relatives. We will empirically analyse whether wages, commutes and changing home are relevant margins of labour adjustment for workers who have been displaced.

2.3 Institutional Background and Data

2.3.1 Institutional Setting of the Dutch Labour Market

In the Netherlands, workers who are collectively displaced, for example through mass-layoffs, are protected by the Law Collective Redundancy Act (in Dutch: Wet Melding Collectief Ontslag). The Law Collective Redundancy Act does not hold for dismissals if the firm is declared bankrupt, as job displacement due to firm bankruptcy concerns an urgent case of displacement. This restriction has two implications for workers whose labour contract is terminated due to bankruptcy of the firm, and these implications are the reason we exploit the quasi-experimental design involving job displacement due to firm bankruptcy.[8]

First, the notification requirement to displaced workers, which is specified in the Law Collective Redundancy Act, applies only at the request of the Public Employment Service. Therefore, in general, no advance notification is required from bankrupt firms to displaced workers. Second, if a firm goes bankrupt, no severance or transition payments are provided by the firm to the displaced worker. These implications are important, as heterogeneity in the advance notification and severance pay can have a confounding effect on the post-displacement length of the unemployment spell and earnings (Addison and Portugal, 1987; Nord and Ting, 1991).

In the context of non-culpable job displacement such as displacement due to firm bankruptcy, the eligibility with respect to the length and

[8] For more information, see articles 3.1, 4.6 and 5a.2 of the Law Collective Redundancy Act on `http://wetten.overheid.nl/BWBR0003026/2018-01-01`.

amount of unemployment benefits is relevant. In the Netherlands, workers are eligible for one month of unemployment benefits for each year of job tenure in which the workers was employed for at least 208 hours. The first two months of unemployment benefits are equal to 75 per cent of the average wage a worker has earned over the last twelve months. After two months of unemployment benefits, the unemployment benefits are paid out at 70 per cent.

2.3.2 Institutional Setting of the Dutch Housing Market

The institutional setting of the Dutch housing market has various characteristics that make it different from, e.g., the institutional setting of the U.S. housing market. First, in the Netherlands there is no formal down-payment requirement when buying a new home. Consequently, the probability of having a liquidity constraint to buy a new home is less likely. Second, Dutch homeowners are characterised by a relatively high LTV, as they can borrow more than the property value of the home. Only in August 2011 there was a binding code of conduct for mortgages introduced, which set an LTV limit at 106 per cent. Since 2012, the binding LTV limit of 106 per cent has been decreasing by 1 percentage point a year to 100 per cent in 2018. Third, all mortgage loans in the Netherlands are full recourse loans. Hence, the phenomenon of strategic defaults, i.e. walking away from a home with negative home equity, does not exist. Fourth, homeowners in the Netherlands, but also in the U.S., can deduct home mortgage interest to reduce their taxable income. The deductibility of home mortgage interest stimulates mortgage debt. Finally, the social rented sector of the Dutch housing market is relatively large (CBS, 2018). In 2012, there were about 7,141,000 Dutch households. Of these households, approximately 59.32 per cent were owner-occupied and 40.68 per cent were rented. Of the households that rent, 88.47 per cent rent social housing and 11.53 per cent rent private housing.

2.3.3 Data Sets

We used various administrative data sets, retrieved from Statistics Netherlands, to create a monthly panel. The data sets we used contain data on individuals, households and firms over the time period 2006 to 2014. We selected displaced workers whose job was terminated due to bankruptcy between July 2007 and December 2011. Each individual employee was followed for eighteen months prior until thirty-six months after the actual or potential job displacement. The actual month of job displacement of a displaced worker is equivalent to the potential month of displacement of a matched non-displaced worker. In Appendix 2.A, we provide additional information on the data construction and sample selections that were applied to create the appropriate sample of individual employees.

2.3.4 Key Dependent Variables

The key dependent variables, which represent the four margins of adjustment, were operated as follows. Employment status was expressed as a zero-one indicator variable that equals one if the individual is employed. The natural logarithm of the hourly wage was constructed by taking the natural logarithm of the monthly contractual gross wage relative to the monthly number of contractual hours worked. The commuting distance was calculated by taking the absolute distance between the central business district (CBD) of the work municipality and the CBD of the neighbourhood of the home. Note that the hourly wage and commuting distance of workers are observed conditional on employment. Unfortunately, the commuting distance is not observed for workers if they were not in the job in December, as the work location is administered only in December. The number of observations that are missing for the model in which commuting distance is the dependent variable, can be observed by the comparison with the model on hourly wage.[9] See Table 2.C.1 in Appendix 2.C for the within change in the hourly wage and commuting distance for displaced and non-displaced workers. Home change was expressed as an indicator variable and equals one if the household has relocated.

2.3.5 Independent Variables and Covariates

The set of key independent variables consists of variables that represent the treatment status, post-displacement status and housing state. The set of covariates consists of various demographic characteristics and job characteristics. All variables were expressed as zero-one indicator variables.

[9] See Table 2.D.1 for the results using a sample where all individuals have complete information on commuting distance. We find similar results.

The treatment status is time-constant and equals one for workers who have been displaced. The post-displacement status equals one if the month under observation is after the month of job displacement. To assess the time dimension of the displacement effects, the post-displacement variable was in some of the models replaced by fifty-five indicator variables. The indicator variables reflect the time gap in months of the period since job displacement and range from minus eighteen to plus thirty-six, respectively. An indicator variable equals one if the period since job displacement corresponds to the given time gap.

Workers' housing state was used to approximate the workers' degree of household leverage and it was represented by six indicator variables. We distinguished between tenants and homeowners, where owners were categorised based on their LTV. Note that we were not able to distinguish between tenants who rent social or private housing. However, most tenants in our sample rent social housing, as 88.47 per cent of all Dutch tenants do so. The LTV, which is expressed as a percentage, was constructed by the observed mortgage loan relative to the property value of the home. The six indicator variables equal one if the worker is a tenant (base category), is an outright owner (i.e. an owner who has an LTV equal to 0 per cent on the home) and has an LTV between 0-33, 33-66, 66-100 and 100-133 per cent, respectively.[10] Note that the so-called underwater homeowners have an LTV over 100 per cent on the home.

The set of demographic characteristics consists of indicators for gender, Dutch nationality, and age (21-30, 30-40, 40-50 and 50-60 years). The set of job characteristics represents the worker's time-invariant job characteristics in the month of job displacement, and consists of indicators for job tenure (3-6, 6-12, 12-18 and over 18 years) and industry (manufacturing and services sector). In addition, the set of covariates includes indicators for the year of job displacement, having children aged 18 or lower, having a partner, and the number of household members (1, 2, 3-4 and more than 4 members).

[10]Unfortunately, the data on mortgage debt do not include the asset side in endowment mortgages. Hence, the levels of the mortgage debt were likely to be overestimated and the effect of workers' housing state is likely to be biased towards zero. To reduce the potential of the attenuation bias in the LTV, we operated housing state as a categorical variable. Table 2.B.5 in Appendix 2.B shows the results for an indicator variable that represents homeownership and a continuous variable that represents owners' LTV. The main conclusions of Chapter 2 are robust to the use of a continuous variable to represent owners' LTV.

2.4 Identification Strategy

2.4.1 Identification Challenges

For Chapter 2, the endogeneity problems of selection into labour turnover and selection into housing state required particular attention. Note that these selection issues are closely connected. For example, a sorting mechanism based on human capital or preferences for stability could simultaneously influence labour market outcomes and housing state choice (Flatau et al., 2003; Van Leuvensteijn and Koning, 2004; Munch et al., 2006, 2008; Moriizumi and Naoi, 2011; Head and Lloyd-Ellis, 2012; Bricker and Bucks, 2016). In this spirit, the likelihood of having a stable job, unemployment risk and housing state choice are likely to be correlated.

Selection issues are likely as various worker characteristics, for example age and gender (Kuhn, 2002; Von Wachter and Bender, 2006), job tenure (Farber, 1999), and industry and education (Farber et al., 1993), affect the probability and cause of exit into unemployment. In turn, the selection issues could be problematic as the cause of unemployment affects, through signalling, the magnitude of displacement effects on post-unemployment labour market outcomes (Gibbons and Katz, 1991; Stevens, 1997; Kuhn, 2002; Hu and Taber, 2011; Frederiksen et al., 2013). We controlled for the selection issues in various ways.

First, following the literature on job displacement, we deal with selection into labour turnover by the use of data on job displacement as an exogenous negative employment shock that set the stage for a quasi-experimental design (Eliason and Storrie, 2006).[11] This empirical design limits the problem of confounding factors that affect post-displacement labour turnover, as displaced workers have an identical signalling value on labour market outcomes and typically do not get an advance notification or severance pay. In line with the job displacement literature, workers with a job spell of at least three years were selected for the group of displaced and non-displaced workers.[12] Thereby, all workers had prior

[11] In our sample, workers were displaced between July, 2007 and December, 2011. Hence, there is variation over time in the variable that represents treatment of workers. This greatly reduces the potential of standard errors that understate the standard deviation of the estimator (Bertrand et al., 2004).

[12] In the Netherlands, workers are eligible for one month of unemployment benefits for each year of tenure in the job. The selection ensures that all workers were eligible for unemployment benefits for the first three months after job displacement. By having

to job displacement a stable employment pattern and displacement was likely to be unforeseen. Furthermore, this sample selection reduces the potential of selection into housing state based on the belief of having a stable job.

Second, we deal with selection into job displacement based on observables, by applying Coarsened Exact Matching (CEM) that makes displaced and non-displaced workers balanced in covariates. CEM is a member of the class of Monotonic Imbalance Bounding matching methods and dominates the propensity score methods (Iacus et al., 2011). By balancing workers in covariates, the selection bias into displacement based on observables, which can arise from lack of common support, was greatly reduced (Heckman et al., 1997, 1998; Heckman and Smith, 1999).

Third, to deal with bias based on unobserved heterogeneity, we apply the double-differences (DD) and triple-differences (DDD) estimator. DD was used to estimate the displacement effects, i.e. the change in outcome after treatment by comparing matched displaced to non-displaced workers. DDD was applied to examine the sources of differences in the displacement effects among workers who have been displaced. For example, the DDD estimate of a given housing state equals the difference between the DD estimate for the given housing state and the DD estimate for the reference category of housing state. The key identification restriction of the DD and DDD estimator requires that, conditional on observables, the change in outcomes of the displaced workers and non-displaced workers follow parallel paths.[13]

Finally, we controlled for many factors that affect the exit rate into unemployment, likelihood of homeownership and the LTV on the home. For example, we controlled for changes in age and the presence of children aged 18 or lower. Moreover, indicator variables for calendar month (107) and NUTS 3 area (39) were included to capture business cycle effects and area-specific fixed effects, respectively. Individual-specific fixed effects were incorporated to eliminate bias from time-invariant unobserved heterogeneity, e.g., constant skill or worker preferences that might simultaneously affect housing state choice and labour market

---

a minimum benefits duration and controlling for the job tenure of the terminated job, we distance ourselves from the effect of benefits duration on post-unemployment labour market outcomes (Katz and Meyer, 1990; Bover et al., 2002).

[13]See Section 2.5 for further discussion. See Figure 2.D.2 for parallel pre-displacement paths of displaced and non-displaced workers using a placebo treatment of the displaced twelve months prior to actual displacement.

outcomes.

In two robustness checks, which are shown in Appendix 2.B, we controlled for changes in the wealth position and duration of home occupancy, and for education level, respectively, to correct for unobserved heterogeneity driven by human capital. Note that we did include individual-specific fixed effects but did not control for education in the main part of our analysis, because data on education is only available for individuals if they received their diploma after the year 1995. The use of the education data would lead to a substantial selection issue, as less leveraged owners, especially outright owners, are relatively old and received their diploma before the year 1995.

2.4.2 Coarsened Exact Matching Procedure

In a matching algorithm there is the trade-off between efficiency and lower bias, i.e. the choice between complete matching and exact matching (Rubin, 2006; Caliendo and Kopeinig, 2008). Exact matching ensures a high quality of matching as the amount of imbalance between matched treated and controls is controlled and limited. However, CEM does not lead to complete matching. Complete matching is achieved if all treated are matched with at least one control. We performed CEM of treated to controls as we prefer a lower bias to efficiency gains. Moreover, we had the opportunity to exploit rich administrative data with a high number of potential controls. Exact matching on coarsened observables ensured that the treated and controls were observably equivalent.

Workers who were displaced due to firm bankruptcy are referred to as treated. The non-displaced are referred to as controls. In the month of job displacement, the treated were matched with a potential match of the group of controls. The controls were required to stay employed in the month of separation of the treated. Each treated was matched with a maximum of two controls. Note that the potential month of displacement of the matched control is equivalent to the actual month of job displacement of the matched treated. Except for job displacement due to firm bankruptcy, the controls were exposed to similar risks of labour turnover as the treated. Exposure of controls to other reasons for job loss is important to avoid overestimates of the displacement effects (Krolikowski, 2018). These risks represent voluntary labour turnover and involuntary labour turnover. The treated or the matched controls were not allowed to be the counterfactual of another treated worker in the other months under observation. For this reason, the order of months in the period July 2007 to December 2011, in which we separately match treated workers with control workers, was taken randomly.

Before we applied CEM, the non-matched sample consisted of 31,888 treated workers. See Table 2.C.2 in Appendix 2.C for individual summary statistics for the treated and controls based on the non-matched sample. The default set of matching variables we incorporated in the matching process consists of indicator variables for gender, age (21-30; 30-35; 35-40; 40-45; 45-50 and 50-60 years), children aged 18 or lower, partner, Dutch nationality, tenancy, LTV (0; 0-33; 33-66; 66-100 and 100-133 per cent), type of job (full-time or part-time), type of contract (fixed or temporary), job tenure (3-6; 6-12; 12-18 and over 18 years), work location (twelve provinces), firm size (10-49; 50-99; 100-499 and 500 or more employed workers), firm industry (twenty-one ISIC sectors), calendar month and calendar year. The matched sample consisted of 20,152 treated workers, which implies a matching rate of 63 per cent. See Table 2.C.3 in Appendix 2.C for individual summary statistics of the treated and controls based on the matched sample.

The matching procedure we applied to balance treated and controls in covariates was successful. Based on the comparison of Table 2.C.3 to Table 2.C.2 in Appendix 2.C, we observe that the difference in sample means between the treated and controls was smaller after matching and many became economically insignificant. See Table 2.C.4 for an overview of the number of matched individuals by housing state and treatment group. See Table 2.C.5 for firm size and firm sector summary statistics in the month of job actual displacement. See Table 2.C.6 for individual summary statistics distinguished by workers' housing state.

To assess the implications of incomplete matching, we matched on the work location at the NUTS 3 level (forty areas) instead of at the provincial level (twelve areas). The matched sample consisted of 14,284 matched treated workers. The matching rate decreased from 68 per cent to 45 per cent. Table 2.D.3 shows that the results are robust to a difference in the matching rate and matching on the NUTS 3 area.

As a robustness check, we matched not only on the default set of matching variables, but also on the worker's categories of the non-housing wealth position (below 0; 0-5,000; 5,000-25,000; 25,000-75,000 and over 75,000 euro) and duration of home occupancy (0-60; 60-180 and over 180 months). In this case, the number of matched treated was 10,128. In a separate robustness check, we used both the loan-to-income (LTI) ratio and LTV as approximations of household leverage.[14] For this robustness check, we matched not only on the default set of matching variables, but also on the LTI ratio categories (0-1.5; 1.5-3.0, 3.0-4.5; 4.5-6.0 and over 6.0). This approach resulted in 16,222 matched treated workers. As a final robustness check, we matched not only on the default set of matching variables, but also on the skill level (i.e. low, medium and high education). Matching on the skill level resulted in a relatively low number of 5,841 matched treated workers. The low number of matched treated individuals was caused by the selectivity of education data, as the education data were only available for individuals if they received their diploma after 1995. The three robustness checks are discussed in Appendix 2.B.

2.4.3 Margins of Adjustment

For each of the margins of adjustment a generic empirical model is specified to estimate the displacement effect. In what follows, Y represents one of the four margins of adjustment – employment, hourly wage, commuting distance and changing home. The empirical model is

$$\begin{aligned} Y_{irt} = {} & \delta(DISPLACED_i \times POST_{it}) + \rho POST_{it} \\ & + \beta' X_{it} + \alpha_i + N_r + D_t + \varepsilon_{irt} \end{aligned} \tag{2.4}$$

$i \in 1, 2, ..., N$; $r \in 1, 2, ..., 40$; $t \in 1, 2, ..., 108$

where subscripts i, r and t denote the worker, regional NUTS 3 area and month, respectively. The systematic differences in the outcome variables are captured by parameter δ of the two-way (double) interaction term between the scalar indicator variables *DISPLACED* and *POST*. The indicator variable *DISPLACED* is time-constant and equals one for workers who experienced job displacement. Note that the main effect of *DISPLACED* is accounted for by including individual-specific fixed effects. The indicator variable *POST* equals one for the post-displacement period of thirty-six months after job displacement. The base and omitted categories of the variables *DISPLACED* and *POST* are the non-displaced and the pre-displacement period, respectively. The worker's housing state

[14] We prefer the LTV to the LTI ratio as our main approximation of household leverage, because the LTV allows for within variation caused by changes in the property value and the mortgage debt. The LTI ratio does not allow for within variation caused by changes in income, as job displacement generates an artificial loss in income.

and covariates are represented by vector X. The parameters of the covariates are represented by vector β. Individual-specific fixed effects are referred to by α. N_r represents indicators for the home location at the NUTS 3 level.[15] Calendar month indicators are denoted by D. ε refers to the idiosyncratic error term.

The second empirical model, shown in (2.5), is specified to assess whether the displacement effect is persistent over the period since job displacement. The empirical model is

$$\begin{aligned} Y_{irt} = & \sum_{\tau=-18}^{36} [\delta^{\tau} DISPLACED_i \times G_{it}^{\tau} + \rho^{\tau} G_{it}^{\tau}] \\ & + \beta' X_{it} + \alpha_i + N_r + D_t + \varepsilon_{irt} \end{aligned} \tag{2.5}$$

where the time-dependent differences are captured using interaction terms among the indicator variables *DISPLACED* and G^{τ}. The main parameters of interest are referred to by δ^{τ}. Parameter τ is defined as the time gap in months of the period since job displacement, which ranges from minus eighteen to plus thirty-six in increments of one. The negative values of τ correspond to the months prior to job displacement. Parameter τ equals zero in the actual and potential month of displacement for the displaced and non-displaced, respectively. The scalar indicator variable G^{τ}, $\tau = -18, ..., 36$, refers to the τ-th time gap between the month of job displacement and the month under observation. For example, indicator variable $G^{\tau=-12}$, which represents the base category, equals one if the period prior to job displacement is equal to twelve months. We used the twelfth month prior to job displacement as the base category, because workers might experience changes in outcomes close to the month of displacement in anticipation of the job displacement.

2.4.4 Housing State and Margins of Adjustment

We added various interaction terms to assess the sources of differences in the use of margins of adjustment by displaced workers. Workers are distinguished by their housing state, demographic characteristics and job characteristics. The empirical model in (2.6) complements the model in (2.4), by adding multiple three-way (triple) interaction terms among a

[15] For reasons of clarity, please note that N_r refers to the location and N refers to the number of individuals.

vector of worker characteristics X, $DISPLACED$ and $POST$. The vector X includes time-varying variables (housing state and age) as well as time-invariant variables (gender, nationality and characteristics of the terminated job). The empirical model is

$$\begin{aligned} Y_{irt} = &(\kappa' X_{it}) \times DISPLACED_i \times POST_{it} \\ &+ (\gamma' X_{it}) \times DISPLACED_i + (\eta' X_{it}) \times POST_{it} \\ &+ \delta DISPLACED_i \times POST_{it} + \rho POST_{it} \\ &+ \beta' X_{it} + \alpha_i + N_r + D_t + \varepsilon_{irt} \end{aligned} \tag{2.6}$$

where the parameters of interest are represented by vector κ.

The empirical model in (2.7) complements that of (2.5). The model in (2.7) is specified to assess the time dimension of the role of worker characteristics in the displacement effects. The time-dependent differences are captured by multiple three-way interaction terms among the indicator variables X, $DISPLACED$ and G^{τ}. The empirical model is

$$\begin{aligned} Y_{irt} = &\sum_{\tau=-18}^{36} [(\kappa'^{\tau} X_{it}) \times DISPLACED_i \times G_{it}^{\tau} \\ &+ \delta^{\tau} DISPLACED_i \times G_{it}^{\tau} + (\eta'^{\tau} X_{it}) \times G_{it}^{\tau} + \rho^{\tau} G_{it}^{\tau}] \\ &+ (\gamma' X_{it}) \times DISPLACED_i + \beta' X_{it} + \alpha_i + N_r + D_t + \varepsilon_{irt} \end{aligned} \tag{2.7}$$

where the parameters of interest are denoted by vector κ^{τ}.

2.5 Empirical Results

2.5.1 Displacement Effects on the Margins of Adjustment

We examine the displacement effects on the four margins of adjustment (see Eq. (2.4)). Columns (1), (2), (3) and (4) of Table 2.1 show the displacement effects on employment, hourly wage, commuting distance and changing home, respectively. For the variables displacement status ($DISPLACED$) and post-displacement period ($POST$), the omitted categories are the non-displaced workers and the pre-displacement period, respectively.

Table 2.1 shows that displaced workers are 25 percentage points less employed than non-displaced workers over the post-displacement period of thirty-six months. Moreover, re-employed displaced workers, compared with employed non-displaced workers, experience a loss of about 6 per cent in hourly wage and an increase of 3 kilometres in the commuting distance. The increase of 3 kilometres in the commuting distance represents an increase of about 20 per cent, as the average commuting distance for displaced workers in the month of displacement equals 15 kilometres (see the summary statistics presented in Table 2.C.3). Finally, we observe a negative displacement effect on the probability of changing home of 0.06 percentage points. The displacement effect on changing home corresponds to a decrease of about 14 per cent, as the mean changing probability for displaced workers in the month of displacement equals 0.0043.[16]

The negative displacement effect on changing home suggests that displaced workers, compared to non-displaced workers, are less able or willing to change home. This could be explained by the fact that displacement generates a negative income shock, which increases difficulties in financing a new home or signing a new rental agreement. In contrast to our findings, Huttunen et al. (2018) show, using Norwegian administrative data, that job displacement increases the propensity to change home between local labour markets by about 30 per cent.[17] They find that displaced workers who change home experience higher losses in employment and wages than stayers, and argue that workers are likely to change home for non-economic reasons such as family ties. An explanation for the difference in findings could be the fact that the surface area of the Netherlands is ten times as small as the surface area of Norway. This could make Dutch displaced workers relatively unwilling to change home for non-economic reasons such as family ties, as relatives are relatively close. An alternative explanation is the difference in the length of the post-displacement period. The post-displacement period of thirty-six months we used could be too short to observe a displacement effect on household mobility. However, Huttunen et al. (2018) show that the increase in mobility takes place in the first two years after displacement. All in all, we argue that the commuting distance, compared to changing

[16]The results in Table 2.1 are robust to standard errors clustered by firm instead of by individual. The conclusion with respect to the displacement effect on commute is robust to the use of the natural logarithm of the commuting distance. These results are available upon request.

[17]Note that we estimate the displacement effect on changing home together for relocations within and between local labour markets. For an empirical model in which the dependent variable is operated as changing home between NUTS 3 areas, we find no significant displacement effect. This result is available upon request.

home, is a more relevant margin of labour adjustment for workers who have been displaced.

Table 2.1
Displacement effects on the four margins of adjustment (Eq. (2.4)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*	-0.2531^{***}	-0.0625^{***}	3.1854^{***}	-0.0006^{***}
	(0.0026)	(0.0017)	(0.2328)	(0.0002)
Number of parameters	161	161	161	161
Number of individuals	54,229	54,229	54,229	54,229
Number of observations	2,982,595	2,719,570	2,686,298	2,982,595

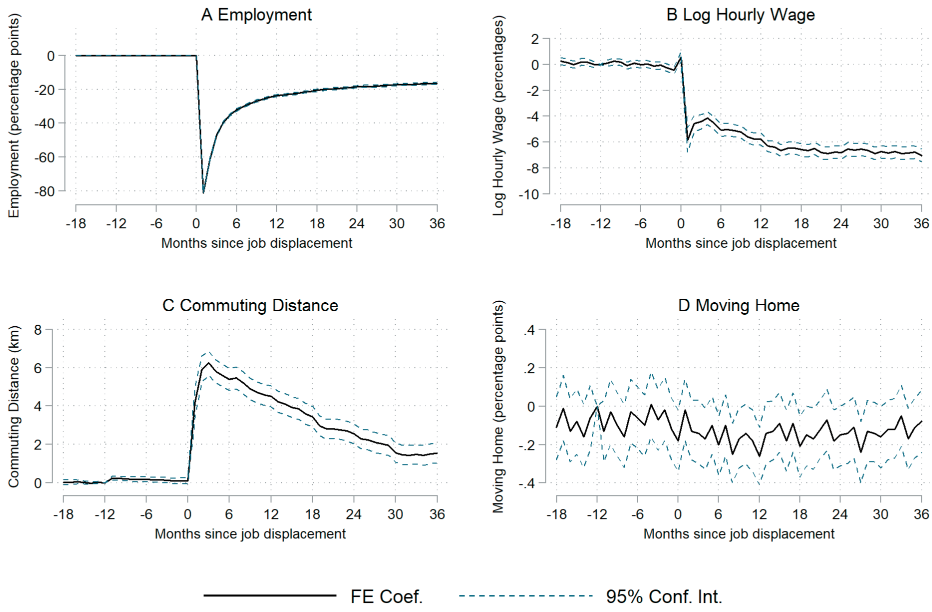
Notes: Each column gives the dependent variable. Parameter estimates of the two-interaction term are reported. Clustered (by individual) standard errors are in parentheses. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. The reference categories of *DISPLACED* and *POST* consist of the non-displaced workers and pre-displacement period, respectively. The regression analyses include individual-specific fixed effects and indicator variables for *POST*, housing state (5), age (3), children aged 18 or lower, partner, the number of household members (3), the NUTS 3 location of the household (39) and calendar month (107). The period under observation is from January 2006 to December 2014, in which displaced and non-displaced workers are observed for 18 months prior until 36 months after the actual and potential month of job displacement, respectively. Parameter estimates of the covariates are not reported.

Figure 2.1 shows the context of changes in the outcome variables of matched displaced and non-displaced workers over the pre- and post-displacement period (see Eq. (2.5)).[18] The fixed effects coefficients on employment and changing home are provided on the y-axis in percentage points (pp) in Figure 2.1A and Figure 2.1D, respectively. The fixed effects coefficients on wages and the commuting distance are provided on the y-axis in percentages (%) and kilometres (km) in Figure 2.1B and Figure 2.1C, respectively. The x-axis registers the time gap between the month of observation and the month of job displacement. The time gap equals zero in the month of actual and potential job displacement, for the displaced and non-displaced, respectively. Observe in Figure 2.1 parallel pre-displacement trends for the displaced and non-displaced workers in the outcome variables, which satisfy the identification restriction of the DD estimator.[19]

[18]The figures in this dissertation are based on the graphic schemes by Bischof (2017).

[19]Figure 2.1B shows a small hump in the monthly wage upon job displacement. Dutch salaries are supplemented with a holiday bonus and year-end bonus that workers

Fig. 2.1. Time-dependent displacement effects on employment (A), log hourly wage (B), commuting distance (C) and changing home (D) (Eq. (2.5)).



Notes: Reference group is the group of non-displaced workers. Reference month is the twelfth month prior to job displacement. The 95% confidence intervals are computed using clustered standard errors by individual. All four fixed effects regression models include 266 parameters of which there are 54 two-way interaction terms. See Table 2.1 for additional notes and statistics.

In Figure 2.1A, the vertical line between months zero and one reveals the exit rate out of employment by the displaced workers. Twelve to twenty-four months after job displacement, the loss in employment is about 27 and 20 percentage points, respectively. This finding is consistent with those reported in the job displacement literature. For example, Schwerdt (2011) finds an effect of 23 percentage points over a post-

typically get paid in May and December, respectively. The hump in wage can be explained by these bonuses, as they are paid upon displacement. Note that this hump underscores the importance of using the twelfth month prior to displacement as the reference month. See Deelen et al. (2014) for a similar finding using Dutch administrative data.

displacement period of five years. Ichino et al. (2017) find a loss of 27 percentage points over the post-displacement period of twelve to twenty-four months.

Figure 2.1B shows a decrease in the wage loss between one and four months after job displacement. This decrease in wage loss could be explained by the low number of re-employed displaced workers, i.e. only 20 per cent of the displaced workers, one month after job displacement (see Figure 2.1A). Note that Chapter 2 is one of the few studies in the literature on job displacement that uses monthly data. For papers that use quarterly or annual data, such a pattern cannot be observed as the monthly variation is smoothed out. Importantly, after the fourth month since job displacement, the loss in wage increases over the post-displacement period and ranges between 4 and 7 per cent (see Figure 2.1B). The estimates are in line with studies that examine the displacement effect on wage for Europe. For example, Schwerdt (2011) finds a wage loss due to job displacement, conditional on re-employment, of about 6 per cent. Huttunen et al. (2011) find a loss of 3 per cent in wage after 7 years. Note that studies that use U.S. data find higher wage losses due to the more centralised wage system (e.g., Couch and Placzek, 2010).

Figure 2.1C shows that displaced workers experience an increase in the commuting distance. However, after three months since job displacement, the increase in commute becomes smaller over the post-displacement spell. Note that Figure 2.1B and Figure 2.1C show composition effects that are caused by workers who exit unemployment, workers who experience job-to-job transitions and workers who change home. Interestingly, the composition effects are almost entirely driven by workers who exit unemployment and take up their first job since displacement. We show that this observation holds in Figure 2.D.1, where we present estimates based on a sample in which we select workers who are in their first job since displacement and did not change home. Hence, the estimates shown in Figure 2.1B and Figure 2.1C reveal a novel pattern: displaced workers who have longer unemployment duration experience smaller increases in commute but higher losses in wage. To the best of our knowledge, this pattern has not been demonstrated in the literature.

We observe a small negative displacement effect, after six months since job displacement, on the probability of changing home (see Figure 2.1D). The finding suggests a delay in the impact of job displacement on the willingness or ability to change home. This delay can be explained by the fact that household relocations to a different home are characterised by a time gap between the month of transaction and the month of the actual home change, as the actual home change is realised once the home buyer registers his or her relocation at the local municipality. Note, however, that the 36 two-way interaction terms that capture the post-displacement effect on changing home are jointly insignificant.

2.5.2 Housing State and the Use of Margins of Adjustment

We examine the role of workers' housing state in the displacement effects on the margins of adjustment. To show the relative importance of workers' housing state, we also document the role of other sources of differences among workers in the displacement effects. The estimation results are displayed in Table 2.2 (see Eq. (2.6)). Workers' housing state is represented by *LTV* indicators, and the reference category consists of workers who are tenants. The set of demographic characteristics includes *FEMALE*, *AGE* and *DUTCH NATIONALITY*, and the reference categories consist of workers who are male, aged 21 to 30 years and non-Dutch, respectively. The set of job characteristics includes *TENURE* and *MANUFACTURING*, and the reference categories consist of workers who are 3 to 6 years in the job and active in the service sector, respectively.[20]

The estimates presented in Column (1) show the sources of differences in the displacement effect on employment. Compared with displaced workers who are tenants, underwater owners are about 7 percentage points more employed after displacement. Tenants incur a loss in employment comparable to outright owners. Interestingly, our findings are consistent with papers that find positive homeownership and mortgage effects on the unemployment duration, but at odds with the paper by Valletta (2013) who finds no significant effect of negative home equity on unemployment duration. The difference in results could be explained by our quasi-experimental design that reduces the potential of selection into labour turnover. Finally, the estimates in Column (1) show that displaced workers who are female, older, non-Dutch and high-tenured experience a relatively high loss in employment. The role of workers' housing state in

[20] We positively evaluate the external validity of our analyses, as we find a comparable role of demographic and job characteristics in the losses in employment and wage to that in the literature (Madden, 1987; Carrington, 1993; Carrington and Zaman, 1994; Chan and Stevens, 1999, 2001; Eliason and Storrie, 2006; Hijzen et al., 2010; Tatsiramos, 2010; Hardoy and Schøne, 2014; Hellerstein et al., 2016; Farber, 2017).

the displacement effect on employment is substantial and comparable to that of other characteristics, such as gender and job tenure.

The parameter estimates in Column (2) highlight the role of worker characteristics in the displacement effect on hourly wage. Compared with displaced tenants and owners who have an LTV between 0 and 33 per cent, displaced outright owners and owners with an LTV over 33 per cent experience a 1 to 2 percentage points higher loss in hourly wage. The estimates suggest that more leveraged owners are less selective in wages when choosing jobs after job displacement. The results indicate that the loss in employment is decreasing and the loss in wages is increasing in the worker's LTV. This observation is supported by the results in Table 2.B.1 of Appendix 2.B, where we control for workers' non-housing wealth and duration of home occupancy. Moreover, the estimates in Column (2) show that the loss in hourly wage increases with age and tenure in the job, and is higher for displaced workers who are active in the manufacturing sector compared with workers who are active in the service sector. To the best of our knowledge, Chapter 2 is the first to demonstrate the importance of workers' housing state for the post-displacement losses in employment and hourly wage.

Column (3) shows that various worker characteristics lead to a difference in the displacement effect on the commuting distance. Although displaced owners, compared with displaced tenants, do not experience a highly different displacement effect on commute, there is a significant difference within the group of displaced owners. Specifically, displaced outright owners experience an increase of about 3 kilometres in commute higher than underwater owners. However, note that the three-way interaction terms for workers' housing state on commute are jointly insignificant. In addition, displaced women experience a significantly lower increase in the commuting distance than men. Moreover, the parameter estimates indicate that displaced workers who are high-tenured and active in the manufacturing sectors experience a relatively high increase in commute.

The parameter estimates in Column (4) show that the displacement effect on changing home does not vary greatly in worker characteristics. For example, housing state seems to have no role in the displacement effect on changing home. The finding that the displacement effect on changing home does not differ between tenants and homeowners, could be explained by the fact that Dutch tenants typically are in social rented housing. Although social tenants have generally lower transaction costs of changing home than homeowners, rent control and the waiting list for social rented housing reduce home changes (Munch and Svarer, 2002; Kattenberg and Hassink, 2017). In addition, the estimates suggest that for underwater owners, job displacement does not function as a trigger of default. In the housing literature, the double trigger theory of default predicts that households are likely to experience a default and change home if two conditions are met: (i) the household is underwater and (ii) the household wants or needs to change home (Foote et al., 2008). In the context of Chapter 2, job displacement could be considered as the potential second trigger that prompts a move out of the house. However, we do not find evidence of this. Our finding is in contrast with the studies of Niu and Ding (2015); Bricker and Bucks (2016), who show that job loss increases the foreclosure rate in the U.S., especially for workers with negative home equity. The disparity between the findings can be explained by the different institutional setting, as in the Netherlands the share of housing going into default and foreclosure is relatively low. The main reason for the low rate of default is, among others, the fact that all mortgage loans are full recourse loans.

We illustrate in Figure 2.2 the time-dependent differences in the importance of workers' housing state for the displacement effects (see Eq. (2.7)).[21] We observe parallel pre-displacement trends in the outcome variables except for changing home. The pre-displacement trend in changing home is relatively stable given the low number of monthly workers who change home (see Table 2.C.3). Importantly, for each of the four margins of adjustment, the F-test on the joint significance of the eighteen pre-displacement three-way interaction terms is statistically insignificant. These results are available upon request.

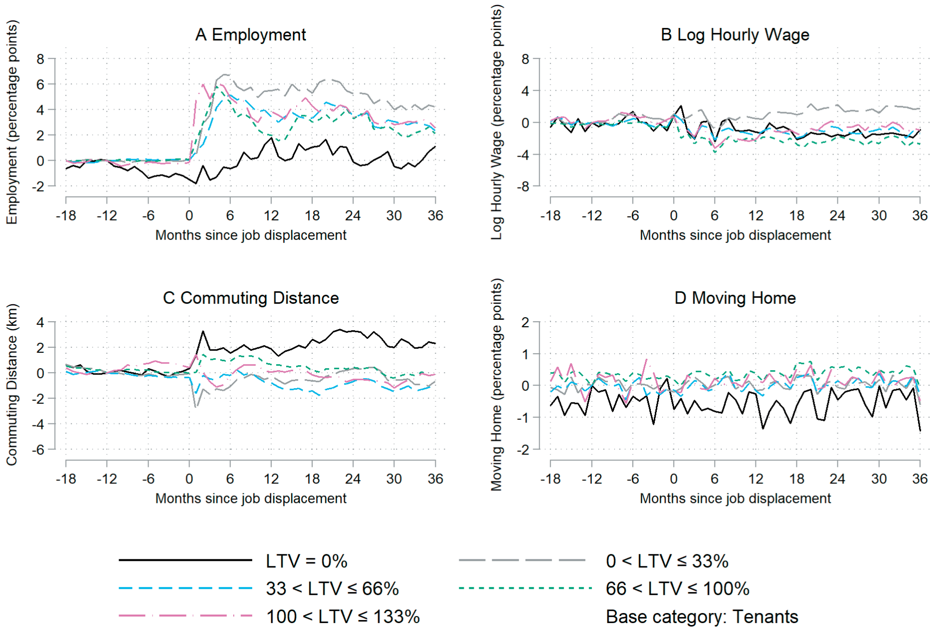
[21]See Figures 2.E.1-2.E.6 in Appendix 2.E for a comprehensive overview of the time-dependent differences in the importance of demographic and job characteristics for the displacement effects.

Table 2.2
The role of worker characteristics in the effects of job displacement (Eq. (2.6)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*×				
Housing state:				
$LTV = 0\%$	0.0149	-0.0101	2.1910*	-0.0008
	(0.0153)	(0.0093)	(1.2563)	(0.0013)
$0 < LTV \leq 33\%$	0.0326***	0.0085	-0.3285	0.0006
	(0.0089)	(0.0057)	(0.7629)	(0.0006)
$33 < LTV \leq 66\%$	0.0231***	-0.0110**	-0.5991	0.0009
	(0.0077)	(0.0051)	(0.6997)	(0.0006)
$66 < LTV \leq 100\%$	0.0255***	-0.0208***	0.3057	0.0011*
	(0.0075)	(0.0049)	(0.6861)	(0.0006)
$100 < LTV \leq 133\%$	0.0688***	-0.0156**	-1.0898	0.0002
	(0.0093)	(0.0062)	(0.8926)	(0.0009)
Joint significance of *LTV*	11.51***	7.14***	1.55	1.17
Demographic characteristics:				
FEMALE	-0.0677***	-0.0051	-5.2966***	0
	(0.0067)	(0.0046)	(0.6522)	(0.0004)
$30 < AGE \leq 40$ years	-0.0358***	-0.0257***	-1.5445*	0.0007
	(0.0094)	(0.0068)	(0.9334)	(0.0012)
$40 < AGE \leq 50$ years	-0.0665***	-0.0498***	-0.801	0.0012
	(0.0097)	(0.0070)	(0.9374)	(0.0011)
$50 < AGE < 60$ years	-0.1641***	-0.0603***	-1.0592	0.0013
	(0.0107)	(0.0075)	(1.0195)	(0.0012)
DUTCH NATIONALITY	0.0972***	-0.0097	0.3658	-0.0020**
	(0.0133)	(0.0083)	(0.9422)	(0.0008)
Joint significance of *AGE*	101.90***	27.53***	1.10	0.70
Job characteristics:				
$6 < TENURE \leq 12$ years	-0.0076	-0.0065	2.7982***	0
	(0.0065)	(0.0044)	(0.6340)	(0.0005)
$12 < TENURE \leq 18$ years	-0.0249***	-0.0142***	2.6314***	0.0005
	(0.0073)	(0.0049)	(0.6922)	(0.0005)
$TENURE > 18$ years	-0.0569***	-0.0443***	3.1450***	0.0004
	(0.0080)	(0.0054)	(0.7662)	(0.0004)
MANUFACTURING	0.0115**	-0.0225***	1.7336***	0
	(0.0053)	(0.0036)	(0.5083)	(0.0004)
Joint significance of *TENURE*	18.92***	24.81***	8.57***	0.96
Number of parameters	220	220	220	220
Number of individuals	54,229	54,229	54,229	54,229
Number of observations	2,982,595	2,719,570	2,686,298	2,982,595

Notes: Parameter estimates of the three-way interaction terms and F-Values of the joint significance of the three-way interaction terms are reported. Loan-to-value (LTV) represents five indicator variables for homeowners' LTV expressed as a percentage. The reference category of the *LTV* categories consists of workers who are tenants. The reference categories of *FEMALE*, *AGE* and *DUTCH NATIONALITY* consist of workers who are male, aged 21 to 30 years and non-Dutch, respectively. The reference categories of *TENURE* and *MANUFACTURING* consist of workers who are 3 to 6 years in the job and active in the service sector, respectively. The parameter estimates of the main and two-way interaction terms of the aforementioned independent variables are not reported. The regressions include individual-specific fixed effects and zero-one indicator variables for children aged 18 or lower, partner, the number of household members (3), the year of job displacement (4), the NUTS 3 location of the household (39) and calendar month (107). The estimates of the main, two-way interaction and three-way interaction terms of children aged 18 or lower, partner, number of household members and the year of job displacement are not reported. The main effects of the NUTS 3 location and calendar month are not reported. See Table 2.1 for additional notes.

Fig. 2.2. Housing state differential in the time-dependent displacement effects (Eq. (2.7)).



Notes: All four fixed effects regression models include 2,763 parameters. See Figure 2.1 for additional notes and Table 2.2 for additional statistics.

Figure 2.2A shows that displaced tenants and outright owners experience a higher loss in employment than other displaced homeowners. This difference is relatively stable over the post-displacement period. Figure 2.2B illustrates that displaced workers who have an LTV between 0 and 33 per cent experience a relatively low loss in wage. Figure 2.2C indicates that outright owners experience a relatively high increase in the commuting distance. We do not find significant differences in the displacement effect on changing home for workers who vary in housing state (see Figure 2.2D). All in all, the results indicate that the importance of workers' housing state for displacement effects on employment, wages and commute is relatively persistent over the period since displacement. The displacement effects by workers' housing state suggest that workers who are displaced choose among adjustment at the employment, wage and commuting margins. In this regard, we argue that these margins of adjustment, next to endogenous search (Morescalchi, 2016), are key to explaining the role of workers' housing state in employment outcomes.

2.6 Conclusion

The Great Recession that started in 2008 resulted in strong negative developments in the Dutch housing market and labour market. In Chapter 2, we have examined workers' resilience to job displacement by focusing on the use of margins of adjustment that are related to space and the importance of workers' housing state. We used Dutch administrative data, which were analysed with a quasi-experimental design involving job displacement that eliminates selection into labour turnover. Our conclusions are twofold.

First, we conclude that the spatial structure of homes and jobs reflects relevant margins of adjustment in response to job displacement. Our novel finding is that displaced workers experience an average increase of about 20 per cent in commuting distance and a decrease of about 14 per cent in changing home over the post-displacement period of thirty-six months. Interestingly, the patterns of adjustment change over the worker's post-displacement period – a longer time since job displacement is associated with a lower loss in employment, a smaller increase in commute, and a higher loss in hourly wage. We showed that the patterns in wage and commute are caused by displaced workers who take up their first job since displacement. The results suggest a remarkable pattern: workers who experience a longer time of unemployment since job displacement prefer a smaller increase in the commuting distance to a lower loss in hourly wage. The results indicate that the commuting distance is a more relevant margin of labour adjustment than changing home for workers who have been displaced. Hence, there are societal costs, in addition to losses in employment and wages, due to an increase in the commuting distance. Our findings are relevant for policies that aim to limit the impact of negative employment shocks.

Second, we conclude that workers' housing state is a substantial and persistent source of heterogeneity in the displacement effects on employment, wages and commute, but not on changing home. We find that displaced underwater homeowners, compared to other displaced owners and tenants, experience a lower loss in employment. Moreover, highly leveraged owners experience a relatively high loss in wage and become employed relatively close to home. In addition, outright owners experience the highest increase in the commuting distance. The results suggest that more leveraged workers have a stronger incentive to become employed. Moreover, it seems that the geographic immobility of displaced homeowners (relative to tenants) and underwater owners (relative to other owners) does not hinder post-displacement employment outcomes. The geographic immobility of Dutch workers could be relatively unimportant for the functioning of the labour market, as the surface area of the Netherlands is relatively small. Further research is needed to understand whether this observation also holds for other countries. Importantly, our findings are relevant for policies that subsidise homeownership or stimulate mortgage debt, as high homeownership and mortgage debt does not seem to hamper displaced workers' labour market outcomes.

All in all, we showed that workers' housing state affects the post-displacement outcomes in employment, wage and commuting distance, but not in changing home. Consistent with financial incentive structures, this evidence suggests that more leveraged displaced owners prioritise sooner re-employment over lower wage losses. In this regard, endogenous job search intensity and efficiency alone does not fully explain the difference in the unemployment duration of workers who vary by housing state. Based on the evidence, however, we cannot discuss the extent to which displaced workers make trade-offs between the use of margins of adjustment. Further research, based on a structural approach, is needed to better understand this issue.

Acknowledgements

Thanks to seminar participants at the 21st IZA Summer School in Labor Economics, Melbourne Institute of Applied Economic and Social Research, Free University Amsterdam Department of Spatial Economics, IAAEU Workshop on Labour Economics 2017, RES PhD Meetings 2017, MIT Sloan School of Management ESWG, 11th Meeting of the Urban Economics Association, 28th European Association of Labour Economists Conference, Urban Economic Association Doctoral Summer School, CPB Netherlands Bureau for Economic Policy Analysis and Utrecht University School of Economics. We wish to thank editor Stuart Rosenthal and two anonymous reviewers whose valuable comments improved the quality of Chapter 2. We also thank Bram van Besouw, Roberto Fernandez, Maarten

Goos, Thomas van Huizen, Zornitza Kambourova, Mark Kattenberg, Bastian Ravesteijn, Vincent Schippers and Ahmad-Reza Sharifi for their insightful comments. We are grateful to Statistics Netherlands for giving us access to the administrative data.

2.A Data Construction and Sample Selections

All individuals, firms and household addresses were uniquely identified on the basis of an encrypted Randomised Identification Number (RIN). We used the data set Bankruptcy Job Endings Register (*Failontslagtab*), which records the worker's RIN, the job's RIN and the date the firm entity is declared bankrupt for individuals who had a job at a firm where at least one entity of the firm experiences bankruptcy. Consequently, we possibly incorporated the so-called false-positives, i.e. we labelled voluntary job terminations in the bankrupted or non-bankrupted entity of a firm as a displacement due to firm bankruptcy. To limit the scope of false-positives we applied various data selections, which are discussed below.

Jobs that ended in year t or $t+1$ surrounding a bankruptcy of a firm entity were registered in the Bankruptcy Job Endings Register. The time span of year t to $t+1$ was chosen as jobs are recorded from firm payrolls that can continue after the verdict of bankruptcy. We selected workers for the group of displaced workers if the date of the job ending was earlier than one year after the date of bankruptcy or later than six months prior to the date of bankruptcy. This restriction ensured that the early leavers, who may have anticipated the plant closure, were incorporated in the analysis (see Schwerdt, 2011).

The Bankruptcy Job Endings Register was combined with multiple other registers. The Job Register (*Baankenmerkenbus*) was used to incorporate the date of job openings, the date of job endings and the RIN of the firm in which the worker was an employee. The Main Job Register (*Hfdbaanbus*) was used to distinguish between the main job and secondary job of an individual. The worker's main job, observed on a monthly basis, is the job with the highest wage. The Job and Wages Register (*Polisbus*) records monthly data based on income statements of employees to the tax office administration, including type of job (full-time or part-time), type of contract (fixed or temporary), number of hours worked and gross wage. The data set Work Location Register

(*Gemstplbus*) was used to incorporate data on the municipality in which the worker was employed.[22] The number of municipalities changed over the period under observation as various municipalities were aggregated. We used the set of 403 municipalities that existed in the year 2014. The Firm Register (*Betab*) was used to incorporate annual data on firm size and firm sector. Firm sectors were classified in 21 sectors according to a five-digit code (in Dutch: Standaard Bedrijfs Indeling), which is based on the International Standard Industrial Classification of All Economic Activities (ISIC). The extraterritorial organisations and bodies sector was excluded as no displaced worker was employed in this sector.

Registers that are based on municipal and tax office administration were used to incorporate personal, home and household information. The Population Register (*Gbapersoontab*, *Gbahuishoudensbus*, *Gbaburgerlijkestaatbus*, *Gbaadresgebeurtenisbus*) contains monthly data on the date of birth, gender, marital status, number of household members and changing home. The Address Object Register (*Gbaadresobjectbus, Vslgwbtab*) contains data on individuals' house address and home location at the neighbourhood level. The Highest Education Register (*Hoogsteopltab*) was used and contains data on workers' highest level of attained education. Education is categorised in lower, secondary and tertiary education based on the International Standard Classification of Education (ISCED). The Integrated Household Income data set (*Integraal Huishoudens Inkomen*), which is based on data measured on the 31st of December retrieved from the tax office, was used to incorporate data on housing state and household income. In the case of changing home, data on housing state was used from the year prior to that of changing. As an example, for household relocations to different homes in 2006 we used data from 2005. The Integrated Capital data set (*Integraal-Vermogensbestand*), which

[22] Work location is not complete nor consistent as the CBS has only data that is measured in December on the number of plants of a firm, the location of each firm plant and the number of employees at each specific plant. Work location is imputed by the CBS using data on the location of the workplace and residential home. Each resident is linked to the closest plant of a firm, conditional on not exceeding the number of workers employed at that specific plant. Hence, we do not observe the work location of workers if they were not in the job in December. We assessed the consequences of the incompleteness and inconsistency of the variable work location by applying two robustness checks. First, we excluded all workers whose firm location is not completely observed for all jobs in the period 2006 to 2014 (see Table 2.D.1). Second, we ran a robustness check with firms that consist of 49 employees at maximum to ensure a low number of firm plants (see Table 2.D.2). We find similar results.

consists of annual records from the tax office measured on the 1st of January, was used to incorporate data on the mortgage debt, non-housing wealth and property value of the home. In the case of changing home during the year, data on the mortgage debt and property value were used from the year after the home change. As an example, for household relocations in 2014 we used data from the year 2015.

The following selections were made to attain an appropriate sample for our analysis. To keep the employment history of a worker tractable, all job spells that were not identified as the main job were excluded. Moreover, we excluded groups of individuals for various reasons. First, we excluded all individuals who were not active in the labour market (e.g., disabled individuals, students and early retirees), who had no administered employment history (e.g., self-employed and long-term unemployed), or were aged below 21 or over 60 years. Second, our data do not distinguish between a bankrupt or restarted firm. Hence, we excluded workers from the group of displaced workers if more than 40 per cent of the displaced workers who were employed at the bankrupt firm became re-employed at another but identical firm. Third, all workers who had ever an LTV over 133 per cent during the period under observation were excluded from the sample, as a higher LTV suggests an administrative error. Finally, all workers with three or more home changes in one calendar year were excluded from the sample, as this would create the problem where we could not observe data on all homes. We kept individuals that experienced two home changes in one year, as on many occasions households change home to one temporary accommodation following the sale of their home.

Prior to Coarsened Exact Matching, individuals were excluded from the group of displaced or non-displaced for various reasons. First, we excluded all workers whose hourly wage or housing state was not completely observed for all jobs and homes in the period 2006 to 2014. In multiple cases this was possible, as we did not perfectly link all the information of the Job Register to the Job and Wages Register and the Housing registers. In addition, we excluded individuals whose hourly wage is equal to or lower than one euro. Second, we excluded all workers with an employment spell shorter than three years. An employment spell of at least three years allows us to incorporate workers who had a stable job and who experienced an unexpected and involuntary job displacement. Thereby, the likelihood of false-positives was reduced. Third, we excluded all workers who, in the month prior to job displacement, worked at a firm with less than ten employees or who worked less than 64 hours in that month. Finally, we had to randomly exclude about 70 per cent of the non-displaced workers (controls) due to computational limitations.

After the process of matching, if the displaced or non-displaced worker of a matched pair was not under observation for the entire period of fifty-five months, the matched pair was excluded. The matched pairs were excluded as the incomplete data imply data gaps due to, e.g., immigration, emigration or death. This selection ensured a weakly balanced sample. Only 35 individual workers, which represents about 0.23 per cent of the full sample, were not observed for the entire period. In total, 123 workers were excluded.

2.B Robustness Checks for the Empirical Models on Worker Characteristics

As discussed in Section Identification Strategy, we created a new matched sample for each of the three robustness checks. The matched samples for each robustness check were created based on a different set of matching variables, which included indicator variables for the non-housing wealth position and duration of home occupancy, LTI ratio, and skill level, respectively.

First, we assess whether the estimates of the interaction terms between job displacement and LTV are robust to the inclusion of approximations of non-housing wealth and duration of home occupancy. The non-housing wealth position of the displaced worker can be of importance for post-displacement outcomes, as it can aid job search through increased mobility or deter job search through decreased job search activity (Henley et al., 1994; Goss and Phillips, 1997). The duration of home occupancy approximates the willingness to change home and is, consequently, an important driver behind the ability to become employed in a distant labour market. Moreover, we capture further unobserved heterogeneity in human capital by controlling for non-housing wealth and duration of home occupancy. Human capital is expected to be positively correlated to non-housing wealth and negatively correlated to duration of home occupancy, as high-skilled workers earn a relatively high income and are characterised by a relatively high geographical mobility (e.g., Bowles, 1970).

The non-housing wealth is represented by five zero-one indicator variables that equal one for non-housing household wealth below 0 (base category), between 0-5,000; 5,000-25,000; 25,000-75,000 and over 75,000 euro, respectively. The duration of home occupancy is represented by three zero-one indicator variables that equal one if the period in the home equals 0-60 (base category), 60-180 and over 180 months, respectively.

Table 2.B.1 shows the role of workers' housing state, non-housing wealth and duration of home occupancy in the displacement effects on the margins of adjustment. The estimates indicate that workers who have positive non-housing wealth experience a relatively low loss in employment. The results suggest that the loss in employment is decreasing and the loss in wage is increasing in workers' LTV. This result supports the results of the model in which we do not control for non-housing wealth and duration of home occupancy (See Table 2.2). Note that by simultaneously controlling for housing state, non-housing wealth and duration of home occupancy, the fixed effects coefficients of the two highest LTV categories on hourly wage become slightly higher. Also, the role of housing state in the displacement effect on commuting distance is more pronounced. Compared with displaced underwater owners, displaced outright owners experience an increase in the commuting distance of about 7 kilometres higher.

Table 2.B.1
The role of workers' housing state, non-housing wealth and duration of home occupancy in the displacement effects (Eq. (2.6)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
$DISPLACED \times POST \times$				
Housing state:				
$LTV = 0\%$	0.0249	-0.0033	4.9238**	-0.0006
	(0.0233)	(0.0141)	(1.9547)	(0.0018)
$0 < LTV \leq 33\%$	0.0195	0.0134	0.4676	0.0006
	(0.0133)	(0.0088)	(1.1557)	(0.0009)
$33 < LTV \leq 66\%$	0.0233**	-0.0192**	-1.2478	0.0002
	(0.0118)	(0.0079)	(1.0096)	(0.0008)
$66 < LTV \leq 100\%$	0.0380***	-0.0312***	-0.125	0.0001
	(0.0117)	(0.0078)	(1.0390)	(0.0009)
$100 < LTV \leq 133\%$	0.0836***	-0.0320***	-1.795	-0.0001
	(0.0145)	(0.0115)	(1.3998)	(0.0016)
Joint significance of LTV	6.93***	8.14***	2.63**	0.28
Non-housing wealth:				
$0 < WEALTH \leq 5,000$ euro	0.0482**	0.0146	-1.9366	-0.0040*
	(0.0192)	(0.0135)	(1.4872)	(0.0021)
$5,000 < WEALTH \leq 25,000$ euro	0.0540***	0.0242*	-2.4711*	-0.0025
	(0.0195)	(0.0136)	(1.4941)	(0.0020)
$25,000 < WEALTH \leq 75,000$ euro	0.0639***	0.0192	-2.1289	-0.0027
	(0.0204)	(0.0144)	(1.6134)	(0.0021)
$WEALTH > 75,000$ euro	0.0319	0.0004	-3.3393**	-0.002
	(0.0213)	(0.0155)	(1.6823)	(0.0021)
Joint significance of $WEALTH$	3.99***	2.87**	1.11	1.71
Duration of home occupancy:				
$60 < DURATION \leq 180$ months	-0.0115	-0.0092	1.3681	0.0003
	(0.0098)	(0.0068)	(0.8824)	(0.0011)
$DURATION > 180$ months	0.0362**	-0.0165*	0.8031	-0.0003
	(0.0151)	(0.0097)	(1.1801)	(0.0014)
Joint significance of $DURATION$	7.91***	1.51	1.31	0.37
Number of parameters	243	243	243	243
Number of individuals	24,950	24,950	24,950	24,950
Number of observations	1,372,250	1,246,029	1,230,402	1,372,250

Notes: The regression analyses include, besides the covariates, multiple three-way interaction terms. Three-way interaction terms are included among the variables $DISPLACED$, $POST$ and LTV, among the variables $DISPLACED$, $POST$ and $WEALTH$ position, among the variables $DISPLACED$, $POST$ and $DURATION$ and among the variables $DISPLACED$, $POST$ and all other covariates. LTV represents five indicator variables for homeowners' LTV expressed as a percentage. The reference category of each LTV indicator consists of workers who are tenants. The reference category of $WEALTH$ consists of workers who have negative non-housing wealth. The reference category of $DURATION$ consists of workers who live between zero and sixty months in their home. See Table 2.2 for additional notes.

Second, we assess whether our results are robust to the inclusion of the LTI ratio as an additional approximation of the financial incentive to work. The LTI ratio is constructed by the mortgage loan of year t relative to the gross household income of year t-1, where year t is the year of job displacement. The LTI ratio is time-invariant, to prevent the situation where a large share of variation in the LTI ratio is caused by changes in household income following job displacement. The LTI ratio is operated as five zero-one indicator variables, which equal one if the LTI ratio ranges between 0-1.5 (base category), 1.5-3.0, 3.0-4.5, 4.5-6.0 and over 6.0, respectively.

Table 2.B.2 shows the role of workers' housing state and LTI ratio in the displacement effects. We find that the fixed effects coefficients of the LTV on wage become higher if we include variables that represent the LTI ratio. However, the fixed effects coefficients of the LTV on employment become smaller and less significant. Compared with displaced workers who have an LTI ratio between 0 to 1.5, displaced workers with an LTI ratio over 6.0 experience a lower loss in the hourly wage. Note, however, that the results in Table 2.B.2 are relatively sensitive, as the categorical variables LTV and LTI ratio are highly correlated, i.e. a correlation equal to 0.88.

Third, we assess whether our results are robust to the inclusion of the worker's skill level (see Tables 2.B.3 and 2.B.4). The impact of skill on the post-displacement losses is theoretically ambiguous. On the one hand, high-skilled workers have a higher job offer arrival rate. The higher job offer arrival rate is driven by the higher willingness to commute and relocate (Zax, 1991). Consequently, the distribution of job offers is increasing in the skill level of the displaced worker. Moreover, the market power of employers is decreasing in the skill level of the displaced worker. Hence, the losses in employment and wages for high-skilled displaced workers are likely to be relatively low. On the other hand, high-skilled workers have a relatively high wage premium due to their firm-specific human capital. If high-skilled workers invested more in human capital than low-skilled workers, the displacement effect on wage would be higher for high-skilled workers. Hijzen et al. (2010) show that high skilled workers have higher initial losses in wage than unskilled workers, but two years after job displacement the skill difference in wage losses becomes statistically insignificant. Farber (2017) shows that a higher number of years in education decreases the losses in employment and earnings.

The data we used on skill level is based on the international standard

Table 2.B.2
The role of workers' housing state and loan-to-income ratio in the effects of job displacement (Eq. (2.6)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*×				
Housing state:				
$LTV = 0\%$	0.0182	-0.0157	1.2188	-0.0015
	(0.0171)	(0.0098)	(1.4024)	(0.0009)
$0 < LTV \leq 33\%$	0.0462***	-0.0058	-0.8464	-0.0005
	(0.0109)	(0.0070)	(0.9083)	(0.0005)
$33 < LTV \leq 66\%$	0.0278*	-0.0313***	0.6234	-0.0001
	(0.0168)	(0.0109)	(1.4444)	(0.0007)
$66 < LTV \leq 100\%$	0.0181	-0.0499***	1.4075	-0.0011
	(0.0192)	(0.0128)	(1.7113)	(0.0009)
$100 < LTV \leq 133\%$	0.0318	-0.0436***	1.1386	-0.0009
	(0.0218)	(0.0150)	(2.0181)	(0.0014)
Joint significance of *LTV*	4.32***	3.85***	0.80	1.19
Loan-to-income:				
$1.5 < LTI \leq 3.0$	0.0096	-0.0013	-0.9248	-0.0003
	(0.0140)	(0.0089)	(1.1891)	(0.0005)
$3.0 < LTI \leq 4.5$	0.0078	0.017	-2.8291*	0.0008
	(0.0173)	(0.0115)	(1.5171)	(0.0007)
$4.5 < LTI \leq 6.0$	0.028	0.0318**	-2.0954	0.0005
	(0.0193)	(0.0128)	(1.7196)	(0.0009)
$LTI > 6.0$	0.0243	0.0437***	-1.9796	0.0011
	(0.0202)	(0.0136)	(1.8190)	(0.0011)
Joint significance of *LTI*	0.99	4.62***	1.05	1.36
Number of parameters	218	218	218	218
Number of individuals	41,359	41,359	41,359	41,359
Number of observations	2,274,745	2,065,937	2,040,254	2,274,745

Notes: Parameter estimates of the three-way interaction terms among *DISPLACED*, *POST* and *LTI*, and among *DISPLACED*, *POST* and *LTI* are displayed. The reference categories of *LTV* and *LTI* consist of workers who are tenants and who have an LTI ratio between 0 and 1.5, respectively. The parameter estimates of the covariates and the two-way interaction terms are not reported. See Table 2.2 for additional notes.

classification of education 1997, and is represented by three variables that equal one if the skill level is low (base category), medium, and high,

Table 2.B.3
The role of workers' housing state in the effects of job displacement, sample of Table 2.B.4 (Eq. (2.6)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*×				
Housing state:				
$LTV = 0\%$	-0.0092	0.0187	2.1841	0.0013
	(0.0308)	(0.0191)	(2.8782)	(0.0033)
$0 < LTV \leq 33\%$	0.0137	0.0136	0.6967	0.0018
	(0.0180)	(0.0117)	(1.4101)	(0.0014)
$33 < LTV \leq 66\%$	0.0007	-0.0016	-1.3845	0.0028**
	(0.0156)	(0.0104)	(1.4063)	(0.0013)
$66 < LTV \leq 100\%$	0.0068	-0.0211**	0.5587	0.0011
	(0.0147)	(0.0096)	(1.3440)	(0.0013)
$100 < LTV \leq 133\%$	0.0684***	0.0060	-1.2052	0.0016
	(0.0169)	(0.0121)	(1.6898)	(0.0018)
Joint significance of *LTV*	4.39***	2.65**	0.84	1.23
Number of parameters	219	219	219	219
Number of individuals	14,028	14,028	14,028	14,028
Number of observations	771,54	686,218	674,767	771,54

Notes: The regression analyses include, besides the covariates, multiple three-way interaction terms. Three-way interaction terms are included among the variables *DISPLACED*, *POST* and *LTV* and among the variables *DISPLACED*, *POST* and all other covariates. See Table 2.2 for additional notes.

respectively. We only incorporated the skill level in a robustness check, because the education data is highly selective as it is only available for individuals who received their diploma after the year 1995.

The fixed effects coefficients of LTV on employment that are shown in Table 2.B.3 and Table 2.B.4 are based on the sample in which we matched on the default set of matching variables and workers' skill level. The coefficients of the three-way interaction terms including the LTV without controlling for workers' skill level (see Table 2.B.3) are similar to the coefficients of the model including covariates for workers' skill level (see Table 2.B.4). Hence, we argue that the results are robust to the inclusion of variables that represent the skill level.

Note, however, that compared to the parameter estimates provided in

Table 2.B.4
The role of workers' housing state and skill level in the effects of job displacement (Eq. (2.6)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*×				
Housing state:				
$LTV = 0\%$	-0.0105	0.0185	1.4114	0.0011
	(0.0309)	(0.0190)	(2.9060)	(0.0033)
$0 < LTV \leq 33\%$	0.0120	0.0154	0.2910	0.0017
	(0.0181)	(0.0118)	(1.4346)	(0.0014)
$33 < LTV \leq 66\%$	-0.0007	-0.0006	-1.9666	0.0027**
	(0.0157)	(0.0105)	(1.4490)	(0.0013)
$66 < LTV \leq 100\%$	0.0090	-0.0207**	-0.4087	0.0009
	(0.0150)	(0.0099)	(1.3951)	(0.0014)
$100 < LTV \leq 133\%$	0.0733***	0.0042	-2.3556	0.0014
	(0.0172)	(0.0123)	(1.7555)	(0.0019)
Joint significance of *LTV*	4.90***	2.61**	1.01	1.14
Skill:				
MEDIUM SKILL	0.0175	-0.0087	1.7206	0.0004
	(0.0126)	(0.0081)	(1.1172)	(0.0009)
HIGH SKILL	-0.0165	0.0058	4.7758***	0.0009
	(0.0176)	(0.0115)	(1.7802)	(0.0013)
Joint significance of *SKILL*	3.45**	1.52	3.60**	0.22
Number of parameters	223	223	223	223
Number of individuals	14,028	14,028	14,028	14,028
Number of observations	771,54	686,218	674,767	771,54

Notes: The regression analyses include, besides the covariates, multiple three-way interaction terms. Three-way interaction terms are included among the variables *DISPLACED*, *POST* and *LTV*, among the variables *DISPLACED*, *POST* and *SKILL* and among the variables *DISPLACED*, *POST* and all other covariates. The reference category of *SKILL* consists of workers who attained lower education. See Table 2.2 for additional notes.

Table 2.2 of the model in which we do not match and control for the skill level, the coefficients shown in Table 2.B.3 and Table 2.B.4 are different. Specifically, by matching on and controlling for housing state and skill level, the coefficients of the two highest LTV categories on hourly wage become smaller. Hence, Table 2.B.3 and Table 2.B.4 suggest that the difference in results is completely driven by the difference in sample, which can be explained by the fact that the education data is highly selective.

Table 2.B.5
The role of homeownership and a continuous loan-to-value ratio in the effects of job displacement (Eq. (2.6)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*×				
Housing state:				
HOMEOWNERSHIP	0.0193**	0.0048	-0.0153	0.0005
	(0.0088)	(0.0057)	(0.7604)	(0.0006)
CONTINUOUS LTV	0.0002**	-0.0003***	-0.0044	0
	(0.0001)	(0.0001)	(0.0089)	(0)
Number of parameters	208	208	208	208
Number of individuals	54,229	54,229	54,229	54,229
Number of observations	2,982,594	2,719,569	2,686,297	2,982,594

Notes: The regression analyses include, besides the covariates, multiple three-way interaction terms. Three-way interaction terms are included among the variables *DISPLACED*, *POST* and *HOMEOWNERSHIP* and among the variables *DISPLACED*, *POST* and *CONTINUOUS LTV* and all other covariates. The reference category of *HOMEOWNERSHIP* consists of workers who are tenants. See Table 2.2 for additional notes.

2.C Summary Statistics

Table 2.C.1 provides multiple statistics that improve our understanding of the within change in hourly wage and commuting distance for the displaced and non-displaced. The within change is calculated by taking the difference between the values of each variable eighteen months after job displacement and the month of potential or actual job displacement.

The displaced are characterised by a decrease in the hourly wage and the non-displaced by an increase in hourly wage. Half of all displaced workers experience no or a modest decline in commuting distance. For the displaced, the within hourly wage change follows a distribution with a long tail to the left. For the non-displaced, the within hourly wage change follows a distribution with a long tail to the right.

Half of all displaced workers experience a sharp increase in commuting distance after job displacement. The mean within change in the commuting distance for the non-displaced is close to zero. Only the bottom and top 5 per cent experience a relatively small decrease and increase, respectively. The within commuting change has a substantial skewness and follows an asymmetrical distribution with a long tail to the right, especially for the displaced.

Table 2.C.1
The within change in hourly wage and commuting distance for displaced and non-displaced workers.

	Hourly wage (log)		Commuting distance (km)	
	Displaced	Non-displaced	Displaced	Non-displaced
Mean	-0.0245	0.0414	3.6006	0.2691
St. Dev.	0.3116	0.1786	32.3825	14.8189
Variance	0.0971	0.0319	1,048.6236	219.6010
Skewness	-0.9973	1.1729	1.1568	0.2277
Kurtosis	31.2475	55.6833	14.4843	54.4108
1th percentile	-0.9594	-0.4967	-95.3368	-49.6782
5th percentile	-0.4577	-0.1572	-38.1269	-6.6214
25th percentile	-0.1247	-0.0010	-1.8662	0
50th percentile	0.0018	0.0338	0	0
75th percentile	0.0979	0.0871	8.9507	0
95th percentile	0.3351	0.2462	51.5134	9.3275
99th percentile	0.7473	0.5854	129.5867	54.6074
Number of observations	15,196	32,693	14,551	32,483

Notes: The individual summary statistics are based on the within change, measured by the difference in the values of each variable between the eighteenth month after job displacement and the month of job displacement.

Table 2.C.2

Individual summary statistics using the non-matched sample.

	Non-displaced		Displaced		
	Mean	St. Dev.	Mean	St. Dev	t-statistic
Employment (=1)	1	0	1	0	
Hourly wage (log)	2.8579	0.3796	2.7951	0.4176	29.48***
Hourly wage (€)	18.8743	11.261	18.6003	41.4605	4.31***
Commuting distance (km)	14.8073	21.1692	17.5219	24.8161	-22.88***
Home change (=1)	0.004	0.0632	0.0044	0.0661	-1.07
LTV (%)	59.5615	32.8507	60.6925	33.4721	-5.23***
LTI ratio	2.9782	28.1622	2.8718	3.4081	0.67
Mortgage debt (€)	120,461	115,406	111,089	113,978	14.49***
Property value (€)	215,537	222,552	197,057	284,727	14.81***
Non-housing wealth (€)	66,307	277,218	55,221	219,700	7.14***
Annual household income (€)	44,152	21,141	42,026	22,134	17.95***
Age (in years)	43.8438	8.5612	43.0404	8.7444	16.74***
Female (=1)	0.4381	0.4962	0.2692	0.4435	60.77***
Dutch (=1)	0.9116	0.2839	0.9022	0.2971	5.90***
Partner (=1)	0.6223	0.4848	0.5971	0.4905	9.30***
No child (=1)	0.5192	0.4996	0.5316	0.499	-4.42***
Household members (#)	3.0033	1.3089	2.9828	1.3043	2.81***
Fixed contract (=1)	0.9539	0.2097	0.9166	0.2766	31.77***
Full-time job (=1)	0.6039	0.4891	0.7215	0.4483	-42.92***
Tenure in the job (in months)	148.9688	94.051	132.5316	87.3091	31.19***
Manufacturing sector (=1)	0.2497	0.4328	0.4759	0.4994	-93.24***
Duration of home occupancy (in months)	106.8996	58.4576	108.0195	59.7186	-3.42***
Number of individuals (#)	24,043,844		31,888		

Notes: The individual summary statistics, provided for the month of actual or potential displacement, are based on the sample prior to matching. The time period under observation is from July 2007 to December 2011. Sample means with standard deviations are provided for the treatment group and control group. The t-statistic is provided to assess whether the mean and standard deviation of each variable for the groups of displaced and non-displaced workers are statistically different from each other. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. By construction, all displaced and non-displaced were employed in the month of actual or potential displacement.

Table 2.C.3
Individual summary statistics using the matched sample.

	Non-displaced		Displaced		
	Mean	St. Dev.	Mean	St. Dev	t-statistic
Employment (=1)	1	0	1	0	
Hourly wage (log)	2.8426	0.3686	2.8314	0.4101	3.30***
Hourly wage (€)	18.4981	8.8651	19.1812	38.6269	-3.13***
Commuting distance (km)	15.0456	21.5756	17.2607	24.4967	-10.98***
Home change (=1)	0.0043	0.0651	0.0036	0.0597	1.22
LTV (%)	59.4587	31.7348	59.7254	31.9765	-0.81
LTI ratio	2.8357	2.9655	2.8737	3.3361	-1.38
Mortgage debt (€)	115,651	110,932	113,492	111,257	2.19**
Property value (€)	208,041	201,155	205,066	334,756	1.29
Non-housing wealth (€)	59,137	197,274	53,153	181,760	3.51***
Annual household income (€)	43,279	20,449	42,812	22,963	2.45**
Age (in years)	43.4753	8.7785	43.537	8.7269	-0.79
Female (=1)	0.2093	0.4069	0.2137	0.4099	-1.19
Dutch (=1)	0.9525	0.2127	0.9457	0.2267	3.54***
Partner (=1)	0.6439	0.4789	0.6369	0.4809	1.63
No child (=1)	0.513	0.4998	0.5227	0.4995	-2.17**
Household members (#)	3.0727	1.3076	3.0375	1.2991	3.03***
Fixed contract (=1)	0.979	0.1432	0.9719	0.1652	5.29***
Full-time job (=1)	0.7946	0.404	0.784	0.4115	2.93***
Tenure in the job (in months)	143.0305	90.6646	142.8202	91.7574	0.26
Manufacturing sector (=1)	0.5266	0.4993	0.5209	0.4996	1.28
Duration of home occupancy (in months)	110.0975	59.4983	111.1839	59.5814	-2.05**
Number of individuals (#)	34,077		20,152		

Notes: The individual summary statistics, provided for the month of actual or potential displacement, are based on the sample after matching. The time period under observation is from July 2007 to December 2011. Sample means with standard deviations are provided for the treatment group and control group. The t-statistic is provided to assess whether the mean and standard deviation of each variable for the groups of displaced and non-displaced workers are statistically different from each other. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. By construction, all displaced and non-displaced were employed in the month of actual or potential displacement.

Table 2.C.4
The number of coarsened exactly matched individuals.

	Number of individuals		
	Displaced	Non-displaced	All
	(1)	(2)	(3)
Housing state:			
Tenant	5,256	8,804	14,060
$LTV = 0\%$	537	812	1,349
$0 < LTV \leq 33\%$	2,959	5,124	8,083
$33 < LTV \leq 66\%$	4,843	8,347	13,190
$66 < LTV \leq 100\%$	4,796	8,193	12,989
$100 < LTV \leq 133\%$	1,761	2,797	4,558
Total	20,152	34,077	54,229

Notes: The number of matched individuals is provided for each housing state and treatment group.

Table 2.C.5
Firm summary statistics in the month of job displacement.

	Firms			
	Bankrupt firms		Non-bankrupt firms	
	Mean	St. Dev.	Mean	St. Dev.
Firm size:				
1-9 employees (=1)	0	0	0	0
10-49 employees (=1)	0.7247	0.4467	0.6258	0.4839
50-99 employees (=1)	0.1114	0.3147	0.1400	0.3470
100-499 employees (=1)	0.1025	0.3033	0.1702	0.3758
500 or more employees (=1)	0.0614	0.2401	0.0640	0.2447
Firm sector:				
Agriculture, forestry and fishing (=1)	0.0099	0.0991	0.0042	0.0644
Mining and quarrying (=1)	0	0	0	0
Manufacturing (=1)	0.2415	0.4280	0.3474	0.4762
Electricity, gas, steam and air conditioning supply (=1)	0	0	0	0
Water supply; sewerage, waste management and remediation activities (=1)	0.0010	0.0311	0.0002	0.0136
Construction (=1)	0.1951	0.3963	0.2010	0.4008
Wholesale and retail trade; repair of motor vehicles and motorcycles (=1)	0.1999	0.4000	0.2207	0.4147
Transportation and storage (=1)	0.0691	0.2537	0.0515	0.2209
Accommodation and food service activities (=1)	0.0121	0.1093	0.0041	0.0636
Information and communication (=1)	0.0372	0.1893	0.0163	0.1268
Financial and insurance activities (=1)	0.0321	0.1764	0.0274	0.1634
Real estate activities (=1)	0.0041	0.0640	0.0012	0.0343
Professional, scientific and technical activities (=1)	0.0851	0.2790	0.0542	0.2265
Administrative and support service activities (=1)	0.0645	0.2457	0.0315	0.1746
Public administration and defence; compulsory social security (=1)	0	0	0	0
Education (=1)	0.0065	0.0805	0.0052	0.0717
Human health and social work activities (=1)	0.0317	0.1751	0.0314	0.1743
Arts, entertainment and recreation (=1)	0.0034	0.0581	0.0011	0.0338
Other service activities (=1)	0.0068	0.0820	0.0027	0.0518
Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use (=1)	0	0	0	0
Activities of extraterritorial organisations and bodies (=1)	0	0	0	0
Number of firms (#)	4,137		27,109	

Notes: Means and standard deviations are provided at the firm level. The time period under observation is from July 2007 to December 2011. Bankrupts firms consist of all distinct firms of which an entity is declared bankrupt and a worker is displaced in the month of actual displacement. Non-bankrupt firms consist of all distinct firms where matched non-displaced workers work in the month of potential displacement.

Table 2.C.6

Individual summary statistics for each housing state using the matched sample.

	Housing state					
		LTV				
	Tenant	0%	0-33%	33-66%	66-100%	100-133%
Employment (=1)	1	1	1	1	1	1
Hourly wage (log)	2.6505	2.8474	2.8877	2.9223	2.9296	2.8711
Hourly wage (€)	14.8504	18.4154	19.1421	20.0920	20.2513	18.9327
Commuting distance (km)	12.2651	13.3215	13.9457	16.0412	17.1179	17.2717
Home change (=1)	0.0081	0.0049	0.0023	0.0016	0.0032	0.0068
LTV (%)	/	0	19.8261	49.5076	83.0918	109.7967
LTI ratio	0.0399	0	1.4240	3.3449	5.3308	6.2244
Mortgage debt (€)	1,467	0	62,635	140,532	215,205	239,894
Property value (€)	/	328,333	330,791	285,318	260,397	219,108
Non-Housing Wealth (€)	23,576	219,632	94,881	66,394	56,354	45,496
Annual household income (€)	34,575	52,589	48,824	46,945	44,78	42,478
Age (in years)	42.3176	46.4894	48.4207	46.1283	41.1604	36.0476
Female (=1)	0.2299	0.1268	0.1501	0.2155	0.2212	0.2242
Dutch (=1)	0.8772	0.9938	0.9971	0.9835	0.9663	0.9632
Partner (=1)	0.4048	0.6170	0.8300	0.7944	0.6687	0.5409
No child (=1)	0.6984	0.6995	0.5246	0.4222	0.3969	0.4655
Household members (#)	2.5125	3.1293	3.4461	3.3830	3.1767	2.9042
Fixed contract (=1)	0.9626	0.9840	0.9865	0.9865	0.9856	0.9743
Full-time job (=1)	0.8128	0.8399	0.7982	0.7657	0.7873	0.8259
Tenure in the job (in months)	126.0956	175.5739	174.0796	162.0679	132.7751	103.2345
Manufacturing sector (=1)	0.5034	0.7020	0.6472	0.5401	0.4748	0.4394
Duration of home occupancy (in months)	101.7157	144.5369	149.8603	127.7999	89.5446	61.0132
Number of observations (#)	5,256	537	2,959	4,843	4,796	1,761

Notes: Sample means, based on the sample after matching, are provided for each housing state of the treatment group in the month of actual displacement. The time period under observation is from July 2007 to December 2011. By construction, all displaced and non-displaced are employed in the month of actual or potential displacement. The LTV and property value is observed if the worker is homeowner and not if the worker is tenant. Tenants can have a mortgage debt if they owned a home prior to their current rented home.

2.D Robustness Checks for the Empirical Models on Margins of Adjustment

Table 2.D.1
Displacement effects for sample of workers with full information on the firm location (Eq. (2.4)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*	−0.2577***	−0.0512***	3.0087***	−0.0006***
	(0.0033)	(0.0020)	(0.2824)	(0.0002)
Number of parameters	161	161	161	161
Number of individuals	36,530	36,530	36,530	36,530
Number of observations	2,009,150	1,839,235	1,839,235	2,009,150

Notes: Each column gives the dependent variable. Clustered (by individual) standard errors are in parentheses. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. The reference category of *DISPLACED* and *POST*, consists of the non-displaced workers and pre-displacement period, respectively. The regression analyses include individual-specific fixed effects and indicator variables for housing state (5), age (3), children aged 18 or lower, partner, the number of household members (3), the NUTS 3 location of the household (39) and calendar month (107). The period under observation is from January 2006 to December 2014. The parameter estimates of the covariates and the main effect of *POST* are not reported.

Table 2.D.2
Displacement effects for sample with firms that have 49 employees at maximum (Eq. (2.4)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*	−0.2338***	−0.0515***	3.9953***	−0.0007***
	(0.0035)	(0.0023)	(0.2860)	(0.0002)
Number of parameters	161	161	161	161
Number of individuals	27,375	27,375	27,375	27,375
Number of observations	1,505,625	1,375,974	1,358,117	1,505,625

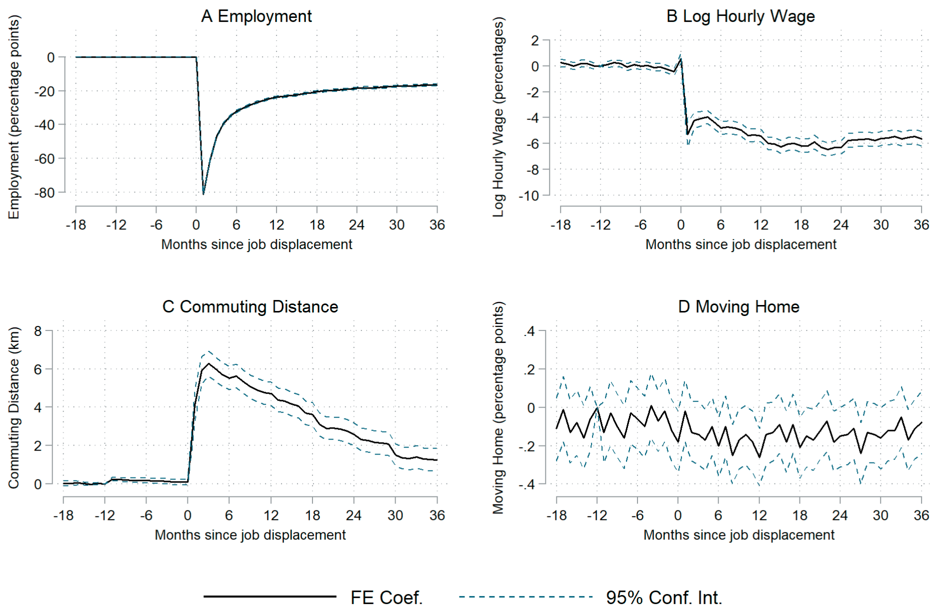
Notes: Each column gives the dependent variable. Clustered (by individual) standard errors are in parentheses. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. The reference category of *DISPLACED* and *POST*, consists of the non-displaced workers and pre-displacement period, respectively. The regression analyses include individual-specific fixed effects and indicator variables for housing state (5), age (3), children aged 18 or lower, partner, the number of household members (3), the NUTS 3 location of the household (39) and calendar month (107). The period under observation is from January 2006 to December 2014. The parameter estimates of the covariates and the main effect of *POST* are not reported.

Table 2.D.3
Displacement effects for sample of workers who are matched on forty NUTS 3 areas (Eq. (2.4)).

	Employment (=1)	Hourly wage (log)	Commuting distance (km)	Home change (=1)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*	-0.2526^{***}	-0.0649^{***}	3.0940^{***}	-0.0003
	(0.0031)	(0.0020)	(0.2790)	(0.0002)
Number of parameters	161	161	161	161
Number of individuals	35,756	35,756	35,756	35,756
Number of observations	1,966,580	1,782,424	1,760,123	1,966,580

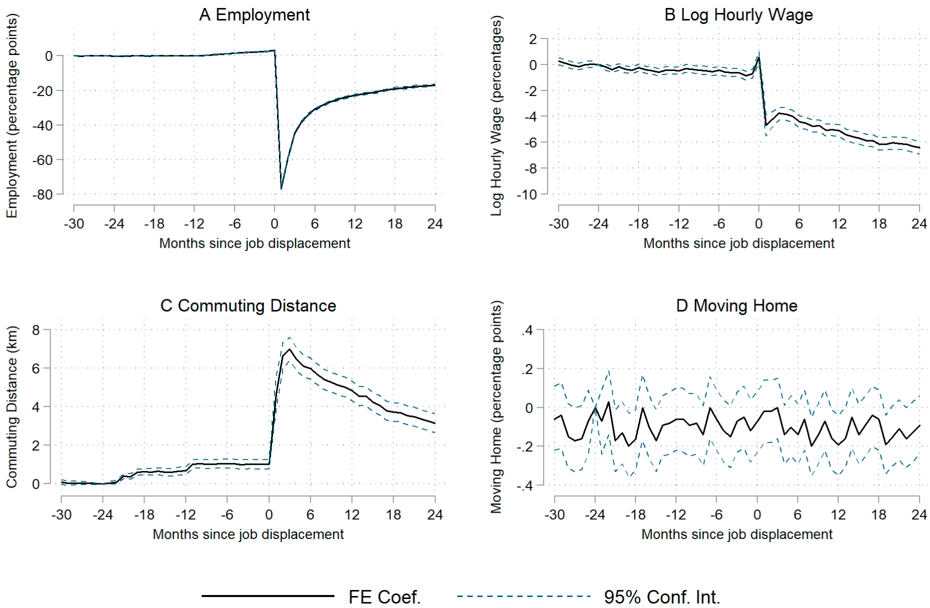
Notes: Each column gives the dependent variable. Clustered (by individual) standard errors are in parentheses. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. The reference category of *DISPLACED* and *POST*, consists of the non-displaced workers and pre-displacement period, respectively. The regression analyses include individual-specific fixed effects and indicator variables for housing state (5), age (3), children aged 18 or lower, partner, the number of household members (3), the NUTS 3 location of the household (39) and calendar month (107). The period under observation is from January 2006 to December 2014. The parameter estimates of the covariates and the main effect of *POST* are not reported.

Fig. 2.D.1. Displacement effects conditional on being in the first post-displacement job and not changed home (Eq. (2.5)).



Notes: Displacement effects on employment (A), log hourly wage (B), commuting distance (C) and changing home (D). For B and C, the post-displacement observations are included conditional on being in the first post-displacement job and not changed home. Reference group is the group of non-displaced workers. Reference month is the twelfth month prior to job displacement. The 95% confidence intervals are computed using standard errors clustered by individual. All four fixed effects regression models include 266 parameters of which there are 54 two-way interaction terms. See Table 2.1 for additional notes and statistics.

Fig. 2.D.2. Placebo treatment on employment (A), log hourly wage (B), commuting distance (C) and changing home (D) (Eq. (2.5)).



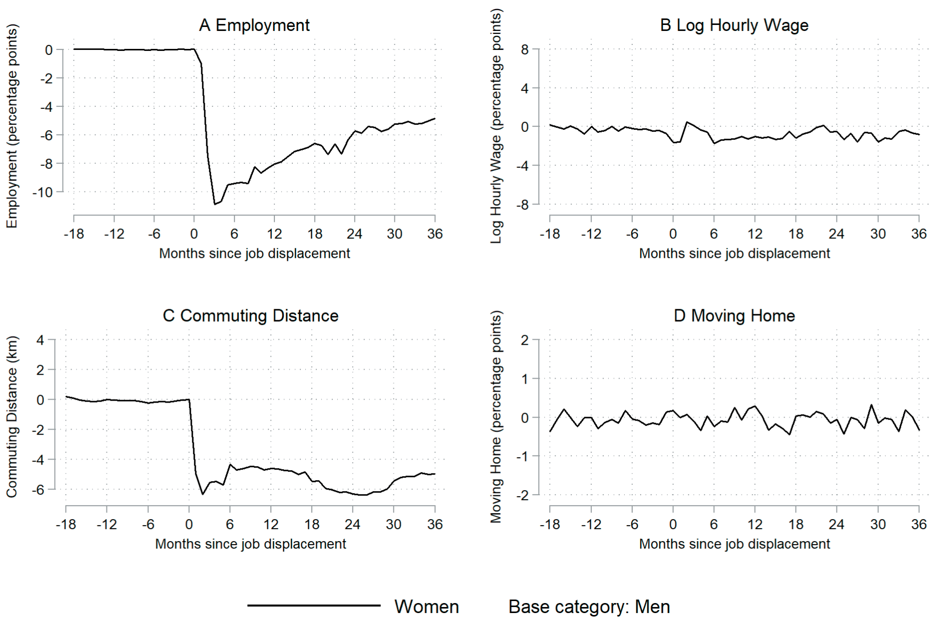
Notes: Displaced workers are matched to non-displaced workers in the month of placebo treatment, i.e. the twelfth month prior to actual displacement. Reference group is the group of non-displaced workers. Reference month is the twenty-fourth month prior to job displacement. The 95% confidence intervals are computed using standard errors clustered by individual. All four fixed effects regression models include 266 parameters of which there are 54 two-way interaction terms. See Table 2.1 for additional notes and statistics.

2.E Time-Dependent Differences in the Importance of Worker Characteristics for the Displacement Effects

Figures 2.E.1-2.E.6 show the importance of demographic characteristics for the displacement effects. Figure 2.E.1 highlights the gender differential in displacement effects. On the one hand, women experience a higher loss in employment than men. Importantly, the difference in the loss in employment diminishes over time since job displacement. On the other, women experience a smaller increase in the commuting distance than men. Figure 2.E.2A and Figure 2.E.2B show that age directly increases the loss in employment and hourly wage. The age differentials in the displacement effect on employment and hourly wage are relatively persistent over the post-displacement period. Figure 2.E.3 shows that the nationality differential in displacement effects varies across time since job displacement. The estimates show that displaced workers who have Dutch nationality experience relatively modest losses in employment.

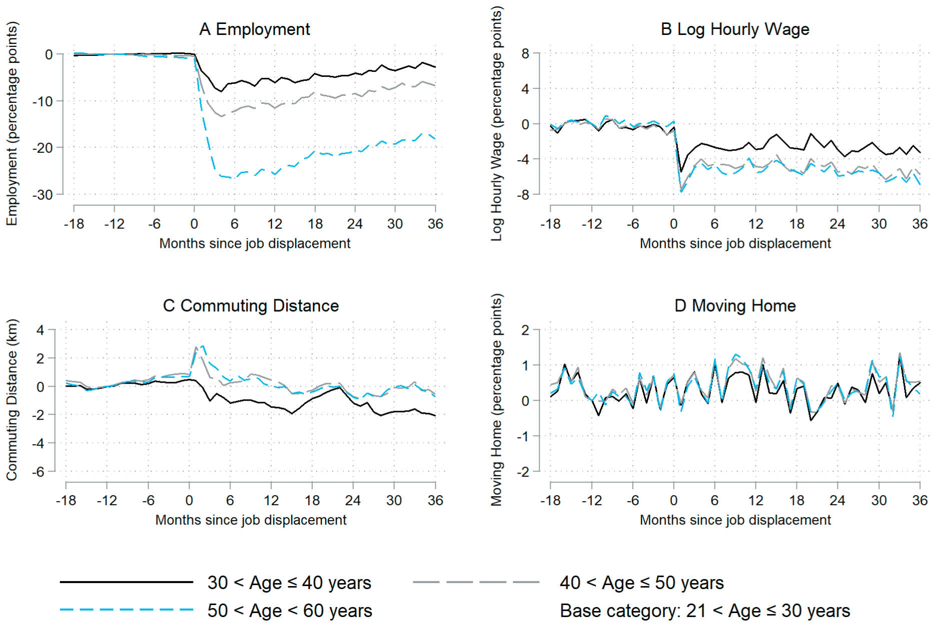
Figures 2.E.4-2.E.6 highlight the role of job characteristics in the displacement effects. Figure 2.E.4A and Figure 2.E.4B show that the worker's length of tenure in the displaced job, especially in the case of a job tenure higher than 18 years, increases the loss in employment and hourly wage, respectively. Figure 2.E.4C shows that workers who had a relatively short job tenure experience the smallest increase in the commuting distance. Figure 2.E.5 shows that displaced workers who were active in the manufacturing sector, as compared with the service sector, experience a substantial and persistent higher loss in the hourly wage of about 2-3 percentage points and a higher increase in the commuting distance of about 2 kilometres. The estimates shown in Figure 2.E.6A indicate that workers who are displaced later in time experience a higher loss in employment and hourly wage. Figure 2.E.6C shows that workers who were displaced in 2007 experience a relatively large increase in the commuting distance. This finding can be explained by a potential seasonality effect, as the workers who were displaced in 2007 were displaced between July to December. The other displacement cohorts consist of workers who were displaced in any month of the calendar year.

Fig. 2.E.1. Gender differential in displacement effects (Eq. (2.7)).



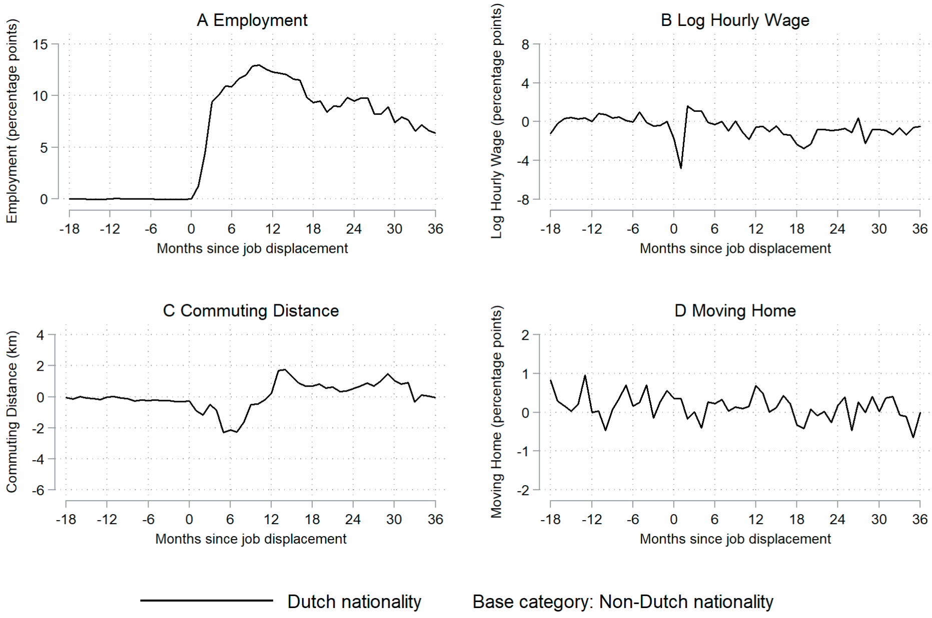
Notes: See Figure 2.1 and Table 2.2 for additional notes.

Fig. 2.E.2. Age differential in displacement effects (Eq. (2.7)).



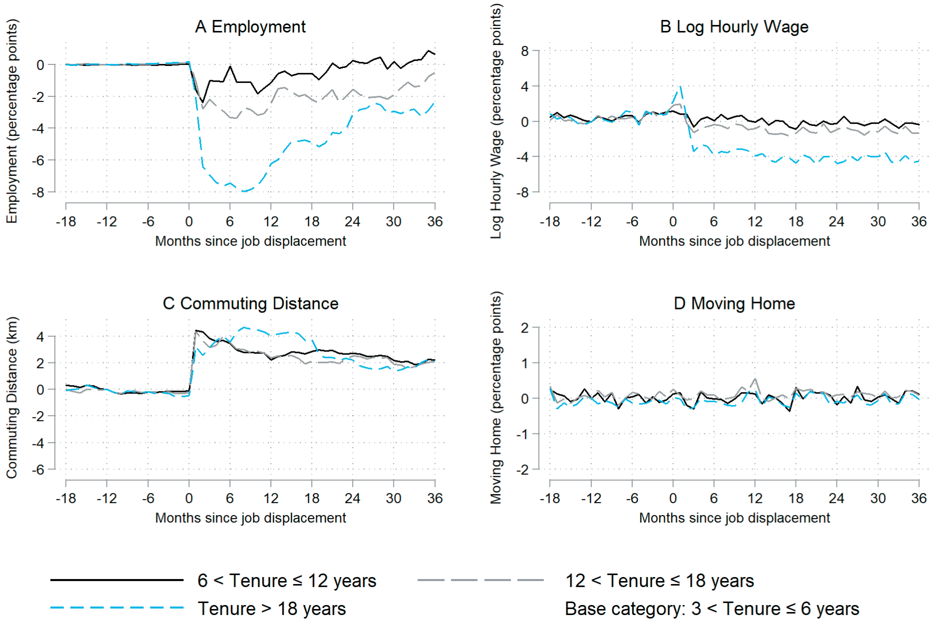
Notes: See Figure 2.1 and Table 2.2 for additional notes.

Fig. 2.E.3. Nationality differential in displacement effects (Eq. (2.7)).



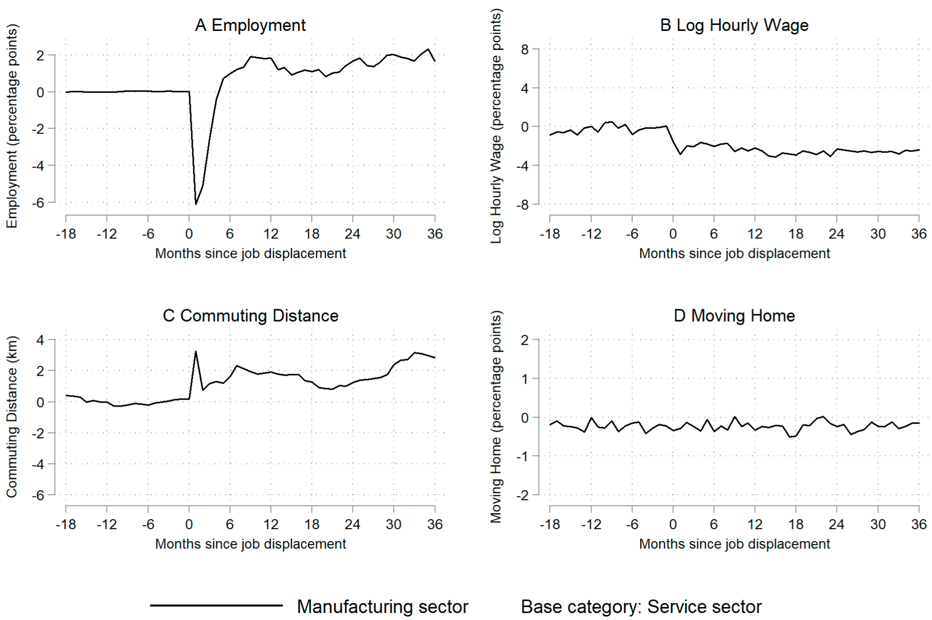
Notes: See Figure 2.1 and Table 2.2 for additional notes.

Fig. 2.E.4. Job tenure differential in displacement effects (Eq. (2.7)).



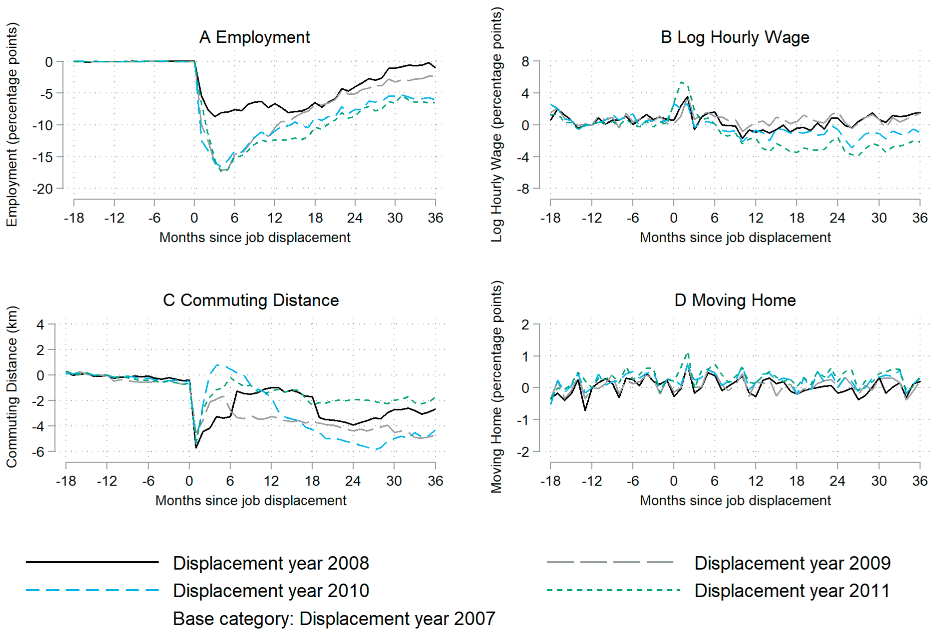
Notes: See Figure 2.1 and Table 2.2 for additional notes.

Fig. 2.E.5. Industry differential in displacement effects (Eq. (2.7)).



Notes: See Figure 2.1 and Table 2.2 for additional notes.

Fig. 2.E.6. Displacement year differential in displacement effects (Eq. (2.7)).



Notes: See Figure 2.1 and Table 2.2 for additional notes.

CHAPTER 3

flowbca: A Flow-Based Cluster Algorithm in Stata[1]

3.1 Introduction

In this article, we introduce the Stata implementation of a flow-based cluster algorithm, `flowbca`, written in Mata.A flow variable registers the total change of the variable from one entity to another entity during a specific period of time. The entity can be a region, firm, or person, and during the process of clustering, they will be grouped according to the size of the bilateral flows. Currently, flow-based cluster algorithms available in Stata focus on visualizing social networks (e.g. Corten, 2011; Miura, 2012). However, these algorithms lack the ability to flexibly aggregate units into clusters based on relational data of flows. The main motivation to write `flowbca` is that there is a need in many statistical applications, also in research fields other than social network analysis (SNA), for an algorithm to flexibly aggregate non-overlapping units into clusters. Specifically, as it provides a choice of how to operate clusters in empirical analyses and allows a researcher to compare alternative sets of clusters.

Given the increasing availability and use of relational data of various types of flows, `flowbca` can be of use to a variety of research fields. For

[1] A paper based on Chapter 3 is published in the Stata Journal entitled “flowbca: A flow-based cluster algorithm in Stata” (Meekes and Hassink, 2018b).

example, the field of economic geography makes use of flows to cluster regional units into regional clusters of economic activity (Coombes, Green, and Openshaw, 1986; Brezzi, Piacentini, Rosina, and Sanchez-Serra, 2012).[2] Alternatively, industrial input-output analysis is based on trade linkages that register the flows of goods that are produced in one production chain and used as input in another production chain (Leontief, 1986; Timmer, Dietzenbacher, Los, Stehrer, and de Vries, 2015). Finally, SNA detects communities (Fortunato, 2010) and defines flow networks (Ford Jr and Fulkerson, 1962; Beguerisse-Daz, Garduo-Hernndez, Vangelov, Yaliraki, and Barahona, 2014) in graphs as connected groups based on the strength of flows between nodes.

`flowbca` is the implementation in Stata of a so-called agglomerative hierarchical clustering algorithm (Fortunato, 2010) to define clusters based on relational data of flows, which has been used in the aforementioned research fields. The key difference between `flowbca` and other agglomerative hierarchical clustering algorithms currently available in Stata is the focus on flow-based clustering instead of distance-based clustering.[3] In general terms, the flow-based cluster algorithm behind `flowbca` can be described as follows. It starts from a set of K disjoint units. The algorithm aggregates two units into one, considering the bilateral flows between the units. Clusters are defined by iteratively repeating this procedure. In each iteration of the algorithm, the decision criterion for aggregating two units into one is based on an optimisation function selecting the maximum flow out of all bilateral flows. The source unit from which the largest flow starts is aggregated to the destination unit.

The algorithm is flexible in various aspects. First, the optimisation function can be based on two definitions of flows, directed and undirected. The former refers to the maximum of the single directed flows that are flowing from one unit to another; the latter denotes the maximum of the sum of two bilateral flows. Second, the optimisation function can be based on absolute flows and relative flows, which are computed by taking

[2]Global regional clusters of economic activity are defined using trade flows (Smith and White, 1992), foreign direct investment flows (Bathelt and Li, 2014) or multinational firms relocation flows (Chen and Moore, 2010) Within-country regional clusters such as local labour markets or local housing markets can be defined using commuting flows from place of residence to place of work, job-to-job turnover flows, household migration flows or job search flows (e.g. Duranton, 2015).

[3]Distance-based clustering uses the distance between a pair of units as a measure of similarity. Similar units are grouped into the same cluster and dissimilar units into separate clusters.

each absolute flow relative to the unit-specific total of outgoing flows. Third, the algorithm allows for flexibility in the stopping criterion by allowing for five optional ex ante user choices different from the default one. Using different options the researcher can thus create different sets of clusters.

After the algorithm has been terminated, the researcher could evaluate the choice of optimisation function and stopping criterion by analysing the level of self-containment of the set of clusters. The level of self-containment is approximated by the average of the internal relative flows. A higher average of the internal relative flows means there is a stronger connectivity within each cluster and a weaker connectivity to outside clusters.

3.2 Flow-Based Cluster Algorithm

3.2.1 The Algorithm

The main inputs of the algorithm is a K-dimensional square matrix that contains the absolute flows between K different units. Effectively, each row represents a different source unit and each column represents a different destination unit.

The algorithm consists of the following five steps. The steps of the algorithm are provided given the default options of the algorithm that will be described in Section 3.3.

Step 1: the absolute flows between K different units are rewritten as a K-dimensional square matrix (i.e. an adjacency matrix in graph theory) of absolute flows $F^{(K)}$:

$$F^{(K)} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1K} \\ f_{21} & f_{22} & \cdots & f_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ f_{K1} & f_{K2} & \cdots & f_{KK} \end{bmatrix} \tag{3.1}$$

where f_{ij} $(i{\neq}j)$ represents the flow in absolute term from source unit i to destination unit j. Flows f_{ii} are defined as the internal absolute flows.

Step 2: the matrix $F^{(K)}$ is rewritten in terms of relative flows as a K-dimensional square matrix $G^{(K)}$. Relative flows are computed by taking each absolute flow relative to the unit-specific total of outgoing flows. $G^{(K)}$ can be expressed as:

$$G^{(K)} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1K} \\ g_{21} & g_{22} & \cdots & g_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ g_{K1} & g_{K2} & \cdots & g_{KK} \end{bmatrix} \tag{3.2}$$

where

$$g_{ij} = f_{ij} / \sum_{t=1}^{K} f_{it} \qquad i, j = 1, 2, \ldots, K$$

Note that the matrix is row-normalised.

Step 3: the optimisation function selects the arguments of the maximum directed relative flow from one unit to another, for $i \neq j$, of all $K \times (K-1)$ pairs of i and j:

$$(r, s) = \arg\max_{\substack{i,j \\ i \neq j}} g_{ij} \tag{3.3}$$

where units r and s are defined as the source unit and destination unit, respectively. If $g_{rs} = 0$ or $K = 1$, the default stopping criterion of the procedure is met and the algorithm is terminated.

Step 4: source unit r will be aggregated to destination unit s. The core of the cluster is defined as the receiving unit, i.e. destination unit s. To be able to adjust matrix $F^{(K)}$, a $K \times (K-1)$-dimensional matrix $C^{(K)}$ is specified. $C^{(K)}$ can be expressed as:

$$C^{(K)} = (e_1, e_2, \ldots, e_{r-1}, e_r + e_s, e_{r+1}, \ldots, e_{s-1}, e_{s+1}, \ldots, e_K) \tag{3.4}$$

where e_i represents the i-th unit column vector. For sake of convenience in the exposition of this algorithm, matrix $C^{(K)}$ is based on the assumption that the identifier value of unit r is larger than the identifier value of unit s.

Step 5: the absolute flows to and from units r and s will be added. The new matrix $F^{(K-1)}$ can be expressed as:

$$F^{(K-1)} = (C^{(K)})^T F^{(K)} C^{(K)} \tag{3.5}$$

where T refers to the transpose operator. $F^{(K-1)}$ is now a square matrix of dimension $(K-1)$. The algorithm continues with step 1 starting with

$F^{(K-1)}$ as an input.

After the stopping criterion of step 3 has been met, the algorithm is terminated and K^* clusters are returned. The matrix of absolute flows between the K^* clusters, $F^{(K^*)}$, equals

$$F^{(K^*)} = C^T F^{(K)} C \tag{3.6}$$

where matrix C is a matrix product of the matrices $C^{(K)} \ldots C^{(K^*)}$, which can be expressed as $C = C^{(K)} C^{(K-1)} C^{(K-2)} \ldots C^{(K^*)}$.

3.2.2 Stopping Criteria

In step 3, the stopping criterion of the algorithm is defined as $g_{rs} = 0$ or $K = 1$. The algorithm allows for five alternative stopping criteria, and each of them is a modification of the stopping criterion mentioned in step 3.

First, the researcher could specify a flow threshold q. The threshold q represents the minimum level of interaction at which a source unit is aggregated to a destination unit. The algorithm is terminated in step 3 if $g_{rs} < q$.

Second, the researcher could specify a minimum number of clusters k. The algorithm is terminated in step 3 if the number of units have reduced to this minimum, i.e. if $k = K^*$.

Third, the researcher could specify a minimum average of the internal relative flows l_a. The average of the internal relative flows L_a is defined as equal to the sum of the internal relative flows g_{ii} relative to the number of clusters K^*:

$$L_a = \frac{1}{K^*} \sum_{i=1}^{K^*} g_{ii} \tag{3.7}$$

The algorithm is terminated in step 3 if $l_a \leq L_a$.

Fourth, the researcher could specify a minimum weighted average of the internal relative flows l_w. The weighted average of the internal relative flows L_w is defined as equal to the sum of the internal absolute flows relative to the sum of all absolute flows:

$$L_w = \frac{1}{N} \sum_{i=1}^{K^*} f_{ii} \tag{3.8}$$

for which the sum of all absolute flows equals $N = \sum_{i=1}^{K^*} \sum_{j=1}^{K^*} f_{ij}$. The algorithm is terminated in step 3 if $l_w \leq L_w$.

Finally, the researcher could impose a minimum internal relative flow l_m that all of the clusters must satisfy. The minimum of the internal relative flows L_m is defined by:

$$L_m = \min_i g_{ii} \tag{3.9}$$

The algorithm is terminated in step 3 if $l_m \leq L_m$.

3.2.3 An Alternative Optimisation Function

The algorithm provides two different optimisation functions, which are based either on the directed or on the undirected flows approach. The optimisation function based on the directed flows selects arg max g_{ij}, considering the maximum of the directed flows from one unit to another. In contrast, the optimisation function based on the undirected flows approach selects arg max $g_{ij} + g_{ji}$, considering the maximum of the sum of two bilateral flows, which can be expressed as:

$$(r, s) = \underset{\substack{i,j \\ i \neq j}}{\arg\max}\, g_{ij} + g_{ji} \tag{3.10}$$

Using the undirected flows approach, the algorithm is terminated if $g_{rs} + g_{sr} = 0$ or $K = 1$. Otherwise, the procedure will continue with step 4 of the algorithm. Note that the new cluster gets the identification number of the unit with the largest incoming flow, which represents the core of the new cluster.

3.2.4 Some Caveats

In step 3 it might be the case that $\arg\max g_{ij}$ holds for multiple pairs of units i, j. Consequently, the pair r, s will not be unique. The following rules are imposed to close the algorithm:

1. If there are two or more source units r, e.g. r_1 and r_2, that both have the maximum flow to the same destination unit s, r_1 is aggregated to s if r_1 has the highest incoming flow from the other source unit(s) r.

2. If source unit r has identical flows to two or more units s, e.g. s_1 and s_2, r is aggregated to s_1 if s_1 has the highest incoming flow from the other destination unit(s) s.

3. If both r and s are not unique, e.g. there are two pairs r_1,s_1 and r_2,s_2, the algorithm aggregates r_1 to s_1 if s_1 has the highest incoming flow from the other destination unit(s) s.

4. For the iterations where a unique pair is still not defined, the algorithm picks one pair r, s of all pairs that correspond to the maximum flow.

3.3 Flowbca Stata Command

3.3.1 Syntax

`flowbca` *varname varlist* [`, q(#) k(#) la(`*fraction*`) lw(`*fraction*`) lm(`*fraction*`) opt_f(#) save_k`]

3.3.2 Description

`flowbca` implements the algorithm that is discussed in Section 3.2 and performs it in Mata. The main inputs for `flowbca` are the variables listed in *varname* and *varlist*. *varname* contains one variable representing the source unit identifier. This variable should be numerical, as string variables are ignored by the `flowbca` command. *varlist* contains a set of variables, one variable for each distinct destination unit, which represents the absolute flows from the source units to the destination unit.

Effectively, the destination unit variables represent the columns of a K-dimensional square matrix of flows between the K units. For example, the value of the first observation of a destination unit variable represents the absolute flow from the first source unit to the corresponding destination unit. The source and destination units should be numbered such that if they are sorted and ordered in a sequential order, the diagonal elements of the K-dimensional square matrix represent the internal absolute flows. If the flow data of the researcher is only available in an $K \times 3$-dimensional matrix in which there are three columns that represent the source unit identifier, destination unit identifier and absolute flows between the units, respectively, the data should be reshaped by the user into a K-dimensional square matrix.

3.3.3 Options

`q(#)` sets the flow threshold. To set a relative flow threshold, place a fraction in parentheses after `q`. To set an absolute threshold, place an integer number in parentheses after `q`. If the threshold is higher than the maximum of all flows, the stopping criterion of the procedure has been met and the algorithm is terminated. The default is to have a flow threshold equal to zero.

`k(#)` specifies the number of distinct clusters the algorithm should define. The default is to define one cluster.

`la(`*fraction*`)` specifies the minimum average of the internal relative flows. If the fraction is lower than or equal to the average of the internal relative flows, the stopping criterion of the procedure is met. The default is no minimum average of the internal relative flows.

`lw(`*fraction*`)` specifies the minimum weighted average of the internal relative flows. If the fraction is lower than or equal to the weighted average of the internal relative flows, the stopping criterion of the procedure is met. The default is no minimum weighted average of the internal relative flows.

`lm(`*fraction*`)` specifies the minimum internal relative flow. If the fraction is smaller than or equal to the minimum value of the internal relative flows, the stopping criterion of the procedure is met. The default is no minimum internal relative flow.

`opt_f(#)` specifies the optimisation function. Four optimisation functions can be chosen. The default is `opt_f(1)`, which implements the directed relative flows approach. The other functions are: `opt_f(2)` implementing the undirected relative flows approach; `opt_f(3)` implementing the directed absolute flows approach; `opt_f(4)` implementing the undirected absolute flows approach.

`save_k` is an option to save the cluster_setk data sets (see below). For each k, the data set contains the absolute flows between the remaining k units (i.e. the matrix $F^{(K)}$ for each k). To save the data sets, specify `save_k`.

3.3.4 Output

`flowbca` saves three data sets.[4]

1. cluster_set contains variables that characterise the defined clusters, including variables that represent the cluster identifier (clusterid), the cluster-specific internal relative flow (internal), the average of the internal relative flows (La), the weighted average of the internal relative flows (Lw), the minimum of the internal relative flows (Lm), the cluster-specific total value of outgoing flows (rowflows), the total value of all flows (N) and a set of variables that represents the flows among all clusters (destinationunit).

2. unit_set contains variables that provide information on each starting unit, including variables that represent the source unit identifier (sourceunit), the cluster to which the unit is aggregated (clusterid), the relative flow at which the source unit is aggregated to a destination unit (g), the number of distinct clusters that are remaining after aggregating the source unit (round) and a zero-one indicator variable that equals one if the unit is the core of a cluster and zero otherwise (core).

3. cluster_setk contains the source unit identifier variable (sourceunit), and one variable for each destination unit representing the absolute flow (destinationunit). If the researcher uses the option save_k, the cluster_setk data sets will be saved.

3.4 Examples

3.4.1 Example 1: Within-country Regional Clusters Based on Commuting Flows

In the first example of a statistical application, a researcher uses `flowbca` to construct regional clusters based on individuals′ commuting flows from municipality of residence to municipality of work. The researcher aims to compare the levels of self-containment of forty NUTS 3 areas and twelve provinces to the levels of self-containment of forty and twelve clusters

[4] We suggest that the researcher creates a data set that consists of the variable sourceunit and a variable that represents the source unit labels. This data set could be merged to the data sets cluster_set and unit_set.

defined using `flowbca`, respectively. Note that a higher level of self-containment means there is a stronger connectivity within each regional cluster and a weaker connectivity to outside regional clusters. That is, clusters that are relatively self-contained are characterised by relatively many individuals who both live and work in the identical cluster. The Dutch NUTS 3 areas offer an interesting point of comparison, as they were defined, in 1971, based on journey-to-work and place-of-work statistics that reflected the employment outcomes and commuting behaviour of the Dutch population. Moreover, in research on European countries, NUTS 3 areas are often used as the regional classification to operate regional clusters (e.g. Ciccone, 2002).

For this example, aggregate data on 7,131,000 commuting flows in 2014 were used, at the municipality level, retrieved from the CBS StatLine open databank of Statistics Netherlands (CBS, 2018).[5] The algorithm starts from a set of 398 municipalities (K).[6] Note that this example uses the option `k()` of `flowbca`, as the researcher aims to define a specific number of clusters. The optimisation function is based on the directed relative flows approach. The directed flows approach was used, as commuting flows are by nature directed in the sense that they flow from one unit to another. Relative flows are preferred to absolute flows, as relative flows function as weights to account for the relative importance of a unit that allows smaller source units to be able to aggregate to bigger destination units. To visualise the defined clusters, the Stata commands mergepoly (Picard and Stepner, 2015) and spmap (Pisati, 2008) were used.

Before we discuss the results, we illustrate the main steps of the specific code used to create Figure 3.1a and Figure 3.1b.

```
. /* A loop is used to define forty and twelve regional clusters
 (Fig. 3.1b and Fig. 3.2b, respectively) */
. local numbers 40 12
. local a=1
. foreach numbs of local numbers {
. /* Open the data set that contains commuting flows across the Dutch
 municipalities, retrieved from Statistics Netherlands CBS Statline. */
```

[5] Note that the researcher could also exploit micro data to construct clusters. For instance, subgroup-specific clusters could be defined using subgroup-specific flows (Farmer and Fotheringham, 2011).

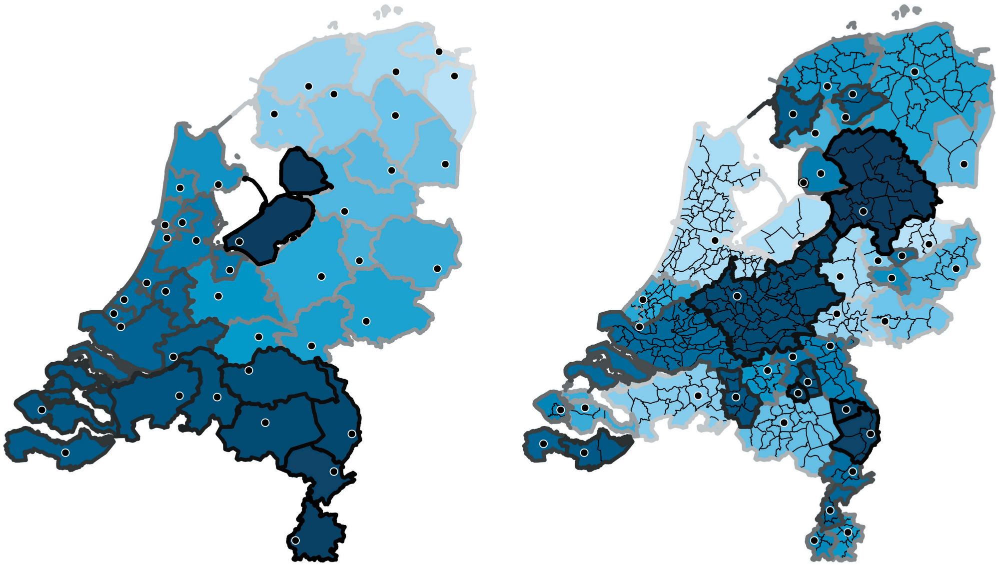
[6] Note that five municipalities, which represent small Wadden islands in the northern part of the Netherlands, were removed. These municipalities were removed as they would be defined as small self-contained clusters that artificially increase the average of the internal relative flows (L_a).

```
. use "CBS_COMM_Flow.dta", clear
. /* Apply flowbca. "homemun" is the source unit identifier;
 "workmun*" are the destination unit identifiers */
. flowbca homemun workmun*, k(`numbs´)
  (output omitted)
. /* Get the summary statistics of the La, Lw, Lm variables */
. display as text "Values of La, Lw, Lm for Figure `a´b"
. sum La Lw Lm
. /* Syntax lines 43 to 61 are specified to merge the cluster labels
 (names of the regions) to the cluster_set data set and the unit_set
 data set (see footnote 4) */
. use unit_set, clear
. rename sourceunit homemun
. merge 1:1 homemun using "Ex1label.dta"
. keep if _merge==3
. drop _merge
. rename homemun sourceunit
. save unit_set_Fig`a´b, replace
/* Open the shape boundary database file "Ex1nldb", which was
 generated using the ESRI shapefile of the Netherlands */
. use "Ex1nldb.dta", clear
. /* Create an identifier */
. gen ID=_n
. order ID
. /* Merge "Ex1nldb" to the codings data set "Ex1id.dta" */
. merge 1:1 ID using "Ex1id.dta"
. /* Merge the data set to "unit_set_Fig`a´b.dta" that includes
 the cluster identifier */
. merge 1:1 sourceunit using "unit_set_Fig`a´b.dta"
. save "Ex1part1", replace
. /* Generate a data set "Ex1pointcoord.dta" that contains
 the point coordinates of the cores of the clusters that
 are returned by the algorithm */
. use "Ex1part1", clear
. keep if core==1
. keep clusterid x_centroid y_centroid
. save "Ex1pointcoord.dta", replace
. /* Mergepoly: the mergepoly command is used to merge adjacent
 polygons from a shape boundary file:
 see http://fmwww.bc.edu/RePEc/bocode/m/mergepoly.html */
. /* We use the coordinate file "Ex1nlcoord" to merge the
 polygons of the units that are in the same cluster
 (given by the variable "clusterid") */
. use "Ex1part1", clear
. mergepoly id_shape using Ex1nlcoord, coordinates(Ex1nlcoord2) replace
by(clusterid)
. /* The output is the "Ex1nlcoord2" data set, which contains
 the coordinates of each cluster. This data set will be used to
 draw the thick border of the cluster in the map below */
. /* Now draw the map with the spmap command */
. use "Ex1part1", clear
. spmap clusterid using Ex1nlcoord, id(id_shape) clmethod(unique)
osize(thin) fcolor(Blues2) legenda(off) polygon(data("Ex1nlcoord2.dta")
ocolor(Greys2) osize(thick ..) by(_ID)) point(data("Ex1pointcoord.dta")
x(x_centroid) y(y_centroid) by(clusterid) size(medium) ocolor(white ..))
```

```
. graph export "Figure`a´b_clusters.png", replace width(5000)
. local a=`a´+1
}
Values of La, Lw, Lm for Figure 3.1b

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          La |        40    .6947013           0   .6947013   .6947013
          Lw |        40     .804838           0    .804838    .804838
          Lm |        40    .3852459           0   .3852459   .3852459

Values of La, Lw, Lm for Figure 3.2b

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          La |        12    .8688188           0   .8688188   .8688188
          Lw |        12    .9008133           0   .9008133   .9008133
          Lm |        12    .7762533           0   .7762533   .7762533
```

Figure 3.1 shows that the weighted average of the internal relative flows (L_w) of the forty defined regional clusters (see graph (b)) is higher than the weighted average of the NUTS 3 areas (see graph (a)). This means there are more individuals who live and work in their defined regional cluster (about 80.5 per cent) than in their NUTS 3 area (about 74.3 per cent). Figure 3.2 shows that the average of the internal relative flows (L_a), the weighted average of the internal relative flows (L_w) and the minimum of the internal relative flows (L_m) are higher in the case of the twelve defined clusters than of the twelve pre-defined administrative provincial areas. All in all, this example shows that `flowbca` can be used to define meaningful regional clusters that are characterised by a relatively high level of self-containment.

Fig. 3.1. NUTS 3 areas and forty defined clusters.

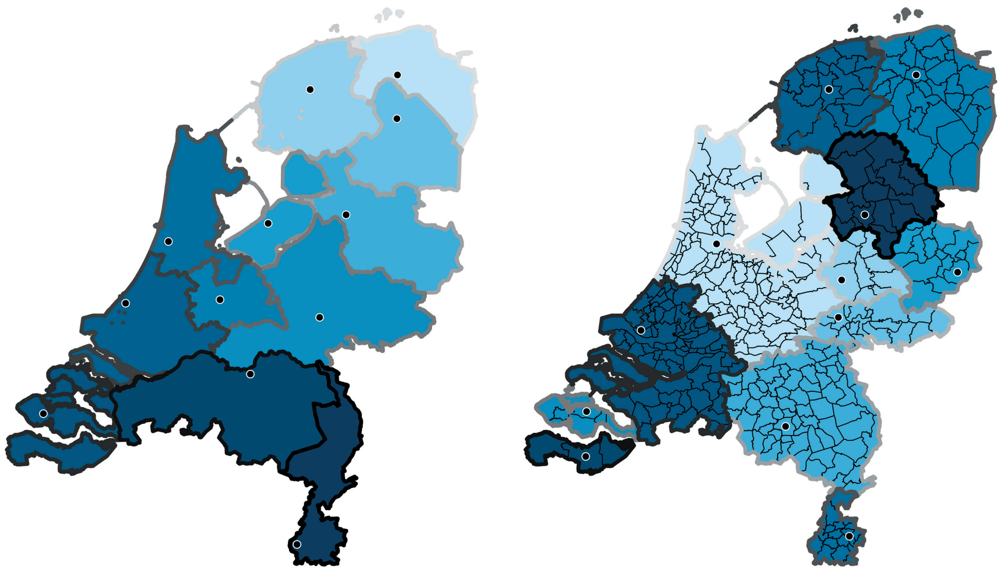


(a) NUTS 3 areas. $L_a = .699$; $L_w = .743$; $L_m = .462$; $K^* = 40$; $K = 40$

(b) Forty clusters. $L_a = .695$; $L_w = .805$; $L_m = .385$; $K^* = 40$; $K = 398$

Notes: A commuting flow registers the number of workers who commute from a municipality of residence to a municipality of work. The NUTS 3 cores (the black dots with a white circle) are defined as the municipality with the highest number of residents. The cores of the defined regional clusters are returned by the algorithm. Each distinct cluster is surrounded by a thick border and highlighted by a different colour. Note that the colour of a cluster does not provide any further information. L_a, L_w and L_m are returned by the algorithm, and refer to the average of the internal relative flows (Eq. (3.7)), the population-weighted average of the internal relative flows (Eq. (3.8)), and the minimum of the internal relative flows (Eq. (3.9)), respectively. K^* and K refer to the number of defined regional clusters and the number of distinct starting units, respectively.

Fig. 3.2. Provinces and twelve defined clusters.



(a) Provincial areas. $L_a = .833$; $L_w = .866$; $L_m = .570$; $K^* = 12$; $K = 12$

(b) Twelve clusters. $L_a = .869$; $L_w = .901$; $L_m = .776$; $K^* = 12$; $K = 398$

Notes: A commuting flow registers the number of workers who commute from a municipality of residence to a municipality of work. The provincial cores (the black dots with a white circle) are the capital cities. The cores of the defined regional clusters are returned by the algorithm. Each distinct cluster is surrounded by a thick border and highlighted by a different colour. Note that the colour of a cluster does not provide any further information. L_a, L_w and L_m are returned by the algorithm, and refer to the average of the internal relative flows (Eq. (3.7)), the population-weighted average of the internal relative flows (Eq. (3.8)), and the minimum of the internal relative flows (Eq. (3.9)), respectively. K^* and K refer to the number of defined regional clusters and the number of distinct starting units, respectively.

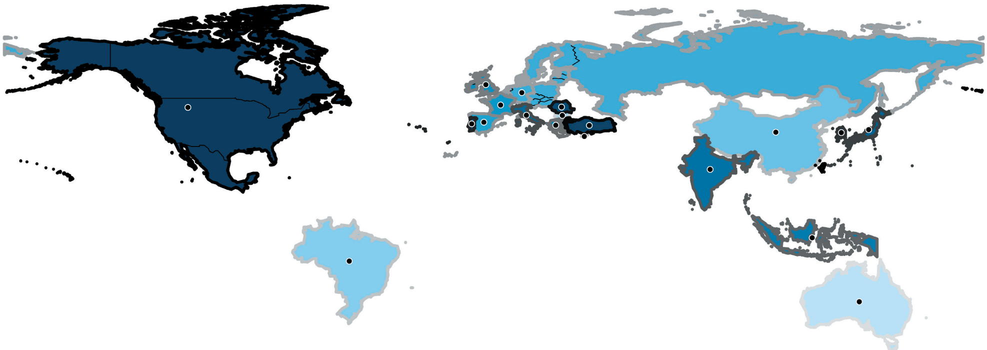
3.4.2 Example 2: Global Regional Clusters Based on National Trade Flows

In the second example, a researcher uses `flowbca` to construct global regional clusters based on trade flows that are defined as the size of the annual trade from an exporting to an importing country. The researcher aims to examine how interrelated countries are in terms of trade, and whether this relatedness changed over time. The researcher defines global clusters of economic activity for the years 1995 and 2011, given a minimum flow threshold which is set equal to five per cent.

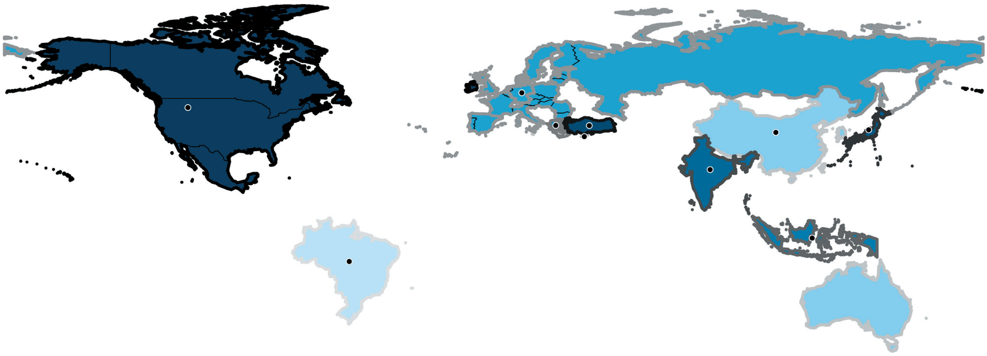
For this example, the World Input-Output Database (WIOD) was used (Dietzenbacher, Los, Stehrer, Timmer, and Vries, 2013). This data set consists of data on the trade flows between forty countries in the period 1995-2011. The algorithm starts from a set of forty countries (K). Note that this exercise uses the option `q()` of `flowbca`. The optimisation function is based on the directed relative flows approach.

The set of clusters in 1995 and 2011 are compared to examine the change in global trade clusters over time. Figure 3.3 shows, given the flow threshold of five per cent, that the number of distinct global clusters (K^*) decreases from nineteen to ten over the period 1995-2011. Consistent with globalization, the decrease in the number of defined clusters suggests that the trade flows within countries decreased relative to the trade flows between countries. Another observation is that the size of the two clusters in which China and Germany is the core, respectively, became larger over time.

Fig. 3.3. Global clusters based on trade flows.



(a) Global trade clusters in 1995. $L_a = .920$; $L_w = .943$; $L_m = .854$; $K^* = 19$; $K = 40$



(b) Global trade clusters in 2011. $L_a = .938$; $L_w = .947$; $L_m = .898$; $K^* = 10$; $K = 40$

Notes: A trade flow registers the size of annual trade from an exporting country to an importing country. The cores of the defined clusters (the black dots with a white circle) are returned by the algorithm. Each distinct cluster is highlighted by a different colour. Note that the colour of a cluster does not provide any further information. Global clusters were defined based on trade flows expressed in millions of dollars between countries from the 1995 and 2011 WIOD data. The flow threshold `q` was set equal to five per cent. Trade flows between countries were computed by aggregating all within-country flows. Nine countries with zero or negative flows were removed. L_a, L_w and L_m are returned by the algorithm, and refer to the average of the internal relative flows (Eq. (3.7)), the population-weighted average of the internal relative flows (Eq. (3.8)), and the minimum of the internal relative flows (Eq. (3.9)), respectively. K^* and K refer to the number of defined global clusters and the number of starting countries, respectively.

3.4.3 Example 3: A Social Network Based on Friendship Ties

In the third example, a researcher uses `flowbca` to detect groups of prison inmates based on friendship ties. A friendship tie could be considered as a binary flow variable from one inmate to another, which is one in case of a friendship. If two inmates indicate a friendship with each other, the ties will flow in both directions. The researcher aims to detect groups of inmates in which each group should have a minimum internal relative flow of at least fifty per cent. The minimum internal relative flow of fifty per cent means that, in each group, at least fifty per cent of the inmates′ friendship ties should be with inmates in their own group.

For this example, the Gagnon and MacRae prison friendship data set was used (MacRae, 1960). The level of interaction between inmates is approximated by multiple zero-one indicator variables that represent friendship ties, which equal one if a given "source" inmate indicates a friendship with a given "destination" inmate and zero otherwise. The algorithm starts from a set of 67 inmates (K). All inmates could indicate as few or as many friendship ties as desired. Inmates could not indicate a friendship with themselves. Note that this example utilises the option `lm()` of `flowbca`. Relative flows were used to have the relative importance of each tie.

Before we discuss the results, we illustrate the main steps of the specific code used to construct Table 3.1.

```
. /* Directed relative flows approach */
. use "Ex3_Prison.dta", clear
. /* Apply flowbca */
. flowbca sourceunit destinationunit*, lm(.5) opt_f(1)
. /* The option lm(fraction) is used to specify the minimum internal
 relative flow. If the fraction is smaller than or equal to the
 minimum value, the algorithm is terminated */
. /* The option opt_f() is specified to use the directed relative
 flows approach */
. /* Drop the inmate (number 35) who is isolated */
. drop if internal==.
(1 observation deleted)
. /* Get the summary statistics of the La, Lw, Lm variables */
. sum La Lw Lm
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          La |         5    .8470662           0   .8470662   .8470662
          Lw |         5    .8461539           0   .8461539   .8461539
          Lm |         5    .7575758           0   .7575758   .7575758
. use unit_set, clear
```

```
. /* Generate the number of inmates in each cluster */
. bysort clusterid: gen n=_N

. tab n if core==1 & n!=1
          n |      Freq.     Percent        Cum.
------------+-----------------------------------
          4 |          1       20.00       20.00
          5 |          1       20.00       40.00
          7 |          1       20.00       60.00
         12 |          1       20.00       80.00
         38 |          1       20.00      100.00
------------+-----------------------------------
      Total |          5      100.00

. /* Undirected relative flows approach */
. use "Ex3_Prison.dta", clear

. /* Apply flowbca */
. flowbca sourceunit destinationunit*, lm(.5) opt_f(2)
. /* The option opt_f() is specified to use the undirected relative
 flows approach */

. /* Drop the inmates (numbers 19, 25, 26 and 35) who are isolated */
. drop if internal==.
(4 observations deleted)

. /* Get the summary statistics of the La, Lw, Lm variables */
. sum La Lw Lm
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          La |        12    .6894538           0   .6894538   .6894538
          Lw |        12    .6758242           0   .6758242   .6758242
          Lm |        12          .5           0         .5         .5

. use unit_set, clear

. /* Generate the number of inmates in each cluster */
. bysort clusterid: gen n=_N

. tab n if core==1 & n!=1
          n |      Freq.     Percent        Cum.
------------+-----------------------------------
          2 |          2       16.67       16.67
          3 |          1        8.33       25.00
          4 |          2       16.67       41.67
          5 |          3       25.00       66.67
          6 |          1        8.33       75.00
          7 |          1        8.33       83.33
          9 |          1        8.33       91.67
         11 |          1        8.33      100.00
------------+-----------------------------------
      Total |         12      100.00
```

Table 3.1 provides information about the groups of inmates that were detected using `flowbca` for both the directed and undirected flows approach, respectively. The results show that the directed flows approach, compared to the undirected flows approach, leads to fewer and bigger groups of inmates. Another observation is that there are more isolated inmates if the optimisation function is based on the undirected flows approach. Effectively, the undirected flows approach puts more weight on the situation where two inmates indicate each other as friend, and leads to more sparse groups.

Table 3.1
Number and size of the detected groups of inmates.

Directed flows approach		Undirected flows approach	
Number of groups for a given size	Size (in # of inmates)	Number of groups for a given size	Size (in # of inmates)
1	38	1	11
1	12	1	9
1	7	1	7
1	5	1	6
1	4	3	5
		2	4
		1	3
		2	2
$L_a = .847$		$L_a = .689$	
$L_w = .846$		$L_w = .676$	
$L_m = .758$		$L_m = .5$	
$K^* = 5$		$K^* = 12$	
$K = 67$		$K = 67$	

Notes: The connected groups of inmates are based on friendship ties between inmates from the Gagnon and MacRae prison data set. A friendship tie registers a friendship as a flow from one inmate to another. The minimum internal relative flow `lm`, which each group should satisfy, was set equal to fifty per cent. The raw prison data set contains 67 inmates. No inmate was disconnected, i.e. the situation where an inmate did not specify nor was specified as friend by another inmate. However, one inmate and four inmates were detected as isolated using the directed and undirected flows approach, respectively. The isolated inmates were removed. L_a, L_w and L_m are returned by the algorithm, and refer to the average of the internal relative flows (Eq. (3.7)), the population-weighted average of the internal relative flows (Eq. (3.8)), and the minimum of the internal relative flows (Eq. (3.9)), respectively. K^* and K refer to the number of defined regional clusters and the number of distinct starting units, respectively.

3.4.4 Example 4: Industrial Clusters Based on Input-output Flows

The final example is a case where a researcher uses `flowbca` to define five U.S. industrial clusters based on input-output flows of goods between U.S. industries. A flow of goods registers the size of the goods delivered by the industry of input to the industry of output. The researcher examines whether the relatedness between industries changed over time, by comparing the set of industrial clusters in 1995 to the set of clusters in 2011. The WIOD was used to define U.S. industrial clusters in the years 1995 and 2011. The algorithm starts from a set of thirty-five industries (K). Note that this exercise utilises the option `k()` of `flowbca`. The optimisation function is based on the directed relative flows approach.

Table 3.2 presents the results of example 4. The sectors "Construction" and "Public Administration and Defence; Compulsory Social Security" are the largest industrial clusters in 1995 and 2011, respectively. Hence, in 2011, the industry "Public Administration and Defence; Compulsory Social Security" uses relatively more goods that are produced in other production chains than the industry "Construction". Interestingly, `flowbca` defines the industries "Food, Beverages and Tobacco", "Textiles and Textile Products" and "Transport Equipment" as the core of a cluster in both 1995 and 2011, which suggests that these clusters have been relatively self-contained over time.

As Table 3.2 shows, the largest cluster is composed of many units and the other clusters are composed of very few units. Example 4 highlights the main limitation of `flowbca`. The main limitation of `flowbca` is that the algorithm defines one relatively big cluster that is composed of many units, if the network is not sparse enough and thus not composed of multiple sub-networks. For example, consider the case that a researcher aims to define clusters using random flows between units. It is likely that in each iteration a source unit is aggregated to the identical destination unit, as the destination unit represents a relatively big cluster due to the aggregations in the earlier iterations.

All in all, to define meaningful clusters, the network should be sparse enough, e.g. by distance (in the case of within-country regional clusters), input-output flows of goods (in the case of industrial clusters) or social interaction (in the case of SNA). Otherwise, the algorithm will define one relatively big cluster and several distinct clusters of very few units. Note that the use of more disaggregated units in the starting set of units would improve the accuracy of the cluster algorithm, as more detailed flow data is used.

Table 3.2
Industrial clusters based on U.S. input-output flows.

1995		2011	
Core of cluster	Size (in # of units)	Core of cluster	Size (in #) of units)
Construction	28	Public Administration and Defence; Compulsory Social Security	30
Food, Beverages and Tobacco	3	Food, Beverages and Tobacco	2
Textiles and Textile Products	2	Chemicals and Chemical Products	1
Chemicals and Chemical Products	1	Basic Metals and Fabricated Metal	1
Transport Equipment	1	Transport Equipment	1
$L_a = .602$		$L_a = .567$	
$L_w = .828$		$L_w = .837$	
$L_m = .335$		$L_m = .287$	
$K^* = 5$		$K^* = 5$	
$K = 35$		$K = 35$	

Notes: U.S. industrial clusters based on U.S. input-output flows of goods expressed in millions of dollars between thirty-five ISIC industries from the WIOD data. The minimum number of clusters `k` was set equal to five. L_a, L_w and L_m are returned by the algorithm, and refer to the average of the internal relative flows (Eq. (3.7)), the population-weighted average of the internal relative flows (Eq. (3.8)), and the minimum of the internal relative flows (Eq. (3.9)), respectively. K^* and K refer to the number of defined regional clusters and the number of distinct starting units, respectively.

3.5 Concluding Remarks

In this article, we have introduced and illustrated the `flowbca` Stata command that can be used to define clusters based on relational data of flows between disjoint units. Four examples of statistical applications in a wide range of research fields were provided to illustrate `flowbca` cluster identification capabilities. Given the increasing availability of relational data of various types of flows, `flowbca` can be of use to a variety of research fields. `flowbca` is flexible, because it allows for various optimisation functions and stopping criteria. The programme is accessible for the researcher, which will hopefully lead to a further development of the algorithm. Overall, the programme is robust, user-friendly and well able to define clusters that are characterised by a high level of self-containment.

Acknowledgements

We wish to thank an anonymous reviewer whose valuable comments improved the quality of Chapter 3. In addition, we are grateful for the comments of seminar participants at the Utrecht University School of Economics. We also thank Rense Corten, Elena Fumagalli, and Bastian Westbrock for insightful comments.

CHAPTER 4

Endogenous Local Labour Markets, Regional Aggregation and Agglomeration Economies[1]

4.1 Introduction

Economists and geographers use the concept of a local labour market (LLM) to describe a self-contained regional area of residence and work activity. LLMs have received much attention from researchers and policy-makers, as they reveal regional differences in population and employment density, employment opportunities, productivity, wages and housing prices (Topel, 1986; Bhaskar et al., 2002; Moretti, 2011). The regional differences across LLMs are explained by agglomeration economies, which refer to benefits derived from spatial concentration of economic processes. The urban economics literature generally makes a distinction between two types of agglomeration economies. The first type is urbanisation economies, which corresponds to externalities from overall economic activity and diversity. The second type is localisation economies, which corresponds to externalities from specialisation of firms located in proximity of each other. The theoretical micro-foundations of agglomeration

[1] Chapter 4 is published as TKI Discussion Paper 18-03 “Endogenous local labour markets, regional aggregation and agglomeration economies” (Meekes and Hassink, 2018a).

economies include improved matching of employers to workers and other inputs, sharing of resources and risk, and learning through the generation, diffusion and accumulation of knowledge (Duranton and Puga, 2004). The structure of an LLM differs among subgroups of the population (Farmer and Fotheringham, 2011), which could lead to subgroup differentials in agglomeration economies. Positive spillovers from agglomeration economies may lead to higher wages and better employment prospects for workers.

There is no general consensus, however, on the spatial scale of LLMs at which agglomeration economies take place (Rosenthal and Strange, 2001; Combes and Gobillon, 2015). From a policy perspective, a better understanding of the structure of workers' LLM is important to improve the effectiveness of regional policies that aim to stimulate agglomeration benefits and regional productivity growth. On the one hand, the returns to agglomeration could be decreasing in the spatial scale. For example, the transfer of knowledge might be more prevalent at a low spatial scale if the transmission of knowledge requires face-to-face contact or happens at accidental encounters. Arzaghi and Henderson (2008) show that for advertising firms, knowledge spillovers from density are large but attenuate rapidly with geographical distance. On the other, the returns to agglomeration are increasing in the spatial scale. Sharing of resources and market access for final and intermediate goods could be more prevalent at a high spatial scale, as state-level transportation modes may affect the location patterns of industries that are sensitive to shipping costs (Rosenthal and Strange, 2001). In addition, the matching mechanism could be prevalent at a high spatial scale, as for example workers with a higher education often become employed far away from their home. An approach to examine the role of spatial scale in the returns to agglomeration in wages is introduced and applied by Rosenthal and Strange (2003, 2008), who examine the attenuation of agglomeration benefits by drawing concentric rings at various distances around the worker's Work Public Use Micro Area. They find that the effect of agglomeration on wages rapidly attenuates with distance.

The main aim of Chapter 4 is to examine the importance of regional aggregation for the returns to agglomeration. So far, studies use different types of administrative regional classifications to assess the implications of geographical aggregation (Briant et al., 2010; Lindo, 2015). In many countries, there are generally two to three regional classifications available at different levels of regional aggregation. We introduce a new approach to assess the role of the spatial scale of LLMs in agglomeration economies by using a flow-based cluster algorithm to define LLMs at various levels of regional aggregation. A higher level of regional aggregation and spatial scale is represented by fewer distinct LLMs. Additionally, two economic applications of the returns to agglomeration are analysed. The first application concerns the effect of agglomeration externalities on wages, which is referred to as the urban wage premium (UWP). We examine the differences in the UWP across LLMs by using various levels of regional aggregation, to assess whether the matching, sharing and learning micro-foundations of agglomeration economies are more prevalent at different spatial scales. The second application focuses on the returns to agglomeration through one specific micro-foundation: whether a dense labour market facilitates better job matching between workers and employers. To examine the role of agglomeration economies in job matching, we use an empirical design involving job displacement due to firm bankruptcy (hereafter: job displacement). A priori, the role of agglomeration in job matching is ambiguous. Denser labour markets are characterised by lower search costs that facilitate matching, but also by increased search complexity and congestion that hinders matching (Wheeler, 2001). We focus on the role of spatial scale and its effect of job displacement on two margins of labour adjustment: the wage margin that represents a heterogeneity effect and the employment margin that represents a quantity effect.

Another aim of Chapter 4 is to examine whether and how the returns to agglomeration differ by subgroup of workers. So far, the literature on LLMs (Manning and Petrongolo, 2017; Monte et al., 2018) and agglomeration economies (for an overview, see Combes and Gobillon (2015)) has been using pre-defined regional classifications to operate aggregate LLMs at different spatial scales, which could explain the mixed results when trying to answer the question whether agglomeration economies are gender-biased and skill-biased. We will define subgroup-specific LLMs and analyse three different ways in how the returns to agglomeration may differ among subgroups. First, the magnitude of the returns to agglomeration could differ among subgroups through differences in the ability to exploit the positive spillovers from agglomeration economies. The vast majority of the literature focuses on this mechanism, assessing gender differentials and education differentials in the agglomeration benefits for wages (e.g., see Phimister (2005), Di Addario and Patacchini (2008) and Rosenthal and Strange (2008)). Second, the role of regional aggregation in the returns to agglomeration might differ among subgroups. This holds if the prevalence of agglomeration economies depends on the worker's demographic characteristics and the worker's spatial scale of LLMs. An example would be that spillovers from agglomeration improve job matching of only high-educated workers and only at a relatively high spatial scale. Third, the returns to agglomeration could differ among subgroups of workers if there are differences in the structure of workers' LLMs. The differences in the opportunity costs of commuting through financial and time constraints suggest that workers vary in LLM size. Indeed, Farmer and Fotheringham (2011) show that the worker's LLM structure is endogenous to worker characteristics. We assess the structure of workers' LLM and its economic consequences by examining whether the benefits from agglomeration economies and the patterns over the level of regional aggregation differ among subgroups of workers.

Our empirical analysis is based on rich administrative linked employee-employer data sets that cover the period 2006 to 2014. We follow the literature by focusing on differences in workers' LLM through differences in workers' commuting flows from residence to workplace (e.g., see Farmer and Fotheringham (2011); Brezzi et al. (2012)).[2] We use flowbca, which is a new flow-based cluster algorithm that can be used to define LLMs at various levels of regional aggregation (Meekes and Hassink, 2018b).[3] The algorithm is flexible and able to define LLMs for different subgroups of workers. We endogenise workers' LLMs to commuting outcomes, gender and education level. The use of endogenous LLMs prevents the problem that holds for pre-defined regional classifications, which is that these are defined in line with administrative needs instead of economic relevance. To be able to compare with the majority of the studies on agglomeration economies, we apply the concept of endogenous LLMs to the analysis of static externalities of agglomeration.[4] Specifically, we analyse the impact of employment density on wages and employment

[2] Note that alternative ways to model differences in workers' LLM can be based on differences in job search behaviour by workers or employers, such as focusing on job-to-job flows (e.g., see Nimczik (2018)).

[3] So far, the literature has used distance-based clustering or concentration indices based on densities that are non-directional by nature (e.g., see Duranton and Overman (2005); Delgado et al. (2016)). We use a flow-based cluster algorithm to examine the structure of LLMs, which is directed by nature as the main input is directional data of commuting flows.

[4] More recent studies examine the dynamic externalities of agglomeration that may benefit local productivity (e.g., see De La Roca and Puga (2016)).

using separate reduced-form regressions.[5] We estimate the returns to agglomeration using multiple sets of aggregate local labour markets (ALLM) and subgroup-specific local labour markets (SLLM) that vary in the level of regional aggregation. We also use pre-defined regional classifications to facilitate a comparison with the literature, including 398 Dutch municipalities, forty European Nomenclature of Territorial Units for Statistics (NUTS) 3 areas and thirty-five public employment services (PES) areas.

Our analysis provides two sets of novel results. First of all, we show that estimates of the UWP are increasing by a magnitude of two to three in the level of regional aggregation, for example by using forty Dutch NUTS 3 areas instead of the 398 Dutch municipalities. For empirical models in which employment density is operated by using fewer distinct LLMs, estimates of the returns to agglomeration in wages are higher. This pattern holds if we use the aggregate LLMs returned by flowbca. Hence, the results suggest that a large share of the benefits from agglomeration – including improved matching, sharing and learning – takes place at a relatively high spatial scale. In this regard, our results imply that the matching and sharing mechanisms are more important for agglomeration benefits than the learning mechanism, as the former mechanisms are considered to take place at a relatively high spatial scale (Rosenthal and Strange, 2001). In addition, we show that at a relatively high spatial scale, the loss in hourly wage for displaced workers who reside in relatively dense LLMs, compared to displaced workers who reside in more sparse LLMs, is significantly more modest. The returns to agglomeration in post-displacement wages are increasing by a magnitude of two in the level of regional aggregation. We find no returns to agglomeration in employment for workers who have been displaced. Our findings suggest that for displaced workers, the positive returns to agglomeration allow them to become more selective in wages with a comparable probability of re-employment. We contribute to the urban economics literature by focusing explicitly on the matching mechanism of agglomeration economies. Our findings are consistent with the geographical matching-function literature, which shows that market scale effects lead to higher wages but not to

[5]Note that we focus on urbanisation effects of employment density and area size, not on localisation effects of specialisation. We emphasise urbanisation economies as they play a more important role in the returns to agglomeration, and the use of variables to approximate localisation economies leads to more serious endogeneity concerns (Combes et al., 2008; Groot et al., 2014; Combes and Gobillon, 2015).

more rapid re-employment (Petrongolo and Pissarides, 2006).

Additionally, we show that the returns to agglomeration are more pressured downwards for workers who are characterised by relatively large LLMs, such as male and high-educated workers. We find that the UWP for high-educated workers compared to low-educated workers is about 100 per cent higher, holding the level of regional aggregation and thus the number of distinct LLMs constant. This finding is consistent with Manning (2003), as it suggests that workers who are characterised by thinner and larger LLMs benefit more from denser labour markets because their wage offer curve is more steep. We find no systematic difference between the UWP of men and women. Importantly, the literature that focuses on subgroup differentials in the returns to agglomeration uses regional classifications that represent relatively large areas that do not differ among subgroups of workers. On the contrary, we show in our descriptive analysis that the structure of workers' LLM differs among subgroups. Specifically, the LLM of low-educated workers and of female workers is relatively small. This descriptive finding is consistent with theoretical mechanisms that suggest that workers vary in the opportunity costs of commuting. In this regard, our results suggest that the returns to agglomeration are substantially overestimated for workers who are characterised by relatively small LLMs. Therefore, we contribute to the literature on subgroup differentials in agglomeration effects by examining the role of the endogenous structure of workers' LLM in economic outcomes.

4.2 Previous Research and Conceptual Setting

4.2.1 The Returns to Agglomeration

The seminal paper on the UWP is by Ciccone and Hall (1996), who focus on the static agglomeration effects on local productivity. For the U.S., Ciccone and Hall (1996) argue that the average labour productivity of a county increases by 6 per cent if the employment density doubles. Empirical urban economists devoted much attention to deal with individual-level endogeneity and local-level endogeneity, including the issues of skill-biased sorting across LLMs and endogenous local determinants that could bias the estimate of the UWP. Glaeser and Maré (2001) were the first to exploit micro data using individual-specific fixed effects to eliminate the potential of sorting more able workers into larger LLMs. Combes et al. (2008) introduce a two-step procedure to control for correlations between local-time unobservables and individual covariates. More recent studies take care of endogenous sorting across LLMs, using more advanced frameworks based on structural models (Gould, 2007; Baum-Snow and Pavan, 2012). See Rosenthal and Strange (2004) and Combes and Gobillon (2015) for comprehensive overviews.

An important question on the returns to agglomeration is to what extent the matching, sharing and learning mechanisms are relevant (Combes and Gobillon, 2015). We examine the importance of one specific mechanism: whether denser markets lead to better job matching between employers and workers in terms of employment and wage outcomes. Specifically, we assess the returns to agglomeration for workers who have been displaced, focusing on whether the externalities from denser LLMs affect the post-displacement employment probability and post-displacement wages. We are not the first to examine the role of employment density and the matching mechanism in employment and wage outcomes. However, most studies ignore various selection problems that arise due to correlations among workers' exit rate into unemployment, cause of unemployment and location choice. For example, the intensity of job-to-job search and the probability of labour turnover is relatively low for workers with relatively high wages (Bhaskar et al., 2002) and low commutes (Crane, 1996).

The theoretical mechanisms underlying the role of agglomeration in post-displacement outcomes are ambiguous. A relatively dense market is characterised by low search frictions, more job opportunities and a high distribution of accepted wages. In turn, the higher number of workers and firms may lead to more productive job matching and lower market power of firms over employees (Manning, 2003). In addition, the literature suggests that workers in dense labour markets experience a greater degree of assortative matching, which leads to better job matches (Helsley and Strange, 1990; Wheeler, 2001; Duranton and Puga, 2004; Andersson et al., 2007). Thereby, displaced workers in dense markets may experience a relatively modest loss in post-displacement employment. Alternatively, a relatively high number of job opportunities may lead workers to become more selective in wages (Duranton and Puga, 2004; Berliant et al., 2006; Petrongolo and Pissarides, 2006). In turn, the positive spillovers from denser markets may not improve employment prospects, but instead increase workers' reservation wage leading to a more modest wage loss. It ultimately is an empirical question whether displaced workers who reside in denser markets experience more modest losses in employment and wages.

4.2.2 Gender Differentials and Education Differentials in the Returns to Agglomeration

The differences in agglomeration economies among subgroups of workers are driven by the theoretical micro-foundations that are internalised by the worker, not by the firm. For this reason, differences in agglomeration benefits among subgroups of workers are likely to be driven by improved job matching and learning of workers instead of improved sharing of risk and resources by firms. The gender differentials in the returns to agglomeration are theoretically ambiguous. The agglomeration benefits could be lower for female workers than for male workers, as women generally work closer to home due to a difference in opportunity costs of commuting through financial and time constraints (White, 1977; Madden, 1981; White, 1986; Hanson and Pratt, 1988; Zax, 1991; Clark et al., 2003; Fernandez and Su, 2004; Roberts et al., 2011). Therefore, on average, women have a smaller LLM. In turn, smaller LLMs are associated with a less productive matching of jobs (Petrongolo and Pissarides, 2006; Di Addario, 2011), leaving women in less productive jobs and lower wages. In addition, women are characterised by a lower labour market attachment and a lower number of working hours, which may reduce the ability to internalise knowledge spillovers. Alternatively, female workers may not experience hindrance from relatively small LLMs, as they have less concentrated labour market opportunities (Madden and Chiu, 1990). Moreover, female workers could experience higher returns to agglomeration, as the benefit of locating in relatively dense areas with more job opportunities and better job matching diminishes the negative effect of having a relatively low willingness to commute and thus a small LLM (Phimister, 2005).

The role of education level in the returns to agglomeration is also theoretically ambiguous. High-educated workers are likely to have higher returns to agglomeration, as larger markets are characterised by lower search costs. Lower search costs enable high-productive firms to conduct more efficient searches for high-productive workers, which generate better job matches and leaves low-educated workers in less productive jobs (Wheeler, 2001). Moreover, high-educated workers might be more able than low-educated workers to exploit knowledge spillovers, because they have the capacity to do so (Rosenthal and Strange, 2008). Alternatively, high-educated workers can have lower returns to agglomeration if larger markets are characterised by a strong amenity advantage for which high-educated workers have a higher marginal willingness to pay.[6] The higher willingness to pay will increase the labour supply of high-educated workers and decrease their wages (Black et al., 2009). Moreover, low-educated workers might be more able than high-educated workers to learn from human capital in close proximity, as they have a greater potential left to benefit (Rosenthal and Strange, 2008). In the literature, the results are mixed. Several studies find that the returns to population density increase with the number of years of education (Wheeler, 2001; Rosenthal and Strange, 2008; Bacolod et al., 2009; Carlsen et al., 2016). Gould (2007) argues that white-collar workers benefit more from working in larger LLMs than blue-collar workers. Other studies find that the returns to population fall with education level (Adamson et al., 2004; Di Addario and Patacchini, 2008; Black et al., 2009; Lee, 2010).

4.2.3 Regional Aggregation and the Relevance of Endogenous Local Labour Markets

For economic applications in the field of economic geography, the spatial scale of LLMs is operated by using a specific regional classification. So far, most studies operate LLMs by using pre-defined "exogenous" regional classifications. Examples of exogenous regional classifications include the U.S. counties (Hoynes, 2000), the U.S. Standard Metropolitan Statistical Area (MSA) (Glaeser and Maré, 2001), the U.S. commuting zone (CZ) (Autor et al., 2013), European NUTS areas (Ciccone, 2002), and UK Travel-to-Work-Area (TTWA) (Petrongolo and Pissarides, 2006; Manning, 2009; Faggio et al., 2016). Although many studies in the field of economic geography use various regional classifications, it is surprising that they generally neglect the role of regional aggregation in the returns to agglomeration. A notable exception is the paper by Briant et al. (2010), who focus on the Modifiable Areal Unit Problem (MAUP). They examine the importance of the size and shape of regional areas for estimates of, among others, the UWP. The authors use six regional classifications and argue that specification issues are the most important to limit biases, and the heterogeneity in the shape of LLMs is of higher-order importance.

[6] Note that the amenity advantage from denser markets especially holds for European cities (Brueckner et al., 1999). In several U.S. metropolitan areas, disamenities from denser areas, e.g., crime and pollution, may outweigh positive amenities.

Interestingly, the authors show that heterogeneity in the size of LLMs is more relevant for the returns to agglomeration than the shape of LLMs: the use of small zoning systems instead of large zoning systems lowers the estimate of the UWP by a magnitude of two. Other studies also reveal stronger agglomeration spillovers when a regional classification at a relatively high level of regional aggregation is used to operate LLMs (e.g., see Ellison and Glaeser (1997) and Rosenthal and Strange (2001)).

The advantage of using a pre-defined exogenous regional classification is that within-country differences in economic outcomes can easily be investigated while research outcomes remain comparable across studies and through time. However, the validity of using exogenous regional classifications is questionable for two reasons. First, the exogenous regional classifications are based on old definitions and are potentially outdated. Over the last decades, the labour force composition experienced changes caused by technological, demographic and economic shifts on a global level. These shifts led to a decrease in the demand for low-educated or average-educated workers and greater wage dispersion (Fernandez, 2001; Goos et al., 2009, 2014), a rise in women's labour force participation (Costa, 2000) and an increase in commuting from home to workplace (Van der Laan and Schalke, 2001; Crane, 2007). Second, the exogenous regional classifications represent non-overlapping areas in the sense that they only vary between areas and not among individuals within areas. A recent paper on the effective size of LLMs is by Manning and Petrongolo (2017), who aim to limit mismeasurement of UK workers' LLM by using a continuous nature of geographic space that allows for overlapping LLMs of two workers who reside in an administratively different but geographically close location. They develop a job search across space framework, in which the worker's size of the LLM depends on the cost of distance. Importantly, Manning and Petrongolo (2017) do not distinguish between different types of workers and model low-skill labour markets that tend to be relatively local, as their data sample contains workers with a relatively low education level.

We define LLMs endogenous to demographic characteristics and commuting outcomes, which allows us to operate workers' LLM at various levels of regional aggregation for different subgroups of workers. By allowing for differences in workers' LLM when they meet different characteristics, we provide an alternative view of overlapping LLMs. We argue that endogenising workers' LLM to worker characteristics is relevant, as theory suggests that the LLM structure is worker-specific.

The LLM structure differs among workers, as workers face differences in utility functions and opportunity costs. An example would be that women put a higher value on working closer to home than men do (White, 1977; Roberts et al., 2011). The gender difference in valuation of leisure and commuting can be explained by various reasons (Fernandez and Su, 2004), including women's dual role as a mother and worker, involving responsibilities for the household and children (Madden, 1981; Crane, 2007). Alternatively, more educated workers have thinner labour markets and are therefore characterised by a higher commuting distance (Manning, 2003).[7] Workers in thinner and larger labour markets are likely to benefit more in terms of wages from longer commuting, as the wage offer curve is raised when the skill-level rises (White, 1988). Consequently, high-educated workers face lower opportunity costs of commuting and work further away from home. Both examples suggest that the LLM structure varies in different subgroups of workers. In turn, differences in the LLM structure among workers may lead to differences in the returns to agglomeration.

4.2.4 Conceptual Setting

The model shown in (4.1) is specified to display the implications of the level of regional aggregation for estimates of the returns to agglomeration.

$$\text{Returns to Agglomeration} = \alpha + \beta \times \text{Regional Aggregation} \tag{4.1}$$

The parameter α represents the baseline returns to agglomeration. The role of the regional aggregation in the returns to agglomeration is represented by β. The parameter β equals zero if estimates of the returns to agglomeration do not depend on the level of regional aggregation. The vast majority of the literature that examines the returns to agglomeration focuses on the estimation of α and implicitly assumes that β equals zero. Also, the literature examines whether there are gender- and education differentials in the level of the returns to agglomeration α. Notably, the

[7]In contrast, other literature argues that high-skilled workers have a relatively high value of time and are therefore characterised by shorter commutes (e.g., see Brueckner et al. (2002)). In the context of the Netherlands, we argue that the mechanism of thinness on the labour market outweighs the mechanism of a higher valuation of time. This observation is consistent with the empirical finding that commutes and education are positively correlated.

literature generally neglects the role of β in the returns to agglomeration. We hypothesise that β is not equal to zero. For example, estimates of the returns to agglomeration could be increasing in the level of regional aggregation, if agglomeration economies are more prevalent at a higher spatial scale. We contribute to the literature by examining the role of regional aggregation in the returns to agglomeration, by examining whether the role of regional aggregation differs among subgroups of workers, and by examining whether subgroups of workers are characterised by a different level of regional aggregation to represent the structure of their LLM.

4.3 Background, Data and Flowbca

4.3.1 Background on the Dutch Regional Classifications

In the Netherlands, the COROP regional classification, defined in 1971, was set out to identify economically and socially integrated areas (CBS, 2018). COROP literally stands for the Coordination Commission Regional Research Programme (in Dutch: Coördinatiecommissie Regionaal Onderzoeksprogramma). The COROP classification is equivalent to the European concept of NUTS 3 areas and comparable to the U.S. concept of Commuting Zones and the UK concept of Travel-To-Work-Areas. The COROP areas (hereafter: NUTS 3 areas) were defined based on journey-to-work and place-of-work statistics that reflected the typical commuting outcomes of Dutch employed workers. In total, there are forty NUTS 3 areas: each NUTS 3 area consists of a core and hinterland area, while the provincial borders are never crossed. Besides administrative units such as provinces and municipalities, Statistics Netherlands and the Dutch government use the forty NUTS 3 areas and the thirty-five public employment services (PES) areas for analytical and political purposes. We use the 398 Dutch municipalities, forty NUTS 3 areas and thirty-five PES areas as reference sets of LLMs, which facilitate a comparison to the alternative sets of defined aggregate LLMs and subgroup-specific LLMs.

4.3.2 Data Sets

We used various administrative micro data sets, retrieved from Statistics Netherlands, covering the period of 2006 to 2014. The micro data sets contain data of individuals, households and firms. The data set Work

Location Register (*Gemstplbus*) was used to incorporate data on the geographical employment location of employees at the municipality level. We used a set of 403 distinct Dutch municipalities that existed in 2014. For the sake of convenience, we removed five municipalities that represent the small and isolated Wadden Islands in the northern part of the Netherlands. The work location is observed annually in December. The Population Register (*Gbapersoontab*, *Gbahuishoudensbus*, *Gbaburgerlijkestaatbus*, *Gbaadresgebeurtenisbus*), which is based on municipal and tax office administration, was used to incorporate data on individuals' date of birth, gender, marital status, number of household members and changing home. We removed observations of workers who were aged below 18 or over 65 years. The Address Object Register (*Gbaadresobjectbus*, *Vslgwbtab*) was used to incorporate data on individuals' home address and location at the municipality level. The Job and Wages Register (*Polisbus*), which is based on income statements of employees to the tax office administration, was used to incorporate data on the type of job (full-time or part-time), type of contract (fixed or temporary), economic sector, number of hours worked and gross wage. We removed observations of workers who were employed less than 0.8 full-time equivalent or 128 hours a month, to make the labour market outcomes of workers who differ in especially gender more comparable. Moreover, we removed observations of workers who earned an hourly wage lower than 3 euro. The Main Job Register (*Hfdbaanbus)* was used to select the main job of the worker, which is the job with the highest annual wage. The Bankruptcy Job Endings Register (*Failontslagtab)* was used to incorporate data on the worker, firm and date of workers' job displacement due to firm bankruptcy. The Highest Education Register (*Hoogsteopltab*) was used to incorporate data on workers' highest level of attained education. The highest level of attained education contains three groups, i.e. low, average and high educational attainment. This categorisation is based on the International Standard Classification of Education (ISCED) and corresponds to lower, secondary and tertiary education, respectively.

4.3.3 Key Variables and Covariates

The key dependent variables include hourly wage and employment. The worker's hourly wage was constructed by taking the natural logarithm of the monthly contractual gross wage relative to the number of contractual hours worked per month. Note that for the urban wage premium annual data set, we constructed workers' hourly wage of the month of December. Thereby, the hourly wage and commuting distance were constructed based on data of the same job of the month of December. The job displacement data set contains monthly data. The worker's employment status was represented by a zero-one indicator variable that equals 1 if the worker is employed, and zero otherwise. The key independent variables can be divided into two sets.

The first set of key independent variables was used to construct the aggregate LLMs and subgroup-specific LLMs, containing a cross-section of commuting flows across municipalities in the year 2014.[8] This set of variables was used for the descriptive analysis. For convenience, we used the cross-section of flows in the year 2014, as the number of distinct municipalities decreased in the period 2006 to 2014. Unfortunately, we were not allowed by Statistics Netherlands to export commuting flows across municipalities between 1 to 9 workers. These flows represent about 3 per cent of the total number of flows and were omitted. Aggregate LLMs were defined based on a set of commuting flows across municipalities of all workers together.[9] The subgroup-specific LLMs were defined using separate sets of commuting flows for workers who differ in gender and education.

The second set was used to approximate agglomeration spillovers and consists of variables that represent the natural logarithm of employment density and the natural logarithm of area size. This set of variables was used for the empirical analysis. Workers' employment density was constructed by taking the number of employed workers in the LLM relative to the area size in kilometres of the LLM. Various regional classifications were used to represent the worker's LLM, including the

[8] We examined the temporal changes in the sets of commuting flows over the period 2006 to 2014, which were relatively small. For the sake of convenience, we use time-invariant LLMs.

[9] Unfortunately, the work location is not entirely consistent as Statistics Netherlands has only data on the number of firm plants, the plant locations and the number of employees at each specific plant. The individual's work location is imputed by Statistics Netherlands using data on the place of work and place of home. Each individual is, based on the home location, linked to the closest firm plant, conditional on not exceeding the number of individuals who were employed at that specific plant. Hence, the amount of commuting interaction between municipalities is likely to be underestimated, in particular for subgroups who are characterised by relatively large LLMs. Consequently, the variation between subgroups in the size of the LLM is also likely to be underestimated.

Dutch municipalities, NUTS 3 areas, PES areas, aggregate LLMs and subgroup-specific LLMs. For a given worker, each regional classification gives different values of the employment density and area size. For a specific number of distinct aggregate LLMs, the employment density and area size differ between the LLMs, but not between workers who reside in the same LLM.[10] For a specific number of distinct subgroup-specific LLMs, the employment density and area size may differ between workers if they reside in the same LLM but vary in gender or education level.

A set of covariates that was used for the empirical analysis contains zero-one indicator variables that represent female, highest attained education (low, average and high education), Dutch nationality, age (18-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-55, 55-60 and 60-65 years), having children aged 18 or lower, having a partner, number of household members (1, 2, 3-4 and more than 4 members), economic sector of the firm (66 categories), the size of the firm (1-9 employees, 10-49 employees, 50-99 employees, 100-499 employees and more than 499 employees), job tenure (3-6, 6-12, 12-18 and over 18 years) and year of job displacement (2007, 2008, 2009, 2010 and 2011). Note that the variables job tenure and displacement year are only used in the empirical analyses on the returns to agglomeration for workers who have been displaced.

4.3.4 Flow-Based Cluster Algorithm

We use flowbca, discussed by Meekes and Hassink (2018b) and extending earlier work by Duranton (2015), which is a flow-based agglomerative hierarchical cluster algorithm that is able to cluster regional units into meaningful self-contained LLMs. Flowbca is flexible and able to define LLMs for different subgroups of workers at various levels of regional aggregation. From a theoretical point of view, the functional criterion to pair two regional units into one depends on the level of interaction. In our analysis, the level of interaction between regional units is approximated by relative commuting flows from residence to workplace. The main

[10]For the year 2014, the minimum, maximum, median and average number of employed workers in a NUTS 3 area (40 distinct units) equals 13,960, 753,749, 109,372, and 170,759 workers, respectively. For the year 2014, the minimum, maximum, median and average number of employed workers in a PES area (35 distinct units) equals 52,194, 722,819, 141,689, and 195,153 workers, respectively. For the year 2014, the minimum, maximum, median and average number of employed workers living in an aggregate LLM (35 distinct units) defined by our algorithm equals 9,452, 1,576,821, 77,836 and 194,855 workers, respectively.

input for the algorithm is a set of commuting flows across municipalities. Alternative sets of aggregate LLMs were constructed at a number of distinct LLMs between 398 and 7. A higher level of regional aggregation leads to fewer distinct LLMs. Subgroup-specific LLMs were defined by separately using commuting flows of subgroups of workers, which include groups of both female workers and male workers, varying in three education levels.

The algorithm that we used to define LLMs can be described as follows. LLMs were defined by iteratively aggregating two regional units into one. A higher number of iterations implies a higher level of regional aggregation. In each iteration, the algorithm selects two units that will be aggregated based on an optimisation function. The optimisation function identifies the maximum relative commuting flow out of all bilateral commuting flows. The "source" unit from which the largest relative commuting flow starts is aggregated to the "destination" unit. This process is repeated until a stopping criterion is met. Examples of stopping criterions are if a specific number of distinct LLMs has been defined or if a specific average level of self-containment has been met. After the algorithm is terminated, the level of self-containment of an LLM is approximated by its population weighted local employment rate. The population weighted local employment rate is constructed by dividing the total number of workers who work local – i.e. those who work and live in their LLM – by the total number of employed workers. A higher local employment rate implies a stronger connectivity within the LLM and a weaker connectivity to outside LLMs. See the Chapter 2 for more information about flowbca.

Two limitations of the algorithm require additional attention. First, Dutch municipalities that are relatively large in terms of population, for example Amsterdam and Utrecht, dominate the absolute number of outgoing flows to other municipalities. This leads to a situation where the algorithm aggregates the larger units to smaller destination units. Consequently, the larger regional units will not be defined as the core of an LLM. To overcome this limitation, we used relative flows that function as weights to account for the relative importance of a unit. Second, if the interaction based on commuting flows across regional units is not sparse enough, the algorithm defines several relatively large LLMs and many small isolated LLMs. For a given LLM, this limitation could lead to the situation where multiple municipalities in the LLM are hardly connected to one another. Figure 4.7 shows that flowbca is able to define meaningful LLMs at various levels of regional aggregation.

4.4 Descriptive Results

We apply flowbca to define LLMs for various subgroups at different levels of regional aggregation. Commuting flows across municipalities are used as the main input for the algorithm to define LLMs. In the first part of the descriptive results, we document the changes in commute over the last decades and we explain why we focus on the six subgroups that vary in gender and education. In the second part of the descriptive results, we show the application of flowbca. We document to what extent the level of self-containment of a set of LLMs depends on the level of regional aggregation. Moreover, we visualise LLMs for workers who vary in gender or education level.

4.4.1 Commuting

We provide Table 4.1 to provide a better understanding of which worker characteristics explain the largest share of variation in workers' commuting distance. Table 4.1 displays the quantile regressions of commuting distance in kilometres on various worker characteristics. The 0.05, 0.25, 0.50, 0.75 and 0.95 quantile regression are provided in Columns (1) to (5), respectively. Table 4.1 shows that female workers and low-educated workers are characterised by a relatively short commuting distance. Moreover, Table 4.1 reveals that the estimates for gender and education, compared to other worker characteristics, are relatively economically significant. This observation holds in particular for the regressions of the 75th percentile and above. Also, the differences among the commuting quantiles is highest for gender and education. The difference in commuting outcomes among subgroups of workers suggest that subgroups are characterised by a different LLM structure. We particularly focus on gender- and education subgroups, because these worker characteristics explain the largest share of variation in commuting outcomes.

Table 4.1
Quantile regressions of commuting distance on worker characteristics.

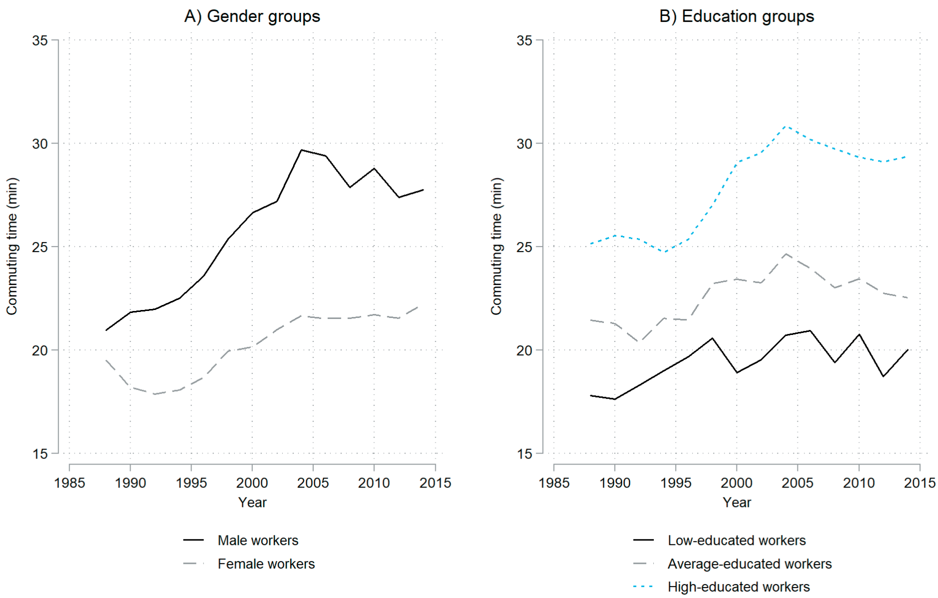
	Commuting distance (km)				
	q05	q25	q50	q75	q95
	(1)	(2)	(3)	(4)	(5)
FEMALE	-0.0614***	-0.3460***	-1.2343***	-2.8691***	-6.2561***
	(0.0066)	(0.0131)	(0.0309)	(0.0569)	(0.2389)
AVERAGE-EDUCATED	0.1264***	0.5218***	1.2518***	2.3795***	4.7847***
	(0.0095)	(0.0183)	(0.0407)	(0.0758)	(0.3389)
HIGH-EDUCATED	0.2982***	1.0884***	3.5444***	7.6499***	13.6635***
	(0.0101)	(0.0166)	(0.0554)	(0.0925)	(0.4143)
$25 < AGE \leq 30$ years	0.0345***	0.1091***	0.5168***	1.2896***	3.0203***
	(0.0117)	(0.0147)	(0.0316)	(0.0917)	(0.3867)
$30 < AGE \leq 35$ years	0.0998***	0.2356***	0.8445***	1.8653***	3.6084***
	(0.0095)	(0.0180)	(0.0481)	(0.0692)	(0.3601)
$35 < AGE \leq 40$ years	0.1556***	0.4728***	1.2471***	2.2744***	4.5206***
	(0.0112)	(0.0240)	(0.0440)	(0.1023)	(0.4297)
$40 < AGE \leq 45$ years	0.1472***	0.4714***	1.2491***	2.3537***	5.6296***
	(0.0151)	(0.0232)	(0.0623)	(0.0748)	(0.4839)
$45 < AGE \leq 50$ years	0.1233***	0.3591***	0.9420***	1.7746***	4.6354***
	(0.0146)	(0.0215)	(0.0456)	(0.1055)	(0.4929)
$50 < AGE \leq 55$ years	0.1078***	0.3371***	0.6551***	1.3828***	4.7094***
	(0.0199)	(0.0280)	(0.0678)	(0.1200)	(0.6200)
$55 < AGE \leq 60$ years	0.1094***	0.2407***	0.4272***	1.1703***	3.9197***
	(0.0165)	(0.0245)	(0.0480)	(0.1123)	(0.4151)
$60 < AGE \leq 65$ years	0.0637**	0.0758	0.1959*	0.7882***	5.2194***
	(0.0294)	(0.0512)	(0.1063)	(0.2138)	(0.9823)
DUTCH NATIONALITY	0.0934***	0.0714***	-0.0281	-0.1698***	-1.8535***
	(0.0108)	(0.0191)	(0.0438)	(0.0628)	(0.2567)
NO CHILDREN	-0.0062	0.0871***	0.1331***	0.3740***	1.5561***
	(0.0094)	(0.0163)	(0.0322)	(0.0550)	(0.3455)
PARTNER	0.0500***	0.2213***	0.3664***	0.2964***	-0.8860***
	(0.0083)	(0.0144)	(0.0352)	(0.0706)	(0.2888)
Number of observations	946,043	946,043	946,043	946,043	946,043

Notes: The dependent variable is the commuting distance measured in kilometres. Parameter estimates of the covariates are reported. Bootstrapped standard errors are in parentheses. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. The reference categories of *FEMALE*, *EDUCATED*, *AGE*, *NATIONALITY*, *NO CHILDREN*, *PARTNER*, consist of workers who are male, low-educated, aged between 20 and 25, have a non-Dutch nationality, children and no partner, respectively. The quantile regression analyses include indicator variables for the number of household members (3), firm economic sector (66), firm size (4), the NUTS 3 location of the household (39) and the calendar year (8). The period under observation is from 2006 to 2014. Data set: the administrative data from Statistics Netherlands. Sample: a five per cent random sample.

Figure 4.1 is the only figure in Chapter 4 that is not based on data retrieved from Statistics Netherlands. We use data from the Dutch SCP labour supply panel (in Dutch: *SCP Arbeidsaanbodpanel*) to observe differences in commuting over the last decades (SCP, 2015). Figure 4.1

shows that for men and women, the average commuting time increased in the period from 1988 to 2014. The increase in commuting time is the most severe for high-educated workers. Moreover, Figure 4.1 shows that workers' commuting time from place of residence to place of work differs among subgroups. Men, compared to women, and high-educated workers, compared to low-educated workers, commute longer. The change in commute over the last decades indicates that regional classifications that have been defined a long time ago, for example the NUTS 3 areas, might be outdated. Moreover, the findings suggest that workers' LLM has become larger over the last decades.

Fig. 4.1. Subgroup-specific changes in the average commuting time over the period 1988 to 2014.



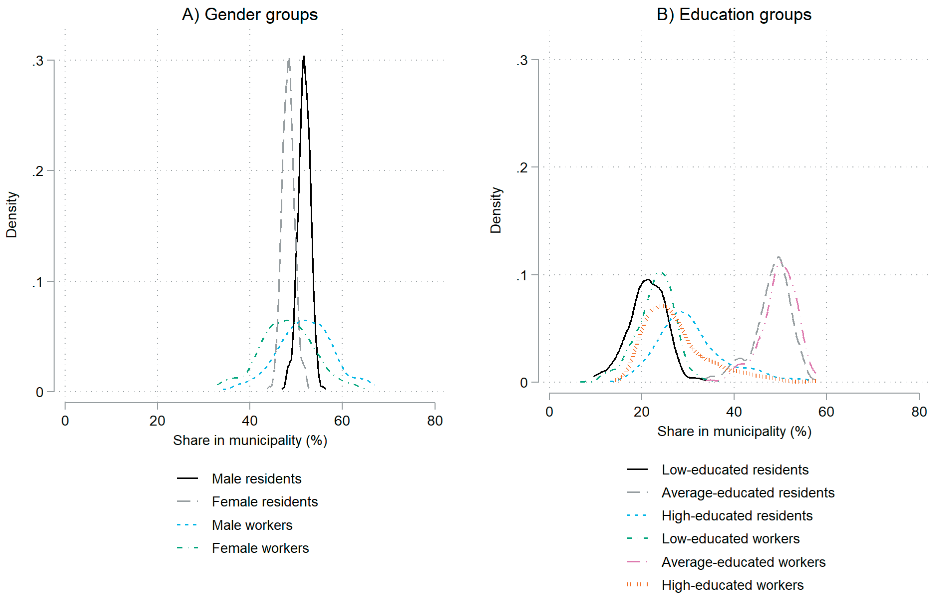
Notes: Data set: the SCP labour supply panel. Sample size: 41,275 observations.

Figure 4.2 shows the density plots of the gender shares (Fig. 4.2A) and education shares (Fig. 4.2B) across 398 municipalities. The shares are separately given for employed individuals in their home municipality and work municipality. Figure 4.2A provides us with several insights. First, there are on average more men than women in the sample. This observation can be explained by the fact that there are more men employed than women. Second, for both men and women, the distribution of workers is much wider than the distribution of residents. A wider distribution suggests higher concentration ratios in specific municipalities. Male and female workers are relatively concentrated in specific municipalities, but male and female residents are more evenly distributed across municipalities. This observation suggests that there exists substantial regional mismatch between the home location and employment location of both male and female workers.

The distribution of high-educated workers is relatively wide (see Fig. 4.2B), which implies that high-educated workers are more concentrated in specific municipalities than low-educated workers. Moreover, Figure 4.2B reveals that the distributions do not differ between residents and workers who belong to the identical education group. Hence, there is not much education-biased regional mismatch between home and employment locations. However, the differences in the concentration ratios between education categories suggest that there is substantial education-biased sorting across municipalities.

Overall, in this subsection, we have shown that workers' gender and education explain the largest share of variation in commuting distance. Moreover, we have shown that workers' commuting time has been increasing over the last decades. This finding underscores the relevance of defining LLMs with more recent data on commuting flows. In addition, the results suggest that there is substantial regional mismatch between workers' residence and work location for both women and men. Also, the results indicate substantial education-biased sorting of workers across regional areas. Our descriptive results motivate the use of subgroup-specific LLMs for workers who differ in gender and education, as these demographic characteristics are the most important for commuting outcomes.

Fig. 4.2. Distribution plot of gender and education shares across municipalities.



Notes: The gender and education shares are constructed by taking the subgroup-specific fraction, separately for residents and workers, in the municipality. The sample contains fractions for 398 distinct municipalities. Data set: the administrative data from Statistics Netherlands.

4.4.2 Endogenous Local Labour Markets

In this subsection, we show how the aggregate LLMs and subgroup-specific LLMs, defined with flowbca, vary in the local employment rate. The local employment rate refers to the relative number of workers who live and work in their LLM. Moreover, we visualise LLMs of men and women separated by three education levels.

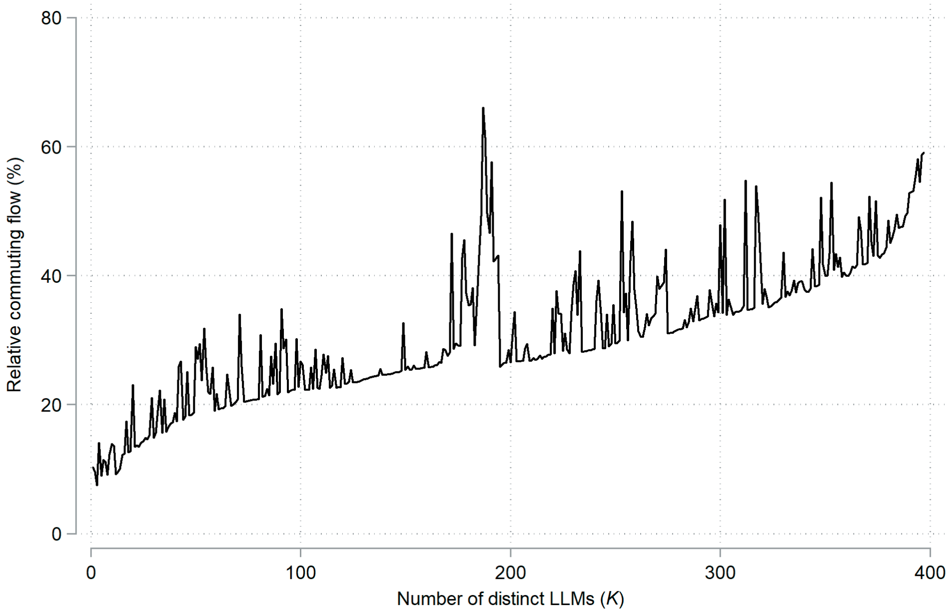
Figure 4.3 shows the maximum relative commuting flow in each iteration of flowbca. The algorithm that we used to define LLMs iteratively aggregates a regional unit to another regional unit, based on the maximum relative commuting flow out of all bilateral flows. The starting set of units contains 398 distinct municipalities. After each iteration of the algorithm, the number of distinct LLMs (K) decreases by one.

Figure 4.3 shows that the relative commuting flow at which units are aggregated is decreasing in the number of iterations. This observation holds as with fewer distinct LLMs there is more connectivity within a given LLM and less connectivity to outside LLMs. However, observe that the relative commuting flow at which units are aggregated is not uniformly decreasing in the number of iterations. This observation can be explained by the following example. Consider three regional units: A, B and C. Unit C has a relative flow of about 25 per cent to unit A and also to unit B. However, unit A is aggregated to unit B as the relative flow from A to B, which is the maximum of all relative flows, equals 30 per cent. After A has been aggregated to unit B, unit C will be aggregated to the combination of A and B, as C has a relative flow of 50 per cent to the new LLM that consists of A and B together.

Figure 4.4 shows the maximum relative commuting flow at which regional units were aggregated to construct the subgroup-specific LLMs for each of the six subgroups. Two observations are in place. First, when aggregating from about 10 to 100 distinct LLMs, women are characterised by a lower relative commuting flow than men. This observation suggests that women work closer to home than men. For a higher number of distinct LLMs, this distinction is less obvious. Second, high-educated workers have generally higher values of the relative commuting flows at which regional units are aggregated. This observation suggests that high-educated workers, compared to low-educated workers, work more often outside their LLM. Figure 4.4 suggests that the extent to which a regional classification reflects workers' LLM strongly depends on the worker's gender and education.

Figure 4.5 shows the population weighted average local employment expressed as a percentage, based on the aggregate LLMs, NUTS 3 areas and PES areas. The local employment rate represents the rate at which workers live and work in the same LLM. For the aggregate LLMs, the local employment rate varies over the number of distinct LLMs. Figure 4.5 shows that local employment decreases in the number of distinct LLMs. This is not surprising, as after two units are aggregated the workers who commute between the two aggregated units will work locally. It is important to note that the local employment rate of the aggregate LLMs is much higher, with an identical number of distinct LLMs, than that of the forty NUTS 3 areas and thirty-five PES areas. This observation suggests that the algorithm that is used to cluster regional units does relatively well in constructing self-contained regional areas of residence

Fig. 4.3. Relative commuting flow at which two units are aggregated.



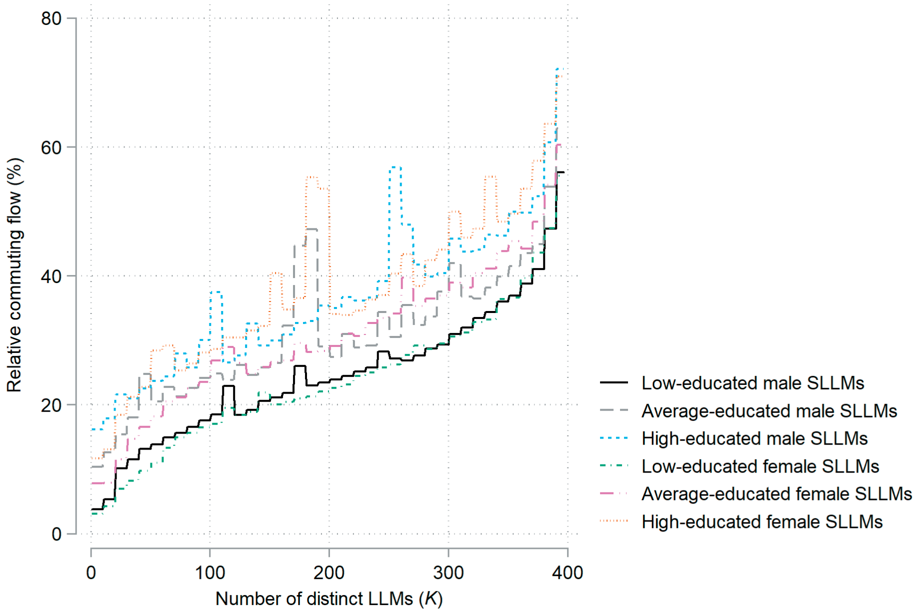
Notes: In each iteration, starting from a set of 398 distinct municipalities, the cluster algorithm selects the regional unit with the highest relative flow and aggregates the source unit to the receiving destination unit. The relative commuting flows are computed by taking each absolute commuting flow from source unit to destination unit relative to the source unit's total of absolute outgoing flows. In total 7,291,815 commuting flows were used. Data set: the Dutch administrative data from Statistics Netherlands.

and work activity.

Figure 4.6 reveals the extent to which the local employment varies over the number of distinct subgroup-specific LLMs. Both male workers and high-educated workers are characterised by lower local employment compared to female and low-educated workers, respectively. This observation suggests that male and high-educated workers are characterised by a relatively high commuting distance and a large LLM, which is consistent with the results in Table 4.1.

Figure 4.7 visualises the LLMs of male and female workers separated by the three educational groups. The stopping criterion of the algorithm

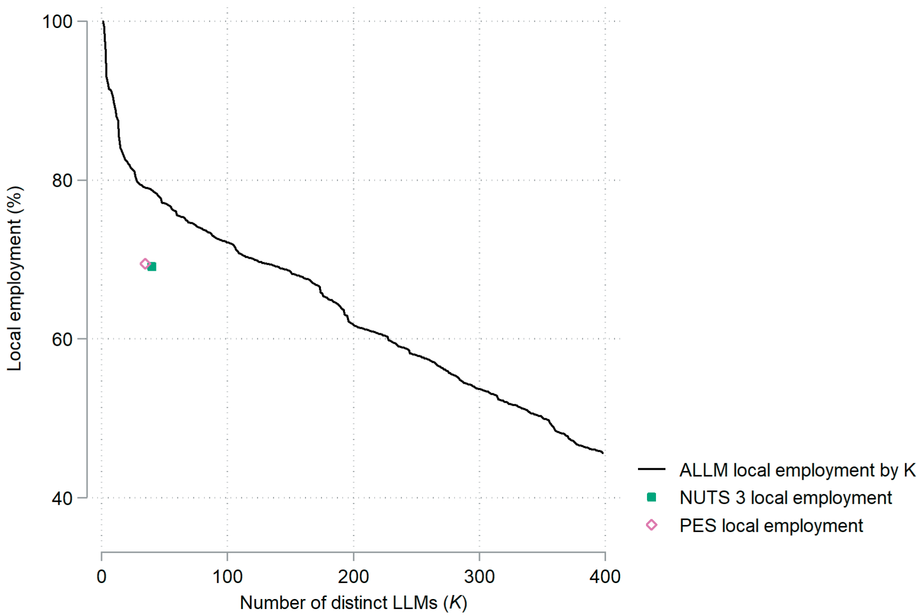
Fig. 4.4. Subgroup-specific relative commuting flow at which two units are aggregated.



Notes: The median of the relative commuting flows, in increments of ten, is given to smooth out the lines and to provide visible patterns. The values of the relative commuting flow, in each iteration, are available upon request. See Figure 4.3 for additional notes.

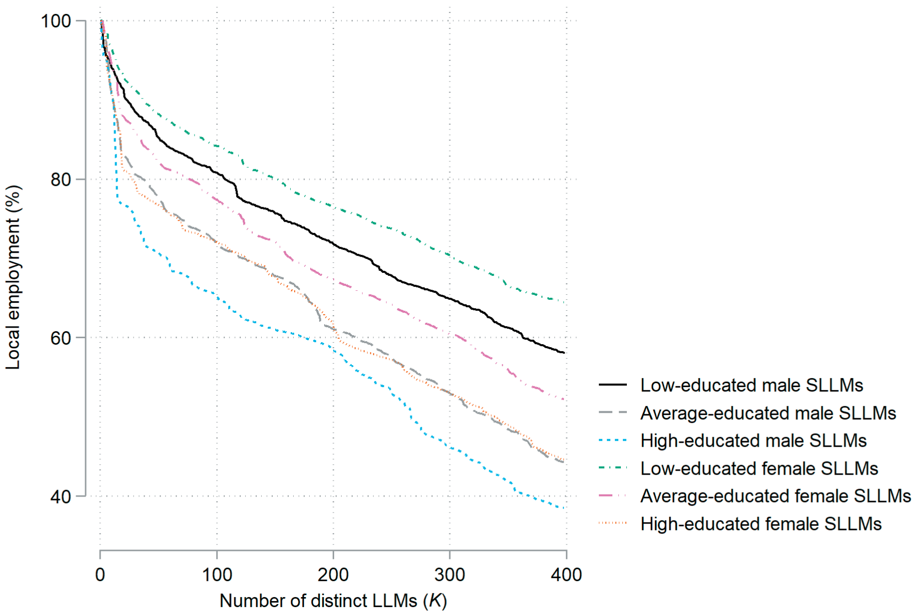
was set equal to a local employment rate of 80 per cent. That is, if 80 per cent of the workers live and work in their LLM, the algorithm is terminated. The stopping criterion of local employment equal to 80 per cent is arbitrary. Importantly, the differences in LLMs between subgroups of the population also hold for stopping criteria with other levels of local employment.

Fig. 4.5. Local employment by regional classification.



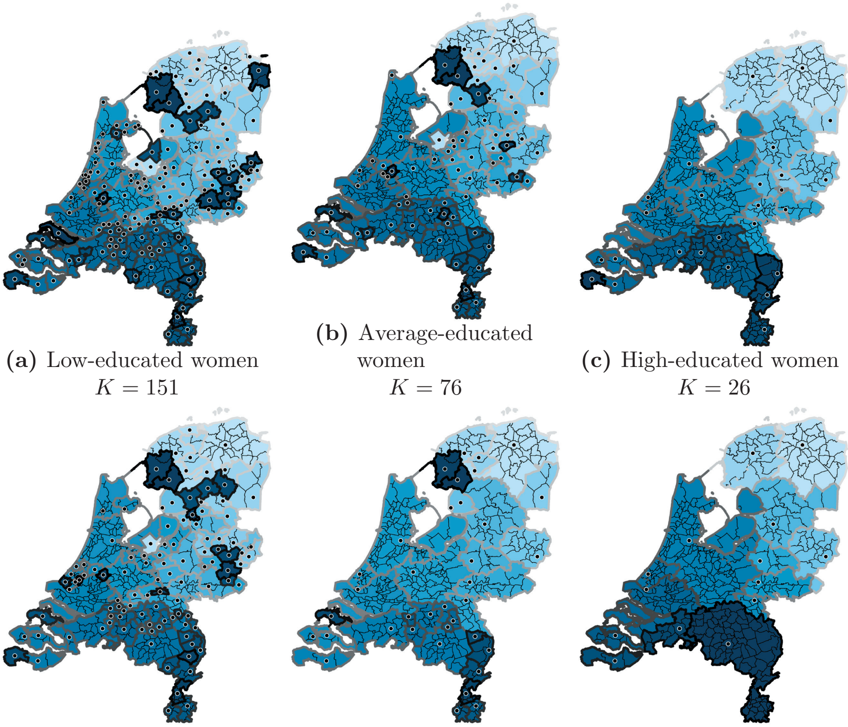
Notes: The local employment rate is constructed by taking the number of workers who live and work in their LLM relative to the total number of workers. See Figure 4.3 for additional notes.

Fig. 4.6. Subgroup-specific local employment by regional classification.



Notes: Local employment by subgroup and regional classification. See Figure 4.5 for additional notes.

Fig. 4.7. Subgroup-specific local labour markets.



(a) Low-educated women $K = 151$

(b) Average-educated women $K = 76$

(c) High-educated women $K = 26$

(d) Low-educated men $K = 107$

(e) Average-educated men $K = 36$

(f) High-educated men $K = 14$

Notes: The stopping criterion of the cluster algorithm is set to a minimum local employment rate of 80 per cent. The number of distinct LLMs is represented by K. The LLMs and its cores (the black dots with a white circle) are returned by flowbca. Each distinct LLM is surrounded by a thick border and highlighted by a colour. Data set: the Dutch administrative data from Statistics Netherlands.

Figure 4.7 shows that the number of distinct LLMs is decreasing in the education of workers and is lower for men. In this regard, high-educated workers and male workers are characterised by an LLM that is relatively less self-contained. This suggests that exogenous regional classifications are generally too large for low-educated and female workers, but too small for high-educated and male workers. Significantly, Figure

4.7 suggests that workers' education is more important for the LLM structure than workers' gender, as differences in the structure of LLMs are more pronounced between education levels.

Overall, our findings are relevant for research that focuses on quantifying regional differences in economic outcomes, as they suggest that the mismeasurement in workers' LLM strongly depends on the characteristics of the data sample. For example, the magnitude of mismeasurement in workers' LLM is very different for a data sample of women compared to a sample of men. The descriptive results in this subsection show that the extent to which a regional classification reflects a worker's LLM strongly depends on the worker's geographical location, gender and education. For this reason, we focus on the roles of aggregate and subgroup-specific LLMs in the returns to agglomeration in wages and employment. Moreover, we focus on gender- and education differentials in the agglomeration externalities.

4.5 Methodology

In this section, we provide the empirical models that we use for the two economic applications discussed in the empirical analyses. The two economic applications we consider include the estimation of the UWP and the impact of job displacement. After introducing the empirical models, we will discuss the main identification challenges that required our particular attention.

4.5.1 Urban Wage Premium

An empirical model, shown in (4.2), is specified to estimate the returns to agglomeration. The dependent variable is the natural logarithm of the hourly wage and the model is given as

$$w_{irt} = \delta J_{rt} + \beta' X_{irt} + \alpha_i + D_t + \varepsilon_{irt} \tag{4.2}$$

$i \in 1, 2, ..., N$; $r \in 1, 2, ..., R$; $t \in 2006, 2007, ..., 2014$

where subscripts i, r and t denote the worker, regional employment area and year, respectively. The main parameter of interest is referred to by δ, which measures the agglomeration benefits for wages by increasing either the local number of employed workers or the local employment density. Equation (4.2) presents a generic empirical model, which is estimated for both the OLS estimator (without the individual-specific fixed effects term α) and the FE estimator.[11] The model is separately estimated using various regional classifications. For each regional classification and regional aggregation level, the values of the variable employment density J in regional area r are different, as the area of the worker's LLM is different. The vector X represents a set of covariates, including demographic characteristics, job characteristics and the area size of the worker's LLM. We include the area size to capture the agglomeration benefits for wages by increasing the spatial extent of a regional area. Individual-specific fixed effects are referred to by α. Annual dummies are denoted by D. ε refers to the idiosyncratic error term.

4.5.2 Job Displacement

A generic empirical model is specified to estimate both the displacement effect on employment and the natural logarithm of hourly wage, respectively.[12] The empirical model is given as

$$\begin{aligned} Y_{irt} = \delta(\mathit{DISPLACED}_i \times \mathit{POST}_{it}) + \rho \mathit{POST}_{it} \\ + \beta' X_{it} + \alpha_i + N_r + D_t + \varepsilon_{irt} \end{aligned} \tag{4.3}$$

$i \in 1, 2, ..., N$; $r \in 1, 2, ..., R$; $t \in 1, 2, ..., 108$

where subscripts i, r and t denote the worker, regional home area and month, respectively.[13] Note that workers are distinguished by their home location instead of employment location, to prevent the problem where we would not observe a worker's geographical employment location during an unemployment spell. The displacement effects on the outcome variables are represented by parameter δ of the two-way (double) interaction term between the scalar indicator variables *DISPLACED* and

[11] Please note that in each of the specifications that are shown in (4.2), (4.3) and (4.4), all parameters refer to a different estimate.

[12] See Chapter 2 for more information on our quasi-experimental empirical design involving job displacement due to firm bankruptcy.

[13] Note that we use annual data for the analysis of the urban wage premium and monthly data for the analysis of job displacement. The time period under observation t for the job displacement data sample ranges from 1 to 108, which refers to January 2006 and December 2014, respectively. Note also that vector X contains a different set of covariates in the urban wage premium data sample and job displacement data sample, respectively.

POST. The time-constant variable *DISPLACED* equals one for workers who have been displaced, and zero otherwise. Note that the main effect of *DISPLACED* is taken care of by including individual-specific fixed effects. The time-varying indicator variable *POST* equals one for the post-displacement period of thirty-six months. The base and omitted reference categories of *DISPLACED* and *POST* are the non-displaced and the period prior to displacement, respectively. The worker's covariates, including demographic characteristics and job characteristics, are represented by vector X. The parameters of the covariates are referred to by vector β. Individual-specific fixed effects are represented by α. N_r represents indicators for the geographical home location at the NUTS 3, PES, aggregate LLM or subgroup-specific LLM level. The aggregate LLMs and subgroup-specific LLMs are returned by flowbca. Calendar month indicators are denoted by D. ε refers to the idiosyncratic error term.

We added various interaction terms to assess the role of agglomeration economies, which is represented by employment density J, in the displacement effects on employment and hourly wage. The empirical model in (4.3) complements the model in (4.2) by adding various three-way (triple) and two-way interaction terms among employment density J, *DISPLACED* and *POST*. Moreover, we included interaction terms among a vector of worker characteristics X, *DISPLACED* and *POST*. The vector X includes time-varying variables (area size and other characteristics) as well as time-invariant variables (female, education and other characteristics of the terminated job). The empirical model is

$$\begin{aligned}
Y_{irt} = {} & (\theta J_{irt}) \times \mathit{DISPLACED}_i \times \mathit{POST}_{it} \\
& + (\iota J_{irt}) \times \mathit{DISPLACED}_i + (\nu J_{irt}) \times \mathit{POST}_{it} \\
& + (\kappa' X_{irt}) \times \mathit{DISPLACED}_i \times \mathit{POST}_{it} \\
& + (\gamma' X_{irt}) \times \mathit{DISPLACED}_i + (\eta' X_{irt}) \times \mathit{POST}_{it} \\
& + \delta \mathit{DISPLACED}_i \times \mathit{POST}_{it} + \rho \mathit{POST}_{it} \\
& + \mu J_{irt} + \beta' X_{irt} + \alpha_i + N_r + D_t + \varepsilon_{irt}
\end{aligned} \tag{4.4}$$

where the main parameter of interest is represented by θ. The parameter θ measures the role of employment density in the displacement effects on the dependent variable.

4.5.3 Identification Challenges

In our study on the returns to agglomeration in wages and employment, three identification challenges required particular attention. The challenges include the MAUP[14], individual-level endogeneity in employment density and local-level endogeneity in employment density.

The first identification challenge concerns the MAUP (Fotheringham and Wong, 1991; Burger et al., 2008; Briant et al., 2010). The MAUP relates to the issue that results and conclusions of empirical analyses are sensitive to the operation of the size, structure and location of regional units. In the literature on agglomeration economies, a wide range of regional classifications are used to operate the worker's LLM. The type of regional classification that is used is important, as it affects the values of variables that approximate the degree of agglomeration – represented by the employed relative to the area size – or the degree of tightness – represented by the vacancies relative to the unemployed. The worker's employment density is the mean of the true size, given that the classification represents the LLM of a "typical" worker. Under a random (classical) measurement error, the mismeasurement leads to a parameter estimate attenuated towards zero. However, the mismeasurement in workers' LLM structure might be non-random. Specifically, there could be a worker-specific component in the structure of workers' LLM, as workers who live in the same neighbourhood are not likely to have identical LLMs. For example, low-educated workers are likely to have a smaller LLM than the mean of the true size, whereas high-educated workers are likely to have a larger LLM. Under a non-classical measurement error, the degree of attenuation is lower and the mismeasurement could lead to a sign reversal of the estimated coefficient. We assess the implications of this identification challenge by using flowbca to define aggregate and subgroup-specific LLMs, which allows workers' LLM to depend on their commuting outcomes, gender and education level.

The second challenge concerns the endogeneity in employment density at the individual level, which is caused by non-random location choices of workers. For example, unobserved characteristics like ability or reasons not related to agglomeration economies might affect the location choice and labour market outcomes (Combes et al., 2012; Combes and Gobillon, 2015). We limit the potential bias from individual-level endogeneity by controlling for many factors that affect location, home change and em-

[14] See Appendix 4.A for a detailed discussion of the implications of the MAUP.

ployment decisions. For example, education level and changes in age are included to control for regional sorting based on skill and age. Moreover, we included individual-specific fixed effects to control for other potential confounding effects of time-constant variables such as abilities and knowledge other than education. In addition, for the empirical analyses on the returns to agglomeration in wages and employment following job displacement, we used a quasi-experimental design involving job displacement. This design is useful to examine the returns to agglomeration, as job displacement results in a non-culpable and unforeseen negative employment shock. By using this design, we remove potential confounding effects on post-unemployment outcomes caused by heterogeneity in the hazard rate into unemployment, signalling value, advance notification and severance pay. Moreover, the use of job displacement ensures a low number of residential relocations, because workers, following a negative employment shock, are less able or willing to relocate home (Meekes and Hassink, 2019). Thereby, our quasi-experimental design limits the problem of sorting across regional areas based on job or wage offers (Mion and Naticchioni, 2009). We compare the labour market outcomes of displaced workers with the outcomes of a control group that consists of comparable but non-displaced workers. We applied coarsened exact matching that makes displaced workers and non-displaced workers balanced in observables (Iacus et al., 2011).[15] Consequently, the selection bias into displacement based on observables is greatly reduced. The identifying restriction rests on whether displaced and non-displaced workers have parallel trends in the outcome variables prior to the month of actual (for the matched displaced) and potential (for the matched non-displaced) job displacement. In Figure 4.D.1 of Appendix 4.D we show that our design satisfies this restriction.

The third challenge concerns endogeneity at the local level, which is caused by aggregate missing variables. Location choices of firms and

[15] The displaced workers are matched to non-displaced workers in the specific month of the job displacement. For the displaced and non-displaced, this month will be referred to as the actual and potential month of job displacement, respectively. The default set of matching variables we used in the process of matching consists of the following variables: indicator variables for gender, age (21-30; 30-35; 35-40; 40-45; 45-50 and 50-59 years), children aged 18 or lower, partner, Dutch nationality, LLM-specific geographical home location, type of job (full-time or part-time), type of contract (fixed or temporary), job tenure (3-6; 6-12; 12-18 and over 18 years), firm size (10-49; 50-99; 100-499 and 500 or more employed workers), economic sector of the firm (twenty-one ISIC sectors), calendar month and calendar year.

workers can be affected by local productivity and local wage levels, or by differences in production and consumption amenities. For example, the more productive firms may self-select into denser LLMs. In this situation, wage premiums cannot be attributed to positive agglomeration spillovers, but are explained by a higher productivity of firms. A way to control for this endogeneity issue is to include location-specific fixed effects. Unfortunately, there is not sufficient within-individual variation in locations for all sets of LLMs. Hence, to allow for a comparison between the results of the OLS and FE estimator, we have not included the location-specific fixed effects in the empirical analyses on the UWP. See Combes and Gobillon (2015) for a discussion on several other reasons why including location-specific fixed effects does not work out well. Importantly, Combes and Gobillon (2015) argue that the issue of local-level endogeneity is less relevant than individual-level endogeneity, as the estimated effect of employment density changes much more when individual-specific fixed effects are included instead of location-specific fixed effects. As a robustness check, we apply a two-step procedure, in the spirit of Combes et al. (2008), to correct for location-specific differences in amenities, productivity and housing prices over time. Combes et al. (2008) introduce the two-step procedure to limit correlations between local-time unobservables and individual covariates. Thereby, the identification challenge involving endogenous sorting of worker quality across regional areas based on wage expectations would be addressed, and home change to a given LLM when expecting a high wage is no longer a source of bias. However, endogeneity would still be an issue if the location choice of a worker depends on the exact wage offered and obtained. See Appendix 4.B for the application of the two-step procedure.

4.6 Empirical Results

4.6.1 Urban Wage Premium

We examine the agglomeration effects on wages (see Eq. (4.2)). Figure 4.8 shows the results of the regressions of the natural logarithm of hourly wages on employment density, demographic characteristics and job characteristics. Figure 4.8A and Figure 4.8B display the results of the OLS and FE regressions, respectively. The estimates are provided for various sets of aggregate LLMs, in which employment density and area size varies by the number of distinct LLMs (K). A lower number of distinct LLMs implies a higher level of regional aggregation. The estimates of the UWP based on the NUTS 3 classification and PES classification, which contain 40 and 35 distinct areas, respectively, are also provided. These estimates do not depend on the number of distinct regional units, but allow for a point of comparison. Note that when K is equal to 398, the regional classification that is used to operate employment density is identical to the set of Dutch municipalities.

Figure 4.8A and Figure 4.8B show that the estimate of the UWP directly decreases in the number of distinct LLMs. For the entire interval of K, the OLS estimates of the UWP ranges between 2.6 and 6.6 per cent. More urbanised LLMs are characterised by a substantial UWP: if the employment density doubles, the increase in wages is about 2.6 to 6.6 per cent. This finding is consistent with those reported by Groot et al. (2014), who also use Dutch data, and find a UWP of 2.1 and 4 per cent using municipalities and NUTS 3 areas to operate LLMs, respectively. Groot et al. (2014) find higher estimates for the instrumental variables estimator. However, under the non-classical measurement error, IV estimates can be biased upward (Hyslop and Imbens, 2001). Wheeler (2001) finds a UWP of 2.7 per cent by using the logarithm of population density as the independent variable of interest at the U.S. MSA level. The UWP estimate is higher in studies that use a dummy variable to differ between urban and rural areas. For example, the studies by Glaeser and Maré (2001) and Yankow (2006) find that American urban workers earn about 25 or 19 per cent more than American rural workers, respectively. Using UK data, D'Costa and Overman (2014) find a UWP of 8.4 per cent.

Following the literature, we also estimate the UWP controlling for individual-specific fixed effects. Individual-specific fixed effects limit the potential of individual-level endogeneity, which is driven by sorting more able workers into larger LLMs. Our FE estimates of the UWP range from 0.3 to 1.4 per cent. The reduction in the UWP by introducing individual-specific fixed effects is consistent with the literature. After including fixed effects, Glaeser and Maré (2001), Yankow (2006) and D'Costa and Overman (2014) find a UWP of 10.9, 5 and 2.3 per cent, respectively.

Fig. 4.8. Aggregate LLM UWP by empirical specification (Eq. (4.2)).

Notes: The dependent variable is the natural logarithm of hourly wage. Each estimate of the natural logarithm of employment density on hourly wage represents a different regression. In each regression, the variables employment density and area size are operated based on a different regional classification. The values at which K is used to operate the aggregate LLMs include 398, 350 to 50 in increments of fifty, 45 to 25 in increments of five, and 22 to 7 in increments of three. K equals 40 and 35 for the NUTS 3 classification and PES classification, respectively. The 95% confidence intervals are constructed using clustered standard errors by LLM. All regression analyses include indicator variables for the worker's gender, education category (2), age group (8), having the Dutch nationality, having a child, having a partner, economic sector of the firm (66), size of the firm (4), number of household members (3) and calendar year (8). The number of estimated parameters for each covariate is provided in parentheses. All regressions include a variable that represents the natural logarithm of the area size of the worker's LLM. The parameter estimates are not reported. Data set: the Dutch administrative data from Statistics Netherlands. The period under observation is from 2006 to 2014. The number of individual-year observations equals 18,882,294.

Observe that the OLS and FE estimates of the UWP based on the forty and thirty-five distinct aggregate LLMs are higher but not significantly different than the NUTS 3 and PES estimates, respectively.

The difference between the OLS and FE estimates in Figure 4.8 suggests that the role of time-constant unobserved heterogeneity in the UWP is substantial. By introducing individual-specific fixed effects, the potential of endogeneity based on omitted variables is more limited. However, an alternative explanation is that the individual-specific fixed effects amplify the attenuation bias (Griliches, 1977; Freeman, 1984; Chowdhury and Nickell, 1985; Griliches and Hausman, 1986), which shifts the line of the aggregate LLM estimates downwards. Hence, we consider the range of 0.3 to 6.6 per cent as the lower and upper boundary of the UWP in the Netherlands, respectively.

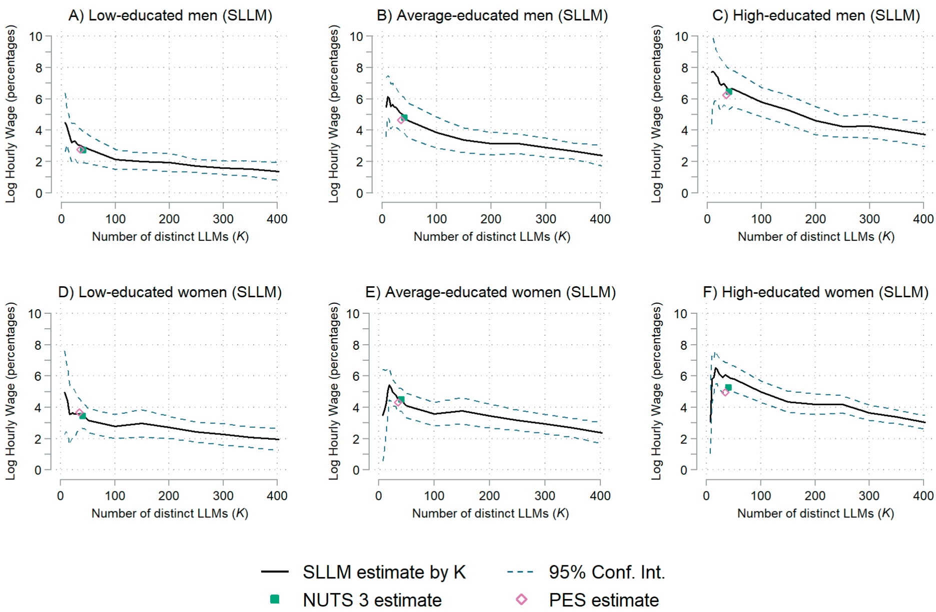
Figure 4.9 shows the UWP for various subgroups, in order to better understand the gender differentials and education differentials in the returns to agglomeration. The subgroup-specific LLMs are used to operate the worker's LLM.[16] In Figure 4.9 and Figure 4.10, graphs A-F consist of six different subgroups. Subgroups A-C and D-F, represent male and female workers, respectively. Subgroups A and D, B and E, and C and F, represent low-educated, average-educated and high-educated workers, respectively. The UWP estimates, using the subgroup-specific LLMs, NUTS 3 and PES regional classifications, reveal that the UWP increases in the attained education level. We find that the UWP is comparable for male and female workers. More importantly, for all subgroups the UWP increases in the level of regional aggregation. This finding suggests that a large share of the returns to agglomeration takes place at a relatively high spatial scale.

Figure 4.10 shows the FE estimates of the UWP for the aforementioned six subgroups.[17] Note on the y-axes that the FE estimates of the UWP are much smaller than the OLS estimates. Consistent with Figure 4.9, Figure 4.10 also shows that the UWP is increasing in the level of regional aggregation and workers' education level. Interestingly, we find weak evidence that the UWP for low-educated and high-educated female workers is underestimated when a pre-defined exogenous regional classification is used, as estimates of the UWP based on the NUTS 3 and PES classification are smaller than estimates based on the subgroup-specific

[16] See Appendix 4.C for the regression analyses using the aggregate LLMs. The estimates of the UWP using the subgroup-specific LLMs are comparable to the estimates of the UWP using the aggregate LLMs, but the difference between the estimate of the UWP among subgroups is slightly larger using the aggregate LLMs.

[17] See Table 4.C.3 for the coefficients and standard errors of the UWP based on FE estimates for the forty NUTS 3 areas and forty subgroup-specific LLMs, respectively.

Fig. 4.9. Subgroup-specific LLM UWP based on OLS estimates (Eq. (4.2)).

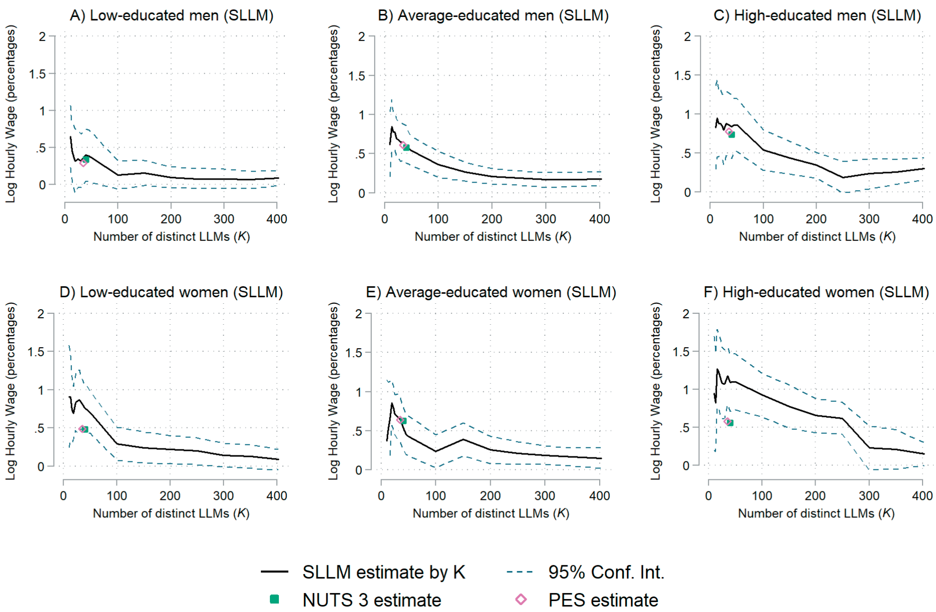


Notes: The employment density and area size of the subgroup-specific LLMs vary in gender and education level. The number of individual-year observations for the subgroups in Figures 4.9A-4.9F equals 2,296,052; 5,400,850; 4,479,115; 864,968; 2,643,962; 3,197,347, respectively. See Figure 4.8 for additional notes.

LLMs holding the regional aggregation level constant. This finding could be explained by the main input of these regional classifications, which include journey-to-work and place-of-work statistics that reflected the typical commuting outcomes of employed individuals that were predominantly male workers. Figure 4.10 suggests that the finding that men enjoy a larger UWP than women depends on the level of aggregation. This observation could explain the mixed evidence in the literature on gender- and education differentials in the returns to agglomeration.

We emphasise several findings. First, we show that at a higher level of aggregation, i.e. at a lower number of distinct LLMs, estimates of the UWP are higher. Moreover, we show that the role of spatial scale in the UWP is similar across the subgroups. It is significant that the descriptive results point out that the measurement error in a worker's

Fig. 4.10. Subgroup-specific LLM UWP based on FE estimates (Eq. (4.2)).



Notes: See Figure 4.8 and Figure 4.9 for additional notes.

LLM depends on the demographic characteristics, since the LLM size and structure depends on the worker's gender and education level. Together, these findings suggest that the MAUP is not the main driver behind the effect of regional aggregation on the returns to agglomeration. Instead, it seems that the micro-foundations of agglomeration economies are more prevalent at a relatively high spatial scale. Then, consistent with the paper of Briant et al. (2010), we find that the empirical specification to estimate the UWP is more important than the regional classification that affects the structure and size of workers' LLM. However, an alternative explanation for differences in estimates of the UWP after introducing individual-specific fixed effects is that the FE estimator amplifies the attenuation bias (Griliches, 1977; Freeman, 1984; Chowdhury and Nickell, 1985; Griliches and Hausman, 1986). Finally, the UWP is increasing in the level of education. We find no gender differential in the UWP. However, our descriptive results suggest that female workers are characterised by smaller LLMs than male workers. In this regard, the empirical analysis suggests that the UWP is overestimated for female workers if a pre-defined regional classification is used to operate LLMs, as pre-defined classifications generally consist of relatively large LLMs.

4.6.2 Job Displacement

We examine to what extent displaced workers' loss of employment and wages depend on the employment density of the LLMs where workers are located. Table 4.2, Table 4.3 and Table 4.4 present the displacement effects on employment and wages for low-educated workers, average-educated workers and high-educated workers, respectively (see Eq. (4.3)). Columns (1) and (2), and (3) and (4), show the displacement effects on employment and hourly wages, respectively. For the variables displacement status ($DISPLACED$) and post-displacement period ($POST$), the omitted categories are the non-displaced workers and the pre-displacement period, respectively.

Table 4.2, Table 4.3 and Table 4.4 show that displaced workers, compared with non-displaced workers, are 20 to 30 percentage points less employed over the post-displacement period of thirty-six months. The negative displacement effect on hourly wage ranges between 3 and 8 per cent. These findings are consistent with those reported in the job displacement literature (Schwerdt, 2011; Ichino et al., 2017). Interestingly, we show that male workers experience a more modest loss in employment than female workers. Moreover, low-educated workers experience a relatively large loss in employment. Also, high-educated workers experience relatively modest losses in both employment and wages. Using the three-way interaction models, we assess the role of local employment density in the displacement effects on employment and wages.

Table 4.2
Displacement effects on employment and hourly wage for low-educated workers (Eq. (4.3)).

	Low-educated men		Low-educated women	
	Employment (=1)	Hourly wage (log)	Employment (=1)	Hourly wage (log)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*	−0.2317***	−0.0715***	−0.3048***	−0.0309**
	(0.0090)	(0.0056)	(0.0206)	(0.0126)
Number of parameters	133	130	119	119
Number of individuals	4,725	4,725	997	997
Number of observations	259,875	224,453	54,835	46,287

Notes: Each column gives the dependent variable. Parameter estimates of the two-way interaction term are reported. Clustered standard errors by aggregate LLM are in parentheses. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. The reference categories of *DISPLACED* and *POST* consist of the non-displaced workers and pre-displacement period, respectively. The regression analyses include individual-specific fixed effects, aggregate home LLM fixed effects and indicator variables for *POST*, age (3), children aged 18 or lower, partner, the number of household members (3), and calendar month (107). Parameter estimates of the covariates are not reported. Data set: the Dutch administrative data from Statistics Netherlands. The period under observation is from January 2006 to December 2014, in which displaced and non-displaced workers are observed for 18 months prior until 36 months after the actual and potential month of job displacement, respectively.

Figure 4.11 reveals to what extent the displacement effects on employment and wages depend on the employment density of LLMs where workers are located (see Eq. (4.4)). The subgroup-specific LLMs are used to operate the worker's LLM. Note that when K is equal to 398, the regional classification that is used to operate employment density is the set of Dutch municipalities. Figure 4.11A shows an insignificant three-way interaction effect of employment density on the post-displacement employment probability. For the models in which employment density is operated using the NUTS 3 or PES regional classification, we find that doubling the employment density of a worker's geographical home location increases the loss in employment by about 1 percentage point. Figure 4.11B shows a positive and significant displacement effect of employment density on hourly wage at a relatively high spatial scale, which include a number of distinct LLMs equal to or lower than 16. Specifically, if the employment density in the geographical home location of displaced workers doubles, the post-displacement loss in wages is about 1 to 2 percentage points lower. The PES estimate of employment density on post-displacement wages is weakly significant and equals 1.4 percentage points. The results suggest that workers who become displaced in dense LLMs, compared to workers in more sparse LLMs, experience a modest loss in wages and an intermediate loss in employment.

Table 4.3
Displacement effects on employment and hourly wage for average-educated workers (Eq. (4.3)).

	Average-educated men		Average-educated women	
	Employment (=1)	Hourly wage (log)	Employment (=1)	Hourly wage (log)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*	−0.2003***	−0.0638***	−0.2754***	−0.0753***
	(0.0054)	(0.0040)	(0.0106)	(0.0078)
Number of parameters	148	148	131	130
Number of individuals	10,995	10,995	3,176	3,176
Number of observations	604,725	543,286	174,680	153,170

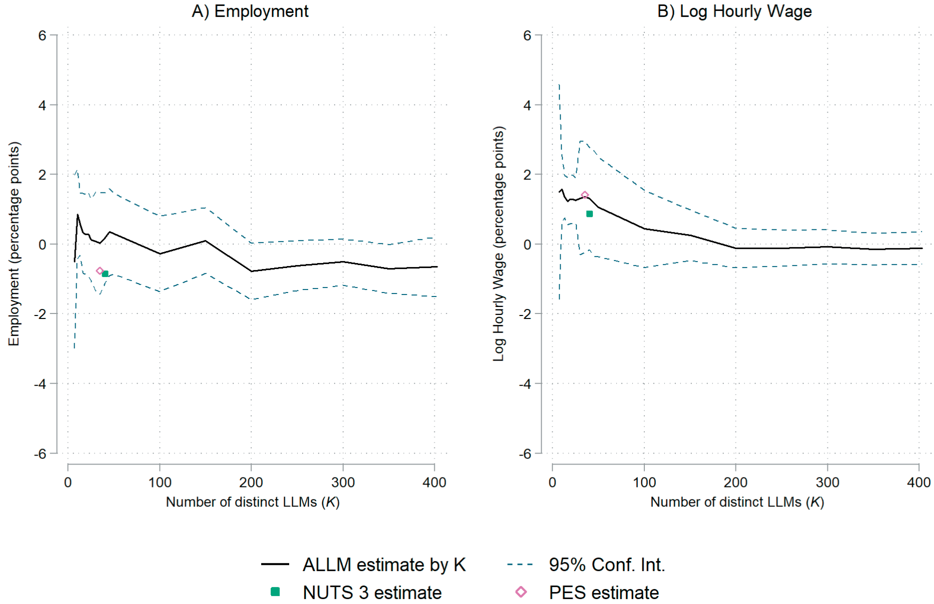
Notes: See Table 4.2 for additional notes.

Table 4.4
Displacement effects on employment and hourly wage for high-educated workers (Eq. (4.3)).

	High-educated men		High-educated women	
	Employment (=1)	Hourly wage (log)	Employment (=1)	Hourly wage (log)
	(1)	(2)	(3)	(4)
DISPLACED × *POST*	−0.2170***	−0.0358***	−0.2649***	−0.0475***
	(0.0113)	(0.0082)	(0.0158)	(0.0117)
Number of parameters	133	133	128	131
Number of individuals	2,599	2,599	1,500	1,500
Number of observations	142,945	131,785	82,500	74,854

Notes: See Table 4.2 for additional notes.

Fig. 4.11. Aggregate LLM displacement effects on employment and wages (Eq. (4.4)).



Notes: Figures 4.11A and 4.11B represent regressions of employment and the natural logarithm of hourly wage, respectively. Parameter estimates of the three-way interaction term, among *DISPLACED*, *POST* and *EMPLOYMENT DENSITY*, are reported. The 95% confidence intervals are constructed using clustered standard errors by LLM. In each regression, the natural logarithms of employment density and area size are operated based on a different regional classification. The values at which K is used to operate the aggregate LLMs include 398, 350 to 50 in increments of fifty, 45 to 25 in increments of five, and 22 to 7 in increments of three. The regressions include main, two-way and three-way interaction terms among *DISPLACED*, *POST*, and each one of the following covariates. The regressions include a variable that represents the area size of the worker's LLM home location and various zero-one indicator variables for gender, age (3), Dutch nationality, job tenure (3), manufacturing, children aged 18 or lower, partner, number of household members (3) and year of job displacement (4). In addition, the regressions include individual-specific fixed effects, calendar-month fixed effects (107) and LLM-specific home location fixed effects (K-1). The period under observation is from January 2006 to December 2014. The number of individual-month equals 1,319,560 and 1,173,835 for the model in which employment and hourly wage is the dependent variable, respectively. See Table 4.2 for additional notes.

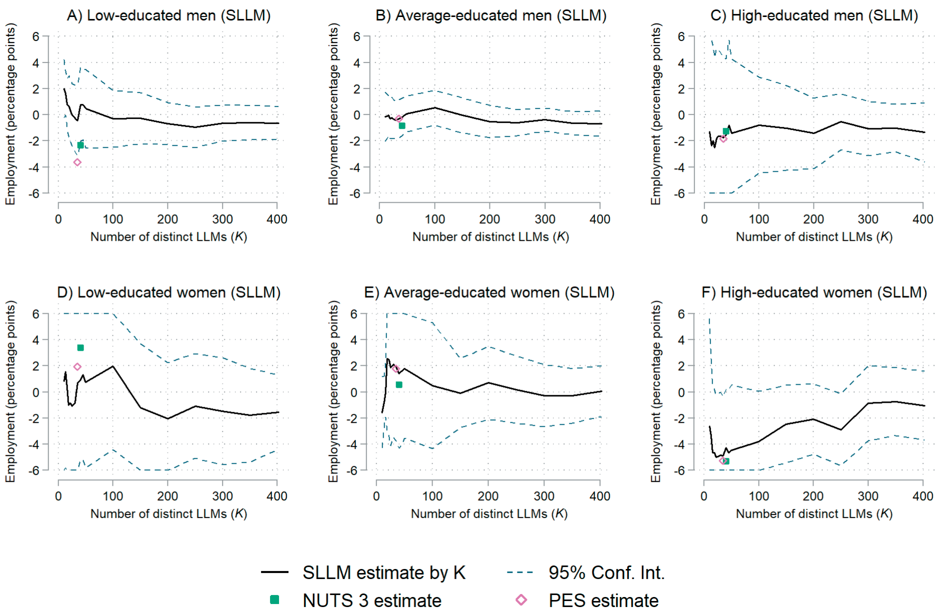
Figure 4.12 and Figure 4.13 reveal the subgroup-specific roles of employment density in the displacement effects on employment and hourly wages, respectively.[18] Figure 4.12 shows that high-educated female workers experience a significant negative effect of employment density, operated by subgroup-specific LLMs, on post-displacement employment. Specifically, for high-educated female workers the loss in employment is about five percentage points higher in a twice as dense location. Also, using the NUTS 3 areas or the PES areas to operate employment density, the loss in employment is two to four percentage points lower for low-educated female workers if they reside in a geographical home location that is twice as large in terms of density.

Figure 4.13 shows a significant effect of employment density, operated by subgroup-specific LLMs, on post-displacement wages for high-educated men and low-educated women. Both subgroups experience more modest losses in hourly wages if they reside in denser LLMs. Using the NUTS 3 or PES areas to operate LLMs, we find a significantly lower loss in wages for high-educated female workers. The results suggest that displacement in a denser LLM would lead to a more modest loss in hourly wage. Note, however, that the empirical evidence is relatively weak as the standard errors are relatively high.

In general, the results suggest that the loss in employment is more modest for low-educated and average-educated workers in more dense labour markets and more pronounced for high-educated workers in more dense labour markets. This could be explained by a more pronounced job search complexity and congestion for high-educated workers in dense labour markets. We find no clear subgroup differentials in the role of employment density in the post-displacement effects on hourly wages for workers who have been displaced.

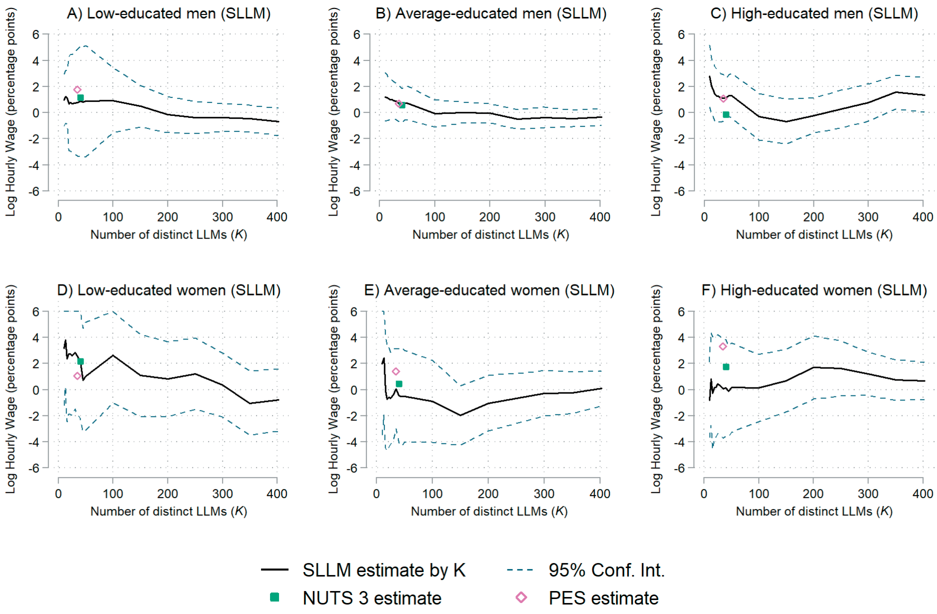
[18] Note that in Figure 4.12 and Figure 4.13, the 95 per cent confidence intervals are in some cases set at a limit of minus six and plus six percentage points to keep the scales of the vertical axes identical. See Table 4.D.5 for the coefficients and standard errors of the subgroup-specific LLM displacement effects for the forty NUTS 3 areas and forty subgroup-specific LLMs, respectively.

Fig. 4.12. Subgroup-specific LLM displacement effects on employment (Eq. (4.4)).



Notes: Figure 4.12 represents regressions of employment. Parameter estimates of the three-way interaction term, among *DISPLACED*, *POST* and *EMPLOYMENT DENSITY*, are reported. The number of individual-month observations for the subgroups in graphs A-F equals 269,060; 612,535; 161,975; 58,905; 172,535; 80,355, respectively. See Figure 4.11 for additional notes.

Fig. 4.13. Subgroup-specific LLM displacement effects on wages (Eq. (4.4)).



Notes: Figure 4.13 represents regressions of the natural logarithm of hourly wages. Parameter estimates of the three-way interaction term, among *DISPLACED*, *POST* and *EMPLOYMENT DENSITY*, are reported. The number of individual-month observations for the subgroups in Figures 4.13A-4.13F equals 232,150; 550,028; 149,369; 49,727; 151,393; 72,919, respectively. See Figure 4.11 for additional notes.

4.7 Conclusion

In Chapter 4, we examine the role of regional aggregation in the returns to agglomeration. A higher level of regional aggregation refers to a regional classification with a higher spatial scale and fewer distinct LLMs. We focus on two economic applications of the returns to agglomeration, including the effect of agglomeration externalities on wages, which is referred to as the UWP, and the returns to agglomeration in wages and employment for workers who have been displaced. We apply a unique flow-based cluster algorithm, using commuting flows across municipalities as the main input, to define alternative sets of aggregate and subgroup-specific LLMs. The subgroup-specific LLMs allow us to examine subgroup differentials in the returns to agglomeration. Our conclusions are twofold.

First, we conclude that the returns to agglomeration strongly depend on the spatial scale of LLMs. We show that estimates of the UWP are increasing by a magnitude of two to three in the level of regional aggregation. This finding suggests that a large share of the positive externalities from agglomeration economies, based on the matching, sharing, and learning micro-foundations, takes place at a relatively high spatial scale. Hence, we argue that improved matching and sharing are more important for agglomeration benefits than improved learning, as agglomeration benefits for matching and sharing take place at a relatively high spatial scale (Rosenthal and Strange, 2001, 2008). In addition, we examine whether agglomeration externalities affect job matching of displaced workers. We show that being located in denser LLMs leads to more modest post-displacement wage losses. Specifically, if a displaced worker is located in an LLM that is a hundred per cent denser, the loss in wage is about 2 percentage points lower. Note that this observation only holds at a relatively high spatial scale. We do not find positive returns to agglomeration in post-displacement employment at any spatial scale. Thus, we argue that the matching mechanism indeed operates at a relatively high spatial scale and leads to heterogeneity effects in job matching through wage differentials, but not to quantity effects in job matching through employment differentials. In line with the matching-function literature (e.g., Petrongolo and Pissarides (2006)), we find that agglomeration spillovers from a denser labour market allow workers to be more selective in wages with a similar re-employment probability.

Second, we conclude that the returns to agglomeration are more attenuated for subgroups of workers who are characterised by large LLMs such as male workers and high-educated workers. Subgroup-specific local labour markets are relevant, as theory suggests that workers differ in their opportunity costs of commuting through time and financial constraints. We focus on workers who differ in gender and education levels, as we show that these demographic characteristics are the most important drivers behind differences in workers' commuting outcomes. For all subgroups, the returns to agglomeration in the UWP increase equally in the level of regional aggregation. This finding is very important as it suggests that the modifiable areal unit problem is not the main driver behind the role of regional aggregation level in agglomeration benefits, because our descriptive analysis shows that the mismeasurement in workers' LLM structure and size depends on the worker's gender and education level. Holding the level of regional aggregation and thus the number of distinct LLMs constant among the different subgroups, we show that the UWP is education-biased and not gender-biased. Compared to low-educated workers, high-educated workers experience a UWP that is about 100 per cent higher. The use of aggregate LLMs or subgroup-specific LLMs provides comparable results of subgroup differentials in agglomeration benefits.

Importantly, our descriptive results show that the structure of workers' LLM is endogenous to the worker's gender, education and geographical location. Female and especially low-educated workers are characterised by small and more distinct LLMs. Given that most studies use regional classifications that represent large areas, we argue that the returns to agglomeration are generally overestimated for workers who are characterised by small LLMs. In this regard, the returns to agglomeration are also gender-biased, as women are characterised by LLMs with relatively low levels of regional aggregation corresponding to smaller returns to agglomeration. We do not find convincing empirical evidence of gender differentials or education differentials in the returns to agglomeration in post-displacement employment and wages. Consequently, our results suggest that the importance of the matching mechanism for subgroup differentials in agglomeration benefits is reasonably small. Overall, the presence of gender- and education differentials in the returns to agglomeration in the UWP is relevant from a societal perspective, as it highlights a trade-off between efficiency and societal wage equality.

Our research gives a better understanding of the structure of workers' LLM and its economic consequences, which is very relevant from a policy perspective. Examples are labour market policies which aim to increase the rate and quality of the job match between worker and employer, or which aim to limit the impact of negative employment shocks (Blumenberg, 2004; Moretti, 2011; Crépon and Van den Berg, 2016). Specifically, we show that denser labour markets lead to more modest losses in wages, but not to variation in the losses in employment. Moreover, our findings are relevant for place-based policies targeted at specific regions or subgroups of the population (Glaeser and Gottlieb, 2008; Neumark and Simpson, 2015). Place-based policies targeted at workers who are characterised by a relatively small LLM, including female workers and low-educated workers, should be focused on smaller geographical locations than policies directed at other subgroups of workers. In this regard, the so-called ripple effect, i.e. the diffusion of the policy impact over the areas next to the targeted location, might be smaller for low-educated workers than for high-educated workers. The role of the spatial scale in the efficiency of policies targeted at different subgroups of the population is a potential area for future research.

All in all, we show how a researcher is able to define meaningful clusters using flowbca, which is done in the context of economic geography by defining aggregate and subgroup-specific LLMs. A key question is whether the use of aggregate and subgroup-specific regional classifications to operate geographic space is important. A main contribution of Chapter 4 is that we show that this is indeed the case, as the regional classification and in particular the level of regional aggregation strongly affects estimates of the agglomeration benefits for wages. The implications of the spatial scale in empirical analyses may be even more important for research on larger regional areas such as the U.S. or Europe, as for these areas there are regional classifications available at a much higher level of regional aggregation such as MSAs. Chapter 4, which deals with the importance of (subgroup-specific) regional aggregation for the empirical analysis of agglomeration economies, could aid with a broader body of research that uses regional classifications to estimate regional differences in economic outcomes.

Acknowledgements

We wish to thank seminar participants at the 32nd Annual Conference of the European Society for Population Economics, and Utrecht University School of Economics.

4.A The MAUP in Workers' Local Labour Market

In this appendix, we provide a detailed discussion about the implications of a measurement error in the size of workers' LLM, also known as the Modifiable Areal Unit Problem (MAUP).

A general model on the UWP is given as

$$Y_{irt} = \beta_0 + \beta_1 J^*_{rt} + \varepsilon_{irt} \tag{4.A.1}$$

where subscripts i, r and t denote the worker, regional unit and year, respectively. The job density is referred to by J^* at regional unit r. Unfortunately, the job density is mismeasured, for example as the size of r is mismeasured. Hence, we do not observe J^* but only J, where J_{rt}=J^*_{rt}+u_{irt}. Parameter u represents the measurement error. The model we estimate, given J^*_{rt}=J_{rt}-u_{irt}, equals

$$\begin{aligned} Y_{irt} = \beta_0 + \beta_1 J^*_{rt} + \varepsilon_{irt} &= \beta_0 + \beta_1 (J_{rt} - u_{irt}) + \varepsilon_{irt} \\ &= \beta_0 + \beta_1 J_{rt} + (\varepsilon_{irt} - \beta_1 u_{irt}) \\ &= \beta_0 + \beta_1 J_{rt} + v_{irt} \end{aligned} \tag{4.A.2}$$

where $v_{irt} = \varepsilon_{irt} - \beta_1 u_{irt}$. The OLS estimator for β_1 is given as

$$\begin{aligned} \hat{\beta}_{1_{OLS}} &= \frac{\sum_{i=1}^{N}(J_{rt}, Y_{irt})}{\sum_{i=1}^{N} J^2_{rt}} \\ &= \frac{\sum_{i=1}^{N}(J^*_{rt}, u_{irt})(\beta_0 + \beta_1 J^*_{rt} + \varepsilon_{irt})}{\sum_{i=1}^{N}(J^*_{rt} + u_{irt})(J^*_{rt} + u_{irt})} \\ &= \frac{\beta_1 cov(J^*_{rt}, J^*_{rt}) + cov(J^*_{rt}, \varepsilon_{irt})}{cov(J^*_{rt}, J^*_{rt}) + 2cov(J^*_{rt}, u_{irt}) + cov(u_{irt}, u_{irt})} \\ &+ \frac{\beta_1 cov(u_{irt}, J^*_{rt}) + cov(u_{irt}, \varepsilon_{irt})}{cov(J^*_{rt}, J^*_{rt}) + 2cov(J^*_{rt}, u_{irt}) + cov(u_{irt}, u_{irt})} \end{aligned} \tag{4.A.3}$$

And the probability limit of $\hat{\beta}_{1_{OLS}}$, given $N \rightarrow \infty$, equals

$$\begin{aligned}plim(\hat{\beta}_{1_{OLS}}) &= \beta\left(\frac{\sigma^2_{J^*}+\sigma_{J^*u}}{\sigma^2_{J^*}+\sigma^2_u+2\sigma_{J^*u}}\right)\\ &\quad+\frac{\sigma_{J^*\varepsilon}+\sigma_{u\varepsilon}}{\sigma^2_{J^*}+\sigma^2_u+2\sigma_{J^*u}}\\ &= \beta\left(1-\frac{\sigma^2_u+\sigma_{J^*u}}{\sigma^2_{J^*}+\sigma^2_u+2\sigma_{J^*u}}\right)\\ &\quad+\frac{\sigma_{J^*\varepsilon}+\sigma_{u\varepsilon}}{\sigma^2_{J^*}+\sigma^2_u+2\sigma_{J^*u}}\end{aligned} \tag{4.A.4}$$

The classical measurement error model holds if $\sigma_{J^*u} = \sigma_{J^*\varepsilon} = \sigma_{u\varepsilon} = 0$, which implies that the covariance between the true employment density and the measurement error term equals zero ($\sigma_{J^*u} = 0$), the covariance between the true employment density and the equation error term equals zero ($\sigma_{J^*\varepsilon} = 0$), and the covariance between the measurement error and the equation error equals zero ($\sigma_{u\varepsilon} = 0$). Under the classical measurement error, the probability limit equals $plim(\hat{\beta}_{1_{OLS}}) = \beta\frac{\sigma^2_{J^*}}{\sigma^2_{J^*}+\sigma^2_u} = \beta\tau$, where $\tau = \frac{\sigma^2_{J^*}}{\sigma^2_{J^*}+\sigma^2_u}$ and τ represents an attenuation bias as it is between zero and one. The bias in the estimate of the coefficient is $plim(\hat{\beta}_{1_{OLS}} - \beta) = \beta\tau - \beta = -\beta(1-\tau) = -\beta\left(1-\frac{\sigma^2_{J^*}}{\sigma^2_{J^*}+\sigma^2_u}\right) = -\beta\frac{\sigma^2_u}{\sigma^2_{J^*}+\sigma^2_u}$.

Note that for papers that operate employment density as a zero-one indicator variable, e.g., D'Costa and Overman (2014), $\sigma_{J^*u} = 0$ does not hold. Consequently, the pool of workers who live in a peripheral area cannot under-report and the pool of workers who live in the urban area cannot over-report. This will amplify the attenuation bias, as the measurement error is negative (positive) if the dummy equals one (zero).

For various reasons it seems highly unlikely that the measurement error in employment density is classical. There exist potential confounding factors (e.g. unobserved ability) that direct high ability workers in denser LLMs (Combes et al., 2012). Moreover, there might be aggregate missing variables that influence local outcomes and local characteristics. For example, better individual outcomes attract more firms and workers in some locations, which in turn affect local characteristics. Generally, the issue of aggregate missing variables is considered to be less important

(Combes and Gobillon, 2015). However, both mechanisms lead to the situation where employment density is correlated to the equation error, i.e. $\sigma_{J^*\varepsilon} \neq 0$. Consequently, the estimates of J are not likely to be consistent as the variable employment density is endogenous for the aforementioned reasons. As a solution to the endogeneity of job density, the literature tends to include more observables and individual-specific fixed effects to correct for unobserved heterogeneity. For the FE estimator, the key identification assumptions are that job changes across LLMs are random and workers' ability is time-constant.

Unfortunately, including more observables or individual-specific fixed effects amplifies the attenuation bias (Griliches, 1977; Freeman, 1984; Chowdhury and Nickell, 1985; Griliches and Hausman, 1986). To explain the role of individual-specific fixed effects in the attenuation bias, we consider the following model

$$Y_{irt} = \beta_0 + \beta_1 J^*_{rt} + a_i + \varepsilon_{irt} \tag{4.A.5}$$

We take the first difference to eliminate the individual-specific fixed effect a_i. Note that the same logic holds for the within estimator.

$$Y_{ir,t} - Y_{ir,t-1} = \beta_1(J^*_{r,t} - J^*_{r,t-1}) + \varepsilon_{ir,t} - \varepsilon_{ir,t-1} \tag{4.A.6}$$

Again, we do not observe $J^*_{r,t}$, but we observe the mismeasured variable $J_{r,t}$. We observe $J_{r,t} = J^*_{r,t} + u_{ir,t}$, which we implement in (4.A.6) to get

$$\begin{aligned} \Delta Y_{ir,t} &= Y_{ir,t} - Y_{ir,t-1} \\ &= \beta_1(J_{r,t} - J_{r,t-1} - u_{ir,t} + u_{ir,t-1}) + \varepsilon_{ir,t} - \varepsilon_{ir,t-1} \\ &= \beta_1(J_{r,t} - J_{r,t-1}) + \varepsilon_{ir,t} - \varepsilon_{ir,t-1} + \beta_1(u_{ir,t-1} - u_{ir,t}) \\ &= \beta_1 \Delta J_{r,t} + v_{ir,t} \end{aligned} \tag{4.A.7}$$

where $v_{ir,t} = \Delta\varepsilon_{ir,t} + \beta_1(u_{ir,t-1} - u_{ir,t})$. The probability limit of $\hat{\beta}_{1_{FD}}$, given $N \to \infty$, equals

$$plim(\hat{\beta}_{1_{FD}}) = \frac{\beta\sigma^2_{\Delta J^*}}{\sigma^2_{\Delta J^*} + \sigma^2_{\Delta u}} \tag{4.A.8}$$

To compute the probability limit of $\hat{\beta}_{1_{FE}}$, we derive the variation in the changes of the true employment density J, i.e. $\sigma^2_{\Delta J^*}$, and the variation in the changes of the measurement error u, i.e. $\sigma^2_{\Delta u}$. This gives

$$\begin{aligned}
\sigma^2_{\Delta J^*_{r,t}} &= var(J^*_{r,t}) + 2cov(J^*_{r,t}, J^*_{r,t-1}) + var(J^*_{r,t-1}) \\
&= \sigma^2_{J^*_{r,t}}(1-\rho) \\
\sigma^2_{\Delta u} &= var(u_{ir,t}) + 2cov(u_{ir,t}, u_{ir,t-1}) + var(u_{ir,t-1}) \\
&= \sigma^2_{u_{ir,t}}(1-r)
\end{aligned} \tag{4.A.9}$$

where $var(J^*_{r,t-1}) = var(u_{ir,t-1}) = 0$ holds under the assumption that $J^*_{r,t-1}$ and $u_{ir,t-1}$ are stationary. The parameter ρ and r represent the first order autocorrelation in J^* and u, respectively. The probability limit equals

$$\begin{aligned}
plim(\hat{\beta}_{1_{FE}}) &= \frac{\beta\sigma^2_{J^*}(1-\rho)}{\sigma^2_{J^*}(1-\rho) + \sigma^2_u(1-r)} \\
&= \frac{\beta}{1 + \dfrac{\sigma^2_{u_{ir,t}}(1-r)}{\sigma^2_{J^*_{r,t}}(1-\rho)}}
\end{aligned} \tag{4.A.10}$$

This equation shows that the attenuation bias will be amplified if r goes to zero (i.e. the measurement error is uncorrelated over time) or ρ goes to one (i.e. the value of the employment density J^* is strongly correlated over time). Note that if the measurement error is time-constant (i.e. $u_{i,t} = u_i$), introducing fixed effects is beneficial as it completely eliminates the attenuation bias. If this is not the case, for example due to changes in the demographic composition of cities or the location of the worker or due to temporal heterogeneity in the level of the data, the attenuation bias is amplified by introducing fixed effects. Effectively, the attenuation bias will increase if the serial correlation in the true values of employment density exceeds the serial correlation in the measurement error. An alternative way to put this is that by introducing individual-specific fixed effects to correct for time-constant sorting across areas, the variation in the measurement error σ_u^2 is increased while the variation in the true employment density σ_{J*}^2 is reduced. This holds for two reasons.

The first being that the relative number of workers with a measurement error in the regional level which is used to operate the LLM is higher with panel data. The higher the number of individual-year observations in the sample, the higher the number of observations that the regional levels are potentially mismeasured. In the context of employment density in a specific regional area, this problem is severe, since the variable is relatively time-constant while there is a measurement error in every period. The second reason is that the agglomeration effects are identified based on a small number of workers who change home across LLMs. Consequently, when using the FE estimator, the coefficient is identified based on a lower number of correct observations. Therefore, the measurement error and sorting of the few households that relocate across LLMs to a different home attenuate the FE coefficients towards zero. Indeed, the literature on agglomeration economies finds lower returns to urbanization in wages when using the FE estimator (Glaeser and Maré, 2001). This finding is typically interpreted as empirical evidence that suggests that the FE estimator limits the potential of time-constant unobserved heterogeneity. An alternative explanation is that the individual-specific fixed effects amplify the attenuation bias.

4.B Two-Step Estimation Procedure

In this appendix, we provide the estimates of the UWP using the two-step procedure in the spirit of Combes et al. (2008). Using the two-step procedure, we limit the potential of sorting of worker quality across LLMs. Moreover, the model controls for differences in amenities and housing prices across LLMs. Figure 4.B.1 shows the estimates of the UWP using the aggregate LLMs to operate workers' LLM.

The first step involves the regression of individual wages on worker covariates and LLM-year FE, expressed as

$$w_{irt} = \sum_{r=1}^{R} \sum_{t=2006}^{2014} [\delta_{rt}(N_r + D_t)] + \beta' X_{it} + \alpha_i + \varepsilon_{irt} \tag{4.B.1}$$

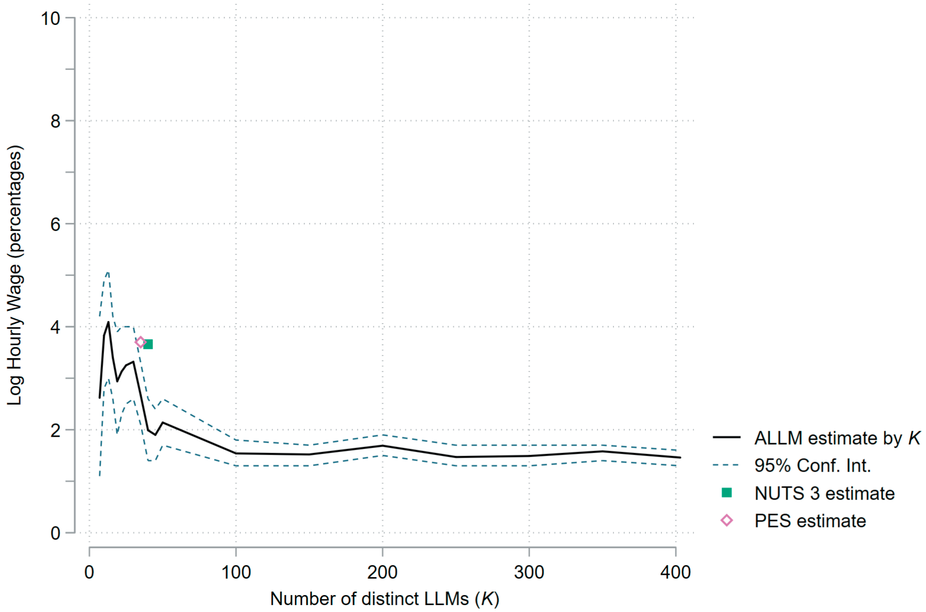
The second step involves the regression of the estimated LLM-year fixed effects on employment density and the annual dummies.

$$\hat{\delta}_{rt} = \beta_1 J_{rt} + D_t + \varepsilon_{rt} \tag{4.B.2}$$

The results of the two-step approach are provided in Figure 4.B.1 and Figure 4.B.2. Compared to the direct approach of estimating the UWP (see Fig. 4.8), the estimates using the two-step approach are lower. This observation suggests that the direct approach leads to an overestimation of the UWP. However, the pattern of the UWP over the number of distinct LLMs is comparable: with fewer distinct LLMs the estimate of the UWP is higher. Combes et al. (2008), using French data and a similar empirical specification, find an estimate of the UWP between 3 and 4 per cent, which is consistent with our findings.

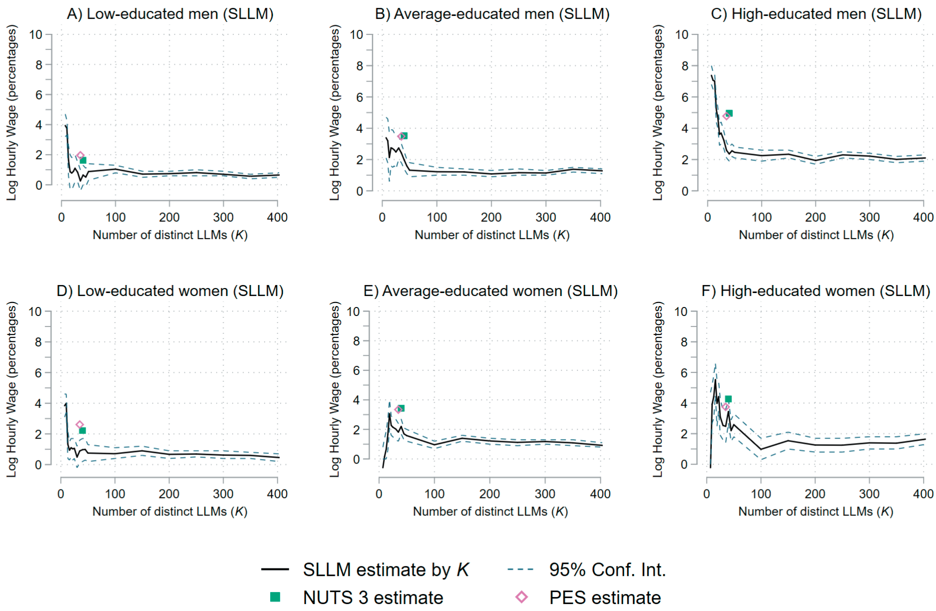
Figure 4.B.2 shows the estimates using the two-step approach and subgroup-specific LLMs. Several observations are in place. First, the returns to agglomeration are increasing in the education level and are higher for men. Second, using the NUTS 3 areas to operate workers' LLM leads to larger differences in the UWP between education levels than when using the subgroup-specific LLMs to operate workers' LLM. Finally, compared to the use of the direct approach, the estimates of the UWP are lower if the two-step approach is used.

Fig. 4.B.1. Aggregate LLM UWP based on the OLS two-step procedure (Eq. (4.B.2)).



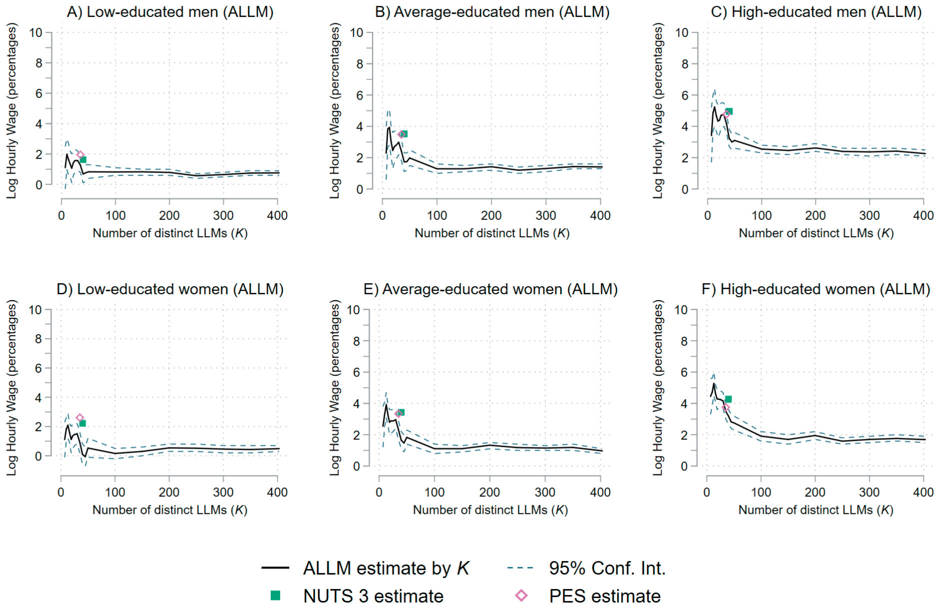
Notes: Estimates of the second stage are provided. See Figure 4.8 for additional notes.

Fig. 4.B.2. Subgroup-specific LLM UWP by subgroup based on the OLS two-step procedure (Eq. (4.B.2)).



Notes: Estimates of the second stage are provided. See Figure 4.8 and Figure 4.9 for additional notes.

Fig. 4.B.3. Aggregate LLM UWP by subgroup based on the OLS two-step procedure (Eq. (4.B.2)).



Notes: Estimates of the second stage are provided. See Figure 4.8 and Figure 4.9 for additional notes.

4.C Urban Wage Premium: Summary Statistics and Robustness Checks

Table 4.C.1
Summary statistics for hourly wage and commuting distance using the urban wage premium data sample.

	Hourly wage (log)	Commuting distance (km)
Mean	2.7685	18.5271
St. Dev.	0.4715	25.2737
Variance	0.2223	638.7611
Skewness	0.1903	3.2560
Kurtosis	4.0985	17.3093
1th percentile	1.5706	0.5493
5th percentile	2.0531	1.2541
25th percentile	2.4652	3.7796
50th percentile	2.7496	9.9365
75th percentile	3.0591	22.0567
95th percentile	3.5525	65.8312
99th percentile	3.9886	132.9291
Number of observations	18,893,075	18,893,075

Notes: The urban wage premium data sample.

Table 4.C.2
Individual summary statistics using the urban wage premium data sample.

	Mean	St. Dev.
Employment (=1)	1	0
Hourly wage (log)	2.7685	0.4715
Hourly wage (€)	17.9216	12.4329
Commuting distance (km)	18.5271	25.2737
Age (in years)	36.2138	11.0393
Female (=1)	0.3552	0.4786
Low-educated (=1)	0.1674	0.3733
Average-educated (=1)	0.4261	0.4945
High-educated (=1)	0.4065	0.4912
Dutch (=1)	0.8995	0.3006
Partner (=1)	0.3901	0.4878
No child (=1)	0.6404	0.4799
Fixed contract (=1)	0.7052	0.4559
Full-time job (=1)	0.7930	0.4051
Manufacturing sector (=1)	0.1913	0.3933
Number of observations	18,893,075	18,893,075

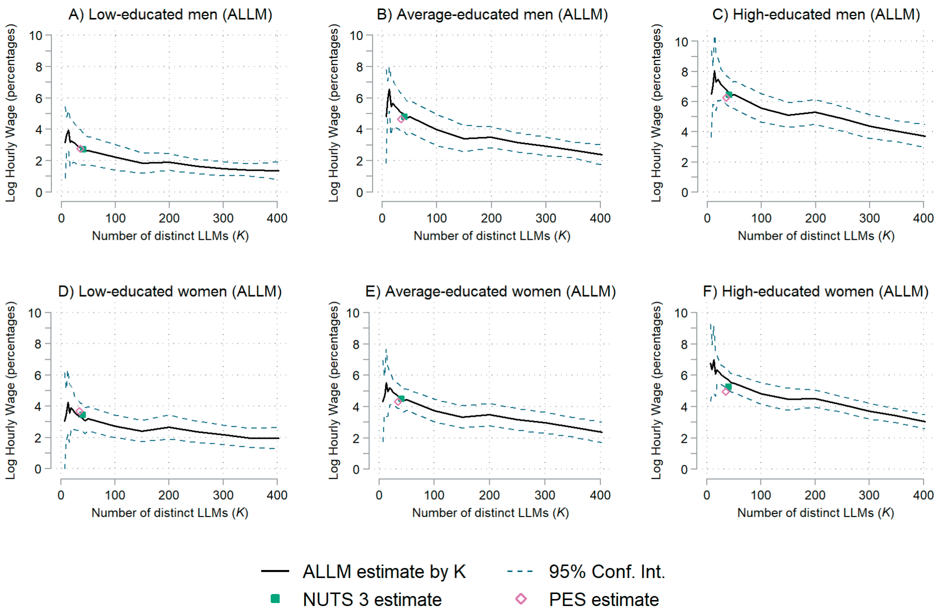
Notes: The urban wage premium data sample.

Table 4.C.3
Coefficients and standard errors of subgroup-specific LLM UWP based on FE estimates (Fig. 4.10, Eq. (4.2)).

	Hourly wage (log)	
Subgroup:	NUTS 3 ($K = 40$)	SLLM ($K = 40$)
Low-educated men	0.0034	0.0040
	(0.0012)	(0.0017)
Average-educated men	0.0058	0.0062
	(0.0010)	(0.0012)
High-educated men	0.0074	0.0084
	(0.0012)	(0.0020)
Low-educated women	0.0048	0.0076
	(0.0015)	(0.0015)
Average-educated women	0.0063	0.0054
	(0.0014)	(0.0013)
High-educated women	0.0056	0.0109
	(0.0015)	(0.0018)

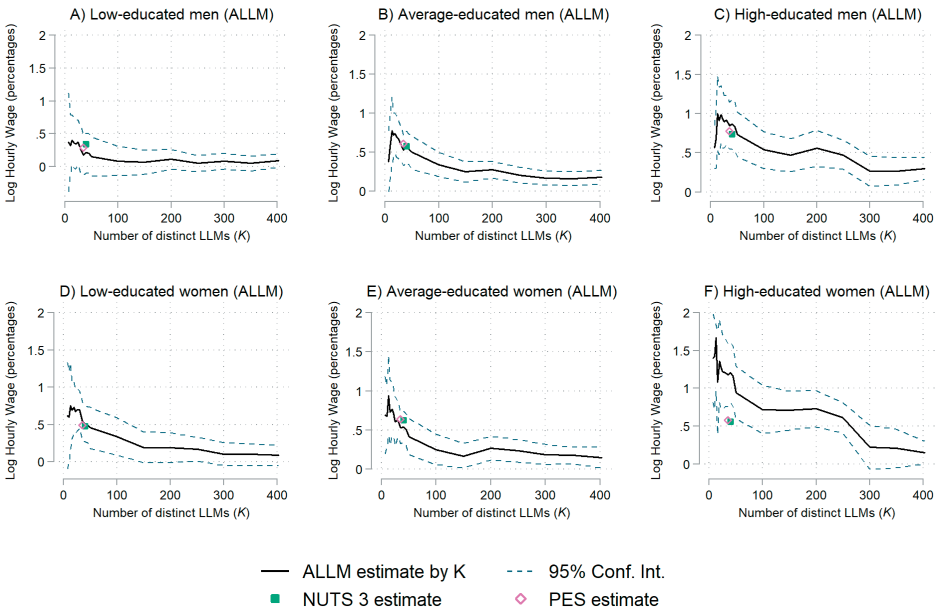
Notes: Each estimate represents a different regression. The coefficients and standard errors are provided for the regressions in which the employment density and area size are operated based on the forty NUTS 3 areas and forty subgroup-specific LLMs, respectively. See Figure 4.10 for additional notes.

Fig. 4.C.1. Aggregate LLM UWP based on OLS estimates (Eq. (4.2)).



Notes: See Figure 4.9 for additional notes.

Fig. 4.C.2. Aggregate LLM UWP based on FE estimates (Eq. (4.2)).



Notes: See Figure 4.10 for additional notes.

4.D Job Displacement: Summary Statistics and Robustness Checks

Table 4.D.1
The within change in hourly wage and commuting distance using the matched job displacement data sample.

	Hourly wage (log)		Commuting distance (km)	
	Displaced	Non-displaced	Displaced	Non-displaced
Mean	-0.0187	0.0490	3.4568	0.5570
St. Dev.	0.3401	0.2113	32.7783	15.7735
Variance	0.1157	0.0446	1074.4167	248.8017
Skewness	-0.3814	3.5438	0.8532	0.3645
Kurtosis	29.9120	106.1860	13.6756	46.0938
1th percentile	-1.0382	-0.5346	-98.9588	-50.3882
5th percentile	-0.4812	-0.1939	-38.6010	-9.3345
25th percentile	-0.1286	-0.0012	-2.8020	0
50th percentile	0.0062	0.0386	0	0
75th percentile	0.1100	0.1031	9.8634	0
95th percentile	0.3712	0.2813	51.4306	14.5955
99th percentile	0.8198	0.6151	127.8004	60.6672
Number of observations	7,248	13,591	6,916	13,390

Notes: The job displacement data set. The individual summary statistics are based on the within change, measured by the difference in the values of each variable between the eighteenth month after job displacement and the month of job displacement.

Table 4.D.2
Individual summary statistics using the non-matched job displacement data sample.

	Non-displaced		Displaced		
	Mean	St. Dev.	Mean	St. Dev	t-statistic
Employment (=1)	1	0	1	0	
Hourly wage (log)	2.8711	0.3903	2.7861	0.4181	32.84***
Hourly wage (€)	19.1870	11.6554	18.6162	50.9130	7.24***
Commuting distance (km)	15.5553	21.9180	17.8662	25.3218	-15.88***
Home change (=1)	0.0059	0.0764	0.0052	0.0718	1.37
Annual household income (€)	44,402	22,164	41,926	22,433	16.80***
Age (in years)	40.6143	9.2440	42.0801	9.1878	-23.90***
Female (=1)	0.4683	0.4990	0.2944	0.4558	52.52***
Low-educated (=1)	0.1723	0.3777	0.3097	0.4624	-54.78***
Average-educated (=1)	0.4153	0.4928	0.5368	0.4987	-37.16***
High-educated (=1)	0.4123	0.4923	0.1535	0.3605	79.29***
Dutch (=1)	0.9107	0.2852	0.9023	0.2969	4.44***
Partner (=1)	0.5376	0.4986	0.5598	0.4964	-6.71***
No child (=1)	0.5282	0.4992	0.5543	0.4971	-7.88***
Household members (#)	2.9257	1.3365	2.9222	1.3101	0.39
Fixed contract (=1)	0.9291	0.2566	0.9068	0.2907	13.12***
Full-time job (=1)	0.5916	0.4915	0.7096	0.4539	-36.21***
Tenure in the job (in months)	118.3416	80.9808	126.3223	86.3185	-14.85***
Manufacturing sector (=1)	0.2093	0.4068	0.4640	0.4987	-94.34***
Number of individuals (#)	10,587,265		22,765		

Notes: The individual summary statistics, provided for the month of actual or potential displacement, are based on the sample prior to matching. The time period under observation is from July 2007 to December 2011. Sample means with standard deviations are provided for the treatment group and control group. The t-statistic is provided to assess whether the mean and standard deviation of each variable for the groups of displaced and non-displaced workers are statistically different from each other. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. By construction, all displaced and non-displaced were employed in the month of actual or potential displacement.

Table 4.D.3
Individual summary statistics using the matched job displacement data sample.

	Non-displaced		Displaced		
	Mean	St. Dev.	Mean	St. Dev	t-statistic
Employment (=1)	1	0	1	0	
Hourly wage (log)	2.8369	0.3786	2.8353	0.4151	0.31
Hourly wage (€)	18.4704	9.2714	19.3620	49.0370	-2.16**
Commuting distance (km)	14.9149	20.5429	17.4778	24.2814	-8.90***
Home change (=1)	0.0060	0.0771	0.0050	0.0707	0.99
Annual household income (€)	45,001	22,597	44,164	21,943	2.87***
Age (in years)	41.1290	9.9092	41.7133	9.5521	-4.59***
Female (=1)	0.2298	0.4207	0.2304	0.4211	-0.10
Low-educated (=1)	0.2330	0.4228	0.2557	0.4363	-4.06***
Average-educated (=1)	0.5821	0.4932	0.5749	0.4944	1.12
High-educated (=1)	0.1849	0.3883	0.1694	0.3752	3.10***
Dutch (=1)	0.9685	0.1747	0.9617	0.1919	2.86***
Partner (=1)	0.5759	0.4942	0.5851	0.4927	-1.44
No child (=1)	0.5548	0.4970	0.5519	0.4973	0.45
Household members (#)	3.0299	1.3294	3.0004	1.3189	1.71*
Fixed contract (=1)	0.9667	0.1794	0.9637	0.1872	1.29
Full-time job (=1)	0.7958	0.4031	0.7873	0.4092	1.60
Tenure in the job (in months)	124.8017	88.2240	129.1366	89.7851	-3.75***
Manufacturing sector (=1)	0.4919	0.5000	0.5078	0.5000	-2.45**
Number of individuals (#)	14,876		9,767		

Notes: The individual summary statistics, provided for the month of actual or potential displacement, are based on the sample after matching. The time period under observation is from July 2007 to December 2011. Sample means with standard deviations are provided for the treatment group and control group. The t-statistic is provided to assess whether the mean and standard deviation of each variable for the groups of displaced and non-displaced workers are statistically different from each other. ***,**,*, correspond to the significance level of 1%, 5%, 10%, respectively. By construction, all displaced and non-displaced were employed in the month of actual or potential displacement.

Table 4.D.4
Firm summary statistics using the matched job displacement data sample.

	Firms			
	Bankrupt firms		Non-bankrupt firms	
	Mean	St. Dev.	Mean	St. Dev.
Firm size:				
1-9 employees (=1)	0	0	0	0
10-49 employees (=1)	0.5881	0.4922	0.7110	0.4534
50-99 employees (=1)	0.1289	0.3351	0.1097	0.3125
100-499 employees (=1)	0.1820	0.3859	0.1067	0.3087
500 or more employees (=1)	0.1010	0.3013	0.0727	0.2596
Firm sector:				
Agriculture, forestry and fishing (=1)	0.0041	0.0638	0.0100	0.0995
Mining and quarrying (=1)	0	0	0	0
Manufacturing (=1)	0.3224	0.4674	0.2540	0.4354
Electricity, gas, steam and air conditioning supply (=1)	0	0	0	0
Water supply; sewerage, waste management and remediation activities (=1)	0.0002	0.0127	0.0007	0.0258
Construction (=1)	0.1988	0.3991	0.1880	0.3908
Wholesale and retail trade; repair of motor vehicles and motorcycles (=1)	0.2112	0.4082	0.2037	0.4028
Transportation and storage (=1)	0.0312	0.1740	0.0503	0.2187
Accommodation and food service activities (=1)	0.0051	0.0714	0.0123	0.1104
Information and communication (=1)	0.0258	0.1585	0.0430	0.2029
Financial and insurance activities (=1)	0.0412	0.1987	0.0360	0.1863
Real estate activities (=1)	0.0014	0.0369	0.0043	0.0657
Professional, scientific and technical activities (=1)	0.0719	0.2584	0.0927	0.2900
Administrative and support service activities (=1)	0.0316	0.1748	0.0537	0.2254
Public administration and defence; compulsory social security (=1)	0	0	0	0
Education (=1)	0.0074	0.0855	0.0060	0.0772
Human health and social work activities (=1)	0.0431	0.2031	0.0353	0.1847
Arts, entertainment and recreation (=1)	0.0022	0.0465	0.0047	0.0682
Other service activities (=1)	0.0026	0.0506	0.0053	0.0728
Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use (=1)	0	0	0	0
Activities of extraterritorial organisations and bodies (=1)	0	0	0	0
Number of firms (#)	3,000		12,487	

Notes: Means and standard deviations are provided at the firm level in the month of job displacement. The time period under observation is from July 2007 to December 2011. Bankrupts firms consist of all distinct firms of which an entity is declared bankrupt and a worker is displaced in the month of actual displacement. Non-bankrupt firms consist of all distinct firms where matched non-displaced workers work in the month of potential displacement.

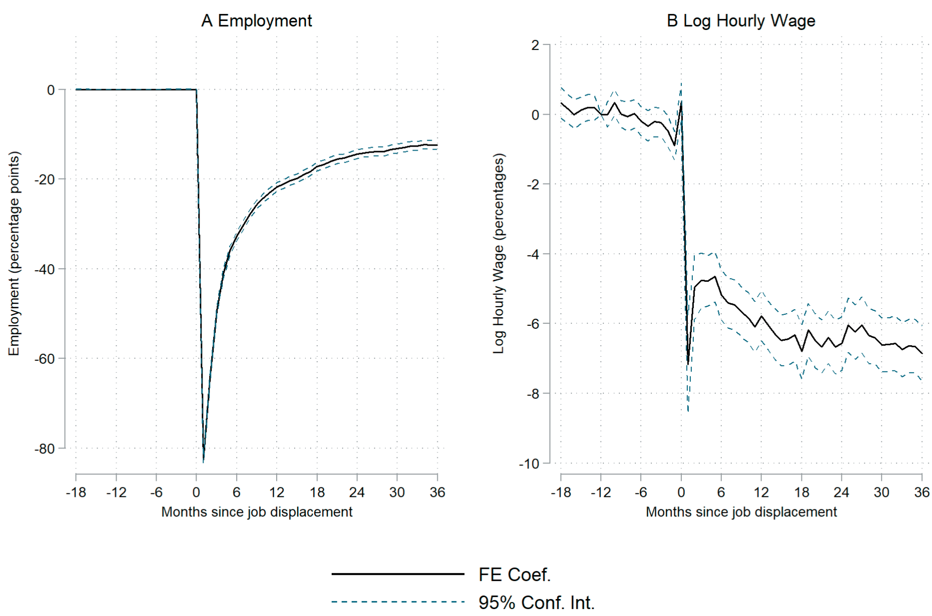
Table 4.D.5

Coefficients and standard errors of subgroup-specific LLM displacement effects (Fig. 4.12 and Fig. 4.13, Eq. (4.4)).

	Employment (=1)		Hourly wage (log)	
Subgroup:	NUTS 3 (K = 40)	SLLM (K = 40)	NUTS 3 (K = 40)	SLLM (K = 40)
Low-educated men	-0.0234	0.0077	0.0115	0.0088
	(0.0126)	(0.0140)	(0.0136)	(0.0209)
Average-educated men	-0.0085	-0.0017	0.0068	0.0053
	(0.0066)	(0.0071)	(0.0055)	(0.0065)
High-educated men	-0.0126	-0.0154	-0.0015	0.0117
	(0.0243)	(0.0284)	(0.0065)	(0.0073)
Low-educated women	0.0340	0.0096	0.0215	0.0211
	(0.0218)	(0.0300)	(0.0165)	(0.0211)
Average-educated women	0.0056	0.0141	0.0045	-0.0046
	(0.0139)	(0.0282)	(0.0141)	(0.0178)
High-educated women	-0.0528	-0.0425	0.0173	0.0013
	(0.0211)	(0.0218)	(0.0170)	(0.0181)

Notes: Each estimate represents a different regression. The coefficients and standard errors are provided for the regressions in which the employment density and area size are operated based on the forty NUTS 3 areas and forty subgroup-specific LLMs, respectively. See Figure 4.12 and Figure 4.13 for additional notes.

Fig. 4.D.1. Time-dependent displacement effects on employment (A) and log hourly wage (B).



Notes: Reference group is the group of non-displaced workers. Reference month is the twelfth month prior to job displacement. The 95% confidence intervals are computed using clustered standard errors by individual. All four fixed effects regression models include 260 parameters of which there are 54 two-way interaction terms. See Table 4.2 for additional notes and statistics.

CHAPTER 5

Conclusion

In the previous chapters of this dissertation, I have examined the structure of workers' local labour markets and its economic consequences. I focused on the topics of local labour markets, job displacement and agglomeration economies. Chapter 2 studied two research questions: "To what extent is the spatial structure of homes and jobs relevant for workers' response to job displacement?", and "What is the role of observed worker characteristics in the effects of job displacement?". Chapter 3 introduced a flow-based cluster algorithm, which allows a researcher to define local labour markets for subgroups of workers at different levels of regional aggregation. This algorithm was necessary to answer the two research questions of Chapter 4: "What is the role of regional aggregation of local labour markets in the returns to agglomeration?", and "To what extent are local labour markets and the returns to agglomeration sensitive to the worker's gender and education level?".

In Section 5.1, I provide a summary and discussion of the main findings of my dissertation. Section 5.2 discusses the scientific contributions to the literature and Section 5.3 elaborates on the main limitations of my research and gives some directions for future research. To conclude the thesis, Section 5.4 discusses the societal relevance of my research, focusing on the policy and real-world implications.

5.1 Summary and Discussion of the Main Findings

5.1.1 Workers' Spatial Response to Job Displacement

Chapter 2 studies the spatial response to job displacement and whether this response depends on the worker's observed characteristics. Dutch administrative data were used and analysed with a quasi-experimental design involving job displacement. The empirical design eliminates the potential of endogenous selection into labour turnover. The estimates show that displaced workers experience, in addition to substantial losses in employment and wage, an increase in the commuting distance and a decrease in the probability of changing home. Specifically, the estimates show that, during the post-displacement period of thirty-six months, displaced workers, on average, (i) are about 25 percentage points less employed, (ii) experience a loss in wage of 6 per cent, (iii) experience an increase in the commuting distance of 3 kilometres and (iv) have a 0.06 percentage points lower probability of changing home. Compared to the mean value, the displacement effects on the commuting distance and changing home represent a positive change of 20 per cent and a negative change of 14 per cent, respectively. Note that changing home is operationalised as such that it equals one if an individual changes home within or between municipalities, and zero otherwise. Importantly, I find no displacement effect on changing home if I use an empirical model in which changing home is operationalised as such that it only equals one if a worker relocates home between NUTS 3 areas.

The findings of Chapter 2 suggest that for workers who have been displaced, the commuting margin is a more relevant margin of labour adjustment than the changing home margin. The negative displacement effect on changing home can be explained by difficulties in getting a new mortgage for an owner-occupied home or a new rental agreement for a rental home after becoming displaced.[1] Thus, although the willingness

[1] Note that it is important to underscore the Dutch-specific institutional setting of the housing market when interpreting the effect of job displacement on changing home, and the role of the worker's housing state. Specifically, the Dutch housing market has: (i) No formal down-payment requirements. (ii) Owner-occupiers are characterised by relatively high loan-to-value ratios. In 2011 there was a binding code of conduct for mortgages introduced, which set the loan-to-value ratio limit at 106 per cent. Since 2012, the loan-to-value ratio limit of 106 per cent has been decreasing by 1 percentage

to change home could be increased after job displacement, for example to enter a distant local labour market, the ability to do so seems to be decreased. However, the second part of Chapter 2 – on the effect the worker's housing state has on the displacement effects – shows that leveraged owners do not experience a different displacement effect on changing home than outright owners or tenants. This observation suggests that displaced workers may not change home for economic reasons but for other reasons such as for family, which is consistent with the study by Huttunen et al. (2018).

Moreover, I show that the patterns of displacement effects change over the worker's post-displacement period. That is, the negative displacement effect on wages becomes more pronounced over the post-displacement period, whereas the increase in the commuting distance diminishes. This finding seems to be driven by displaced workers who have longer unemployment duration and favour shorter commutes over higher wages. The increase in commute shortly after the month of job displacement suggests that displaced workers prefer rapid re-employment to working close to home. This finding could be interpreted as that the worker's local labour market is actually larger than a researcher would observe based on commuting outcomes of typical employed workers.

Also, I examine in Chapter 2 the role of worker characteristics in the effects of job displacement. In particular, I focus on the importance of workers' housing state in the displacement effects. Compared with displaced tenants and outright owners, I find that more leveraged displaced owners are rapidly re-employed and experience a smaller increase in the commuting distance, but experience a higher loss in wage as well.[2] Specifically, I show that, relative to displaced tenants, displaced underwater owners (i.e. owners who face negative home equity) experience a 7 percentage points lower loss in employment. Outright owners experience the highest loss in employment of all displaced homeowners. With respect to the displacement effect on wages, I show that displaced tenants

point a year to 100 per cent in 2018. (iii) Dutch households are characterised by low default and foreclosure rates, potentially due to the fact that all mortgage loans are full recourse loans (Ghent and Kudlyak, 2011). (iv) Homeowners can deduct home mortgage interest from their taxable income. (v) The Dutch social rented sector is relatively large, as it represents over 35 per cent of all Dutch households (CBS, 2018).

[2] Note that in the Netherlands, the rented sector consists for about 88 per cent of social housing. In general, tenants who rent social housing are less likely to change home than those who rent private housing, as social housing is less expensive and more difficult to acquire.

experience a relatively modest loss. In addition, I show that highly leveraged owners who have been displaced become employed relatively close to home. Conversely, displaced outright owners experience the highest increase in the distance between the home and work location. The results suggest that financial incentive structures may drive leveraged displaced owners to prefer rapid re-employment over more modest wage losses.

Another relevant finding is that the displacement effect on commute is relatively large for male workers. Compared to female displaced workers, male displaced workers experience an increase in commute that is about 5 kilometres larger. The gender differential in the displacement effect on commute is relatively stable over the post-displacement period of thirty-six months. This finding suggests that the local labour market of female workers is relatively small and inflexible. In turn, this observation could explain the gender differential in the displacement effect on employment – compared to female displaced workers, male displaced workers experience a loss in employment that is about seven percentage points smaller.

5.1.2 Local Labour Markets and Agglomeration Economies

Chapter 3 introduces a flow-based cluster algorithm. The flow-based cluster algorithm allows a researcher to define meaningful clusters that can be of use to a variety of research fields. In this dissertation, I used the algorithm to define local labour markets based on data of workers' commuting flows from municipality of home to municipality of work. The algorithm allows me to operate workers' local labour market at different levels of regional aggregation. Moreover, it allows me to define local labour markets for different subgroups of workers. Subgroup-specific local labour markets are relevant, as theory suggests that workers differ in their opportunity costs of commuting through time and financial constraints. In Chapter 4, I examine the structures and economic consequences of workers' local labour market. Chapter 4 builds on Chapters 2 and 3 by focusing on the returns to agglomeration in wages and employment of typical employed workers and of workers who have been displaced.

Chapter 4 provides descriptive evidence that underscores the relevance of subgroup-specific local labour markets. The descriptive results reveal that workers' commuting has been increasing in the Netherlands over the period 1988-2014. Moreover, I show that workers' gender and education level explain the largest share of differences in workers' commuting outcomes. This observation is the reason why I focus on gender differentials and education differentials in the returns to agglomeration. Also, the descriptive results suggest that both men and women experience substantial regional mismatch between home and employment locations, as male and female workers are relatively concentrated in specific municipalities while male and female residents are more evenly distributed across municipalities. Moreover, the descriptive results suggest that there is education-biased sorting across municipalities, as there are substantial differences between education groups in the extent to which workers and residents are concentrated in municipalities. Specifically, high-educated workers and high-educated residents are relatively concentrated in municipalities, while low-educated workers and low-educated residents are more evenly distributed across municipalities. Note that the descriptive analysis suggests that there is substantial education-biased sorting across municipalities but not education-biased regional mismatch, as workers and residents within each education group are evenly distributed across municipalities. I apply the flow-based cluster algorithm to endogenise workers' local labour market to their geographical location, commuting outcomes and demographic characteristics. The descriptive results indicate that both male workers and high-educated workers especially are characterised by large local labour markets.

The empirical results in Chapter 4 show that the return to agglomeration in wages, i.e. the urban wage premium, increases by a magnitude of two to three in the level of regional aggregation. This finding suggests that positive spillovers from agglomeration, based on the matching, sharing, and learning micro-foundations of agglomeration economies, are more prevalent at a relatively high spatial scale. Based on different specifications, I find an urban wage premium in the range of 0.3 to 6.6 per cent, which I consider as the lower and upper boundary of the urban wage premium in the Netherlands, respectively. This range is consistent with other studies that use Dutch administrative data to examine the urban wage premium (e.g., see Groot et al. (2014)), but slightly lower than observed in other countries (Combes and Gobillon, 2015). In addition, Chapter 4 analyses the differences of the impact of job displacement across local labour markets. Specifically, I examine the returns to agglomeration in wages and employment for workers who have been displaced. I show that at a relatively high level of regional aggregation, displaced workers in dense local labour markets, compared to displaced workers in more sparse local labour markets, experience more modest losses in wages and comparable losses in employment. If a worker becomes displaced in a local labour market that is hundred per cent denser, the loss in hourly wage is about 2 percentage points lower. I find no positive spillovers from agglomeration economies in employment for workers who have been displaced. In this regard, the returns to agglomeration following job displacement lead to variation in wage differentials, but not to variation in employment differentials. The findings in Chapter 4 suggest that a local labour market that is relatively urbanised makes workers more resilient to job displacement. That is, a denser local labour market allows displaced workers to become selective in wages with a comparable probability of re-employment.

The final part of the empirical analysis focuses on subgroup differentials in the returns to agglomeration. The returns to agglomeration could vary among subgroups in three ways. First of all, the magnitude of the returns to agglomeration could depend on the worker's gender or education level, because of differences in the ability to exploit spillovers from agglomeration economies. An example would be that the urban wage premium is increasing in the worker's education, as relatively high-educated workers benefit most from the improved matching, sharing and learning. Secondly, the role of regional aggregation in the returns to agglomeration could differ among subgroups. An example would be that for high-educated workers only, the prevalence of agglomeration economies is increasing in the spatial scale of local labour markets. Finally, the size and structure of a worker's local labour market could depend on worker characteristics, which would lead to differences in the returns to agglomeration. An example would be that male workers are characterised by a larger local labour market than female workers. From an empirical point of view, if the latter example holds, it would not be justified to use an identical regional classification to operate both male and female workers' local labour market for a study that examines the gender differential in agglomeration benefits.

The empirical results in Chapter 4 show that given a specific regional classification, the urban wage premium for high-educated workers, compared to low-educated workers, is about 100 per cent higher. I find no systematic difference between men and women in the urban wage premium. Moreover, I show that the returns to agglomeration are similarly increasing in the level of regional aggregation for all subgroups of workers. Specifically, the role of regional aggregation in the returns to agglomeration seems comparable among different subgroups of workers. Notably, I show that the local labour market is relatively large for male workers and more educated workers. Consequently, holding the number of distinct local labour markets constant among the subgroups of workers, I argue that the returns to agglomeration are underestimated for workers who are characterised by a large local labour market and overestimated for workers who are characterised by a small local labour market. I do not find robust empirical evidence of gender- or education differentials in the agglomeration benefits for post-displacement employment and wages.

5.2 Scientific Contributions

This dissertation has studied the structures of workers' local labour markets and its economic consequences for labour market outcomes. The research combines labour economics with urban and regional economics, and incorporates the topics of local labour markets, job displacement and agglomeration economies. The contributions are fivefold.

The first contribution of this dissertation is mainly to the field of labour economics, as it shows that the spatial structure of homes and jobs represents relevant margins of adjustment for workers who have been displaced. I examine how workers respond to job displacement by focusing on spatial margins of adjustment, including the commuting distance and changing home. So far, the argument of literature on job displacement, based on human capital theory, goes that displaced workers lose their human capital and wage premiums (Jacobson et al., 1993). Consequently, displaced workers experience substantial losses in post-displacement employment and wage outcomes. The focus on losses in human capital, however, ignores the way displaced workers could use margins of adjustment that are related to space. I focus on commuting patterns and relocation decisions, as these spatial margins are key to employment outcomes and wage dispersion (e.g., see Manning (2003); Van Ommeren and Fosgerau (2009)). I show that displaced workers, in addition to substantial losses in employment and wage, experience an increase in the commuting distance and a small decrease in the probability of changing home. The results suggest that commuting is a more relevant margin of labour adjustment for workers who have been displaced, while changing home is more relevant for family reasons (Huttunen et al., 2018). The patterns in the post-displacement effects change over the worker's post-displacement period – the negative displacement effect on wages becomes more pronounced, whereas the increase in the commuting distance diminishes. These patterns in wage and commute are caused by workers who entered their first job after displacement. This finding suggests that displaced workers who have longer unemployment duration prefer shorter commutes to higher wages. From a methodological point of view, my research contributes also to the field of urban economics by using a quasi-experimental empirical design involving job displacement.

The second contribution is to highlight the importance of workers' characteristics for the displacement effects. Surprisingly, the job displacement literature neglects the importance of workers' housing state. The role of workers' housing state in displacement effects is interesting, as the effect of the displaced worker's housing state on employment is theoretically ambiguous. On the one hand, relative to tenants, homeowners search more efficiently for jobs through better search methods (Morescalchi, 2016). Moreover, negative home equity owners and mortgage owners compared to outright owners have a greater search effort due to differences in payment obligations, as they have a stronger financial incentive to become employed. This could increase employment prospects. On the other, homeowners relative to tenants, and mortgage owners compared to outright owners, are less able to relocate home, because of higher transaction costs of changing home and a more severe equity constraint, respectively (Chan, 2001). This could hinder employment prospects. Alternatively, workers who differ in housing state may have a different reservation wage or different valuation of commuting costs. Consequently, displaced workers who face a severe home equity constraint may be more willing to accept a job with a lower wage or that is further away in distance. I show that workers' housing state affects the post-displacement outcomes in employment, wages and commute, and the magnitude is comparable to that of other relevant worker characteristics such as gender. From a theoretical point of view, I contribute to the literature by showing that more leveraged displaced owners experience modest losses in employment, but large losses in wages. The results suggest that workers who are more leveraged have a stronger financial incentive to become employed – they prefer sooner re-employment to more modest wage losses. Moreover, I show that displaced outright owners and male workers experience a relatively high increase in the commuting distance, which indicates that their local labour market is relatively large and flexible in structure.

The third contribution is from a methodological point of view, as I developed a Stata command entitled flowbca. Flowbca is a flow-based agglomerative hierarchical clustering algorithm in Stata. So far, the flow-based cluster algorithms available in Stata focus on visualising social networks (e.g., see Corten (2011); Miura (2012)). A limitation of these algorithms is that they are not able to flexibly aggregate units based on relational data of flows. Other agglomerative hierarchical clustering algorithms that are available in Stata are distance-based instead of flow-based. The key difference between distance-based and flow-based clustering is that distance-based clustering uses the distance in for example location or unemployment statistics between two regional units as a measure of similarity instead of relational data of flows. The flowbca command is written in Stata's development language Mata. The use of the Mata programming language ensures that the algorithm is faster and more efficient compared to matrix operations done in Stata without making use of Mata. The flowbca command might be very useful for researchers in various research disciplines, as in many disciplines the availability of relational data of flows has been increasing over the last years. Examples of applications I introduced include (i) the clustering of regional units into local labour markets based on relational data of commuting flows, which is relevant for the field of economic geography, (ii) the clustering of production chains into industries based on data of trade flows, which is relevant for input-output analysis, and (iii) the clustering of individuals into communities based on data of friendship ties, which is relevant for social network analysis. I argue that the flowbca command is user-friendly and well equipped to define meaningful clusters.

The fourth contribution is mainly to the field of urban and regional economics, as it highlights the relevance of the regional aggregation level and subgroup-specific local labour markets for the estimation of agglomeration economies. I define aggregate local labour markets endogenous to the worker's geographical location and commuting outcomes. Moreover, I define subgroup-specific local labour markets endogenous to the worker's geographical location, commuting outcomes, and gender and education level. Endogenising the worker's local labour market is relevant, as a number of papers suggest that workers' commuting behaviour changed over time and is different among subgroups of workers (Green et al., 1986; Manning, 2003; Farmer and Fotheringham, 2011). I show that the estimate of the urban wage premium increases by a magnitude of two to three in the level of regional aggregation. From a conceptual point of view, this finding suggests that the positive externalities from agglomeration economies take place at a relatively high spatial scale of local labour markets. This finding is robust to aggregate and subgroup-specific local labour markets, and different empirical specifications. Importantly, the role of regional aggregation in the returns to agglomeration in wages is similar among the six subgroups of workers that vary in gender and education level. This observation indicates that the modifiable areal unit problem is not the main driver behind the results, as the measurement error in workers' local labour market of using a pre-defined aggregate regional classification likely differs between subgroups. I argue that the importance of the level of regional aggregation that is used to operate local labour markets for the empirical results and conclusion might be larger for research on larger geographical areas. Specifically, larger geographic areas such as the U.S. or Europe can be characterised by regional classifications that vary more in the level of regional aggregation, particularly because these areas could be characterised by a higher level of regional aggregation. My results indicate that the use of a higher level of regional aggregation increases the possibility of overestimating agglomeration benefits, as I show that estimates of agglomeration benefits are increasing in the level of aggregation.

The fifth contribution is to underscore that agglomeration benefits for wages are both education-biased and gender-biased. Holding the level of regional aggregation constant among the different subgroups, I show that the urban wage premium of high-educated workers, compared to low-educated workers, is about 100 per cent higher. I find no gender differential in the urban wage premium if I hold the level of regional aggregation constant. Importantly, the descriptive analysis in Chapter 4 shows that female workers are characterised by a smaller local labour market than male workers. This observation suggests that male workers experience higher agglomeration benefits than female workers, as the agglomeration benefits are increasing in the level of regional aggregation of local labour markets. One of the main points made in this dissertation is that the structure of a local labour market could differ across regions and, notably, also among different subgroups of workers. Given that most studies use aggregate pre-defined regional classifications that represent large areas (see De Groot et al. (2016) for a comprehensive overview), I argue that the agglomeration benefits are generally overestimated for workers who are characterised by small local labour markets such as female and low-educated workers, and underestimated for workers who are characterised by large local labour markets such as male and high-educated workers.

5.3 Limitations and Directions for Future Research

In this section, I will discuss the main limitations of the research that is documented in this dissertation.

The first limitation relates to the analysis of the spatial response to job displacement and the role of worker characteristics in the displacement effects. The main limitation is that I do not examine to what extent the spatial response and labour response to job displacement is jointly determined. That is, for workers who are displaced, changes in wage, commute and home, are jointly determined. Note that the results suggest that displaced workers who have longer unemployment duration prefer smaller increases in commute to higher losses in wage. However, this finding is inferred from separate reduced-form models instead of structural models. I recognise this limitation and I suggest further analysis of the joint decisions of the four margins of labour adjustment in response to job displacement as a direction for future research.

Another limitation is that I do not analyse the precise causes of differences in spatial outcomes, such as the causes of differences in the spatial response to job displacement. Moreover, I do not examine the causes of differences in the structure of workers' local labour markets. Based on observed worker characteristics, I am able to reveal the extent to which (i) the response to job displacement, (ii) the local labour market, and (iii) the agglomeration benefits, differ among subgroups of workers. However, the underlying micro-foundations of differences in workers' spatial outcomes remains an important research topic, particularly in the field of economic sociology (e.g., see Fernandez and Su (2004)). A related direction for future research is the analysis of the role of agglomeration economies in the spatial response to job displacement, as I only focus on the importance of agglomeration economies for the labour response to job displacement.

The main limitation of my research on the structure of workers' local labour markets is that I do not simultaneously allow for continuous local labour markets (e.g., in the spirit of Manning and Petrongolo (2017)) and for differences in local labour markets among different types of workers (e.g., subgroup-specific local labour markets). Manning and Petrongolo (2017) recognise the continuous nature of geographic space and develop a model of job search that allows for overlapping local labour markets, which avoids arbitrary border effects. These effects are driven by the drawback that local labour markets are operated such that they are non-overlapping. For example, two workers could be based in different administrative areas, as they live across borders, even though they are closely located to each other. However, Manning and Petrongolo (2017) model low-skill labour markets that tend to be more local, as their data sample contains workers with a relatively low education level. My dissertation uses an alternative approach by analysing subgroup-specific local labour markets, in which local labour markets differ among and not within subgroups of workers.

A last limitation, related to the analysis of agglomeration benefits, is that I do not disentangle the roles of urbanisation economies and localisation economies in agglomeration benefits. In the empirical analyses on the importance of regional aggregation for the returns to agglomeration in wages and employment, only the variables employment density and area size are included in the empirical specification. Note that employment density generally plays a more important role in the returns to agglomeration than the more endogenous variables that approximate localisation economies (Combes et al., 2008; Groot et al., 2014; Combes and Gobillon, 2015). Importantly, a worker or industry may be more productive due to higher total economic activity and due to higher industry specialisation. Thus the urbanisation economies and localisation economies are correlated to some extent, as total economic activity is likely to be positively correlated to higher industry specialisation (Combes and Gobillon, 2015). Thereby, the variables that I include to capture agglomeration economies are likely to capture a substantial part of urbanisation externalities and localisation externalities. Another related research topic that I do not touch upon in this dissertation concerns the co-location behaviour of heterogeneous firms, which affects the magnitude of agglomeration spillovers among firms and across local labour markets (Van Oort et al., 2012; Knoben et al., 2016; Faggio et al., 2016). I leave the structure of firms' local labour markets and its economic consequences for future research.

5.4 Societal Relevance: Policy and Real-World Implications

In this section, I discuss the societal relevance of my research from both a policy perspective and a real-world perspective.

One of the main goals of this dissertation is to get a better understanding of the structures and economic consequences of workers' local labour markets. My research reveals the differences in the structure of local labour markets of subgroups of workers who differ in gender and education level. From a policy perspective, a better understanding of the differences in the structure of local labour markets is relevant for multiple socio-economic reasons. This would provide opportunities for increasing the efficiency of policy decisions, including the effectiveness of

(i) regional policies that aim to stimulate agglomeration benefits and regional productivity growth through improved co-operation between city-regions and geographical upscaling of economic activities (OECD, 2007, 2014). Over the last decades, the Dutch government increased decentralisation by giving more responsibilities to municipalities (e.g., see Kattenberg and Vermeulen (2018)). Conversely, my research shows that the agglomeration benefits are more prevalent at a high regional aggregation level, which suggests urban and regional policy should tend to be more generic and centralised. Indeed, the OECD (2014) argues that the Netherlands lacks a comprehensive urban and regional policy framework at the national level. Moreover, my research on the structure of local labour markets is relevant for the topics of upscaling and rescaling of provinces and municipalities (OECD, 2014).

(ii) place-based policies targeted at specific regions or subgroups of the population (Glaeser and Gottlieb, 2008; Neumark and Simpson, 2015), should tend to be more specific, local and decentralised. For example, the subgroup-specific local labour markets returned by the flow-based cluster algorithm underscore that the structure and size of targeted areas should depend strongly on the targeted subgroup of workers and degree of urbanisation of the location.

(iii) labour market policies which aim to increase the matching quality of worker to employer or which limit wage inequality (Moretti, 2011; Crépon and Van den Berg, 2016). See the discussion below on the project in collaboration with the temporary work and human resource consulting services company Randstad Holding for examples of insights that may improve job matches.

(iv) transportation policies directed at decreasing congestion (Small and Verhoef, 2007). For example, a better understanding of commuting flows across the Netherlands might benefit infrastructure planning. According to OECD (2007), congestion in the so-called "Randstad" regional area costs about 2 billion euro per year.

The analysis on the response to job displacement is relevant for policies that aim to limit the impact of negative employment shocks, as it provides a better understanding of how workers respond to job displacement by revealing the use of margins of labour adjustment. I show that job displacement is costly from a societal perspective – besides severe losses in employment and wages, displaced workers experience an increase in the commuting distance of about 20 per cent compared to the mean commute. In addition, the results suggest that displaced workers make trade-offs that affect their post-displacement labour market outcomes. The findings indicate that displaced workers who have been unemployed longer prefer more modest increases in commuting distance over more modest losses in wages. Importantly, the commute margin seems a more relevant margin of adjustment than changing home, as job displacement decreases the probability of household relocation. This observation can be explained by the fact that the surface area of the Netherlands is relatively small. In this regard, the housing decision is not always driven by the location of the job as is suggested by Van Leuvensteijn and Koning (2004). Conversely, displaced workers search locally for jobs that are within their commuting range. Overall, the housing market may hamper displaced workers' labour market outcomes through a relatively low geographical mobility, but the commuting margin of labour adjustment helps them to become resilient to a negative employment shock.

I also analyse whether there is a strong role of worker characteristics in the displacement effects, focusing in particular on workers' housing state. This research is relevant for policy purposes, as it gives a better understanding of whether homeownership or high mortgage debt hampers workers' post-displacement labour market outcomes. The worker's housing state received a lot of attention recently, because of the Great Recession that started in 2008. This recent crisis led to an increase in the share of homeowners that face negative home equity – also known as underwater homeowners – as transaction prices and property values declined. The results show that more leveraged owners experience a relatively modest loss in employment and high loss in wage. Thus my research suggests that workers who are more leveraged have a stronger incentive to become re-employed, and realise re-employment at the expense of higher losses in wages. Moreover, geographic immobility of homeowners (compared to tenants) and of negative home equity owners (compared to other owners) does not hinder post-displacement employment outcomes. Note, however, that over 85 per cent of Dutch tenants rent social housing, which makes them relatively immobile. From a labour economics perspective, the loan-to-value ratio limit of 106 per cent that has been introduced in 2011 might not be beneficial, as my research suggests that having a large share of underwater homeowners does not seem to hinder the functioning of the labour market. In fact, underwater owners have a stronger financial incentive to become employed and, indeed, become rapidly re-employed. Moreover, the commute margin of labour adjustment could be used to become employed if workers are not able to relocate to a different home. From a societal point of view, however, it may not be preferable to have a large share of underwater homeowners, as high debt may lead to negative emotional effects.

From the perspective of real-world implications, my research already made some impact. I started a project together with Randstad Holding to examine differences in the size and structure of local labour markets of temporary workers in the Netherlands. Randstad Holding is a global leader in flexible work and human resources services. The goal of the project was to construct local labour markets, using the flow-based cluster algorithm. The main input was commuting flows from home to work location of temporary workers that are employed by Randstad Holding. Attention is paid to differences in local labour markets of workers who differ in, among others, (i) gender, (ii) skill, (iii) occupation and (iv) Randstad Holding business divisions. The Dutch business divisions consist of Randstad NL, Tempo-Team and Yacht. The research project provided insight into the core and structure of local labour markets of temporary workers. These insights are relevant for various reasons, including the distribution of Randstad Holding firm offices, recruitment of workers and clients (clients are the firms that employ the temporary workers), and the ability to advise clients of Randstad Holding about locations of new offices. In addition, the project gives insights about dispersion of regional scarcity of a specific type of workers. Finally, the insights of the research project may lead to an increase in the matching rate and matching quality of workers to clients, as individual-specific local labour markets could be built in as a feature in the process of job matching.

By gaining a better understanding of the winners and losers of the increasing urbanisation of urban areas, my research touches upon an important societal trade-off between equality and efficiency. The results show that workers' wages are increasing in the density of their local labour market. Furthermore, my research reveals that workers' resilience to a negative employment shock depends on the geographical location and the density of the local labour market. Specifically, I show that displaced workers who reside in relatively dense local labour markets, compared to displaced workers who reside in more sparse local labour markets, experience a more modest loss in wages but a comparable loss in employment. In this regard, living in a dense local labour market provides economic value for workers, as it leads to higher wages although it does not affect workers' individual employment outcomes. Moreover, I show that the economic value of living in a dense local labour market is subgroup-specific. The urban wage premium is higher for more educated workers and male workers. Conversely, the returns to agglomeration in employment and wage outcomes for workers who have been displaced is less subgroup-specific. This finding suggests that the importance of the matching mechanism of agglomeration for subgroup differentials in agglomeration benefits is reasonably small.

My research also shows that the use of a level of regional aggregation to operate local labour markets affects the empirical results and conclusion of research. This finding is relevant for regional policies that are based on conclusions of research in the field of economic geography. For example, the OECD (2007, 2016) concludes that the urban core of the Netherlands – i.e. the Randstad regional area – under-performs in terms of labour productivity growth. The OECD (2007, 2016) argues that the low productivity growth is due to a relatively high regional fragmentation in the Netherlands. Policy advice is to strive for geographical up-scaling of economic activities and policies (Van Oort et al., 2015), bringing the economies of municipalities closer together. My findings suggest that a substantial share of the positive externalities from agglomeration economies takes place at a high spatial scale. A question is whether the conclusions of OECD (2007, 2016) are robust to the use of a regional classification that uses a higher level of regional aggregation to operate self-contained economic areas. All in all, I argue that policy makers should recognise the endogeneity of workers' local labour market in their work by focusing on the level of regional aggregation, subgroup of workers and the structure and size of local labour markets at which policies are targeted.

References

Adamson, D. W., D. E. Clark, and M. D. Partridge (2004). Do urban agglomeration effects and household amenities have a skill bias? *Journal of Regional Science 44*(2), 201–224.

Addison, J. T. and P. Portugal (1987). The effect of advance notification of plant closings on unemployment. *Industrial and Labor Relations Review 41*(1), 3–16.

Andersson, F., S. Burgess, and J. I. Lane (2007). Cities, matching and the productivity gains of agglomeration. *Journal of Urban Economics 61*(1), 112–128.

Andersson, F. and T. Mayock (2014). How does home equity affect mobility? *Journal of Urban Economics 84*, 23–39.

Arzaghi, M. and J. V. Henderson (2008, October). Networking off Madison Avenue. *Review of Economic Studies 75*(4), 1011–1038.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The geography of trade and technology shocks in the United States. *American Economic Review: Papers and Proceedings 103*(3), 220–225.

Bacolod, M., B. S. Blum, and W. C. Strange (2009). Skills in the city. *Journal of Urban Economics 65*(2), 136–153.

Baert, S., F. Heylen, and D. Isebaert (2014). Does homeownership lead to longer unemployment spells? The role of mortgage payments. *De Economist 162*(3), 263–286.

Bathelt, H. and P.-F. Li (2014). Global cluster networksforeign direct investment flows from Canada to China. *Journal of Economic Geography 14*(1), 45–71.

Battu, H., A. Ma, and E. Phimister (2008). Housing tenure, job mobility and unemployment in the UK. *The Economic Journal 118*(527), 311–328.

Baum-Snow, N. and R. Pavan (2012). Understanding the city size wage gap. *Review of Economic Studies 79*(1), 88–127.

Beguerisse-Daz, M., G. Garduo-Hernndez, B. Vangelov, S. N. Yaliraki, and M. Barahona (2014). Interest communities and flow roles in directed networks: the Twitter network of the UK riots. *Journal of The Royal Society Interface 11*(101), 1–12.

Berliant, M., R. R. Reed, and P. Wang (2006). Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics 60*(1), 69–95.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bhaskar, V., A. Manning, and T. To (2002). Oligopsony and monopsonistic competition in labor markets. *Journal of Economic Perspectives 16*(2), 155–174.

Bischof, D. (2017). New graphic schemes for Stata: plotplain & plottig. *Stata Journal 17*(3), 748–759.

Black, D., N. Kolesnikova, and L. Taylor (2009). Earnings functions when wages and prices vary by location. *Journal of Labor Economics 27*(1), 21–47.

Blumenberg, E. (2004). En-gendering effective planning: Spatial mismatch, low-income women, and transportation policy. *Journal of the American Planning Association 70*(3), 269–281.

Bover, O., M. Arellano, and S. Bentolila (2002). Unemployment duration, benefit duration and the business cycle. *The Economic Journal 112*(479), 223–265.

Bowles, S. (1970). Migration as investment: Empirical tests of the human investment approach to geographical mobility. *Review of Economics and Statistics 52*(4), 356–362.

Brezzi, M., M. Piacentini, K. Rosina, and D. Sanchez-Serra (2012). Redefining urban areas in OECD countries. In *Redefining "Urban"*, pp. 19–58. Organisation for Economic Co-operation and Development.

Briant, A., P. P. Combes, and M. Lafourcade (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics 67*(3), 287–302.

Bricker, J. and B. Bucks (2016). Negative home equity, economic insecurity, and household mobility over the Great Recession. *Journal of Urban Economics 91*, 1–12.

Brown, J. and D. A. Matsa (2017). Locked in by leverage: Job search during the housing crisis. SSRN Scholarly Paper ID 2880784, Social Science Research Network, Rochester, NY.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (1999). Why is central Paris rich and downtown Detroit poor? An amenity-based theory. *European Economic Review 43*(1), 91–107.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (2002). Local labor markets, job matching, and urban location. *International Economic Review 43*(1), 155–171.

Burger, M. J., F. G. Van Oort, and G. A. Van der Knaap (2008). A treatise on the geographical scale of agglomeration externalities and the Modifiable Areal Unit Problem. ERIM Report Series Research in Management ERS-2008-076-ORG.

Caliendo, M., A. C. Gielen, and R. Mahlstedt (2015). Home-ownership, unemployed's job search behavior and post-unemployment outcomes. *Economics Letters 137*, 218–221.

Caliendo, M. and S. Kopeinig (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys 22*(1), 31–72.

Carlsen, F., J. Rattsø, and H. E. Stokke (2016). Education, experience, and urban wage premium. *Regional Science and Urban Economics 60*, 39–49.

Carrington, W. J. (1993). Wage losses for displaced workers: Is it really the firm that matters? *Journal of Human Resources 28*(3), 435–462.

Carrington, W. J. and A. Zaman (1994). Interindustry variation in the costs of job displacement. *Journal of Labor Economics 12*(2), 243–275.

CBS (2018). Centraal Bureau voor de Statistiek, Den Haag/Heerlen.

Chan, S. (2001). Spatial lock-in: Do falling house prices constrain residential mobility? *Journal of Urban Economics 49*(3), 567–586.

Chan, S. and A. H. Stevens (1999). Employment and retirement following a late-career job loss. *American Economic Review: Papers and Proceedings 89*(2), 211–216.

Chan, S. and A. H. Stevens (2001). Job loss and employment patterns of older workers. *Journal of Labor Economics 19*(2), 484–521.

Chen, M. X. and M. O. Moore (2010). Location decision of heterogeneous multinational firms. *Journal of International Economics 80*(2), 188–199.

Chowdhury, G. and S. Nickell (1985). Hourly earnings in the United States: Another look at unionization, schooling, sickness, and unemployment using PSID data. *Journal of Labor Economics 3*(1), 38–69.

Ciccone, A. (2002). Agglomeration effects in Europe. *European Economic Review 46*(2), 213–227.

Ciccone, A. and R. E. Hall (1996). Productivity and the density of economic activity. *American Economic Review 86*(1), 54–70.

Clark, W. A. V., Y. Huang, and S. Withers (2003). Does commuting distance matter? Commuting tolerance and residential change. *Regional Science and Urban Economics 33*(2), 199–221.

Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics 63*(2), 723–742.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica 80*(6), 2543–2594.

Combes, P.-P. and L. Gobillon (2015). Chapter 5 - The empirics of agglomeration economies. In G. Duranton, J. V. Henderson, and W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, Volume 5, pp. 247–348. Elsevier.

Coombes, M. G., A. E. Green, and S. Openshaw (1986). An efficient algorithm to generate official statistical reporting areas: The case of the 1984 travel-to-work areas revision in Britain. *Journal of the Operational Research Society 37*(10), 943–953.

Corten, R. (2011). Visualization of social networks in Stata using multidimensional scaling. *Stata Journal 11*(1), 52–63.

Costa, D. L. (2000). From mill town to board room: The rise of women's paid labor. *Journal of Economic Perspectives 14*(4), 101–122.

Couch, K. A. and D. W. Placzek (2010). Earnings losses of displaced workers revisited. *American Economic Review 100*(1), 572–589.

Coulson, N. E. and L. M. Fisher (2002). Tenure choice and labour market outcomes. *Housing Studies 17*(1), 35–49.

Coulson, N. E. and P. L. E. Grieco (2013). Mobility and mortgages: Evidence from the PSID. *Regional Science and Urban Economics 43*(1), 1–7.

Crane, R. (1996). The influence of uncertain job location on urban form and the journey to work. *Journal of Urban Economics 39*(3), 342–356.

Crane, R. (2007). Is there a quiet revolution in women's travel? Revisiting the gender gap in commuting. *Journal of the American Planning Association 73*(3), 298–316.

Crépon, B. and G. J. Van den Berg (2016). Active labor market policies. *Annual Review of Economics 8*(1), 521–546.

D'Costa, S. and H. G. Overman (2014). The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics 48*, 168–179.

De Groot, H. L., J. Poot, and M. J. Smit (2016). Which agglomeration externalities matter most and why? *Journal of Economic Surveys 30*(4), 756–782.

De La Roca, J. and D. Puga (2016). Learning by working in big cities. *Review of Economic Studies*, 1–37.

Deelen, A., M. De Graaf-Zijl, and W. Van den Berge (2014). Labour market effects of job displacement for prime-age and older workers. CPB Discussion Paper 285.

Delgado, M., M. E. Porter, and S. Stern (2016). Defining clusters of related industries. *Journal of Economic Geography 16*(1), 1–38.

Demyanyk, Y., D. Hryshko, M. J. Luengo-Prado, and B. E. Sørensen (2017). Moving to a job: The role of home equity, debt, and access to credit. *American Economic Journal: Macroeconomics 9*(2), 149–181.

Di Addario, S. (2011). Job search in thick markets. *Journal of Urban Economics 69*(3), 303–318.

Di Addario, S. and E. Patacchini (2008). Wages and the city. Evidence from Italy. *Labour Economics 15*(5), 1040–1061.

Dietzenbacher, E., B. Los, R. Stehrer, M. Timmer, and G. d. Vries (2013). The construction of world input-output tables in the WIOD project. *Economic Systems Research 25*(1), 71–98.

DiPasquale, D. and E. L. Glaeser (1999). Incentives and social capital: Are homeowners better citizens? *Journal of Urban Economics 45*(2), 354–384.

Duranton, G. (2015). Delineating metropolitan areas: Measuring spatial labour market networks through commuting patterns. In T. Watanabe, I. Uesugi, and A. Ono (Eds.), *The Economics of Interfirm Networks*, Number 4 in Advances in Japanese Business and Economics, pp. 107–133. Springer Japan.

Duranton, G. and H. G. Overman (2005). Testing for localization using micro-geographic data. *Review of Economic Studies 72*(4), 1077–1106.

Duranton, G. and D. Puga (2004). Chapter 48 - Micro-foundations of urban agglomeration economies. In J. V. Henderson and J.-F. Thisse

(Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2063–2117. Elsevier.

Eliason, M. and D. Storrie (2006). Lasting or latent scars? Swedish evidence on the longterm effects of job displacement. *Journal of Labor Economics 24*(4), 831–856.

Ellison, G. and E. L. Glaeser (1997). Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy 105*(5), 889–927.

Faggio, G., O. Silva, and W. C. Strange (2016). Heterogeneous agglomeration. *Review of Economics and Statistics 99*(1), 80–94.

Fang, C. and D. Yu (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning 162*, 126–136.

Farber, H. S. (1999). Chapter 37 - Mobility and stability: The dynamics of job change in labor markets. In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume III, pp. 2439–2483. Elsevier.

Farber, H. S. (2017). Employment, hours, and earnings consequences of job loss: US evidence from the displaced workers survey. *Journal of Labor Economics 35*(S1), S235–S272.

Farber, H. S., R. Hall, and J. Pencavel (1993). The incidence and costs of job loss: 1982-91. *Brookings Papers on Economic Activity. Microeconomics 1993*(1), 73–132.

Farmer, C. J. Q. and A. S. Fotheringham (2011). Network-based functional regions. *Environment and Planning A 43*(11), 2723–2741.

Fernandez, R. M. (2001). Skill-biased technological change and wage inequality: Evidence from a plant retooling. *American Journal of Sociology 107*(2), 273–320.

Fernandez, R. M. and C. Su (2004). Space in the study of labor markets. *Annual Review of Sociology 30*(1), 545–569.

Ferreira, F., J. Gyourko, and J. Tracy (2010). Housing busts and household mobility. *Journal of Urban Economics 68*(1), 34–45.

Flatau, P., M. Forbes, P. H. Hendershott, and G. Wood (2003). Homeownership and unemployment: The roles of leverage and public housing. NBER Working Paper 10021.

Foote, C. L., K. Gerardi, and P. S. Willen (2008). Negative equity and foreclosure: Theory and evidence. *Journal of Urban Economics 64*(2), 234–245.

Ford Jr, L. R. and D. R. Fulkerson (1962). *Flows in networks*. Princeton University Press.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports 486*(35), 75–174.

Fotheringham, A. S. and D. W. S. Wong (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A 23*(7), 1025–1044.

Frederiksen, A., R. Ibsen, M. Rosholm, and N. Westergaard-Nielsen (2013). Labour market signalling and unemployment duration: An empirical analysis using employer–employee data. *Economics Letters 118*(1), 84–86.

Freeman, R. B. (1984). Longitudinal analyses of the effects of trade unions. *Journal of Labor Economics 2*(1), 1–26.

Ghent, A. C. and M. Kudlyak (2011). Recourse and residential mortgage default: Evidence from US states. *Review of Financial Studies 24*(9), 3139–3186.

Gibbons, R. and L. F. Katz (1991). Layoffs and lemons. *Journal of Labor Economics 9*(4), 351–380.

Glaeser, E. L. and J. D. Gottlieb (2008). The economics of place-making policies. *Brookings Papers on Economic Activity 39*(1 (Spring)), 155–253.

Glaeser, E. L., M. E. Kahn, and J. Rappaport (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics 63*(1), 1–24.

Glaeser, E. L. and D. C. Maré (2001). Cities and skills. *Journal of Labor Economics 19*(2), 316–342.

Goos, M., A. Manning, and A. Salomons (2009). Job polarization in Europe. *American Economic Review: Papers and Proceedings 99*(2), 58–63.

Goos, M., A. Manning, and A. Salomons (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review 104*(8), 2509–26.

Goss, E. P. and J. M. Phillips (1997). The impact of home ownership on the duration of unemployment. *Review of Regional Studies 27*(1), 9–27.

Gould, E. D. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *Review of Economic Studies 74*(2), 477–506.

Green, A. E., M. G. Coombes, and D. W. Owen (1986). Gender-specific local labour market areas in England and Wales. *Geoforum 17*(3), 339–351.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica 45*(1), 1–22.

Griliches, Z. and J. A. Hausman (1986). Errors in variables in panel data. *Journal of Econometrics 31*(1), 93–118.

Groot, S. P., H. L. De Groot, and M. J. Smit (2014). Regional wage differences in the Netherlands: Micro evidence on agglomeration externalities. *Journal of Regional Science 54*(3), 503–523.

Hamermesh, D. S. (1987). The costs of worker displacement. *The Quarterly Journal of Economics 102*(1), 51–75.

Hanson, S. and G. Pratt (1988). Spatial dimensions of the gender division of labor in a local labor market. *Urban Geography 9*(2), 180–202.

Hardoy, I. and P. Schøne (2014). Does the impact of plant closure on labour market attachment differ between immigrants and native workers across the business cycle? *Empirical Economics 46*(1), 229–252.

Head, A. and H. Lloyd-Ellis (2012). Housing liquidity, mobility and the labour market. *Review of Economic Studies 79*, 1559–1589.

Heckman, J. J., H. Ichimura, J. Smith, and P. Todd (1998). Characterizing selection bias using experimental data. *Econometrica 66*(5), 1017–1098.

Heckman, J. J., H. Ichimura, and P. E. Todd (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies 64*(4), 605–654.

Heckman, J. J. and J. A. Smith (1999). The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies. *The Economic Journal 109*(457), 313–348.

Hellerstein, J. K., M. J. Kutzbach, and D. Neumark (2016). Labor market networks and recovery from mass layoffs before, during, and after the Great Recession. IZA Discussion Paper 9852.

Helsley, R. W. and W. C. Strange (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics 20*(2), 189–212.

Henderson, V., A. Kuncoro, and M. Turner (1995). Industrial development in cities. *Journal of Political Economy 103*(5), 1067–1090.

Henley, A., R. Disney, and A. Carruth (1994). Job tenure and asset holdings. *The Economic Journal 104*(423), 338–349.

Hijzen, A., R. Upward, and P. W. Wright (2010). The income losses of displaced workers. *Journal of Human Resources 45*(1), 243–269.

Hoynes, H. W. (2000). Local labor markets and welfare spells: Do demand conditions matter? *Review of Economics and Statistics 82*(3), 351–368.

Hu, L. and C. Taber (2011). Displacement, asymmetric information, and heterogeneous human capital. *Journal of Labor Economics 29*(1), 113–152.

Huttunen, K., J. Møen, and K. G. Salvanes (2011). How destructive is creative destruction? Effects of job loss on job mobility, withdrawal and income. *Journal of the European Economic Association 9*(5), 840–870.

Huttunen, K., J. Møen, and K. G. Salvanes (2018). Job loss and regional mobility. *Journal of Labor Economics 36*(2), 479–509.

Hyslop, R. and G. W. Imbens (2001). Bias from classical and other forms of measurement error. *Journal of Business & Economic Statistics 19*(4), 475–481.

Iacus, S. M., G. King, and G. Porro (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association 106*(493), 345–361.

Ichino, A., G. Schwerdt, R. Winter-Ebmer, and J. Zweimüller (2017). Too old to work, too young to retire? *Journal of the Economics of Ageing 9*, 14–29.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993). Earnings losses of displaced workers. *American Economic Review 83*(4), 685–709.

Kattenberg, M. and W. Vermeulen (2018). The stimulative effect of an unconditional block grant on the decentralized provision of care. *International Tax and Public Finance 25*(1), 166–199.

Kattenberg, M. A. C. and W. H. J. Hassink (2017). Who moves out of social housing? The effect of rent control on housing tenure choice. *De Economist 165*(1), 43–66.

Katz, L. F. and B. D. Meyer (1990). The impact of the potential duration of unemployment benefits on the duration of unemployment. *Journal of Public Economics 41*(1), 45–72.

Knoben, J., A. Arikan, F. Van Oort, and O. Raspe (2016). Agglomeration and firm performance: One firm's medicine is another firm's poison. *Environment and Planning A 48*(1), 132–153.

Kothari, S., I. Saporta-Eksten, and E. Yu (2013). The (un)importance of geographical mobility in the Great Recession. *Review of Economic Dynamics 16*(3), 553–563.

Krolikowski, P. (2018). Choosing a control group for displaced workers. *ILR Review 71*(5), 1232–1254.

Kuhn, P. (2002). *Losing work, moving on: International perspectives on worker displacement.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Lee, S. (2010). Ability sorting and consumer city. *Journal of Urban Economics 68*(1), 20–33.

Leontief, W. W. (1986). *Input–output economics*. Oxford University Press.

Lindo, J. M. (2015). Aggregation and the estimated effects of economic conditions on health. *Journal of Health Economics 40*, 83–96.

MacRae, D. (1960). Direct factor analysis of sociometric data. *Sociometry 23*(4), 360–371.

Madden, J. F. (1981). Why women work closer to home. *Urban Studies 18*(2), 181–194.

Madden, J. F. (1987). Gender differences in the cost of displacement: An empirical test of discrimination in the labor market. *American Economic Review: Papers and Proceedings 77*(2), 246–251.

Madden, J. F. and L. C. Chiu (1990). The wage effects of residential location and commuting constraints on employed married women. *Urban Studies 27*(3), 353–369.

Manning, A. (2003). The real thin theory: Monopsony in modern labour markets. *Labour Economics 10*(2), 105–131.

Manning, A. (2009). The plant size-place effect: Agglomeration and monopsony in labour markets. *Journal of Economic Geography 10*(5), 717–744.

Manning, A. and B. Petrongolo (2017). How local are labor markets? Evidence from a spatial job search model. *American Economic Review 107*(10), 2877–2907.

Masser, I. and P. J. B. Brown (1975). Hierarchical aggregation procedures for interaction data. *Environment and Planning A: Economy and Space 7*(5), 509–523.

Mayock, T. (2015). Wages, housing prices and commutes. *Real Estate Economics 44*(1), 258–300.

Meekes, J. and W. H. J. Hassink (2018a). Endogenous local labour markets, regional aggregation and agglomeration economies. TKI Discussion Paper 18-03.

Meekes, J. and W. H. J. Hassink (2018b). flowbca: A flow-based cluster algorithm in Stata. *Stata Journal 18*(3), 564–584.

Meekes, J. and W. H. J. Hassink (2019). The role of the housing market in workers' resilience to job displacement after firm bankruptcy. *Journal of Urban Economics 109*, 41–65. `https://doi.org/10.1016/j.jue.2018.11.005`.

Melo, P. C., D. J. Graham, and R. B. Noland (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics 39*(3), 332–342.

Mion, G. and P. Naticchioni (2009). The spatial sorting and matching of skills and firms. *Canadian Journal of Economics/Revue Canadienne d'Économique 42*(1), 28–55.

Miura, H. (2012). Stata graph library for network analysis. *Stata Journal 12*(1), 94–129.

Modestino, A. S. and J. Dennett (2013). Are American homeowners locked into their houses? The impact of housing market conditions on state-to-state migration. *Regional Science and Urban Economics 43*(2), 322–337.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 3855–3890.

Morescalchi, A. (2016). The puzzle of job search and housing tenure: A reconciliation of theory and empirical evidence. *Journal of Regional Science 56*(2), 288–312.

Moretti, E. (2011). Chapter 14 - Local labor markets. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 1237–1313. Elsevier.

Moriizumi, Y. and M. Naoi (2011). Unemployment risk and the timing of homeownership in Japan. *Regional Science and Urban Economics 41*(3), 227–235.

Mulalic, I., J. Van Ommeren, and N. Pilegaard (2014). Wages and commuting: Quasi-natural experiments' evidence from firms that relocate. *The Economic Journal 124*(579), 1086–1105.

Munch, J. R., M. Rosholm, and M. Svarer (2006). Are homeowners really more unemployed? *The Economic Journal 116*(514), 991–1013.

Munch, J. R., M. Rosholm, and M. Svarer (2008). Home ownership, job duration, and wages. *Journal of Urban Economics 63*(1), 130–145.

Munch, J. R. and M. Svarer (2002). Rent control and tenancy duration. *Journal of Urban Economics 52*(3), 542–560.

Neumark, D. and H. Simpson (2015). Chapter 18 - Place-based policies. In G. Duranton, J. V. Henderson, and W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, Volume 5, pp. 1197–1287. Elsevier.

Nimczik, J. S. (2018). Job mobility networks and endogenous labor markets. *Mimeo*.

Niu, Y. and C. Ding (2015). Unemployment matters: Improved measures of labor market distress in mortgage default analysis. *Regional Science and Urban Economics 52*, 27–38.

Nord, S. and Y. Ting (1991). The impact of advance notice of plant closings on earnings and the probability of unemployment. *Industrial and Labor Relations Review 44*(4), 681–691.

OECD (2007). Randstad Holland, Netherlands. OECD Territorial Reviews.

OECD (2010). General assessment of the macroeconomic situation. In *OECD Economic Outlook*, Volume 1 of *2010*, pp. 11–71. OECD Publishing.

OECD (2014). Netherlands. OECD Territorial Reviews.

OECD (2016). The Metropolitan Region of Rotterdam-The Hague, Netherlands. OECD Territorial Reviews.

Openshaw, S. (1983). *The modifiable areal unit problem.* Norwich: Geo Books.

Petrongolo, B. and C. Pissarides (2006). Scale effects in markets with search. *The Economic Journal 116*(508), 21–44.

Petrongolo, B. and C. A. Pissarides (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature 39*(2), 390–431.

Phimister, E. (2005). Urban effects on participation and wages: Are there gender differences? *Journal of Urban Economics 58*(3), 513–536.

Picard, R. and M. Stepner (2015). MERGEPOLY: Stata module to merge adjacent polygons from a shapefile.

Pisati, M. (2008). SPMAP: Stata module to visualize spatial data.

Poorthuis, A., P. Renooy, and P. Tordoir (2015). De veranderende geografie van Nederland. Technical report, Regioplan.

Roberts, J., R. Hodgson, and P. Dolan (2011). It's driving her mad: Gender differences in the effects of commuting on psychological health. *Journal of Health Economics 30*(5), 1064–1076.

Rogerson, R., R. Shimer, and R. Wright (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature 43*(4), 959–988.

Rosenthal, S. S. and W. C. Strange (2001). The determinants of agglomeration. *Journal of Urban Economics 50*(2), 191–229.

Rosenthal, S. S. and W. C. Strange (2003). Geography, industrial organization, and agglomeration. *Review of Economics and Statistics 85*(2), 377–393.

Rosenthal, S. S. and W. C. Strange (2004). Chapter 49 - Evidence on the nature and sources of agglomeration economies. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2119–2171. Elsevier.

Rosenthal, S. S. and W. C. Strange (2008). The attenuation of human capital spillovers. *Journal of Urban Economics 64*(2), 373–389.

Rubin, D. B. (2006). *Matched sampling for causal effects.* Cambridge, UK: Cambridge University Press.

Schwerdt, G. (2011). Labor turnover before plant closure: "Leaving the sinking ship" vs. "Captain throwing ballast overboard". *Labour Economics 18*(1), 93–101.

SCP (2015). SCP arbeidsaanbodpanel 1985–2012. DANS, Sociaal en Cultureel Planbureau, Den Haag.

Simpson, W. and A. Van der Veen (1992). The economics of commuting and the urban labour market. *Journal of Economic Surveys 6*(1), 45–62.

Small, K. A. and E. T. Verhoef (2007). *The economics of urban transportation*. Routledge.

Smith, B. A. (1978). Measuring the value of urban amenities. *Journal of Urban Economics 5*(3), 370–387.

Smith, D. A. and D. R. White (1992). Structure and dynamics of the global economy: Network analysis of international trade 1965–1980. *Social Forces 70*(4), 857–893.

Smith, T. E. and Y. Zenou (1997). Dual labor markets, urban unemployment, and multicentric cities. *Journal of Economic Theory 76*(1), 185–214.

Smith, T. E. and Y. Zenou (2003). Spatial mismatch, search effort, and urban spatial structure. *Journal of Urban Economics 54*(1), 129–156.

Steegmans, J. and W. H. J. Hassink (2018). Decreasing house prices and household mobility: An empirical study on loss aversion and negative equity. *Journal of Regional Science 58*(3), 611–634.

Stein, J. C. (1995). Prices and trading volume in the housing market: A model with down-payment effects. *The Quarterly Journal of Economics 110*(2), 379–406.

Stevens, A. H. (1997). Persistent effects of job displacement: The importance of multiple job losses. *Journal of Labor Economics 15*(1), 165–188.

Tatsiramos, K. (2010). Job displacement and the transitions to reemployment and early retirement for non-employed older workers. *European Economic Review 54*(4), 517–535.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015). An illustrated user guide to the World Input-Output Database: The case of global automotive production. *Review of International Economics 23*(3), 575–605.

Topel, R. (1990). Specific capital and unemployment: Measuring the costs and consequences of job loss. *Carnegie-Rochester Conference Series on Public Policy 33*, 181–214.

Topel, R. H. (1986). Local labor markets. *Journal of Political Economy 94*(3), S111–S143.

Valletta, R. G. (2013). House lock and structural unemployment. *Labour Economics 25*(C), 86–97.

Van den Berg, G. J. (1992). A structural dynamic analysis of job turnover and the costs associated with moving to another job. *The Economic Journal 102*(414), 1116–1133.

Van der Laan, L. and R. Schalke (2001). Reality versus policy: The delineation and testing of local labour market and spatial policy areas. *European Planning Studies 9*(2), 201–221.

Van Leuvensteijn, M. and P. Koning (2004). The effect of home-ownership on labor mobility in the Netherlands. *Journal of Urban Economics 55*(3), 580–596.

Van Ommeren, J. and M. Fosgerau (2009). Workers' marginal costs of commuting. *Journal of Urban Economics 65*(1), 38–47.

Van Ommeren, J., G. J. Van Den Berg, and C. Gorter (2000). Estimating the marginal willingness to pay for commuting. *Journal of Regional Science 40*(3), 541–563.

Van Oort, F., E. Meijers, M. Thissen, M. Hoogerbrugge, and M. Burger (2015). De concurrentiepositie van Nederlandse steden: Van agglomeratiekracht naar netwerkkracht. Technical report, Platform31.

Van Oort, F. G., M. J. Burger, J. Knoben, and O. Raspe (2012). Multilevel approaches and the firm-agglomeration ambiguity in economic growth studies. *Journal of Economic Surveys 26*(3), 468–491.

Von Wachter, T. and S. Bender (2006). In the right place at the wrong time: The role of firms and luck in young workers' careers. *American Economic Review 96*(5), 1679–1705.

Wheeler, C. H. (2001). Search, sorting, and urban agglomeration. *Journal of Labor Economics 19*(4), 879–899.

White, M. J. (1977). A model of residential location choice and commuting by men and women workers. *Journal of Regional Science 17*(1), 41–52.

White, M. J. (1986). Sex differences in urban commuting patterns. *American Economic Review: Papers and Proceedings 76*(2), 368–372.

White, M. J. (1988). Location choice and commuting behavior in cities with decentralized employment. *Journal of Urban Economics 24*(2), 129–152.

Yankow, J. J. (2006). Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics 60*(2), 139–161.

Zax, J. S. (1991). Compensation for commutes in labor and housing markets. *Journal of Urban Economics 30*(2), 192–207.

Nederlandstalige Samenvatting

In deze dissertatie doe ik onderzoek naar (i) de gevolgen van baanverlies vanuit een regionaal perspectief, (ii) de verschillen tussen groepen werknemers afhankelijk van de geografische omvang van hun lokale arbeidsmarkten, en (iii) de omvang van agglomeratievoordelen in lokale arbeidsmarkten met een hoge werkgelegenheidsdichtheid ten opzichte van arbeidsmarkten met een lage werkgelegenheidsdichtheid.

Met baanverlies doel ik specifiek op baanverlies als gevolg van faillissement van het bedrijf waar de werknemer in kwestie in dienst was. De volgende onderzoeksvragen zullen beantwoord worden: hoe vergaat het werknemers na ontslag met betrekking tot de kans om een baan te vinden, het loon, de woon-werkafstand en de kans om te verhuizen? Bestaan er tussen verschillende typen werknemers verschillen in de gevolgen van baanverlies?

Het concept van lokale arbeidsmarkten wordt door onderzoekers gebruikt om een afgebakend regionaal gebied aan te duiden dat qua economische activiteit relatief op zichzelf staat, bijvoorbeeld door pendelstromen tussen lokale arbeidsmarkten te meten. Nederland kan op verschillende manieren worden opgedeeld in lokale arbeidsmarkten waarbij binnen zo'n arbeidsmarkt relatief veel werknemers van hun woonlocatie naar hun werklocatie pendelen en waarbij relatief weinig werknemers de grenzen van die arbeidsmarkten overschrijden. Vaak wordt er in onderzoeken voor een regionale indeling gekozen waarbij 70 tot 90 procent van de werknemers lokaal werkzaam is. Lokaal werkzaam houdt in dat een werknemer werkt in de lokale arbeidsmarkt waar hij of zij woonachtig is. Als onderzoekers voor verschillende typen werknemers lokale arbeidsmarkten definiëren met een specifiek percentage dat lokaal werkzaam is, dan kan

het dus zijn dat er grote verschillen bestaan in de geografische omvang van die lokale arbeidsmarkten en de verschillende typen werknemers die woonachtig zijn in dezelfde locatie. Een van de onderzoekvragen die beantwoord zal worden is: bestaan er verschillen tussen typen werknemers en de geografische omvang van hun lokale arbeidsmarkten?

Agglomeratievoordelen in lokale arbeidsmarkten zijn gedefinieerd als schaalvoordelen voor de productiviteit van werknemers en bedrijven door een hoge bevolkings- of werkgelegenheidsdichtheid. Een hoge dichtheid betekent in dit geval dat er veel individuen wonen en werken in hetzelfde geografisch gebied, bijvoorbeeld in een stad. Dit gaat gepaard met een goede infrastructuur en de aanwezigheid van een brede arbeids- en afzetmarkt. Om binnen een empirische analyse de agglomeratievoordelen te kunnen meten is het van belang dat er lokale arbeidsmarkten gedefinieerd worden. Als men onderzoek wilt doen naar Nederland, houdt dit in dat er gekozen moet worden hoe Nederland wordt opgedeeld in op zichzelf staande lokale arbeidsmarkten. Een onderzoeker kan kiezen uit verschillende regionale classificaties, bijvoorbeeld de 35 werkgebieden van het Uitvoeringsinstituut Werknemersverzekeringen (UWV) of de 403 Nederlandse gemeenten. Als Nederland wordt opgedeeld in minder verschillende gebieden met een grotere oppervlakte, dan houdt dat in dat agglomeratievoordelen worden gemeten op een hoger niveau van regionale aggregatie. Andere onderzoeksvragen die beantwoord zullen worden zijn: hoe beïnvloedt de regionale classificatie de schattingsresultaten van de omvang van agglomeratievoordelen wanneer het gaat om loon? In hoeverre hebben werknemers agglomeratievoordelen met betrekking tot de kans om een baan te vinden en hun loon na ontslag vanwege faillissement van het bedrijf? In hoeverre verschilt de omvang van agglomeratievoordelen tussen verschillende typen werknemers?

In mijn analyses gebruik ik administratieve gegevens van het Centraal Bureau voor de Statistiek (CBS) uit de periode 2006 tot 2014. Deze rijke bron van informatie stelt mij in staat om onderzoek te doen naar de gehele Nederlandse populatie op het gebied van banen, bedrijven en individuen. Een ander voordeel van de CBS-gegevens is dat ik de empirische schattingen kan corrigeren voor allerlei kenmerken die belangrijk zijn voor verschillen in arbeidsmarktuitkomsten, zoals persoonskenmerken, baankenmerken en kenmerken van huishoudens – hierbij kun je denken aan geslacht, opleidingsniveau, leeftijd, nationaliteit, gezinssamenstelling, economische sector en het jaar waarin de gegevens gemeten zijn.

In het onderzoek integreer ik concepten en inzichten uit de arbeidseconomie, de stedelijke economie en de regionale economie. Hieronder volgt een overzicht van de belangrijkste resultaten en beleidsimplicaties.

Baanverlies

In hoofdstuk 2 bestudeer ik hoe werknemers in economisch opzicht reageren op baanverlies. Ontslagen werknemers worden tot drie jaar na hun ontslag gevolgd en vergeleken met vergelijkbare werknemers die niet ontslagen zijn, met als doel om de effecten van baanverlies op vier aanpassingspatronen te analyseren. Specifiek doe ik onderzoek naar het effect van ontslag op de kans op een nieuwe baan, het loon, de woon-werkafstand en de verhuiskans. Ik gebruik administratieve gegevens van het CBS en vergelijk de arbeidsmarktuitkomsten van ongeveer 20.000 ontslagen werknemers met die van 30.000 vergelijkbare werkende personen in de periode 2006 tot 2014.

Het onderzoek laat zien dat werknemers die ontslagen zijn gemiddeld 25 procentpunten minder werkzaam zijn over een periode van 3 jaar na het ontslag. Bovendien ondervinden ontslagen werknemers die weer een baan hebben gevonden gemiddeld een verlies van 6 procent in loon en een toename van 3 kilometer in de woon-werkafstand. Daarnaast laten de schattingsresultaten een afname in de verhuiskans zien van 0,06 procentpunt. Vergeleken met het gemiddelde van de woon-werkafstand en verhuiskans van werknemers in het onderzoeksbestand gaat het hier na het ontslag om een toename van 20 procent in de woon-werkafstand en een afname van 15 procent in de verhuiskans. Hoewel de bereidheid om van woonlocatie te veranderen zou kunnen toenemen na ontslag, bijvoorbeeld om naar een kansrijkere lokale arbeidsmarkt te verhuizen, blijkt juist dat de verhuiskans afneemt. Een uitleg voor deze bevinding kan zijn dat het moeilijker is om een nieuwe hypotheek of een nieuw huurcontract te krijgen na baanverlies. Uit de resultaten blijkt ook dat de aanpassingspatronen veranderen over de periode na ontslag. Ik toon aan dat een langere werkloosheidsduur na het ontslag wordt geassocieerd met een kleinere toename van de woon-werkafstand, maar met een hoger verlies in loon. Deze bevinding lijkt te worden veroorzaakt omdat ontslagen werknemers met een langere werkloosheidsduur een voorkeur hebben om dichterbij huis te werken in plaats van een hoger loon te krijgen. De resultaten van hoofdstuk 2 laten zien dat voor werknemers die zijn ontslagen de woon-werkafstand een relevantere marge van aanpassing is dan verhuizen.

Ook bestudeer ik in hoofdstuk 2 of het effect van baanverlies afhangt van de kenmerken van de werknemer. Ik onderzoek het belang van demografische kenmerken, baankarakteristieken en de status in de woningmarkt van de werknemer. De status in de woningmarkt definieer ik als zes verschillende typen bewoners: huurders, kopers die hun hypotheek volledig hebben afbetaald en kopers onderverdeeld in vier verschillende loan-to-value (LTV) ratio groepen. De LTV-ratio wordt berekend door de hypotheekschuld te delen door de waarde van de woning. De groep die van mij extra aandacht krijgt zijn de huiseigenaren met een huis dat "onder water" staat. Huiseigenaren met een huis onder water hebben een negatief netto huisvermogen, ofwel een LTV-ratio van meer dan 100 procent, aangezien ze een hypotheekschuld hebben die hoger is dan de woningwaarde. Ik besteed hier extra aandacht aan vanwege de economische crisis van 2009, aangezien die leidde tot een sterke daling in de prijzen en waarde van woningen.

Tot slot laat hoofdstuk 2 zien dat het effect van ontslag anders is voor huurders dan voor de verschillende typen huiseigenaren. Vergeleken met andere ontslagen huiseigenaren en huurders, hebben ontslagen huiseigenaren met een huis onder water een baanvindkans die 4-6 procentpunt hoger is. De resultaten suggereren dat een hoog LTV-ratio, ofwel een hoge hypotheekschuld ten opzichte van de waarde van de woning, een sterke prikkel geeft om relatief snel een baan te vinden. Daar staat tegenover dat bij ontslagen werknemers die een hoog LTV-ratio hebben een verlies in loon wordt waargenomen dat 1-2 procentpunt hoger ligt. Bovendien ondervinden huiseigenaren die hun volledige hypotheek hebben afbetaald de grootste stijging in de woon-werkafstand. Hoewel ontslagen werknemers met een huis dat onder water staat mogelijk geografische immobiliteit ervaren, vormt dit geen belemmering voor de baanvindkans na ontslag. Deze resultaten zijn relevant voor beleidsmakers die beslissen over de verdere verlaging van de LTV-limiet van huizenbezitters en het subsidiëren van eigenwoningbezit.

Lokale arbeidsmarkten

In hoofdstuk 3 introduceer ik een algoritme genaamd flowbca, geschreven in het statistische softwareprogramma Stata. Flowbca kan worden gebruikt om entiteiten zoals gemeenten te clusteren op basis van relationele data van stromen tussen entiteiten. De stromen tussen entiteiten geven de mate van interactie tussen die entiteiten aan. Met het algoritme kunnen onderzoekers met verschillende achtergronden clusters definiëren die relevant zijn voor hun specifieke onderzoeksgebied. Een voorbeeld is het gebruik van woon-werkverkeer van werknemers die pendelen van woongemeente naar werkgemeente. Op basis van deze pendelstromen construeer ik lokale arbeidsmarkten die veel interactie hebben op het gebied van woon-werkverkeer binnen de lokale arbeidsmarkt en weinig interactie hebben tussen lokale arbeidsmarkten. In hoofdstuk 3 geef ik ook andere voorbeelden, zoals clusters binnen de context van de input-output literatuur, waarbij gebruik gemaakt wordt van handelsstromen tussen landen en tussen economische sectoren, of clusters ten behoeve van sociale netwerkanalyse, gebaseerd op vriendschapsconnecties tussen individuen.

In hoofdstuk 4 heb ik flowbca gebruikt om lokale arbeidsmarkten te definiëren voor verschillende typen werknemers op basis van gegevens van het woon-werkverkeer tussen gemeenten. Het algoritme heeft me in staat gesteld de lokale arbeidsmarkt van werknemers te definiëren voor verschillende niveaus van regionale aggregatie. Het niveau van regionale aggregatie is hoger naarmate ik Nederland onderverdeel in grotere – en dus in een kleiner aantal – verschillende gebieden. Bovendien kan ik met het algoritme lokale arbeidsmarkten definiëren voor verschillende subgroepen van werknemers. Deze subgroep-specifieke lokale arbeidsmarkten zijn relevant omdat economische theorieën en statistieken suggereren dat de verschillen in woon-werkverkeer tussen werknemers bestaan door onderlinge verschillen in hoeveel tijd en financiële middelen die ze beschikbaar hebben om te pendelen. Een voorbeeld is het feit dat hoogopgeleiden gemiddeld een hogere woon-werkafstand hebben dan laagopgeleiden. Dit heeft als reden dat met name voor hoogopgeleiden de banen die mogelijk een goede match opleveren schaarser en minder homogeen verspreid zijn over Nederland. Voor een hoogopgeleide is het de moeite waard om het aantal potentiële banen groter te maken door verder te willen reizen, omdat dit de kans op een goede match verhoogt, iets wat gepaard gaat met een hoger loon.

In de descriptieve analyse van hoofdstuk 4 laat ik zien dat het geslacht en het opleidingsniveau van werknemers een groot deel van de variatie in woon-werkafstand verklaart. Deze observatie is de reden waarom ik me in de empirische analyse van hoofdstuk 4 richt op de verschillen in agglomeratievoordelen tussen mannen en vrouwen en tussen werknemers met verschillende opleidingsniveaus. Ook suggereren de resultaten dat zowel mannen als vrouwen een aanzienlijke regionale mismatch ondervinden tussen woon- en werklocaties, aangezien mannelijke en vrouwelijke werknemers relatief geconcentreerd zijn in specifieke werkgemeenten, terwijl mannelijke en vrouwelijke inwoners gelijkmatiger verdeeld zijn over woongemeenten. De resultaten laten verder zien dat hoogopgeleide werknemers en inwoners zich relatief gezien concentreren in specifieke gemeenten, terwijl laagopgeleide werknemers en inwoners gelijkmatiger over alle gemeenten verdeeld zijn wanneer het aankomt op hun woon- en werklocatie. Ik pas het clusteralgoritme toe om de lokale arbeidsmarkt van werknemers afhankelijk te maken van hun geografische locatie, woon-werkverkeer en demografische kenmerken. De resultaten geven aan dat mannelijke werknemers en hoogopgeleide werknemers, in vergelijking met vrouwelijke werknemers en laagopgeleide werknemers, worden gekenmerkt door grotere en dus minder unieke lokale arbeidsmarkten. Dit is een belangrijke bevinding aangezien in veel onderzoek en beleid er impliciet vanuit gegaan wordt dat de omvang van de geografische gebieden waar individuen werken of zoeken naar werk gelijk is voor alle verschillende subgroepen binnen de populatie.

Agglomeratievoordelen

In hoofdstuk 4 zet ik flowbca ook in om onderzoek te doen naar de economische gevolgen van verschillen in de geografische omvang van een lokale arbeidsmarkt. Ik concentreer op verschillen in agglomeratievoordelen die tot uitdrukking komen in lonen en werkgelegenheid voor zowel werkenden als ontslagen werknemers. Agglomeratievoordelen komen voor als er toenemende schaalvoordelen zijn in productiviteit door de omvang en dichtheid van werknemers en bedrijven, wat bijvoorbeeld het geval is in de grotere steden. De onderliggende theoretische mechanismen van agglomeratievoordelen door hogere dichtheid en specialisatie zijn (i) het beter matchen van werkgevers en werknemers, (ii) het doeltreffender delen van grondstoffen en risico's, en (iii) het efficiënter ontwikkelen en accumuleren van kennis. Ik analyseer of de agglomeratievoordelen in lonen afhangen van het niveau van regionale aggregatie binnen lokale arbeidsmarkten. Bovendien onderzoek ik of de omvang van agglomeratievoordelen afhankelijk is van het geslacht en opleidingsniveau van de werknemer. Voor dit onderzoek gebruik ik administratieve gegevens van het CBS van ongeveer 4 miljoen werkenden die ik volg over de periode 2006 tot 2014. Belangrijk om te vermelden is dat ik in de empirische schattingen rekening houd met allerlei persoonskenmerken, huishoudenskenmerken en baankenmerken die significant zijn voor het verklaren van de verschillen in loon.

De empirische resultaten in hoofdstuk 4 laten zien dat de agglomeratievoordelen in lonen, de zogenoemde stedelijke loonpremies, toenemen met een factor twee tot drie op het niveau van regionale aggregatie. De stedelijke loonpremie geeft aan hoeveel procent het loon hoger is naarmate de werkgelegenheidsdichtheid verdubbelt in de lokale arbeidsmarkt waar de werknemer werkzaam is. Bijvoorbeeld wanneer de 35 UWV-gebieden en de 403 Nederlandse gemeenten worden gebruikt om de lokale arbeidsmarkt van werknemers te definiëren, de stedelijke loonpremie respectievelijk 4,9 en 2,7 procent is, ceteris paribus. Dit bereik komt overeen met andere onderzoeken die Nederlandse administratieve gegevens gebruiken om de stedelijke loonpremie te analyseren. De hoogste stedelijke loonpremie die gemeten wordt is 6,6 procent op basis van een set van dertien verschillende lokale arbeidsmarkten die werden gedefinieerd met flowbca. Deze bevinding suggereert dat de positieve effecten van agglomeratie, gebaseerd op de mechanismen voor matching, delen en kennis, sterker tot uiting komen op een hoger niveau van regionale aggregatie. Op basis van verschillende regionale classificaties en empirische schattingsmethoden kom ik uit op een stedelijke loonpremie van 0,3 tot 6,6 procent. Dit bereik beschouw ik als de onder- en de bovengrens binnen Nederland. De resultaten benadrukken het belang van het gebruik van regionale classificaties met verschillende niveaus van regionale aggregatie voor het definiëren van lokale arbeidsmarkten in een onderzoek.

Hoofdstuk 4 behandelt bovendien de rol van lokale arbeidsmarkten in het effect van baanverlies. Ik onderzoek in het bijzonder de agglomeratievoordelen voor ontslagen werknemers in loon en werkgelegenheid. Ik laat zien dat ontslagen werknemers die wonen in dichtbevolkte lokale arbeidsmarkten vergeleken met ontslagen werknemers die wonen in minder dichtbevolkte lokale arbeidsmarkten een kleiner verlies in loon en een vergelijkbaar verlies in baankans ervaren. Als een ontslagen werknemer woont in een lokale arbeidsmarkt die een twee keer zo hoge werkgelegenheidsdichtheid heeft, en wanneer daarbij ook gecorrigeerd wordt voor persoonskenmerken, huishoudenskenmerken en baankenmerken, is het verlies in loon ongeveer 2 procentpunten lager. In dit opzicht leiden de agglomeratievoordelen voor ontslagen werknemers tot verschillen in loon, maar niet tot verschillen in baanvindkans. De bevindingen suggereren dat het wonen in een lokale arbeidsmarkt met een relatief hoge werkgelegenheidsdichtheid werknemers weerbaarder maakt tegen negatieve werkgelegenheidsschokken, aangezien ze selectiever kunnen zijn in loon maar dezelfde baanvindkans houden.

Het laatste deel van de empirische analyse richt zich op de verschillen in de omvang van agglomeratievoordelen tussen mannen en vrouwen en tussen werknemers met verschillende opleidingsniveaus. De empirische resultaten laten zien dat alle subgroepen van de populatie, inclusief mannen, vrouwen, laagopgeleiden, middenopgeleiden en hoogopgeleiden, profiteren van een stedelijke loonpremie. Er zijn wel verschillen tussen subgroepen in de omvang van de loonpremie. De stedelijke loonpremie voor hoogopgeleide werknemers in vergelijking met laagopgeleide werknemers is ongeveer twee keer zo hoog bij een willekeurige regionale classificatie. Ik heb geen systematisch verschil ontdekt in de stedelijke loonpremie tussen mannen en vrouwen. Echter, in de descriptieve analyse laat ik met het gebruik van flowbca zien dat de lokale arbeidsmarkt relatief groot is voor mannelijke werknemers en hoogopgeleide werknemers en relatief klein is voor vrouwelijke werknemers en laagopgeleide werknemers.

Het is belangrijk dat in de literatuur vooraf bepaalde regionale classificaties zoals de 35 UWV-gebieden worden gebruikt die hetzelfde zijn in termen van de geografische omvang en het aantal lokale arbeidsmarkten voor verschillende typen werknemers. Vanwege het gebruik van vooraf bepaalde regionale classificaties in onderzoek betoog ik dan ook dat het agglomeratievoordeel wordt onderschat voor werknemers die worden gekenmerkt door een grote lokale arbeidsmarkt en wordt overschat voor werknemers die worden gekenmerkt door een kleine lokale arbeidsmarkt. De empirische bevindingen in hoofdstuk 4 laten zien dat zodra hiermee rekening gehouden wordt, de stedelijke loonpremie hoger is voor mannelijke hoogopgeleide werknemers en mannelijke middenopgeleide werknemers vergeleken met vrouwelijke hoogopgeleide werknemers en vrouwelijke middenopgeleide werknemers. De bevinding van de aanwezigheid van verschillen in de omvang van agglomeratievoordelen tussen mannen en vrouwen en tussen werknemers met verschillende opleidingsniveaus is relevant vanuit een maatschappelijk perspectief, omdat het de afweging tussen economische efficiëntie en sociale ongelijkheid onderstreept.

Beleid

De analyse van de gevolgen van baanverlies vanuit een regionaal perspectief is relevant voor beleid gericht op het verminderen van effecten van negatieve werkgelegenheidsschokken, omdat het inzicht geeft in hoeverre werknemers regionale aanpassingspatronen gebruiken na ontslag. Ik laat zien dat baanverlies kostbaar is vanuit een maatschappelijk perspectief: naast de aanzienlijke verliezen in werkgelegenheid en lonen ervaren werknemers na ontslag een toename in de woon-werkafstand van ongeveer 20 procent. Bovendien suggereren de resultaten dat ontslagen werknemers afwegingen maken tussen verschillende aanpassingspatronen: ontslagen werknemers met een langere werkloosheidsduur hebben een voorkeur voor dichterbij huis werken in plaats van een hoger loon. Belangrijk is ook dat in de context van baanverlies de woon-werkafstand een relevantere marge van aanpassing blijkt te zijn dan verhuizen naar een nieuwe woonlocatie, aangezien baanverlies een positief effect heeft op de woon-werkafstand en een negatief effect heeft op de verhuiskans. Een uitleg voor deze bevinding kan zijn dat de oppervlakte van Nederland relatief klein is, waardoor verhuizen naar een andere lokale arbeidsmarkt minder effectief is voor het verbeteren van arbeidsmarktuitkomsten.

Ontslagen werknemers zoeken dus naar banen die binnen hun lokale arbeidsmarkt liggen. In dit opzicht wordt door werknemers de beslissing voor de woonlocatie niet gemaakt op basis van de werklocatie, maar wordt de beslissing voor de werklocatie gemaakt op basis van de woonlocatie. De woningmarkt kan de arbeidsmarktuitkomsten van ontslagen werknemers belemmeren door een relatief lage geografische mobiliteit, maar de pendelmarge van aanpassing helpt hen weerbaar te zijn tegen een negatieve werkschok. Deze bevinding is relevant voor het beleid gericht op het verminderen van effecten van negatieve werkgelegenheidsschokken, aangezien het suggereert dat werknemers die minder makkelijk kunnen pendelen, bijvoorbeeld in geografische gebieden met een mindere infrastructuur of werknemers die minder toegang hebben tot vervoersmiddelen, kwetsbaarder zijn op de arbeidsmarkt.

Het onderzoek naar de verschillen tussen groepen werknemers en de geografische omvang van hun lokale arbeidsmarkten laat zien dat deze omvang sterk afhankelijk is van de beoogde subgroep van werknemers. Zo laat ik zien dat mannelijke werknemers en hoogopgeleide werknemers, in vergelijking met vrouwelijke werknemers en laagopgeleide werknemers, worden gekenmerkt door grotere en dus minder unieke lokale arbeidsmarkten. Dit is een belangrijke bevinding voor beleidsmakers aangezien het suggereert dat regionaal en lokaal arbeidsmarktbeleid voor verschillende subgroepen van individuen effectiever kan zijn wanneer er verschillende maten van geografische omvang gebruikt worden.

De analyse van de agglomeratievoordelen in lokale arbeidsmarkten met een hoge werkgelegenheidsdichtheid ten opzichte van die met een lage dichtheid laat zien dat de omvang van de stedelijke loonpremies toeneemt wanneer de lokale arbeidsmarkt een grotere dichtheid heeft. Bovendien suggereren de resultaten dat agglomeratievoordelen sterker tot uiting komen op een hoger niveau van regionale aggregatie dan verwacht wordt op basis van de onderliggende mechanismen van agglomeratie, te weten matchen, delen en kennis, die meer lokaal verondersteld worden. Hiermee wil ik zeggen dat de agglomeratievoordelen in een stedelijk gebied zoals Amsterdam niet alleen tot uiting komen binnen de gemeente Amsterdam, maar ook in de naburige gemeenten vanwege allerlei interacties tussen werknemers en bedrijven. Een uitleg hiervoor is de sterke economische connectiviteit tussen Amsterdam en de naburige gemeenten, wat bijvoorbeeld zichtbaar is in de pendelstromen van en naar de stad. Samen met de naburige gemeenten vormt Amsterdam een lokale arbeidsmarkt waarvan Amsterdam de kern is, en alle gemeenten in deze lokale arbeidsmarkt profiteren van de goede infrastructuur en de aanwezigheid van een brede arbeidsmarkt en afzetmarkt. De bevindingen zijn relevant voor de evaluatie van stedelijk en regionaal beleid gericht op het stimuleren van agglomeratievoordelen en regionale productiviteitsgroei.

Al met al pleit ik ervoor dat beleidsmakers de mogelijke verschillen in de geografische omvang van lokale arbeidsmarkten van werknemers in acht nemen. Het is belangrijk om de verschillen in lokale arbeidsmarkten tussen subgroepen van de populatie te erkennen waarop beleid gericht is. Bovendien zal het nuttig zijn om regionaal arbeidsmarktbeleid en woningmarktbeleid te concentreren op grotere gebieden dan gemeenten, omdat, zoals ik laat zien in deze dissertatie, alle gemeenten die deel uitmaken van de Randstad onlosmakelijk met elkaar verbonden zijn op het gebied van pendelstromen.

Curriculum Vitae

Jordy Meekes was born in Lelystad, Flevoland, the Netherlands. He acquired his Bachelor's degree in Economics and Business Economics at Utrecht University (the Netherlands) and, following his Research Master's degree in Multidisciplinary Economics, he did his PhD in Economics at Utrecht University School of Economics. During his PhD, Jordy was a Short-Stay Visiting Fellow at MIT Sloan School of Management (the United States) and he did a research project in collaboration with Randstad Holding, a service company for temporary work and human resource consulting (the Netherlands). He is currently employed as a Research Fellow at the Melbourne Institute of Applied Economic and Social Research at the University of Melbourne (Australia).

U.S.E. Dissertation Series

USE 001 **Bastian Westbrock** (2010): Inter-firm networks: Economic and sociological perspectives

USE 002 **Yi Zhang** (2011): Institutions and international investments: Evidence from China and other emerging markets

USE 003 **Ryan van Lamoen** (2011): The relationship between competition and innovation: Measuring innovation and causality

USE 004 **Martijn Dröes** (2011): House price uncertainty in the Dutch owner-occupied housing market

USE 005 **Thomas van Huizen** (2012): Behavioural assumptions in labour economics: Analysing social security reforms and labour market transitions

USE 006 **Martijn Boermans** (2012): International entrepreneurship and enterprise development

USE 007 **Joras Ferwerda** (2012): The multidisciplinary economics of money laundering

USE 008 **Federico D'Onofrio** (2013): Knowing to transform: Three ways for agricultural economists to observe Italy 1900-1940

USE 009 **Saraï Sapulete** (2013): Works council effectiveness: Determinants and outcomes

USE 010 **Britta Hoyer** (2013): Network formation under the threat of disruption

USE 011 **Coen Rigtering** (2013): Entrepreneurial orientation: Multilevel analysis and consequences

USE 012 **Beate Cesinger** (2013): Context and complexity of international entrepreneurship as a field of research

USE 013 **Jan de Dreu** (2013): Empirical essays on the governance of financial institutions

USE 014 **Lu Zhang** (2013): Industrial specialization: Determinants, processes and consequences

USE 015 **Matthias Filser** (2013): Strategic issues in entrepreneurship and family business research

USE 016 **Mikko Pohjola** (2013): A compilation of studies on innovation in firms: Capabilities, strategies, and performance

USE 017 **Han-Hsin Chang** (2013): Heterogeneity in development

USE 018 **Suzanne Heijnen** (2014): Analyses of sickness absence

USE 019 **Mark Kattenberg** (2014): The economics of social housing: Implications for welfare, consumption, and labor market composition

USE 020 **Daniel Possenriede** (2014): The economics of temporal and locational flexibility of work

USE 021 **Dirk Gerritsen** (2014): The relevance of security analyst opinions for investment decisions

USE 022 **Shiwei Hu** (2014): Development in China and Africa

USE 023 **Saara Tamminen** (2014): Heterogeneous firms, mark-ups and income inequality

USE 024 **Marcel van den Berg** (2014): Does internationalization foster firm performance?

USE 025 **Emre Akgündüz** (2014): Analysing maternal labour supply and child care quality

USE 026 **Jasper Lukkezen** (2014): From debt crisis to sovereign risk

USE 027 **Vesile Kutlu Koç** (2015): Essays on subjective survival probabilities, consumption, and retirement decisions

USE 028 **Brigitte Crooijmans** (2015): Leiden fusies tot efficiëntere woningcorporaties? Een exploratieve studie naar schaalvoordelen in de sociale huisvesting

USE 029 **Andrej Svorenčík** (2015): The experimental turn in economics: A history of experimental economics

USE 030 **Seçil Danakol** (2015): Foreign direct investment, foreign aid and domestic entrepreneurship

USE 031 **Ioana Deleanu** (2015): Anti-money laundering efforts: Failures, fixes and the future

USE 032 **Jaap Oude Mulders** (2016): Organizations, managers, and the employment of older workers after retirement

USE 033 **Malka de Castro Campos** (2016): Private consumption-savings behavior and macroeconomic imbalances

USE 034 **Tahereh Rezaei Khavas** (2016): Fairness concerns and cooperation in context

USE 035 **Joyce Delnoy** (2016): Auctions with competing sellers and behavioral bidders

USE 036 **Krista Bruns** (2017): Emergence and diffusion of institutions and their effect on economic growth

USE 037 **Daan van der Linde** (2017): Democracies under rising inequality: New tests of the redistributive thesis

USE 038 **Swantje Falcke** (2017): On the move: Analyzing immigration determinants and immigrant outcomes

USE 039 **Joep Steegmans** (2017): House prices and household mobility in the Netherlands: Empirical analyses of financial characteristics of the household

USE 040 **Najmeh Rezaei Khavas** (2017): Essays in information economics

USE 041 **Maryam Imanpour** (2017): The role of social networks for combating money laundering

USE 042 **Ye Li** (2018): Hydrogen infrastructure decisions through a real option lens

USE 043 **Li Lin** (2018): Leadership across cultural contexts

USE 044 **Werner Liebregts** (2018): Hidden entrepreneurship: Multilevel analyses of the determinants and consequences of entrepreneurial employee activity

USE 045 **Ian Koetsier** (2018): Government debt: The economic consequences of natural disasters and pension funds' herding

USE 046 **Jordy Meekes** (2019): Local labour markets, job displacement and agglomeration economies

This dissertation focuses on the structure of workers' local labour markets and the resulting economic consequences. A local labour market can be described as a regional area that is relatively self-contained in terms of residence and work activity. The analyses focus on three topics. Firstly, the workers' spatial and labour response to job displacement following firm bankruptcy: to what extent is the spatial structure of homes and jobs relevant for workers' post-displacement outcomes? Secondly, the introduction of a flow-based cluster algorithm, flowbca, in the statistical software package Stata. The main purpose of flowbca is to define local labour markets based on relational data of commuting flows from place of home to place of work. Thirdly, the consequences of the differences between workers' local labour markets: what is the role of the structure of local labour markets and the level of regional aggregation in the agglomeration benefits for employment and wages? These questions lead to an incisive conclusion on local labour markets and the way we study them.

ISBN 978-94-91870-31-6