Spatial disparities in labor and housing markets

Von der Mercator School of Management, Fakultät für Betriebswirtschaftslehre, der

Universität Duisburg-Essen

zur Erlangung des akademischen Grades

eines Doktors der Wirtschaftswissenschaft (Dr. rer. oec.)

genehmigte Dissertation

von

Malte Janosch Borghorst

aus

Aachen

Referent: Prof. Dr. Tobias Seidel

Korreferentin: Prof. Dr. Jens Wrona

Tag der mündlichen Prüfung: 19. Februar 2024

Acknowledgment

I want to express my gratitude to my supervisor, Tobias Seidel, for entrusting me with the freedom to develop my research skills and ideas and for supporting me in my first steps in academic research. I am particularly indebted to my co-authors Jesper Hybel, Ismir Mulalic, and Jos van Ommeren for their unwavering support, frequent discussion, and patience, without which these projects would not have been possible.

I also thank Avtandil Abashishvili, Fabian Bald, Bence Boje-Kovacs, Eyayaw Beze, Fabian Dehos, Marvin Finkemeier, Joschka Flintz, Lorenz Gschwent, Johannes Gallé, Jakob Gutschlhofer, Lea Nassal, Marie Paul, Maximilian Perl, Elias Stapput Knudsen, Karolin Süß, Anna Temel, Lu Wei, Jens Wrona, and all members of the Research Training Group (RTG) Regional Disparities & Economic Policy, as well as colleagues in Duisburg for their helpful comments and fruitful discussions.

I thank Statistics Denmark, Aalborg University, Kraks Fond, University of Duisburg Essen, and the Research Data Center at RWI for generously providing access to the research data. Financial support from the German Research Foundation (DFG) via the RTG Regional Disparities & Economic Policy and Kraks Fond is gratefully acknowledged.

Finally, I would like to point out how grateful I am for the constant non-scientific support I have experienced in recent years. I cannot overstate the contribution of the emotional support from my family and friends.

Contents

List of Figures

List of Tables

Introduction

Spatial disparities persist as a complex and unresolved topic in current research and policy. Understanding individual decisions around spatial frictions is crucial for assessing the general equilibrium effects of policies in their impact on disparities. In this context, urban economics is a field that has gained renewed attention due to recent waves of urbanization, methodological advancements, and the availability of geolocated datasets. This collection of research projects explores individuals' and households' decisions around spatial frictions across commuting choices, the urban wage premium, the role of taxation in housing markets, and potential policy implications.

The first project reveals a significant gender disparity in the influence of commuting distances on job retention. We find that women, especially those with children, are more likely to leave their jobs when facing long commutes—a trend not observed among men. By utilizing a dynamic search model, this study illustrates how commuting costs increase substantially for women after becoming mothers, shedding light on the interplay between gender roles, family responsibilities, and labor market participation. This project contributes to the literature on the value of time, especially for commuting and care work, labor supply, and the gender wage gap.

The second project explores the urban wage premium in Denmark and its gender-specific implications. By analyzing data from the Danish working population, this study identifies an urban wage premium for both wage levels and wage growth. We find that women receive a lower return to work experience from working in cities compared to men, indicating gender disparities in the benefits from agglomeration. These findings underscore the gender dimension of spatial disparities in the labor market, with implications for career trajectories, location choice, and the realization of agglomeration benefits.

The third project shifts the focus to housing markets, examining the impact of real estate transfer taxes on housing affordability, especially in markets with substantial rental share. Drawing from nationwide data in Germany, this study exploits spatial and temporal variations in tax rates and finds that the transfer tax significantly reduces both, house prices and rents. This research also emphasizes spatial disparities in the reaction of house prices and rents, where urban and rapidly growing markets show little-to-none impact compared to rural or declining areas. These insights could guide policymakers seeking to address disparities in the housing market and housing affordability.

In summary, these research projects contribute to our understanding of the impact of spatial frictions on individual and household decisions in labor and housing markets. By doing so, they offer valuable insights for policymakers, urban planners, and researchers aiming to foster more inclusive and equitable urban environments in the face of persistent regional disparities.

Note to future readers: Throughout my dissertation, the COVID-19 pandemic disrupted our lives and prompted a re-evaluation of decisions regarding residential location, commuting patterns, housing choice, and the role of childcare. Although it is essential to acknowledge that my findings primarily describe mechanisms before the onset of the COVID-19 crisis, my research addresses many challenges that we faced during the crisis by exploring the underlying workings of decisions and challenges individuals and policymakers faced, such as the trade-off between childcare and commuting time, the attraction of urban wages and the decisions to move.

CHAPTER 1

Commuting, children and the gender wage gap[1]

Malte Borghorst, Ismir Mulalic and Jos van Ommeren

Abstract: We demonstrate that women with children are much more likely to leave their jobs when they have a long commute, which is not true for men. Interpreting these results through the lens of a dynamic search model, we demonstrate that the costs of commuting increase substantially for women after they have children. For women with children, a 12-kilometer increase in commuting distance induces costs equivalent to about 16% of their wage. At the same time, compensating wage differentials for commuting are low for all workers, also for women with children, implying that women with children are not compensated for their higher commuting costs through higher wages.

[1]Ismir Mulalic (Department of Economics, Copenhagen Business School, Porcelænshaven 16A, DK-2000 Frederiksberg, Denmark, email: imu.eco@cbs.dk) and
Jos van Ommeren (Department of Spatial Economics, VU University, De Boelelaan 1105 1081 HV Amsterdam, email: jos.van.ommeren@vu.nl. Jos van Ommeren is a Fellow of the Tinbergen Institute, Amsterdam.)
The authors thank seminar participants at the 11th European Meeting of the Urban Economics Association, 15th North American Meeting of the Urban Economics Association, 34th Annual Conference of the European Association of Labour Economists, 28th Annual Conference of the Society of Labor Economists, International Transport Economics Association Conference 2020, the European Regional Science Association Workshop on Spatial Dimensions of Labor Economics, RWI – Leibniz-Institut für Wirtschaftsforschung workshop on Regional Inequality, Kraks Fond– Institute for Urban Economic Research, University of Duisburg-Essen, DFG-RTG 2484 - Regional Disparities & Economic Policy, VU Amsterdam and Copenhagen Business School also provided helpful comments. Research support from Kraks Fond – Institute for Urban Economic Research, Copenhagen (kraksfond@kraksfond.dk) is acknowledged. The usual disclaimer applies.

1.1 Introduction

Over the last decades, earnings for men and women have converged due to the reduced gap in education, skills, and labor participation (Altonji and Blank, 1999; Blau and Kahn, 2017; Maasoumi and Wang, 2019; Gallen et al., 2019). However, women still earn substantially less than men, despite decades of equal-pay laws. This gender pay gap has been argued to be essentially a child penalty for women because childbirth induces career interruptions and reduced working hours (Manning and Petrongolo, 2008; Blau and Kahn, 2017; Kleven et al., 2019b; Cortés and Pan, 2020; Card and Hyslop, 2021).

Using administrative register data for the full working population in Denmark for the years 2003-2013 we apply an event study methodology – the birth of the first child – and demonstrate, that women not only earn substantially less but also strongly decrease their commute after the birth of the child relative to men. This finding makes sense, as for many workers, adjusting the length of the commute through a job move is an important behavioral margin to optimize time devoted to labor as they are severely constrained in their choice of working hours (Böheim and Taylor, 2004).

Consistent with this finding, we show that women with a long commute are *several times* more likely to change jobs when they have a child, which is not true for men.[2] We also show that workers with a higher wage are less likely to move jobs. Interpreting these results through the lens of a dynamic search model as in Gronberg and Reed (1994), Van Ommeren and Fosgerau (2009) and Le Barbanchon et al. (2021), we estimate how many workers are willing to trade off wages for a shorter commute, i.e. we estimate the marginal cost of commuting. We show that this cost is the same for men and women before the birth of a child, but after the birth, it is substantially higher for women.

The sudden increase in the commuting cost for women after becoming a mother implies that women with children face a different trade-off between wages and commuting distance. This suggests that women with children may receive a different level of wage compensation for commuting. To investigate this further, we estimate the relationship between commuting distance and wages and use the event of childbirth to assess the role of children in this relationship. We show that compensating wage differentials for commuting are low for all workers, particularly for women with children, implying that women with children are not compensated for the higher cost of commuting through higher wages in the labor market.

[2] Petrongolo and Ronchi (2020) show that women are more likely than men to move job given a long commute, but ignore the role of children. The reduction in commuting distance for mothers has also been documented in Germany, see Skora et al. (2020)

Our study refers to a range of literature. First, our paper refers to a literature emphasizing that women have higher commuting costs, resulting in restrictive job search and shorter commutes (Le Barbanchon et al., 2021; Farré et al., 2020; Petrongolo and Ronchi, 2020). Employing revealed preference data, we demonstrate that women *with children* bear higher marginal costs of commuting. Consistent with that we show that gender differences in the length of the commuting distance come into existence after the birth of the first child.

Second, we contribute to the urban economics literature aiming at estimating the marginal cost of commuting, i.e. the marginal willingness to pay for commuting. Commuting costs are fundamental as they determine the urban spatial structure by influencing the size as well as the structure of cities (Wheaton, 1974; Fujita, 1989; Lucas and Rossi–Hansberg, 2002; Baum-Snow, 2010; Ahlfeldt et al., 2015; Heblich et al., 2020), but surprisingly few estimates of the commuting costs exist.

Third, our paper also relates to a large literature on the value of non-wage job attributes for workers (Ophem, 1991; Gronberg and Reed, 1994; Bonhomme and Jolivet, 2009; Sullivan and To, 2014). Important non-wage job attributes include health insurance (Gruber and Madrian, 2004; Aizawa and Fang, 2020), employer-provided retirement benefits (Altonji and Paxson, 1992), employer-provided cars (Gutiérrez-i Puigarnau and Van Ommeren, 2011), and employer-provided parking (Van Ommeren and Wentink, 2012).

Fourth, our paper refers to a range of theories that aim to understand the relationship between wages and commuting distance. These theories predict that employers compensate workers for commuting (Wheaton, 1974; Madden, 1985; Hwang et al., 1992; Zenou, 2009a). For example, for a labor market characterized by job search and wage dispersion, workers receive an implicit compensation for commuting (Manning, 2003b). For a perfect labor market, employers that are further located from residential areas have to pay higher wages to attract workers (Wheaton, 1974; Madden, 1985) so workers are explicitly compensated.

In the current paper, *(i)* we apply the methodology introduced by Gronberg and Reed (1994) to estimate the marginal cost of commuting derived from information about the effects of commuting distance and wages on job mobility given assumptions on the job search environment (as in Van Ommeren et al. (2000), Manning (2003a), Van Ommeren and Fosgerau (2009) and Le Barbanchon et al. (2021)), and *(ii)* we estimate compensating wage differentials for commuting. We offer several improvements.

Our first, and main, improvement is that we improve the Gronberg-and-Reed methodology to estimate the cost of commuting as applied in Van Ommeren et al. (2000), Manning (2003b) and Van Ommeren and Fosgerau (2009). In essence, this approach estimates the effect of non-wage job characteristics (i.e., commuting distance) and wages on job mobility. The ratio of these effects provides information about the willingness to pay for these non-wage characteristics. The underlying idea is that workers search for a job where the distribution of wages of alternative jobs is given (Pissarides, 2000). Consequently, workers with higher wages are less likely to move jobs, because alternative jobs have become less attractive.

The fundamental econometric problem with this approach is that workers are heterogeneous, so the worker's wage is an increasing function of the productivity level. However, a higher level of productivity shifts the distribution of wage offers to the right. For example, if one observes a worker with a high wage, then it may be the case that this worker is particularly productive (compared to another period), or that this worker had a lucky draw from job offers (Barlevy, 2008). Only in the latter case, there would be a strong incentive not to move to another job. Consequently, not controlling for worker productivity will result in an estimate of the marginal effect of wages which is biased towards zero. This bias may be large because it is generally thought that the relationship between wages and productivity is very tight (and even one-to-one according to fully competitive labor market models without search). The literature is aware of this bias, and in empirical applications, workers' characteristics (e.g. education, age, sector) are used as controls (Manning, 2003b; Van Ommeren and Fosgerau, 2009). However, many characteristics of the worker are still unobserved. In the current paper, we improve on the empirical method by including worker-fixed effects.

The inclusion of worker fixed effects, which controls for time-invariant unobserved heterogeneity, is not sufficient (and may make it even worse): workers' wages strongly vary over time, and if workers' productivity changes, the wage offer function changes over time. For example, it implies that if we observe an associate professor who receives a wage increase from her current employer because of a top-five publication, it is plausible that her wage offer distribution would also be affected by this publication. We solve the econometric problem by combining the worker-fixed effects with an IV approach. In essence, we are looking for an instrument that determines a worker's wage, but not directly the wage offer distribution of this worker, as this would directly affect job mobility.

We use the average wage of other workers with similar positions within the same firm as an instrument, where we control for firm characteristics – sector and firm size – which are known to correlate with nonwage amenities (Oi and Idson, 1999). Hence, the identifying assumption we make is that changes over time in the wage offer distribution of a worker are not related to changes over time in the average wage of other workers in the same firm, conditional on the sector and firm size.[3]

[3]Using an IV approach also reduces other econometric issues, such as measurement error in net income, e.g. because the tax rate on labor income depends on non-labor activities such as house

This assumption can be criticized because of the presence of unobserved nonwage amenities that correlate with the average wage within the firm, but not with sector or firm size. We deal with this by adding many other controls, such as more detailed sector controls, average educational level, the proportion of female employees, and the presence of female top management. These latter controls aim to capture amenities particularly important to female employees with children, such as flexible working hours. Our results remain robust.

Our second improvement is that our study presents a significant advance in data quality compared to previous studies. We use administrative register data for the universe of the working population of Denmark (rather than survey data), and we observe a precise measure of commuting distance. This allows the econometric analysis to control for unobserved time-invariant worker characteristics using worker-fixed effects and calculate our instrument, whereas previous studies rely on cross-section identification.

Our third improvement is in the analysis of the relationship between wages and commuting, where we estimate compensating wage differentials for commuting (Madden, 1985; Zax, 1991). First, we allow this relationship to differ by gender and the presence of children. Second, we include worker and household-by-residence-location fixed effects, where residence location is measured at the parish level. Worker fixed effects are standard and deal with the unobserved time-invariant worker characteristics. The inclusion of household-by-residence-location fixed effects improves identification by dealing with reverse causality from endogenous residence location, where the length of the commuting trip depends on the household income level (Wheaton, 1974; Fujita, 1989; Lucas and Rossi–Hansberg, 2002; Zenou, 2009a). The inclusion of both types of fixed effects implies that we essentially use the information on changes over time in the commuting distance of men and women who belong to the same household and live in the same residence location.

When interpreting our results, we assume that the labor market is characterized by search frictions and, therefore, not competitive. This interpretation is consistent with our main theoretical framework where we estimate the marginal cost of commuting, as this framework relies on a labor market where workers get offers from a wage distribution. This is fundamental because frictions in the matching between workers and jobs imply that the workers' evaluation of these job attributes is *not* equal to the compensating wage differentials of these job attributes (Hwang et al., 1992; Mulalic et al., 2013; Mas and Pallais, 2017). This criticism applies in principle to all job attributes, but in particular to commuting, as job search frictions are thought to be essential to explain commuting outcomes (Manning, 2003a; Le Barbanchon

ownership.

et al., 2021).[4] One of the main consequences, in line with Manning (2003b), is that we interpret the estimates of the compensating wage differentials as correlations between wages and commuting distance that exist given a job search process where workers choose to accept or reject job offers implying a certain wage and a commuting distance, rather than as causal effects of commuting (e.g. due to a relocation of the employer to another location) on wages.

The remainder of the paper is organized as follows. In Section 1.2 we present and describe the data. We then first in Section 1.3 establish the relevance and extent of the gender pay and commuting gaps using an event study methodology, and then in Section 1.4 derive the marginal cost of commuting. We estimate and discuss the marginal cost of commuting in Section 1.5. Section 1.6 deals with the relationship between wage and commuting. Finally, Section 1.7 presents the main conclusions.

1.2 Data

Our sample consists of longitudinal administrative register data for the full working population in Denmark. We observe all workers' demographic information (such as gender, number of children, and education) and labor market outcomes (such as annual wage, occupation, and sector).

We restrict our sample to workers who are employed between 2003 and 2013 and we censor observations of workers who move into non-employment, so all our job moves refer to job-to-job moves. This restriction makes it likely that the job moves (observed by us) tend to be voluntary, which will be a requirement of the approach introduced later on. Furthermore, we select observations of individuals who experience the birth of their first child either in this period or within up to 9 years before or 4 years after this period. This restriction is useful because workers without children may face different labor market conditions. We also impose a standard set of sample selection criteria for workers, i.e. we exclude workers younger than 19 or older than 45, workers who are in ongoing education, teleworkers, workers with an extremely low income (the lowest percentile), and workers with commuting distances exceeding 50 km. Commuting distance is calculated for each worker as the shortest route between the worker's residence and workplace location, taking into account that the shortest distance due to changes in road infrastructure (Börjesson et al., 2019; Mulalic and Rouwendal, 2020). In our analyses, we capture wages using annual *net* labor income, which includes a commuting tax deduction.[5] Commuters

[4]There is also another fundamental reason why hedonic wage models are less informative to estimate the worker's marginal cost of commuting. Commuting is not a pure job attribute as it is worker specific, because it depends on the worker's residence location, as emphasized in the urban economic literature. The consequences of this are discussed in Section 1.5.4.

[5]In 2019, commuters were entitled to deduct 1.96 DKK, about 0.20 US dollars, from gross income

in Denmark are entitled to a tax deduction when the commute exceeds 12 km, which disproportionally benefits male commuters.[6]

We focus on full-time workers, which facilitates the interpretation of our empirical findings because for part-time workers we do not observe the exact number of hours worked. We define job mobility as a move from a (full-time) job to another job (which can be full-time or part-time).[7] We have slightly more than 3 million observations.[8] Due to the childbirth and age selections, we focus on workers at the beginning of their career: workers are, on average, about 28 years in the period before the birth of their first child and about 35 years in the period after.

Table 1.1 shows that the average commute for men and women before the birth of their first child is quite similar: men commute 13.2 km and women commute 12.3 km, so a difference of 1 km, about 8%. After childbirth, however, it increases for men by almost 2 km to 15.0 km, while for women it increases by only 0.4 km to 12.7 km. The average increase in commuting distance for men after the birth of the first child is around 2.3 km, so by about 17%, longer, which is substantial. The time devoted to commuting increases then by approximately 30 minutes per week.[9]

Wages for men exceed wages for women before and after childbirth, but their difference is larger after childbirth: the gender pay gap amounts to 12% before childbirth and 24% after. It further appears that the Danish job market is characterized by high labor turnover and therefore by short job durations (on average 3 years). Around 16% of workers move to another job within a year. Residential moving behavior is particularly important before childbirth (about 19% of workers move residence each year), but this drops to 9% after childbirth. The shares of men and women that move residence or job before and after childbirth are similar. In

per kilometer driven, so about 4 DKK per (one-way) commuting distance. A range of other European countries have similar commuting tax deductions (Potter et al., 2006; Paetzold, 2019).

[6]In terms of deduction incidence, gender differences are moderate: without children, 41% of men and 37% of women receive the deduction, and with children 49% of men and 41% of women. However, on average, the implied subsidy is several times larger for men. With children, the average annual commuting subsidy for men is 652 DKK whereas for women it is only 163 DKK; without children, the annual subsidy is 255 DKK for men and 61 DKK for women. Relative to annual income, this subsidy is small, and even for men with children, it is only 0.17% whereas, for women with children, it is 0.05%. Consequently, the subsidy implied by the deduction is too low to notably affect wage setting and job mobility. 1 DKK $\approx$ 0.15 \$.

[7]Take note that the share of part-time workers is generally low (10-15%) and stays more or less the same before and after having a child, and across genders, see Figure B.3 in Appendix 1.B. Out-of-sample job mobility is also limited at 6-10%. Therefore, the focus on full-time workers is less restrictive in Denmark compared to other countries (Kleven et al., 2019a).

[8]Our original sample consists of about 10 million observations. We exclude observations with commuting distances outside the range (about one million observations), observations not referring to parents (about 4 million observations), part-time (about 0.5 million observations), censoring income (about 0.1 million observations), and observations with missing values (about 0.2 million observations).

[9]In Appendix 1.A, using survey data, we show that the marginal effect of distance (in kilometers) on commuting time (in hours per trip) is about 0.025. We then multiply 0.025 by the increase in commuting distance (2.3 km) x 10 trips.

Table 1.1: Summary statistics by gender and period (birth of first child)

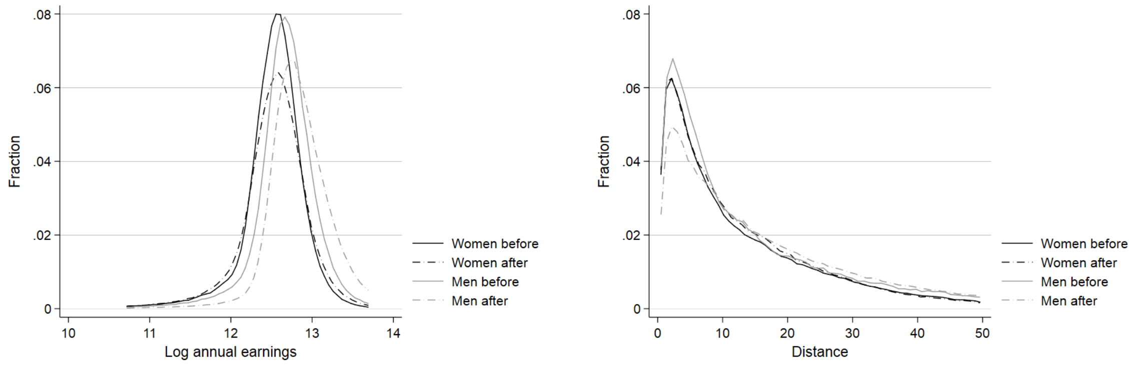
	Men		Women	
	Mean	Std. dev.	Mean	Std. dev.
Before childbirth				
Commute (km)	13.20	11.91	12.26	11.72
Annual net income (DKK)	336,469	114,670	293,826	98,184
Job move	0.18	0.38	0.17	0.37
Residence move	0.19	0.39	0.19	0.39
Job tenure	2.80	2.63	2.43	2.14
Age	28.45	4.99	27.94	4.44
N	501,478		443,408	
After childbirth				
Commute (km)	15.04	12.24	12.65	11.03
Annual net income (DKK)	394,345	137,048	298,310	113,435
Job move	0.15	0.36	0.15	0.36
Residence move	0.09	0.29	0.09	0.28
Job tenure	4.36	3.90	3.94	3.43
Age	36.20	4.30	34.91	4.22
N	1,140,917		1,179,652	

Notes: Full-time workers in the ten years around the birth of the first child. Observations of the year of the childbirth are excluded. 1 DKK ≈ 0.15 $.

Figure 1.1a, we show distributions of log wage by gender and presence of a child. A remarkable feature of the distributions is that they are similar for men and women before the event, but not after: in particular the share of women with low wages increases, while for men the whole distribution moves to the right. In Figure 1.1b we show the commuting distributions by gender and child. It appears that after childbirth the share of men with short commutes strongly drops, while for women this does not occur.

Finally, we have also examined to what extent changes in commuting distance are predominantly due to a residential move or a job move. It appears that the average (absolute) change in commuting distance is about 7.0 km given a residential move, whereas the (absolute) change in commuting distance given a job move is somewhat higher and equal to 9.4 km. Consequently, changes in commuting distance are mainly a labor market phenomenon and originate less in residential moving, as residential moves are mainly local, particularly for households with children.

Figure 1.1: Distribution of wages and commuting distance by gender and first child



(a) Wage **(b)** Commuting distance (km)

1.3 Gender, wage and commuting gap

We first establish the relevance and extent of the gender commuting gap using a standard event study methodology based on the birth of the first child, following studies such as Kleven et al. (2019b). We employ individual-level variation in the timing of the child's birth. Observed sharp changes in wage and commuting for mothers relative to fathers around the birth of the first child are likely orthogonal to unobserved determinants of these outcomes as they evolve smoothly over time. To reduce the selection effects of childbirth, we only select individuals who become a parent for the first time either during the period of observation or in the 10 years before or after childbirth.

Event time is denoted by t (measured in years) and we observe the childbirth at time $t = 0$ (the actual childbirth occurs between -1 and 0). We focus on two outcome variables if worker i: wage and the length of the commute, both denoted by $y^g_{i,s,t}$. We then estimate the effect of the childbirth at $t = 0$ on $y^g_{i,s,t}$, for each gender g separately, controlling for year s and age $h_{i,s}$:

$$y^g_{i,s,t} = \sum_{j \neq t'} \alpha^g_j \cdot \mathbb{I}[j = t] + \sum_k \beta^g_k \cdot \mathbb{I}[k = h_{i,s}] + \sum_l \gamma^g_l \cdot \mathbb{I}[l = s] + v^g_{i,s,t}, \tag{1.1}$$

where event time effects are captured by α^g_j which yields the event time effect in relation to the year of the birth and $\mathbb{I}$ denotes an indicator variable.[10] In (1.1) we exclude α^g_j for $j \neq t'$ which is the reference category. This implies that the event time coefficients measure the impact of the birth of the first child relative to t'. When we focus on commuting distance then $t' = -1$, i.e. the last year before the worker is affected by childbirth. When we focus on wages then $t' = -2$, as we wish to allow for

[10] In our application, α^g_j range from -10 until $+9$. This specification does not include worker-fixed effects but they will be included in later analysis.

reduced wages due to maternity leave in the year before the childbirth. β_k^g captures the effects of a set of age dummies (to control for life cycle), γ_l^g a set of year dummies (to control for time trends), and $v_{i,s,t}^g$ is a (gender-specific) error term (age dummies are important because women are often younger than man when having their first child). The estimated $\tilde{\alpha}_j^g$ are converted to percentage changes by $\tilde{\alpha}_t^g/\tilde{y}_{i,s,t}^g$, where $\tilde{y}_{ist}^g$ is the predicted outcome using the estimated coefficients (while excluding α_j^g), i.e. $\tilde{y}_{i,s,t}^g = \sum_k \tilde{\beta_k^g} \cdot \mathbb{I}[k = h_{i,s}] + \sum_l \tilde{\gamma_l^g} \cdot \mathbb{I}[l = s]$. It captures the event time effect at t as a share of the counterfactual outcome (i.e. no child at t').

In Figure 1.2, we show $\tilde{\alpha}_t^g/\tilde{y}_{i,s,t}^g$ based on the estimates of (1.1). Figure 1.2a shows a gender pay gap of about 15% immediately after childbirth compared to the year before pregnancy. It also shows that the wages of women and men follow the same trend before (and after) birth. Women's wages drop substantially after childbirth, while in contrast men's wages only slightly decrease. Moreover, the figure also shows that the effect of the birth of the first child is very persistent, i.e. it remains at the same level 10 years after the child's birth. These results are not novel to the literature. For example, these results are consistent with Kleven et al. (2019b) who find that the gap remains after 20 years.[11]

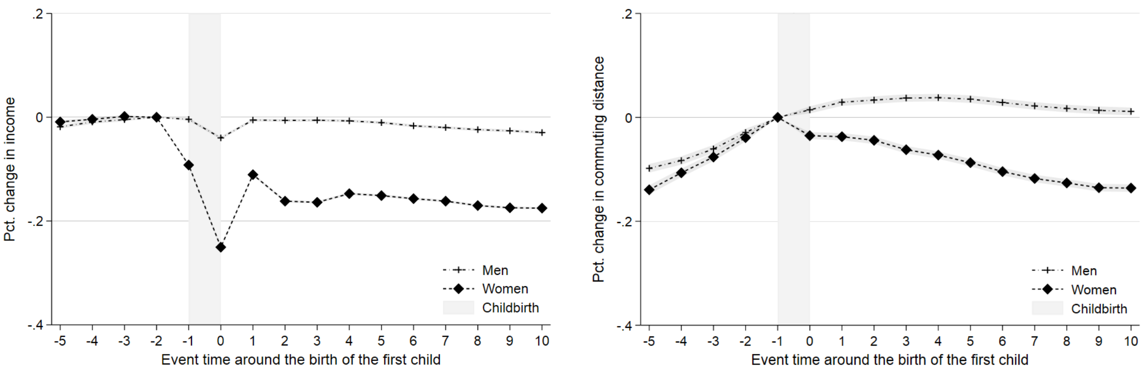
We now focus on the role of childbirth in commuting distance, which is of interest here. Figure 1.2b shows that the commuting distances of women and men follow the same upward trend before the birth of the child, but after childbirth, women's commuting distance gradually reduces, while men's commuting distance uninterruptedly follows the trend a few years after the childbirth and then stagnates. The gender commuting distance gap ranges from about 5% immediately after childbirth (compared to the year before pregnancy) to about 15% ten years after. The resulting difference in commuting patterns after childbirth hints towards an increase in the cost of commuting for women after having a child.

Additionally, we have tested whether the observed gender difference in commuting distance after childbirth is sensitive to additional controls. For example, we have performed the same analysis with two additional control variables: education and the number of workers at the firm level. The results remain robust.

The latter result raises the question of whether the observed gender differences in the commuting distance are predominantly due to residential moving – which implies that households tend to make residential moves which make them locate closer to the workplace location of the new mother rather than the new father – or predominantly due to gender differences in workplace locations when moving job.

[11]When we include part-time workers, our results do not fundamentally change. We also estimate models on sub-samples of workers that either do not move jobs or do not move residence or both (in the period starting 3 years before the birth), see Figure B.1 in Appendix 1.B. The results remain the same.

Figure 1.2: Wage, commuting and the first child



(a) Wage **(b)** Commuting distance

Notes: wage and commuting distance event time effects around the birth of the first child. The grey area marks the time interval of the birth of the first child. The shaded 95 percent confidence intervals are based on robust standard errors.

To investigate this, we focus on sub-samples of workers that either do not move jobs or do not move residence (in the period starting 3 years before the birth), see Figure B.2 in Appendix 1.B. These figures suggest that the gender differences in commuting distance after childbirth are predominantly due to gender differences regarding job location. This is consistent with the notion that residential moving is relatively rare in Denmark.

1.4 Theoretical foundations

In this section, we discuss the theoretical foundation to estimate the marginal cost of commuting and explain that this theoretical foundation implies compensating wage differentials for commuting. We also discuss alternative theoretical perspectives discussed in the literature that imply compensating wage differentials for commuting.

Our theoretical starting point is that we focus on a labor market where workers have to search for jobs located at different locations, and where workers maximize their utility by moving jobs from one location to another location (Manning, 2003b). The essential assumption made is that jobs are characterized by wages and commuting distance and that employers post wage offers drawn from a given wage distribution which does not vary over space.[12] Workers will accept all jobs that offer a utility increase. Workers do not expect to move residence or to have a child which changes their costs of commuting.

[12]This assumption implies that the wage offer distribution does not depend on the wage level of the worker. For heterogeneous workers with different productivity levels, this assumption is unlikely to hold, because the wage offer distribution is a function of the worker's productivity. We deal with this in our empirical application using an IV approach.

1.4.1 The marginal cost of commuting

We are interested in estimating the marginal cost of commuting, defined here as the marginal monetary valuation of commuting distance. Intuitively, one can do so by using information on the effect of commuting distance on voluntary job mobility (Van Ommeren et al., 2000; Van Ommeren and Fosgerau, 2009).

We assume a labor market with jobs that are characterized by wages and commuting distance, and where employers post wages drawn from a wage distribution (Manning, 2003b). Space is homogenous: every point in space has the same level of employment, population, and wage distribution. Space is two-dimensional and workers are not allowed to move residence.

Workers get utility from wages, w, and disutility from distance to work, x. Utility is additive in the *logarithm of wages* and commuting. Hence, $v = log(w) - \alpha x$. Job offers, implying an offer of w^* and x^*, arrive at an exogenous arrival rate λ. Wage offers come from a continuous wage offer distribution $F(w^*)$. Workers will accept all jobs that offer a utility increase. The job moving rate, so the rate of job offers that are accepted, is denoted by θ. Finally, it is assumed that workers maximize lifetime utility V while discounting the future at a given discount rate.

Comparative statics analysis shows that α has an ambiguous effect on the job moving rate (see Appendix 1.C). This makes sense because an increase in commuting costs reduces the utility of the current job. This effect is proportional to the length of the commute, so this effect is weak for short commutes and strong for long commutes. At the same time, an increase in α makes all job offers to become less attractive. Interestingly, the latter effect appears to be convex, as the job moving rate is, conditional on the utility of the current job, inversely proportional to the *square* of α. An increase in α reduces the job moving rate more than proportionately, because space is two-dimensional, suggesting that a higher α (e.g. due to having a child) may have major consequences to restrict job moving as also implied by the study of Le Barbanchon et al. (2021).[13]

We are interested in estimating the value of the instantaneous marginal cost of commuting, MCC, defined by $-(\partial v/\partial x)/(\partial v/\partial w) = \alpha w$. Hence, α can be interpreted as the (relative) marginal cost of commuting, i.e. the marginal cost of commuting relative to the wage. Consequently, the marginal cost of commuting can be estimated by estimation of α. It is straightforward to show that the marginal cost of commuting, MCC, can be derived from information about job mobility (see Appendix 1.C):

$$MCC \equiv -\frac{\partial v/\partial x}{\partial v/\partial w} = -\frac{\partial \theta(w,x)/\partial x}{\partial \theta(w,x)/\partial w} = -\frac{\partial \theta(w,x)/\partial x}{\partial \theta(w,x)/\partial log(w)} w = \alpha w. \tag{1.2}$$

[13] Manning (2003b) assumes that space is one-dimensional, and therefore gets instead that the job moving rate is inversely proportional to α rather than α^2.

Consequently, MCC is equal to αw, where α is equal to the ratio of the marginal effect of commuting distance on job mobility and the marginal effect of log wage on job mobility. This result also holds given less restrictive assumptions such as non-homogeneous space, endogenous job search, business cycles, and job moving costs are allowed (Van Ommeren et al., 2000).

In the current paper, our econometric methodology to estimate MCC is to derive estimates for α using estimates of the effects of log wages and commuting distance on the job moving rate. Our key interest is to examine to what extent α depends on the presence of children, and whether or not this differs by gender.

When interpreting the results, there is a subtle issue regarding the expectations households have about the future cost of commuting. Van Ommeren et al. (2000) show that (1.2) does *not* hold for when a worker expects changes in the cost of commuting unrelated to changes in workplace location. The primary example is that the worker expects to move residence or have a baby. In that case, the ratio of the marginal effects on job mobility is equal to the *expected* marginal cost of commuting, defined by $\mathbb{E}[MCC] \equiv -\frac{\partial V/\partial x}{\partial V/\partial w}$, which is also equal to αw. Hence, there is a subtle difference in interpretation.

When workers expect to move residence, and therefore expect a change in the length of the commute, workers typically care less about their current commute, hence we expect that $\mathbb{E}[MCC] < MCC$. Note however that if the residential move is very local (which applies to many residential moves), the expected change in the commuting distance through a residential move is minimal, then $\mathbb{E}[MCC] = MCC$. In contrast, when workers expect to have a child, it is plausible that the commuting costs go up, so we expect that $\mathbb{E}[MCC] > MCC$. On the other hand, female workers who expect to have a child also anticipate not working during maternity leave (for about 12 months), which suggests that $\mathbb{E}[MCC] < MCC$. To examine whether the difference between the instantaneous and the expected monetary valuation of the commuting distance is important, we examine whether anticipation of residential moves or childbirth plays a role in our estimates.

Above, we have ignored that a worker's *wage* may change without changing job (a phenomenon one frequently observes in the data). One reason for such a wage change might be that the productivity of the worker has changed. It is plausible that the wage offer distribution of this worker also changes. One can easily allow for this by assuming that the current wage and the wage offer distribution both depend on the worker's level of productivity, which is drawn from a given productivity distribution at a given rate. In this case, (1.2) still holds conditional on the level of productivity. However, unconditional on this productivity level, it does not hold, because the wage is correlated to the wage offer. As we do not observe the productivity level, in the empirical analysis, we will deal with this by instrumenting the wage.

Another reason for an on-the-job wage change is that there is productivity shock *at the firm level*. This can also be easily allowed for theoretically by assuming that the worker's wage depends on the firm level of productivity, which is drawn from a given firm productivity distribution function. We emphasize that it must be assumed that this drawing is *not* correlated to the worker wage offer distribution. This assumption makes sense because the latter wage offer distribution is determined by other firms. Given these assumptions, (1.2) still holds, also unconditional on the firm's productivity level. We do not observe the productivity shocks to the firm, but we will use the average wage within the firm as a proxy for this shock, which can be used to instrument changes in the worker's wage.

1.4.2 Compensating differentials

The above assumptions imply that the distribution of accepted wages is increasing in the length of the commute (in the sense of first-order stochastic dominance), see Manning (2003b), i.e. accepted wages and commuting distance are positively correlated. This result is intuitive as workers trade off wages and commuting when accepting job offers, which results in an *implicit compensating wage differential*. When workers get a job offer far away from their home, they are more likely to accept the job offer when the wage offer is high, implying a positive relationship between wages and distance. This compensation is, on average, less than complete: workers who travel longer distances to work tend to be worse off.

These compensating wage differentials may differ by gender and the presence of children. On theoretical grounds, women with children may receive higher or lower levels of compensating wage differentials. If women with children face higher costs of commuting, they may make different trade-offs between wages and commuting distance when searching for jobs, see e.g. Manning and Petrongolo (2008), and more precisely, who only accept jobs far away if they offer a higher wage, implying that their implicit compensation might be higher (Manning, 2003b). On the other hand, women with children tend to work in jobs and sectors, where wage dispersion tends to be less (several times more likely to work in the public sector, less likely in managerial positions). Given lower levels of wage dispersion, implicit wage compensation tends to be less (for example, in the extreme case of no wage dispersion, there is no implicit wage compensation). This suggests that women with children may receive lower levels of implicit wage compensation. Moreover, wage compensation for commuting may be different for women with children, because women with children tend to sort into jobs at different sectors, which may have wage offer distributions that differ from other workers (Blau and Kahn, 2017).

1.4.3 Alternative theoretical perspectives

Spatial variation in wages

Compensating wage differentials, in the spirit of Rosen, refer to wage bonuses that firms provide to marginal workers to accept some adverse job attributes. Commuting distance is particular in this respect as it is not a pure job attribute: it varies at the worker-job level because the residence location is worker-specific. This issue has received a lot of attention in the urban economic literature (Fujita, 1989).

Compensating wage differentials for commuting in the urban economic literature allows for spatial variation in locations of jobs and residences, but typically relies on the assumption of a perfect labor market with complete information – i.e. no job search frictions. Workers' wages would then be equal to their marginal productivity, and compensation differences for commuting for workers employed *at the same workplace location* would not exist. At the same time, compensation differences for commuting for workers residing *at the same residential location* must exist, as employers have to compensate workers to accept jobs further away from their homes (Wheaton, 1974; Fujita, 1989).

This theory indicates therefore that firms may pay higher wages if they are located at a location far from where workers live (or to be more precise, far from where the marginal worker lives). This suggests that the wage of a worker employed at a certain firm is a positive function of the average commuting distance, where the average is calculated for workers employed at the firm. For empirical evidence, we refer to Timothy and Wheaton (2001).

Consequently, workers may receive compensation for commuting, but only if their commuting distance is strongly related to the average commuting distance of the firm (so, the variation of commuting distance of workers within the firm is small). However, because of search imperfections, this not to be true (in our data, the correlation between the worker's commuting distance and the worker's firm-average commuting distance is only 0.27). We show later on that if we control for the average commuting distance, we get almost identical results for the effect of individual commuting distance, implying that spatial variation in wages is not the source of the compensating wage differentials estimated by us.

Bargaining and other forms of monopsony power

We now discuss alternative models where employers have market power (Card, 2022). In our theoretical setup, we have ignored that some workers may bargain about wages which may depend on the commuting distance (Van Ommeren and Rietveld, 2005; Zenou, 2009a; Mulalic et al., 2013; Biasi and Sarsons, 2021), or that employers have other forms of monopsony power (Manning, 2003a). The theoretical studies investigating this indicate that bargaining, as well as other forms of monopsony power, induce employers' wages that are an increasing function of the length of the commute, see e.g. Van Ommeren and Rietveld (2005), Mulalic et al. (2013), and Biasi and Sarsons (2021). This allows for the possibility that women (with children) may be treated differently from other workers, as suggested by Manning (2003a) and Barth and Dale-Olsen (2009). In contrast, several studies have also pointed out that the result that wages increase as a function of distance does not hold when workers are perfectly mobile in the housing market and house prices perfectly compensate workers for the length of their commute. The assumption of perfect mobility and the housing market is essential here. For example, given positive residential mobility costs and compositing house prices, Zenou (2009b) shows that employers still partly compensate workers for their commuting costs. For reviews, we refer to Zenou (2009a) and Mulalic et al. (2013). In our empirical application, we investigate the importance of bargaining.

Productivity and distance

Up to now, we have assumed that commuting distance is not directly related to productivity. This assumption is not in line with the idea that commuting distance is negatively related to productivity as workers with long commutes tend to be more ill (and, in particular when taking public transport, an issue which became clear during the Covid crisis), and have higher levels of fatigue (Koslowsky et al., 2013; Künn-Nelen, 2016). It is also not in line with the idea that workers with long commutes are more likely to shirk, so they are more absent, as the costs of being fired are lower (Zenou, 2002; Ross and Zenou, 2008; Van Ommeren and Gutiérrez-i Puigarnau, 2011). This suggests that employers may pay different, and more likely lower, wages for workers with longer distances. This is particularly relevant in the context of our paper because women (with children) tend to have higher absenteeism rates (Künn-Nelen, 2016; Daly and Groes, 2017). Consequently, in our empirical application, the wage compensation differential for commuting should be interpreted as the *net* wage differential, i.e. the gross wage differential for commuting assuming that productivity does not depend on distance minus the marginal effect of distance on productivity.

1.5 Marginal cost of commuting: empirical application

In this section, we turn to the estimation of the marginal cost of commuting. The first two subsections show how the marginal cost of commuting can be estimated using our econometric approach which is supported by a graphical approach. Subsection 1.5.3 reports our main findings of estimating the marginal cost of commuting and subsection 1.5.4 presents robustness checks.

1.5.1 Econometric approach

We aim to estimate the parameter α to derive the marginal cost of commuting as explained. This is not the first study that exploits information on job mobility to derive the marginal cost of commuting (Van Ommeren et al., 2000; Manning, 2003b; Van Ommeren and Fosgerau, 2009). We make two fundamental contributions. First, we employ a large panel of workers over a long period, so we can identify the parameters of interest using worker-fixed effects, whereas the previous studies essentially rely on strategies identifying parameters of interest without worker-fixed effects. Second, we introduce an instrumental variable approach to deal with the issue that workers' wage offer distribution is unobserved and correlated to their current wage.

We aim to estimate the causal effects of wage and commuting distance on job mobility. We employ a linear probability model, as in Manning (2003b), which offers two advantages. First, it estimates the average causal marginal effect (Angrist and Pischke, 2008, p.93), in which we are interested. Second, we wish to include many fixed effects for a very large dataset, which is computationally cumbersome for the non-linear approaches, such as survival analysis and discrete choice models which have been applied in this context (Van Ommeren et al., 2000; Van Ommeren and Fosgerau, 2009).

We differentiate both effects by gender, g, and the presence of a child, c. One complication, as is common with annual data, is that we observe the commuting distance at the end of the year and the average wage per year. Consequently, in the year that the worker moves, the average wage is a combination of the before-the-move wage and after-the-move wage, which is problematic because we wish to know the effect of the before-the-move wage on job mobility. To deal with this, we define a job move in year t, when the actual move takes place the year after. Given this definition, we use a job moving dummy indicator $J_{i,t}$ which captures whether a worker, i in year, t, moves job. We then use the following linear probability model, to estimate the effects of instrumented log wage and commuting distance on job mobility:

$$J_{i,t} = \alpha_{g,c} \cdot x_{i,t} + \beta \cdot log(w_{i,t}) + \gamma \cdot X_{i,t} + \delta_{g,c} + \lambda_i + \kappa_t + \varepsilon_{i,t}, \tag{1.3}$$

where our main interest is in the marginal effects of commuting distance, $x_{i,t}$ and log wage, $log(w_{i,t})$, which are captured by the coefficients $\alpha_{g,c}$ and β, respectively.

Importantly, $\alpha_{g,c}$ is gender and child-specific. We also include $\delta_{g,c}$, which is a gender and child interaction term, which allows job mobility to change over time for reasons not captured by commuting distance or wages. This is essential, as the literature has shown that wages discretely jump around the birth of a child, a characteristic which also holds in our data, suggesting that other factors than only wage may discretely change.[14]

$X_{i,t}$ consists of a vector of additional controls, which includes marital status, broad sector controls (NACE 1), firm size, the average age of workers at the firm, and job tenure. We include worker λ_i and year κ_t fixed effects, and $\varepsilon_{i,t}$ is an idiosyncratic error term. Standard errors are clustered at the firm level. We emphasize here that we include worker-fixed effects, so we control for time-invariant worker characteristics. Consequently, we examine whether changes in the wage levels of workers affect their job mobility.

In the literature, to deal with the endogeneity of wages, empirical approaches rely on identification by using control variables. Given that we include worker fixed effects this implies that changes in wages are not correlated to changes in the wage offer distribution. This assumption is unlikely to hold. For example, if we observe that a worker receives a higher wage while staying at the same job, it is very plausible that the productivity of this worker has increased, and therefore the wage offer distribution of this worker also has changed.

To address this issue, we use an instrumental variable approach, where we use (log of) the average wage of *similar* workers that work at the same firm as an instrument, where similar is defined as belonging to the group of workers who have children during the observed time interval and who are in the same job position, where we distinguish between 7 broad job positions (e.g., manager). The underlying idea of this instrument is that productivity improvements *at the firm level* reflect in individual workers' wage increases, which do not affect the wage offer distribution of this worker. These productivity improvements at the firm level should be contrasted with the productivity improvements at the individual level, which do affect the wage distribution of a worker. In order to argue that the average wage is exogenous, we take two steps: we exclude the wage of the worker and we only include firms with at least 10 workers, which refers to about 95% of all workers. By excluding small firms, we avoid the inclusion of workers who are owners of the firm rather than employees, or workers who are family members, for which the wage does not reflect market wages.

The underlying assumption to justify the IV approach is that the average wage of the firm does not directly affect individual job-moving decisions, except through

[14]For example, we allow for the situation that women with children receive fewer job offers for unobserved reasons.

its effect on the individual wage of the worker. To minimize the possibility that the average wage is correlated to the presence of unobserved nonwage amenities, we control for a range of firm characteristics, including firm size and sector that are known to correlate to wages and nonwage amenities (Oi and Idson, 1999).

One may argue that also when we control for firm size and sectors, we do not fully address the presence of non-wage amenities.[15] Arguably, there might be more subtle unobserved nonwage characteristics (e.g. training opportunities) that are relevant to workers which correlate with the wage in the firm, and which are not picked up by our set of controls (e.g. flexible working times). To test for this, we will add controls, such as very detailed sector controls (NACE 3), average education, the share of females, and the presence of female top managers which proxies for amenities important to female workers with children.[16]

Another issue with the instrument is when *job-level* wage increases are driven by technological changes that are shared with other jobs in the same firm *and* in other firms.[17] In this case, the exclusion restriction would not hold. To address this, we use an alternative instrument, where we use the average wage of workers at the same firm, who are in job positions that are *not* similar to the worker. Using this instrument, we get similar results, but with a somewhat larger marginal cost of commuting.[18] So our approach that relies on using the wages of workers in similar positions as an instrument is the more conservative estimate.

To challenge the instrumental variable approach, we also examine a range of alternative specifications. For example, we have also examined other specifications with other definitions of ”similar workers”. When we include older workers in the same job position, the first-stage impact of the instrument becomes smaller, but the

[15]Also note that according to the admittedly somewhat outdated US literature, non-wage amenities are hardly important to workers, except for pensions and health care (Turner, 1987). In Denmark, healthcare and pensions are mandatory (Gruber and Lettau, 2004), suggesting that non-wage amenities may not be relevant in our context. Nevertheless, this literature ignores other more recent non-wage amenities, such as childcare benefits, that are likely important for young workers with children. Company cars have been shown to be relevant as these fringe benefits are relevant to workers in other European countries (Gutiérrez-i Puigarnau and Van Ommeren, 2011). We note here that studies for Denmark rule out the importance of these two types of non-wage amenities. In Denmark, company cars are hardly offered, as the Danish tax system offers little advantage of having a company car (Harding, 2014). Also, childcare is rarely supplied by firms as a fringe benefit. In Denmark, only 1-2% of firms offer (paid) childcare or any additional childcare allowance (Galanaki and Papalexandris, 2012).

[16]We do *not* include firm fixed effects. In that case, one effectively uses differences in the average wage growth experienced by the same worker at different firms as an instrument of the wage change. The first-stage effect of the average wage is then close to zero, resulting in an instrument that is either weak or not robust to minor changes in specification.

[17]We are less worried about *firm-level* wage increases that are driven by technological changes that are shared with other firms, as our results remain robust given additional detailed sector controls.

[18]This instrument also addresses the issue that the exclusion restriction will also fail if job mobility decisions of workers directly depend on the wages of co-workers in similar job positions, for example, because of jealousy. If these co-workers receive a raise in wages, and the worker does not, the worker could feel worse off and leave. Then the instrument would be invalid.

results remain robust. Finally, note that if the exclusion restriction does not strictly hold, then the bias in the estimates is unlikely large, as our instrument is strong. Here we use arguments developed by Conley et al. (2012); Angrist and Keueger (1991); Bound et al. (1995) which show that the bias from violation of the exclusion restriction is negatively related to the strength of the instrument.

We also estimate separate models for men and women, so we allow for differential effects in wages as well as other explanatory variables. Such a specification is in line with the labor economics literature, where there is a discussion to what extent the effects of wage on job mobility are gender specific, as these differences might be indicative of monopsony power by firms. A general finding in that literature is that these effects are very similar, for example, Manning (2003a) (an exception is Barth and Dale-Olsen (2009) that differentiates firms based on their gender composition, but not how firms differentiate between workers with a different gender).

Our results support the hypothesis that the effects of wages on job mobility are similar for males and females, and if they differ then these effects are smaller for women (in absolute value), implying that even if our wage coefficient estimate is biased, then the estimate of the *ratio* of MCC of women to men, which is one of our main interests, is an underestimate.

Moreover, our estimates could be biased because of unobserved household characteristics. To deal with these issues, we include household fixed effects, η_h, in some specifications for each worker belonging to the same household, h. The inclusion of this fixed effect is relevant for workers that change household (e.g. through divorce), as it is otherwise subsumed in the worker fixed effect. Hence, we essentially compare (instrumented) changes in wages and commuting distance of men and women workers who belong to the same household before and after they have a child.

1.5.2 Graphical approach

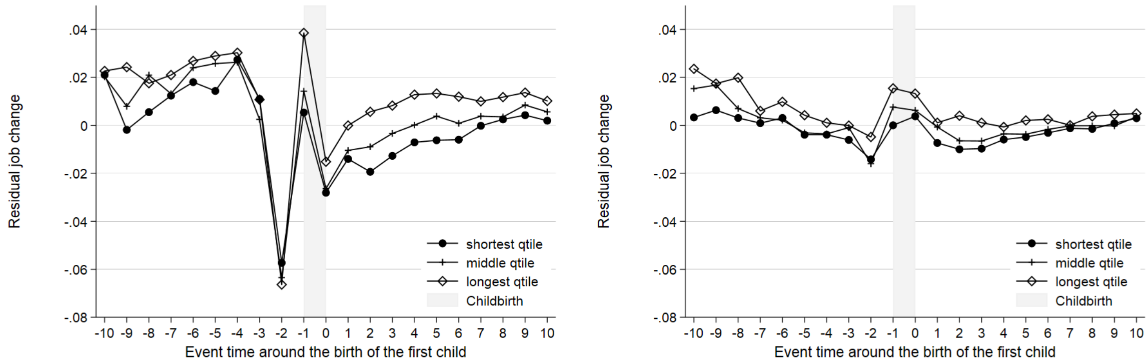
To support our econometric specification, we have examined the effect of commuting distance on job mobility graphically for several distance quantiles definitions (e.g. 3 quantiles, 5 quantiles, etcetera). Here, we control for worker fixed effects and the same controls used in our econometric approach, later on, so we show results for the job move *residuals*.[19] The results for these different quantiles definitions are very similar. In Figure 1.3 we show job mobility for 3 distance quantiles, so we show the job mobility residuals for a group of workers with a short commuting distance, for a

[19] Because we use controls, we apply the following two-step procedure. We first estimate a regression as in (1.3), but where we exclude commuting distance as an explanatory variable:

$$J_{i,t} = \beta \cdot log(w_{i,t}) + \gamma \cdot X_{i,t} + \delta_{g,c} + \lambda_i + \kappa_t + \varepsilon_{i,t}.$$

In the figures, we show the estimated residuals $\hat{\varepsilon}_{i,t}$.

Figure 1.3: Job changes by distance quantiles



(a) Women **(b)** Men

Notes: We estimate a regression as in (1.3), but where we exclude commuting distance as an explanatory variable, i.e. $J_{i,t} = \beta \cdot log(w_{i,t}) + \gamma \cdot X_{i,t} + \delta_{g,c} + \lambda_i + \kappa_t + \varepsilon_{i,t}$. The figures display the estimated job mobility residuals $\hat{\varepsilon}_{i,t}$.

group with a long commuting distance, and for a third group which is in the middle with respect to commuting distance.

There are several messages in this figure. First, and most importantly, workers belonging to the long commuting distance quantile tend to move jobs more, and this effect is particularly visible for women with children. Second, there is an extreme drop in job mobility of females just before the birth, which is likely due to a combination of reasons, including the effect of a Danish law that states that if women announce that they are pregnant, they cannot be fired, which reduces the incentives to search for another job. Third, for women who (expect to) become pregnant in the year after, we do not observe that the job mobility residual is higher for those with a long commuting distance. One possible explanation is that those women realize that during maternity leave they will not commute at all.

1.5.3 Empirical results

Our main results using different specifications to identify the marginal cost of commuting by estimating (1.3) can be found in Table 1.2. As we have seen job mobility around childbirth is extremely volatile, especially for women, which may potentially affect the estimates of the econometric analysis, in this specification, we exclude observations in the year before the birth, the year of the birth, as well as the year after the birth (in all other specifications, we keep these observations). All coefficients are estimated precisely and have expected signs. In all specifications, the wage is instrumented and it appears that the instrument is very strong with high F-values and has the expected positive sign. For example, for the specification shown in column [1], the effect of the log average wage on the individual's log wage is about 0.13, with an F-value equal to 6,640.

In column [1], which is our preferred specification, it is shown that the effects of commuting distance on job mobility are very similar for men and women before they have children, with coefficients equal to 0.0010 and 0.0007, respectively. Hence, given a hypothetical increase of about one standard deviation in the length of the commute, which is equal to almost 12 km, job mobility rates increase by about 0.012. After the birth of the child, the estimated effect of distance is about the same for men, and equal to 0.0010, but for women, the estimated effect is about 0.0025, so almost 3 times the estimated effect for their male counterparts. This supports our claim that gender differences regarding commuting play an important role after the birth of the first child. *Women who have a child are much more likely to leave their job when they have a long commute, which is not true for men*. This result is novel to the literature, as previous studies speculated about this effect, but failed to show this, see e.g. Van Ommeren and Fosgerau (2009).

Focusing on the same column, it appears that the effect of log wage on job mobility is negative, with a coefficient equal to about -0.18. This estimate implies that a 10% increase in the current wage decreases the job mobility rate by roughly 0.02, which is about 12% of the mean job mobility rate of 0.17. The order of magnitude of this estimate seems to make sense intuitively. For example, it suggests that a doubling of the wage in the current job would prevent most workers from leaving voluntarily ($0.17 - ln(2) \times 0.18 \approx 0.05$). This estimate implies a job moving elasticity with respect to the wage of about -1.1 (0.18/0.17), which is in line with the estimates obtained by Barth and Dale-Olsen (2009) for workers in the manufacturing industry in Norway (using a different methodology with different types of instruments), which increases confidence in our results.

We are particularly interested in deriving the marginal cost of commuting, i.e. the marginal willingness to pay for a (one-way) commuting distance of 1 km, using (1.2). The results for MCC are shown in the panel below the estimated coefficients. To improve interpretation, we focus here on a one-way commuting distance increase of 12 km, as this is equal to the mean commuting distance for women, in our data, as a percentage of annual wage.

Our headline results, using (1.2) and our estimates of column [1] of Table 1.2, demonstrate that for men, irrespective of whether or not they have a child, and women before having a child, the MCC given a 12 km increase of commuting distance is about 5-7% of the wage.

When having a child, the MCC given a 12 km increase of one-way commuting distance is substantially higher for women and equal to 16% of the wage. The latter finding is in line with the idea that (full-time) women with children often have more childcare and household responsibilities than men, hence their marginal dis-utility

Table 1.2: Job mobility (IV estimates)

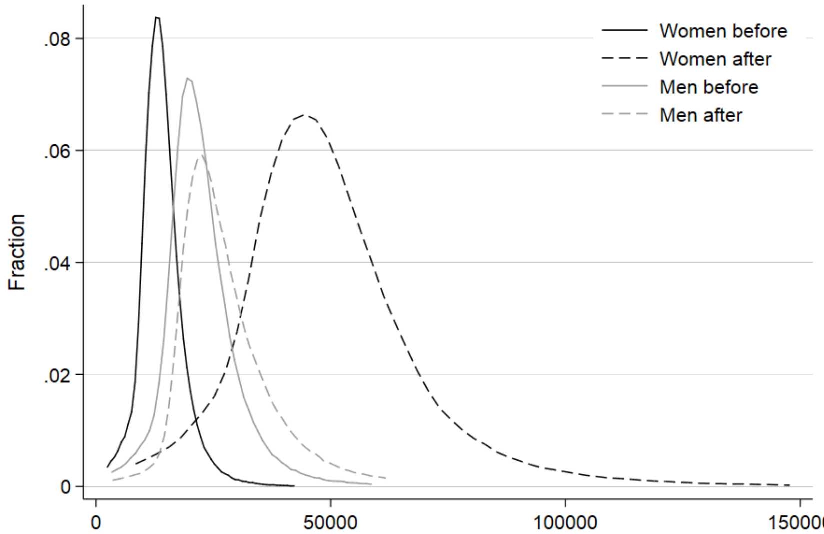
Dependent variable: Job change		[1]	[2] Women	[3] Men	[4] No anticipated residence move	[5] No anticipated childbirth
Distance (km)						
Women	no child	0.0007***	0.0008***		0.0006***	0.0014***
		(0.0001)	(0.0001)		(0.0002)	(0.0001)
	child	0.0025***	0.0025***		0.0027***	0.0025***
		(0.0001)	(0.0001)		(0.0001)	(0.0001)
Men	no child	0.0010***		0.0012***	0.0012***	0.0006***
		(0.0001)		(0.0001)	(0.0002)	(0.0001)
	child	0.0010***		0.0009***	0.0009***	0.0012***
		(0.0001)		(0.0001)	(0.0001)	(0.0001)
Log. wage		-0.182***	-0.125***	-0.215***	-0.117***	-0.173***
		(0.022)	(0.036)	(0.020)	(0.028)	(0.023)
First stage results						
Average wage at firm		0.126***	0.097***	0.151***	0.142***	0.124***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
F statistic for IV		6,640	2,005	7,773	4,865	6,075
Controls		yes	yes	yes	yes	yes
Worker fixed effect		yes	yes	yes	yes	yes
No. of observations		2,243,915	1,136,004	1,107,911	1,436,361	2,086,894
Marginal cost of commuting (% of annual wage) per 12 km increase (1 std. dev.)						
Women	no child	-0.047	-0.075		-0.059	-0.100
		(0.009)	(0.025)		(0.020)	(0.015)
	child	-0.164	-0.238		-0.273	-0.172
		(0.022)	(0.070)		(0.068)	(0.025)
Men	no child	-0.065		-0.067	-0.125	-0.043
		(0.011)		(0.008)	(0.035)	(0.011)
	child	-0.069		-0.051	-0.090	-0.085
		(0 .009)		(0.006)	(0.023)	(0.012)
Average job change		0.1675	0.1637	0.1714	0.1623	0.1698

Notes: The sample consists of full-time workers. We exclude observations in the year before the birth, the year of the birth, as well as the year after the birth. All specifications include the following controls: a child indicator, marital status, job tenure in linear and squared form, number of workers in the firm, the average age of workers at the firm, and year controls. Log wage is instrumented using the average wage of similar workers of the same firm. MCC is estimated using the ratio of the marginal effect of commuting distance on job mobility and the marginal effect of log wage on job mobility, see equation (1.2). Standard errors are clustered at the firm-year level and can be found in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

of commuting will be higher. Clearly, the estimated marginal costs of commuting are very similar for different model specifications shown in Table 1.2.

Our assumption that utility is additive in the logarithm of wages and commuting

Figure 1.4: Distribution of the Marginal Cost of Commuting (MCC)



Notes: The marginal cost of commuting MCC in % of annual wages per 12 kilometers has been computed using the estimated coefficients from model [1] in Table 1.2 and the observed distribution of annual wage.

implies that MCC is proportional to the wage, see (1.2), hence our estimate implies that there is a distribution of marginal commuting costs. Figure 1.4 shows the estimated distributions of the *annual* marginal commuting costs per 12 kilometers (in DKK), using the estimated coefficients from model [1] in Table 1.2 and the distribution of annual wage. It shows that the MCC distributions are very similar for men before and after they have children, with a mean of about 22,400 DKK and 27,300 DKK, respectively. For women, the distributions before and after having a child are quite different: before the birth, the mean is about 14,100 DKK, whereas after the birth of the child, the MCC distribution for women shifts to right with the mean of about 49,700 DKK.

We show now that in all other specifications (including ones not presented), women with children have a much higher MCC. One alternative specification is motivated by the labor economics literature, where it is hypothesized that employers have more monopsony power over women than men suggesting that the wage effect may differ between men and women, which has not been substantiated by empirical research (Manning, 2003a; Barth and Dale-Olsen, 2009). To allow for this possibility, we re-estimate the model separately for women and men, see columns [2] and [3]. We note that the effects of commuting distance are essentially identical, but that the effect of wage on job mobility might be gender specific. A standard t-test of gender differences just rejects the null hypothesis of equality of wage effects at the 5% significance level (the t-value is equal to 2.1). This finding supports the hypothesis by Manning (2003a) that women are less sensitive to wage increases.

However, more importantly for the current study, the main result that the loss in MCC is substantially higher for females with children is also supported by the specification. The calculated effects for the marginal costs of commuting are hardly affected for men (compare columns [1] and [3]), whereas for women we even find somewhat larger estimates (compare columns [1] and [2]). In the remainder of the paper, we continue assuming that there are no gender differences in wage effects, as this provides more conservative estimates.

In the last 2 columns of the table, we estimate models for more selective samples. In Section 1.4.1, we have explained that the interpretation of the marginal cost of commuting as defined by (1.2) somewhat changes if workers expect to change residence or have a child after accepting a new job because the estimate refers then to the expected marginal commuting costs, $\mathbb{E}[MCC]$ which may differ from the MCC.

As is common in revealed preference studies, we do not observe the expectations of households. However, we can investigate this issue by making additional assumptions. We assume that households are completely myopic about the far future, defined here as more than three years. If one then excludes observations of households for an interval of 3 years before a residential move or excludes observations of households for the interval of three years before they have a child, then one essentially has samples of households who do not expect to move residence or expect to have a child.

In column [4], we exclude observations of households before a residential move, so we focus on a subsample of workers who arguably didn't anticipate moving residence. We find that the estimates of commuting distance are hardly affected despite removing a substantial share of the data. The effect of the wage is somewhat smaller now (in absolute sign), suggesting that the current MCC exceeds the $\mathbb{E}[MCC]$. This makes sense, as, by moving residence, it is possible to reduce the current cost of commuting. In contrast, in column [5], we find that the estimates of the wage effect are very similar, but the effect of the commuting distance of workers currently without children has changed. It appears that the MCC of women who do not expect to have children is somewhat higher, but still substantially below the marginal cost of commuting for women with children. One possible reason why we find that the MCC for women who do not expect children exceeds the $\mathbb{E}[MCC]$ for women who do expect children, is that the latter does not expect to commute during the maternity leave period, which is typically 12 months in Denmark. This interpretation is supported by Figure 1.3 which shows that women just before they get pregnant are hardly sensitive to the length of the commuting distance.

In this study, we estimate commuting costs using commuting distance. The main advantage of the latter measure compared to an alternative measure used in the literature, commuting time, is that distance does not depend on the mode of transport, which is endogenously chosen. However, it also has a disadvantage as it does not directly give insight into the marginal cost of commuting *time* (rather than distance), which may be either expressed in terms of (leisure) time lost or in monetary terms, which are also useful measures.

To calculate the marginal cost of commuting time, we have to make additional assumptions. We assume that workers commute each day back and forth between the residence and the workplace (without combining these trips with other trips, e.g. dropping children at school, which may reduce the effective commuting time) and assume that the number of hours worked per day is 7.4 for full-time workers (in line with other studies). Furthermore, we need to have information about the effect of a marginal increase in commuting distance on commuting time. To derive the latter, we use the Danish National Travel Survey (NTS), which provides information on the commuting behavior of about 80,000 randomly selected individuals who fill out a one-day travel diary.

For the population of young workers we are interested in, the marginal effect of distance on one-way commuting time (in hours) is about 0.025, see Appendix 1.A.[20] This estimate implies, given i.e. a 40 km increase, the (one-way) commuting time increases exactly by one hour, which makes sense. It follows that the marginal effect of distance on *daily* commuting time is about 0.050. The implied MCC for one hour of commuting per day before the childbirth is then 52% of the hourly wage for women without children and for men.[21] For female workers with children, the MCC for commuting time is substantially higher, about 1.25 times the hourly wage, i.e. it exceeds the hourly wage. We have assumed that workers commute each day. Note that given the, maybe more plausible, assumption that workers do not commute to work one day a week, e.g. because of working from home or because of a business trip, then the MCC for commuting time is about 25% higher.

How do these estimates compare with the literature? Note that in most previous studies (Ophem, 1991; Van Ommeren and Fosgerau, 2009; Manning, 2003b), com-

[20]According to the speed literature, the effect of travel distance on travel time is diminishing, because the marginal increase in travel time is less for longer distances, see, for example, Couture et al. (2018). In line with that, we estimate the marginal effect of distance on travel time using a log-log specification, see Table A.2 in Appendix 1.A. We find a coefficient of 0.58, almost identical to the estimates reported for the United Kingdom by Van Ommeren and Dargay (2006). For this specification, the average marginal effect is equal to the product of the estimated coefficient and the average inverse speed (the ratio of travel time and travel distance). Given an estimate of 0.58 (see Table A.2) and an average inverse speed of about 0.043 (see Table A.1), it appears that the average marginal effect is 0.025.

[21]Given our estimates of column [1] of Table 1.2, the MCC (per km) is about 0.0034 (-0.0010/0.294) of the daily wage. The MCC for one hour of commuting per day is then 0.068 (0.0034/0.050) of the daily wage, as the marginal effect of distance on daily commuting time is 0.050. Given the typical number of hours worked per day (7.4), the MCC for one hour of commuting per hour worked is exactly half the hourly wage (7.4 * 0.068=0.5).

muting time rather than commuting distance was used as a proxy for commuting costs, so one can only compare with our implied commuting time estimates. Nevertheless, it appears that our implied estimates of MCC for commuting time are *substantially* less than the estimates obtained in those studies (about a factor two). One explanation is that it is plausible that the estimated coefficients of log wage were downward biased in those studies. Another explanation is that we have a sample of young workers, which is in line with our finding that the MCC appears to be higher for older workers with children as indicated by our estimates in column [2] of Table 1.2.

The only study we are aware of that also uses distance (Van Ommeren et al., 2000), finds roughly the same point estimate, but the confidence interval of this estimate is very large, so their point estimate must be interpreted as suggestive. Important for the current study which focuses on the role of children and gender, the current study is the first study that can differentiate between the MCC for men and women and demonstrates the importance of the presence of children with precisely estimated point estimates.

1.5.4 Sensitivity analysis

We have performed several sensitivity analyses of our preferred specification [1] of Table 1.2. First, we have also applied an event time methodology, where we let the distance coefficients vary per year. Second, we focus on the non-linear effects of distance. Third, we examine the importance of additional firm-level controls to examine the robustness of using our firm-level instrument. Fourth, we examine a range of alternative specifications, including where we control for household fixed effects.

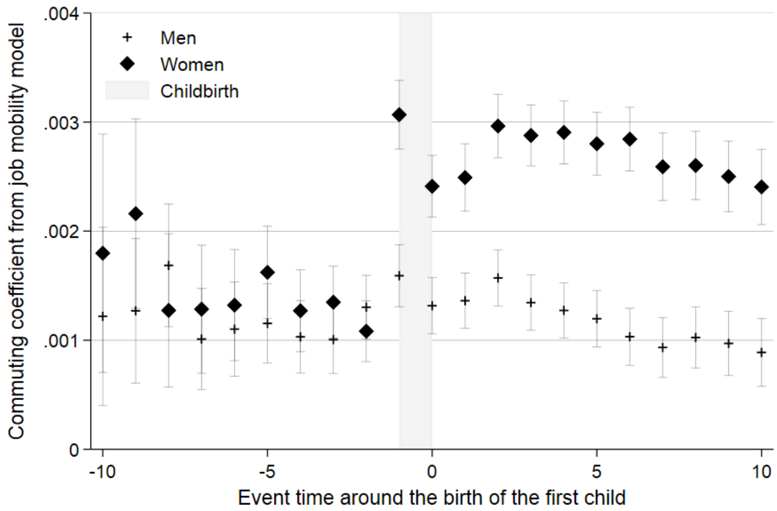
Event time methodology results

In the previous analyses, we have assumed that the estimated distance coefficients discretely jump after the birth of the first child, implicitly assuming that these coefficients do not vary over time otherwise. To investigate this further, we also estimate models that exploit an event time methodology, i.e. we re-estimate our preferred specification, but we allow the (gender-specific) distance coefficients $\alpha_{g,j}$ to vary over time, i.e. these coefficients vary by year j relative to the event of the birth. Consequently, we essentially estimate:

$$J_{i,t} = \sum_j \alpha_{g,j} \cdot x_{i,t} + \beta \cdot log(w_{i,t}) + \gamma \cdot X_{i,t} + \lambda_i + \kappa_t + \varepsilon_{i,t}, \tag{1.4}$$

where we instrument $log(w_{i,t})$.

Figure 1.5: Job mobility: commuting distance coefficients



Notes: Estimated coefficients of commuting distance on job mobility around the birth of the first child when including worker fixed effects and other controls.

In Figure 1.5, we show the estimated distance coefficients for men and women around the year of birth. It clearly shows that the coefficients for men are very similar for the different years before and after the event. In addition, the coefficients of women are indistinguishable from the male coefficients before childbirth but jump discretely after childbirth. Consequently, we believe that the jump in the coefficients for women when they have a child supports our methodology, and therefore our findings. In contrast, there are no good reasons to believe that the effect of wages will also jump discretely around birth. In line with this, when we test for a discrete jump, we do not find any evidence for such an effect (see Figure D.1 in Appendix 1.D).

Non-linear distance effects

We have also investigated whether the distance on job mobility is linear, see Table B.1 in Appendix 1.B. It appears that linearity is a reasonable assumption for our data. For example, when we impose that all distance effects are not gender-child specific and we include the square and the cube of distance, then the latter two terms are statistically insignificant. We have also estimated piecewise linear distance specifications with two knots (at 10 and 20 km), i.e. we estimate separate (gender-child specific) coefficients for short, medium, and long distances. In this case, the distance coefficients are very similar. When we estimate the same model for the different gender-child samples, the coefficients suggest, that linearity cannot be rejected for males (using a standard F test).

Additional firm-level control variables and selections

In our IV approach, we use as an instrument the average wage (of similar workers) within the firm. In these estimations, we control for firm size as well as the average age of the workers belonging to the firm to avoid the criticism that the average wage has a direct effect on individual wages. Nevertheless, one criticism of the estimation procedure is that we do not control sufficiently for firm characteristics, including non-wage amenities, which may invalidate the instrument if these firm characteristics are correlated to the instrument and affect job mobility directly.

To address this issue, we have estimated model specifications with two types of firm-level control variables. First, we add gender-neutral controls: more detailed sector controls, average education shares, and region dummies which aim to control for the confounding bias of unobserved firm characteristics. Second, we control for the share of female workers and a proxy for family friendliness, which aims to control for the confounding bias of unobserved characteristics that tend to be appreciated by female workers with children. For simplicity, we impose that the effect of distance does not vary by gender and child, which is not essential, because we focus on the effect of wage. These results are shown in Table B.2 in Appendix 1.B. Arguably, controlling for sectors is potentially important, as it has been known for many years that wages are structurally higher in certain sectors, whereas there are also substantial job mobility differences between sectors. In case the sectoral wage differences and sectoral job mobility differences are correlated with each other, then the instrument would be invalid. This suggests that controlling for sectors is essential. To address this, we add additional controls for sectors at NACE 2 (88 sectors) and even NACE 3 (272 sectors) levels, respectively, as shown in columns [2] and [3]. The effects of commuting distance remain the same, whereas the effect of wage is slightly less pronounced. Consequently, the estimates are rather insensitive to sector controls, even when we control for sector in a very detailed way.

Similarly, adding controls for the share of workers with a certain educational level or the share of female workers results in almost identical results (see columns [4] and [5]). Column [6] shows that including regional fixed effects (5 regions) does not affect the estimation results. Finally, in column [7], we follow Kleven et al. (2019a) and include a proxy variable to measure the family friendliness, which is based on whether the management team includes women with young children (under 15 years of age). This allows us to proxy many non-wage amenities, such as tolerance for taking sick days off, flexible working hours, and the option of working remotely. Again the overall estimation results remain unchanged. In conclusion, it appears that for all these additional specifications, the effects of commuting distance and wage are robust, reducing the likelihood that our effects are confounded by unobserved non-wage amenities.

Alternative specifications

We have also estimated a range of alternative specifications. First, we have estimated a specification where we add household fixed effects, so we additionally control for unobserved time-invariant household characteristics. This essentially means that we identify the effects of interest by comparing the behavior of men and women within the same household (i.e. a husband and wife). The results reported in Table B.3 of Appendix 1.B demonstrate that the MCC results are almost identical by including these additional controls.

Second, we have investigated the robustness of the results using several other specifications that appear in the literature (no worker fixed effects, household fixed effect rather than worker fixed effects, no instrumenting of wage), see again Table B.3 in Appendix 1.B.

We start with a specification where we do *not* control for worker fixed effects, but replace these fixed effects with a range of control variables including age, gender, and education, see column [3]. In this case, it appears that the effect of distance is robust. In contrast, although the instrument is very strong (with the first-stage coefficient of about 0.56), it appears that there is a positive effect of the wage on job mobility, which doesn't make sense from an economic point of view. Clearly, the instrument is invalid without worker-fixed effects, because of worker sorting. Then, we show a specification where we do not control for worker-fixed effects but replace these fixed effects with household fixed effects, see column [4]. It appears that the estimated effects of wages are quite different. This reinforces our previous conclusion that the average wage is only valid as an instrument given worker fixed effects. Again, the effects of commuting distance remain robust.

In column [5], we show a specification where we do not instrument the wage. It appears now that the estimated effect of wage is about 5 times lower, suggesting that workers are hardly sensitive to wage increases. This makes sense because the latter specification doesn't account for that a wage increase also shifts the wage offer distribution for the worker. Again, we find that the effects of commuting distance remain the same. Hence, in conclusion, it appears that the effects of commuting distance are robust to the methodology used, whereas the effect of wage is *not* and depends on the methodology used. It appears essential not only to use worker-fixed effects but also to instrument the wage. Finally, column [6] includes controls for occupational rank and management status. Note that these controls are potentially endogenous, as changes in occupational rank are frequently closely linked to changes in wages. Nevertheless, also controlling for these factors, the results remain unchanged.

Third, we have examined specifications allowing the wage effect to be different for workers with or without a child. The estimates in Appendix 1.D imply that workers are less responsive to wages after having a child, but the estimates for the MCC, which is our main interest, are almost identical, except for females with a child for which the MCC is even somewhat higher than in previous specifications. Consequently, if anything, these results reinforce our main conclusion that the MCC is substantially higher for female workers with a child.

1.6 Wage and commuting

1.6.1 Hedonic wage regression

We have shown that women with children bear higher costs for commuting. This raises the question of whether they are partially compensated through higher wages. Arguably, wage compensation for commuting may be different for women with children, because women with children tend to sort into jobs at different sectors, firms, and workplace locations (Blau and Kahn, 2017; Timothy and Wheaton, 2001). It may also be different because they face higher costs of commuting, and therefore make different trade-offs between wages and commuting distance when searching for jobs (Manning, 2003b; Manning and Petrongolo, 2008). Alternatively, it may be different, because of differences in bargaining power for other forms of monopsony power of the firms they work for (Barth and Dale-Olsen, 2009), or because the relationship between commuting distance and productivity is different for women with children (e.g. women are more likely to take public transport which increases the likelihood of delays or falling ill).

To estimate the compensating wage differentials for commuting, we apply a standard hedonic wage regression with the logarithm of wage $log(w_{i,h,t})$ of worker i of household j at time t as the dependent variable, where we include a gender-child-specific effect of the logarithm of commuting distance, $log(x_{i,h,t})$:

$$log(w_{i,h,t}) = \beta_{g,c} \cdot log(x_{i,h,t}) + \gamma \cdot X_{i,h,t} + \lambda_i + \kappa_t + \eta_{h,t} + \varepsilon_{i,h,t}, \tag{1.5}$$

where we include worker fixed effects, λ_i, household by residential location fixed effects, $\eta_{h,t}$, year fixed effects, κ_t, and a range of household and labor market control variables $X_{i,h,t}$ (a gender and child interaction term, family status, job tenure, sector, and firm size), and where $\varepsilon_{i,h,t}$ is a standard idiosyncratic error term.

In this setup, we include worker-fixed effects to control for education level and innate ability differences that tend to be correlated to the length of the commute, so we address time-invariant omitted variable bias. In addition, we control for household-by-residential-location fixed effects, where residential location is defined by the household's parish.[22] The inclusion of household-by-residence-location fixed effects improves identification by dealing with reverse causality from endogenous residence location, where the length of the commuting trip depends on the household income level (Wheaton, 1974; Fujita, 1989; Lucas and Rossi–Hansberg, 2002; Zenou, 2009a; Redding and Rossi-Hansberg, 2017). The inclusion of individual and household-by-residence-location fixed effects implies that we essentially use the information on changes over time in the commuting distance of men and women who belong to the same household and live in the same parish and then examine to what extent wages change over time. We also control for a range of household control variables that may change over time (e.g. family status). Finally, we control for firm size and industrial sector to capture differences in fringe benefits that are potentially correlated to wages as well as the length of the commute.

We emphasize that the hedonic wage literature that aims at deriving the workers' willingness to pay for job attributes starts from the assumption that the labor market is fully competitive. Given the assumption of a frictionless labor market, our estimated compensating wage differentials would be equal to workers' marginal cost of commuting thus our estimate would reflect the causal effect of commuting distance on wages as we address omitted variable bias and reverse causality.

In contrast, our starting point is that the labor market is characterized by search frictions and, therefore, not competitive. This is fundamental for the interpretation of our results because frictions in the matching between workers and jobs imply that the workers' evaluation of these job attributes is *not* equal to the compensating wage differentials of these job attributes (Hwang et al., 1992; Mulalic et al., 2013; Mas and Pallais, 2017). This issue applies in principle to all job attributes, but in particular to commuting, as job search frictions are thought to be essential to explain commuting outcomes (Manning, 2003a; Le Barbanchon et al., 2021). Nevertheless, this does not mean that hedonic wage models are not useful in the context of commuting. Given the presence of search frictions, workers will accept job offers and therefore simultaneously accept a commuting distance as well as wage. Consequently, we interpret the estimates of the compensating wage differentials not as causal estimates but as correlations between wages and commuting distance that exist when workers move jobs given a job search process where workers choose to accept or reject job offers. Acceptance of a job offer implies a certain wage and commuting distance combination, rather than as causal effects of commuting on wages.

[22] A parish is an administrative area consisting of several villages or localities originating from the Middle Ages. Only a few alterations to the parishes were made since 1841. There are 2158 parishes in Denmark.

We use these models to estimate the compensating wage differentials for commuting and show that, in line with the above considerations, these wage differentials are below the worker's marginal cost of commuting, so workers, including women with children, are hardly compensated for the higher commuting costs through higher wages.

1.6.2 Results

We have estimated a range of hedonic wage specifications. Our first, and preferred, specification is based on (1.5), i.e. a specification with individual and household-by-residence-location fixed effects and a full range of controls, as reported in column [1] of Table 1.3.

The estimates reported in column [1] have several messages. First, they show that there is a positive relationship between commuting distance and wages for males and females, independent of whether they have a child or not. This result is in line with several economic theories, including those that allow for job search imperfections (Manning, 2003b) and theories that allow for spatial wage differences (Lucas and Rossi–Hansberg, 2002).

Second, the implied marginal effects tend to be small. For men with children, the estimated coefficient is equal to 0.0061. This implies that a 12 km increase in commuting distance (from 6 to 18 km), increases wages by about 0.7% (calculated as $0.0061 \times (log(18) - log(6))$. For women with children, the estimated coefficient is equal to 0.0019. Hence, a 12 km increase in commuting distance increases their wages by 0.25%. *Consequently, the compensating wage differentials for commuting are an order of magnitude less than the marginal cost for commuting.*[23] Our finding that the marginal cost of commuting, i.e. the marginal willingness to pay for commuting, far exceeds the compensating wage differentials is consistent with the notion that job search is essential to explain commuting behavior. In a world without job search imperfections, where workers have full control over the chosen commuting distance, workers would sort into workplace locations where they are fully compensated for longer commutes. We clearly do not live in such a world.

Third, although we have seen previously that women with children bear *higher* cost for commuting, their compensation for commuting does not exceed and is even *lower* than that of their male counterparts, implying an increase in the gender wage gap. We emphasize that the child-induced gender wage gap is rather small. For example, if we increase the commuting distance by 12 km both for men and women with children, then the gender wage gap increases by about 0.45%.

[23]The MCC for 12 km commuting was estimated to be 6% of the daily wage for men with children and 16% of the daily wage for women with children.

Table 1.3: Effect of commuting distance on income by gender and child

	[1] Househ. & indiv. FEs	[2] Indiv. FE	[3] No sector controls	[4] Average distance	[5] Average distance	[6] Firm FE	[7] Workplace mun. FE	[8] Cross section
Women								
No child	0.0032***	0.0030***	0.0043***	0.0026***	0.0032***	0.0017*	0.0026***	0.0033***
	(0.0009)	(0.0007)	(0.0009)	(0.0006)	(0.0006)	(0.0007)	(0.0007)	(.0008)
Children	0.0019*	0.0024***	0.0021**	0.0020***	0.0017***	−0.0008	0.0020***	0.0124***
	(0.0008)	(0.0006)	(0.0008)	(0.0005)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Men								
No child	0.0038***	0.0045***	0.0039***	0.0064***	0.0054***	0.0014*	0.0042***	0.0090***
	(0.0009)	(0.0007)	(0.0009)	(0.0006)	(0.0006)	(0.0007)	(0.0007)	(0.0008)
Children	0.0061***	0.0067***	0.0058***	0.0064***	0.0047***	0.0027***	0.0064***	0.0238***
	(0.0007)	(0.0006)	(0.0007)	(0.0005)	(0.0005)	(0.0006)	(0.0006)	(0.0006)
Time var.	yes	yes	yes	yes	yes	yes	yes	yes
Household FE	yes	no	yes	no	no	no	no	yes
Worker FE	yes	yes	yes	yes	yes	yes	yes	no
Sector	yes	yes	no	yes	yes	no	yes	yes
Firm FE	no	no	no	no	no	yes	no	no
Workplace FE	no	no	no	no	no	no	yes	no
Avg. distance	no	no	no	municipality	firm	no	no	no
No. of obs.	2,298,453	2,298,453	2,298,453	2,287,723	2,288,132	2,265,144	2,293,621	2,288,132

Notes: The Dependent variable is log wage and the main independent variable is log commuting distance. We exclude observations in the year before the birth, the year of the birth, as well as the year after the birth. Time-varying controls include a gender and child interaction term, family status, job tenure, and firm size. Household fixed effects refer to household-by-residential-location fixed effects. Average distance refers to additional average distance controls measured at the municipality as well as firm level. Workplace fixed effects are measured at the municipality of the firm. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

We have investigated the robustness of the results in several specifications. In column [2], we exclude household-by-residence-location fixed effects. We find that the results are almost identical. This result is important because it demonstrates that reversed causation is unlikely an important matter and sorting into different residential locations does not affect the compensation for commuting distance in the labor market, justifying labor market models that typically ignore this issue.

In column [3], we exclude sector controls. These controls were included for several reasons. One of the reasons is that we aim to control for unobserved fringe benefits (e.g., free employer parking, company car, and childcare) that are potentially correlated to commuting distance, and that may differ by gender. We find that excluding sector controls does not have any influence on the results, suggesting that not observing fringe benefits is unlikely to determine our results.

In the next two columns, we investigate the importance of wage compensation due to the location of the firm by adding additional controls. Theory suggests controlling for the commuting distance of the marginal worker at the firm, which we proxy in two ways (Timothy and Wheaton, 2001). In column [4], we control for the average commuting distance in the municipality where the firm is located. In the next column, we control for the average commuting distance of the firm. We find almost identical results. This implies that (explicit) wage compensation, because of the location of the firm does not play a major role.

In column [6], we include additional firm fixed effects. The inclusion of these fixed effects allows us to test the idea that the wage offer distribution is firm-specific. If that is the case, then the inclusion of firm fixed effects must strongly reduce the compensation for commuting. In line with that, we find that wage compensating levels are substantially less, and even strictly zero for women with children. In the last column, we include additional workplace municipality fixed effects. The inclusion of these fixed effects allows us to test the idea that the wage offer distribution is municipality-specific. In line with the results in column [3], where we have included the average distance, rather than the municipality-fixed effects, we find that the results are virtually unchanged. Consequently, our results suggest that wage compensation differentials arise because firms vary in the level of wages paid, but the latter is not systematically related to the location of the firm.

In column [8], we have investigated the importance of including worker-fixed effects. We have investigated this by excluding these fixed effects while adding two other, potentially important, control variables, education, and age, that are not-identified when including these fixed effects. It appears that the results qualitatively hold if we exclude these fixed effects, but the sizes of the estimates are quite different, suggesting that one grossly overestimates the compensation when excluding worker fixed effects. Note that we still include household-by-residence-location fixed effects here, so these results imply that including household fixed effects while excluding worker fixed effects, is not sufficient to deal with omitted variable bias, an issue discussed in the literature (Manning, 2003b). When one also excludes household-by-residence-location fixed effects, as used to be common in the older cross-sectional literature, one finds that the bias in the estimates becomes even more pronounced.

Our empirical analysis of job mobility is based on a theoretical framework that assumes that firms post wages, such that workers face a distribution of wages while searching for another job. A potential criticism of this framework is that workers do not only search for wage offers that are posted but also bargain about wages with firms. To differentiate between bargaining and the distribution of wages explanation is difficult (and not key to our paper). One attempt is to focus on workers who do not change employers, as for these workers changes in wages are more likely due to bargaining than due to inter-firm wage dispersion. Using this idea, we have estimated the effect of commuting distance on wages using the methodology introduced by

Mulalic et al. (2013), which uses firm relocation as a cause of exogenous variation in commuting distance for workers who stay with their firm. We do not find any evidence that wage bargaining related to commuting distance plays a role (Table E.1 col. [1]-[4]).[24]

Furthermore, we have investigated the functional form of commuting distance. It appears that the marginal effect of distance (at the mean commuting distance) is identical for linear, as well as for higher-order polynomial, and logarithm specifications, but at the same time supports the choice of using the logarithm rather than for example a linear specification.[25]

1.7 Conclusion

A large literature shows that the gender wage gap strongly increases after the birth of the first child. We provide complementary analyses of the role of the birth of the first child on gender differences in commuting distance as well as in preferences for commuting distance using administrative register data for the full working population in Denmark.

Employing childbirth as an event for identification, we demonstrate that women with children are much more likely to leave their jobs when they have a long commute – the marginal effect of distance on job mobility is about 3 times higher – which is not true for their male counterparts with children. Furthermore, we apply an IV approach to estimate the effect of wages on job mobility. Employing a dynamic search model, these results imply that the marginal cost of commuting increases substantially for women after the birth of the first child. A 12-kilometer increase in commuting distance induces costs equivalent to about 16% of wages for women with children. Consequently, women with children bear a higher cost of commuting. At the same time, compensating wage differentials for commuting are low for all workers, also for women with children, implying that women with children are not compensated for their higher commuting costs through higher wages.

Our findings are consistent with the notion that gender differences in the costs of commuting are important as argued, for example, by Le Barbanchon et al. (2021). A subtle, but important, contribution here is that we show that these gender differences are only important when children are present.

[24] In contrast to this finding, we find evidence for bargaining for a sample of older workers that are excluded in the current study. These results can be found in Table E.1 col. [5]-[7].

[25] We have also investigated the importance of compensation for commuting for the overall gender pay gap for these polynomial specifications using a decomposition methodology, as introduced by Blinder (1973) and Oaxaca (1973). This decomposition method cannot be applied to the results shown in Table 3, where we use the logarithm of the distance, because it requires that the effect at distance zero is defined. These results confirm the results discussed in our baseline specification.

Appendix

1.A Marginal effect of distance on commuting time

We use the Danish National Travel Survey (NTS) to estimate the marginal effect of distance on commuting time. The NTS provides information on the travel behavior of randomly selected individuals who fill out a one-day travel diary. Information is collected continuously throughout the year. We use NTS for the years 2006-2019 and select individuals (18-70 years old) who report commuting trips and exclude observations with missing information and observations for which the one-way commuting distance exceeds 108 km (99 percentile), the one-way commuting time exceeds 95 minutes (99 percentile), or the average commuting speed is below 3.6 km/h (1 percentile) or above 79.5 km/h (99 percentile). Given these selection criteria, we exclude 6.7% of commuting trips. Our final sample includes 81,577 commuting trips.

Table A.1: Descriptive statistics for Danish national travel survey

	All commuters		Comm. 25-45 years	
	Mean	Std. dev.	Mean	Std. dev.
Trip length (km)	14.20	15.40	14.60	15.58
Trip time (minutes)	21.35	16.51	21.73	16.44
Trip speed (km/h)	35.82	20.03	36.35	20.12
Trip inverse speed (h/km)	0.045	0.040	0.043	0.038
Car (share)	0.65	0.48	0.65	0.48
Public transport (share)	0.09	0.20	0.09	0.29
Walking (share)	0.04	0.04	0.04	0.19
Bicycle (share)	0.21	0.41	0.22	0.41
Male (share)	0.49	0.50	0.49	0.50
Age (year)	43.43	12.08	36.50	5.77
Number of obs. (commuting trips)	81,577		37,524	

Table A.1 provides descriptives. On average, the one-way commuting time is 21 minutes, the one-way commuting distance is about 14 km and the speed is 36 km/h. The mean inverse speed is 0.045. The most popular commuting mode is the car (65%), while only 9% of workers commute by public transport. Bicycle use is very common: more than 21% of workers commute by bicycle. For the sample of workers between 25-45 years, which is the relevant population for our paper, the descriptive statistics are almost identical.

According to the speed literature, the effect of distance on travel time is diminishing, because the marginal increase in travel time is less for longer distances, see, for example, Couture et al. (2018). In line with that, when we regress travel time on travel distance, we use a log-log specification, see Table A.2. In the first model [1], we find a coefficient of 0.58, slightly higher than the value reported for the United Kingdom by Van Ommeren and Dargay (2006). When we estimate the models for the sample of workers between 25-45 years, the estimated coefficients are almost identical, see column [2]. Finally, we re-estimate the latter model separately for women and men, see columns [3] and [4]. Again it appears that the coefficient is about 0.58.

Table A.2: Travel distance and travel time

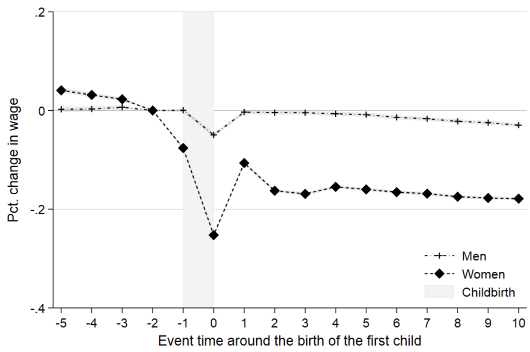
	All commuters	Commuters 25-45 years		
			Man	Woman
Dep. variable	log(time)	log(time)	log(time)	log(time)
	[1]	[2]	[3]	[4]
log(distance)	0.5792***	0.5846***	0.5945***	0.5771***
	(0.0013)	0.0019	(0.0027)	(0.0027)
const.	-2.5231***	-2.5354***	-2.5755***	-2.5028***
	(0.0030)	0.0045	(0.0066)	(0.0062)
R-squared	0.7197	0.7200	0.7316	0.7086
Number of obs.	81,577	37,524	18,443	19,081

Notes: Standard errors are in parentheses, *** $p < 0.01$.

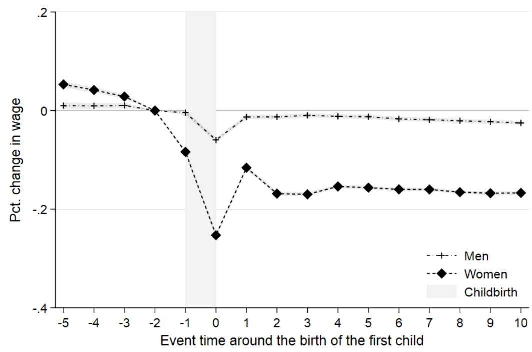
We are interested in the marginal effect of commuting distance (measured in km) on commuting time (measured in hours). Given a log-log specification, the average marginal effect is equal to the product of the estimated coefficient and the average inverse speed (the ratio of travel time and travel distance). Given an estimate of 0.58 (see Table A.2) and an average inverse speed of about 0.045 and 0.043 respectively (see Table A.1), it appears that the mean marginal effect is 0.026 for the full sample and 0.025 for the sample of commuters 25-45 years, receptively.

1.B Job mobility and the birth of the first child

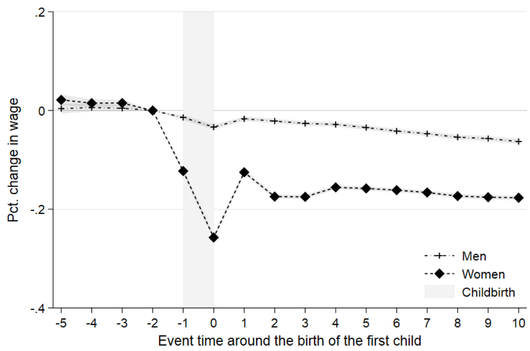
Figure B.1: Event study results for different samples: wage



(**a**) Job fixed and residence flexible



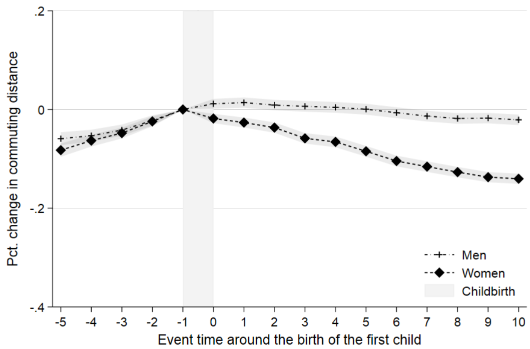
(**b**) Job flexible and residence fixed (t=-3)



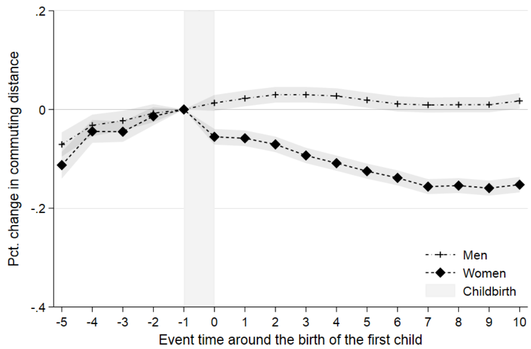
(**c**) Job and residence fixed

Notes: Wage event time effects around the birth of the first child. The gray area marks the time interval of the birth of the first child. The shaded 95 percent confidence intervals are based on robust standard errors.

Figure B.2: The Event study results for different samples: commuting distance



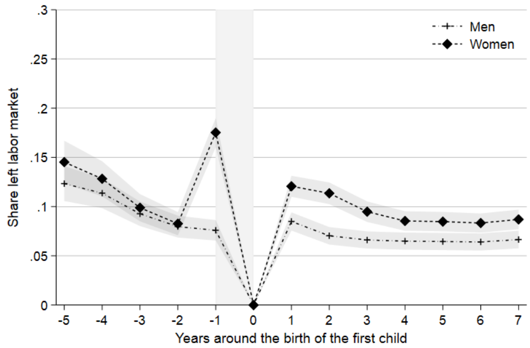
(a) Job fixed (from t=-3)



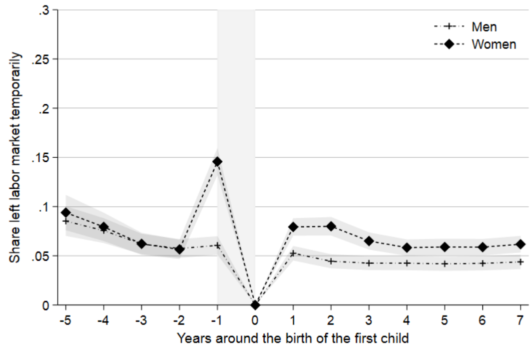
(b) Residence fixed (from t=-3)

Notes: Commuting distance event time effects around the birth of the first child. The grey area marks the time interval of the birth of the first child. The shaded 95 percent confidence intervals are based on robust standard errors.

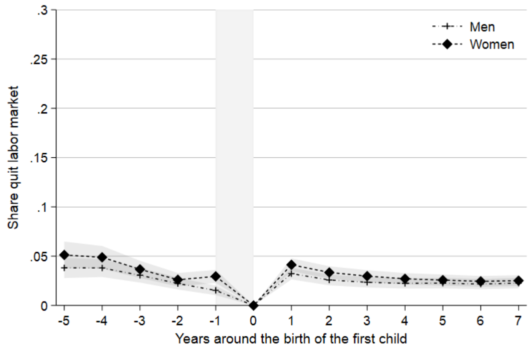
Figure B.3: Share of sample quits and part-time selection



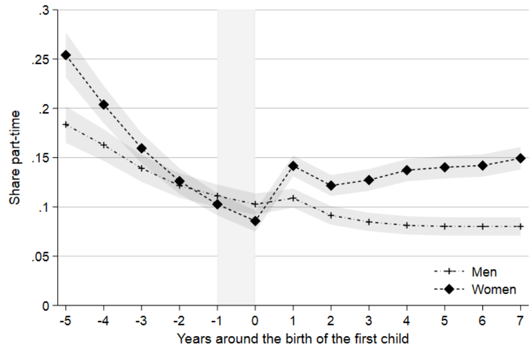
(a) Share of workers leaving the final sample



(b) Share of workers leaving - returning



(c) Share of workers leaving - not returning



(d) Shares part-time before full-time selection

Notes: The gray area marks the birth of the first child and the observation of the birth is the seed of our sample, meaning we observe everyone at t=0. The part-time shares have been calculated before the selection on full-time has been made, and the share of sample quits is calculated on the final sample.

Table B.1: Linear probability job mobility model

	[1] Overall	[2] Women no child	[3] Women child	[4] Men no child	[5] Men child	[6] Polynomial
Spline						
<10km	0.0013***	-0.0003	0.0037***	0.0007	0.0017***	
	(0.0002)	(0.0006)	(0.0004)	(0.0006)	(0.0005)	
10km-30km	0.0010***	0.0002	0.0024***	0.0005**	0.0010***	
	(0.0001)	(0.0003)	(0.0002)	(0.0003)	(0.0002)	
30km-50km	0.0011***	0.0024***	0.0024***	-0.00001	0.0004	
	(0.0002)	(0.0006)	(0.0005)	(0.0005)	(0.0004)	
Polynomial						
distance						0.0012***
						(0.0004)
distance^2						-0.0001
						(0.00002)
distance^3						0.0000001
						(0.0000003)
Number of obs.	2,507,138	353,792	907,456	403,204	842,686	2,507,138
F test for spline	0.47	6.29	3.05	0.41	0.70	
R^2	0.054	0.084	0.069	0.067	0.057	0.054

Notes: The sample consists of full-time workers. All specifications include the following controls: a gender and child interaction term, marital status, job tenure in linear and squared form, number of workers in the firm, average age of workers at the firm, as well as year controls. Log wage is instrumented using the average wage of similar workers of the same firm. MCC is estimated using the ratio of the marginal effect of commuting distance on job mobility and the marginal effect of log wage on job mobility, see equation (1.2). Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.2: Job mobility models (2SLS): Additional firm-level control variables and selection

Dependent variable: Job change	[1] Main specification	[2] Sector FE NACE 2	[3] Sector FE NACE 3	[4] Education	[5] Share men	[6] Regional FE	[7] Family friendliness
Distance (km)	0.0018***	0.0018***	0.0018***	0.0018***	0.0018***	0.0018***	0.0018***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Log wage	-0.241***	-0.214***	-0.197***	-0.185***	-0.209***	-0.245***	-0.241***
	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)	(0.022)	(0.022)
First stage results							
Average wage at company	0.127***	0.123***	0.120***	0.123***	0.123***	0.125***	0.127***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
F statistic for IV	6,811	6,528	6,299	6,355	6,348	6,716	6,812
Controls	yes	yes	yes	yes	yes	yes	yes
Worker FE	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes
No. of observations	2,243,915	2,243,915	2,243,915	2,243,915	2,243,915	2,243,915	2,243,915
Marginal cost of commuting (% of annual wage) per 12 km increase (1 std. dev.)							
	-0.089	-0.102	-0.109	-0.116	-0.103	-0.088	-0.089
	(0.009)	(0.011)	(0.013)	(0.015)	(0.012)	(0.008)	(0.009)
Average job change	0.1675	0.1675	0.1675	0.1675	0.1675	0.1675	0.1675

Notes: The sample consists of full-time workers. All specifications include the following controls: marital status, job tenure in linear and squared form, number of workers in the firm, average age of workers at the firm, and year controls. Specification [7] in addition also includes control for family friendliness, i.e. whether the firm's management includes women with young children (below 15 years of age). Log wage is instrumented using the average wage of similar workers of the same firm. *MCC* is estimated using the ratio of the marginal effect of commuting distance on job mobility and the marginal effect of log wage on job mobility, see equation (1.2). Standard errors are clustered at the firm-year level and can be found in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.3: Alternative specifications of the job mobility model

Dependent variable: Job change	[1] 2SLS Main specification	[2] 2SLS Worker and household FE	[3] 2SLS Cross section	[4] 2SLS Household FE	[5] OLS Worker FE	[6] 2SLS Occupational rank
Distance (km)	0.0018***	0.0020***	0.0012***	0.0020***	0.0017***	0.0018***
	(0.0001)	(0.0001)	(0.00003)	(0.0001)	(0.0001)	(0.0001)
Log wage	-0.241***	-0.228***	0.022***	-0.029**	-0.053***	-0.226***
	(0.022)	(0.026)	(0.005)	(0.007)	(0.002)	(0.024)
First stage results						
Average wage	0.127***	0.112***	0.561***	0.372***	-	0.123***
at company	(0.002)	(0.002)	(0.002)	(0.002)	-	(0.002)
F statistic for IV	6,811	5,052	100,722	36,005	-	5,664
Controls						
Time variant	yes	yes	yes	yes	yes	yes
Age, educ.	no	no	yes	yes	no	no
Worker FE	yes	yes	no	no	yes	yes
Household FE	no	yes	no	yes	no	no
Year FE	yes	yes	yes	yes	yes	yes
No. of obs.	2,243,915	2,243,915	2,243,915	2,243,915	2,243,915	2,243,915
Marginal cost of commuting (% of annual wage) per 12km increase (1 std.)						
	-0.089	-0.102	0.671	-0.797	-0.396	-0.095
	(0.009)	(0.012)	(0.142)	(0.195)	(0.021)	(0.011)
Average job change	0.1675	0.1675	0.1675	0.1675	0.1675	0.1675

Notes: The sample consists of full-time workers. Time variant controls include our standard set of controls: marital status, job tenure in linear and squared form, number of workers in the firm, average age of workers at the firm, and year controls. Occupational rank includes general managers (CEOs), managers, workers at the non-managerial level, and employees (excluding young people and trainees). Log wage is instrumented using the average wage of similar workers of the same firm. MCC is estimated using the ratio of the marginal effect of commuting distance on job mobility and the marginal effect of log wage on job mobility, see equation (1.2). Standard errors are clustered at the firm-year level and can be found in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

1.C Marginal cost of commuting: theory

Comparative statics

We now derive the effect of α on the voluntary job-to-job rate, θ, i.e. the arrival rate of jobs which increases utility. To derive this effect, we introduce $\lambda(v^*)$ which defines the arrival rate of job offers that offer utility v^*. Workers will accept all job offers for which hold that $log(w^*) - \alpha x^* > log(w) - \alpha x$. Note that when the job is at distance x^*, then the (log) wage offer is $v^* + \alpha x^*$. This arrival rate can therefore be written as (space is two-dimensional and we, therefore, multiply the job offer density function with $2\pi x^*$):

$$\lambda(v^*) = \lambda \int_0^\infty f(v^* + \alpha x^*) 2\pi x^* dx^*. \tag{C.1}$$

We change the variable of integration to $log(w^*)$, so we get:

$$\lambda(v^*) = \frac{2\pi\lambda}{\alpha^2} \int_v^\infty (log(w^*) - v^*) f(log(w^*)) dlog(w^*). \tag{C.2}$$

Now consider a worker with a job offering $log(w^*)$ at a distance equal to x^*, i.e. a job that offers exactly utility v^*. This worker will accept all job offers v^* which exceed v. The job moving rate θ is then defined by:

$$\theta(w,x) = \int_v^\infty \lambda(v^*) dv^* = \frac{2\pi\lambda}{\alpha^2} \int_v^\infty \int_v^\infty (log(w^*) - v^*) f(log(w^*)) dlog(w^*) dv^*. \tag{C.3}$$

Equation (C.3) allows us to do comparative statics. Given (C.3), it is straightforward to see that the job moving rate depends negatively on v ($\partial\theta(w,x)/\partial v < 0$). Furthermore, v depends positively on wages while negatively on distance. Consequently, an increase in the current wage or a decrease in the length of the commute will result in a lower job moving rate, i.e. $\partial\theta(w,x)/\partial w < 0$ and $\partial\theta(w,x)/\partial x > 0$. Such a result is in line with intuition.

It also allows us to investigate how α affects the job moving rate. Given (C.3), it appears that v, the job moving rate is inversely proportional to the ratio of the arrival rate λ and the *square* of the marginal cost of commuting α. However, the effect of α on the job moving rate is ambiguous, as it reduces v. For workers with a short commute, an increase in α reduces the job moving rate, whereas for those with a long commute, an increase in α increases the job moving rate.

One can show this formally by differentiating $\theta(w,x)$ with respect to α:

$$\frac{\partial\theta(w,x)}{\partial\alpha} = -\frac{2\theta(w,x)}{\alpha} + \frac{\partial\theta(w,x)}{\partial x}\left(\frac{\partial v}{\partial x}\right)^{-1}\frac{\partial v}{\partial\alpha} = -\frac{2\theta(w,x)}{\alpha} + \frac{\partial\theta(w,x)}{\partial x}\frac{x}{\alpha}. \tag{C.4}$$

For x equals 0, the expression is negative, whereas for large values of x, the second term exceeds the first term, as the first term is bounded.

Deriving marginal cost of commuting

Given (C.3), one can write $\theta(w,x)$ as $\theta(v(w,x))$, so one may derive the marginal cost of commuting, MCC, as:

$$MCC \equiv -\frac{\partial v/\partial x}{\partial v/\partial w} = -\frac{\partial\theta(w,x)/\partial x}{\partial\theta(w,x)/\partial w} = -\frac{\partial\theta(w,x)/\partial x}{\partial\theta(w,x)/\partial log(w)}w = \alpha w. \tag{C.5}$$

Consequently, MCC can be estimated using the ratio of the marginal effect of commuting distance on job mobility and the marginal effect of log wage on job mobility.

1.D Gender-child-specific wage effects

We have estimated models where we also allow $\beta^{w}_{g,c}$ to vary by child and gender:

$$J_{i,t} = \alpha_{g,c} \cdot x_{i,t} + \beta_{g,c} \cdot log(w_{i,t}) + \gamma_{g,c} \cdot X_{i,t} + \delta_{g,c} + \lambda_i + \kappa_t + \varepsilon_{i,t}, \quad \text{(D.1)}$$

so we have 4 endogenous variables, which we instrument. We use 4 instrumental variables in the first stage (the average wage in the firm interacted with the group). The results are shown in Table D.1. We find that workers are somewhat less responsive to wages after having a child, but the MCC is almost identical, except for females with a child for which the MCC is higher than in previous specifications. Consequently, if anything, these results reinforce our main conclusion that the MCC is substantially higher for female workers with a child.

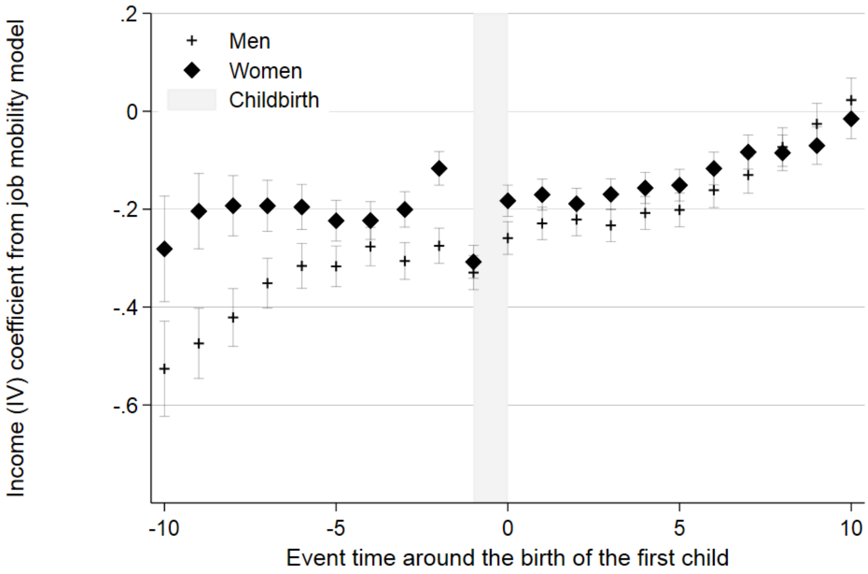
Table D.1: Linear probability job mobility model

		First stage		Second stage		
		Instrument	F-stat	Distance (km) coefficient	log wage coefficient	Calculated MCC
Women	no child	0.585	6,632	0.0008***	-0.275***	-0.037
		(0.003)		(0.0001)	(0.023)	(0.006)
	child	0.400	6,725	0.0024***	-0.115***	-0.254
		(0.002)		(0.0001)	(0.012)	(0.049)
Men	no child	0.599	7,359	0.0011***	-0.275***	-0.049
		(0.003)		(0.0001)	(0.024)	(0.006)
	child	0.408	7,883	0.0010***	-0.117***	-0.010
		(0.002)		(.0001)	(0.022)	(0.020)
No. of observations				2,243,915		

Notes: The sample consists of full-time workers. We exclude observations in the year before the birth, the year of the birth, as well as the year after the birth. All specifications include the following controls: child indicator, marital status, job tenure in linear and squared form, number of workers in the firm, the average age of workers at the firm, year, and worker fixed effects. Log wage is instrumented using the average wage of similar workers of the same firm. MCC is estimated using the ratio of the marginal effect of commuting distance on job mobility and the marginal effect of log wage on job mobility, see equation (1.2). Standard errors are clustered at the firm-year level and can be found in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To test whether this change in the level of the wage coefficient is due to a discrete jump around childbirth or that it varies continuously over event time, we interact the instrumented wage with event time (21 year dummies) and gender. To deal with the issue that we have 42 instruments, we use the procedure discussed in

Figure D.1: Event time specific wage (IV) effects



Notes: wage (IV) event time effects around the birth of the first child. Further controls are specified in section 1.5.4. The grey area marks the birth of the first child. Standard errors are clustered by firm and year.

Balli and Sørensen (2013), and applied in Levkovich et al. (2020), which imposes restrictions in the first stage, by assuming that the relationship between wages and the instrument does not depend on event time, which makes the procedure more efficient. Figure D.1 shows the results from this procedure. It shows that the effect of wages slowly changes over event time, without showing a discrete jump around the birth of the first child.

1.E Hedonic regression

Table E.1: Bargaining effects from company moving

	parents				full danish working population (age<65)		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
log distance ×	log income	log income	log income	log income	log income	log income	log income
women	-0.0107	-0.0676	-0.0107	0.00199			
	(0.0168)	(0.0467)	(0.0168)	(0.0481)			
women children	0.00386	0.0140	0.00386	0.0437			
	(0.0188)	(0.0325)	(0.0188)	(0.0423)			
men	-0.00446	0.0183	-0.00446	0.000415			
	(0.00942)	(0.0175)	(0.00942)	(0.0184)			
men children	-0.00179	0.00183	-0.00179	0.00174			
	(0.00836)	(0.0113)	(0.00836)	(0.0128)			
log distance					0.00795*		
					(0.00382)		
log distance						0.00558	
						(0.00388)	
log distance × mothers						0.0117**	
						(0.00407)	
log distance ×							
age 18-45							0.00192
							(0.00402)
age 30-45							0.00795*
							(0.00380)
age 45-60							0.00894*
							(0.00381)
Observations	10773	6020	10773	4916	32148	32148	32148

Notes: All columns include the years t=-4/-1,3,4 around the company-plant moving and the controls for sector, marriage status, year fixed effects, job tenure and its squared form and company size. Further sample restrictions for the columns are as follows: [1] none, [2] no residence move, [3] only positive distance changes, [4] no residence move and company size larger 10. The sample for col. [5], [6] and [7] consists of the full working population, independent of childbirth until the age of 65 (because of fertility all other analysis are sampled conditional on childbirth between 18 and 45). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

CHAPTER 2

Agglomeration and gender [1]

Malte Borghorst, Jesper Hybel and Ismir Mulalic

Abstract: Women may benefit less from agglomeration because they search for jobs more locally and change jobs less often, which leads to worse matches over their careers, and subsequently reduces the return to experience and the learning channel of agglomeration economies. Using a panel of the full working population in Denmark for the years 2008-2016, we first demonstrate the existence of an urban wage premium not only for the wage level but also for wage growth. We then show that the portable part of the value of experience is lower for women than for men. We demonstrate that women's reduced earnings from work experience are due to smaller extra returns to experience gained and greater advantages from using experience in cities.

[1] Jesper Hybel and Ismir Mulalic (Department of Economics, Copenhagen Business School, Porcelænshaven 16A, DK-2000 Frederiksberg, Denmark, email: imu.eco@cbs.dk).
The authors thank participants of the seminars at Kraks Fond– Institute for Urban Economic Research, University of Duisburg-Essen, DFG-RTG 2484 - Regional Disparities & Economic Policy, Aalborg University and Copenhagen Business School. Research support from Kraks Fond – Institute for Urban Economic Research, Copenhagen (kraksfond@kraksfond.dk) is acknowledged. The usual disclaimer applies.

2.1 Introduction

Spatial disparities are a result of agglomeration effects and have long been a policy concern across regions and cities (e.g., the EU Cohesion Fund). Agglomeration effects exist, because Firms enjoy productivity advantages from denser locations (Combes et al., 2012; Moretti, 2011)[2] which typically arise from improved sharing, matching, or learning in dense labor markets (Duranton and Puga, 2004; Puga, 2010). Consequently, wages tend to be higher in densely populated areas, leading to an Urban wage premium (Henderson, 2003). Moreover, workers in cities accumulate human capital, such as valuable work experience, faster compared to those in thinner labor markets (Roca and Puga, 2017). We know from the literature that these return to work experience varies substantially by gender[3]. This paper identifies differences in gender-specific human capital accumulation and investigates the gender-specific benefits from agglomeration.

In this context, Denmark is an interesting case because of its equality-friendly labor market. Women's participation rate in Denmark is high (70 % in 2007), and women constitute about 50 % of the entire workforce. Using a rich administrative data set for Denmark that follows workers over a decade and across municipalities, we estimate the gender-specific returns to the experience acquired and used at different locations. We first construct worker and time-specific measurements of working experience collected at different job places and urban areas. Then, we use Mincerian wage regressions to learn about place-specific wages corrected for observed worker characteristics and to identify the initial unobserved worker ability as reflected in the worker (individual) fixed effects. Finally, utilizing the panel structure of the available data, we explore if this estimated value of knowledge accumulated in cities is different for men and women and if it persists after relocating. We find that the portable part of the value of experience for women is lower than for men due to lower additional gains from experience accumulated in cities, as well as the larger additional benefits from using experience in the top density areas relative to men.

Our paper relates to the extensive literature on agglomeration economies. Following the seminal contribution by Glaeser and Mare (2001), a large body of empirical literature has identified a significant urban wage premium, see, e.g., Rosenthal and Strange (2004); Puga (2010); Melo et al. (2009); Combes and Gobillon (2015). Sources for this urban wage premium are location fundamentals, worker and firm sorting,

[2] The benefits related to proximity to other economic agents are generally referred to as agglomeration effects and are the subject of extensive literature in spatial economics (Duranton and Puga, 2004; Rosenthal and Strange, 2004; Puga, 2010; Gaubert, 2018).

[3] See Blau and Kahn (2017) for an overview.

and agglomeration economies. Because firms producing tradable goods tolerate high urban wages only if they result in increased productivity (Moretti, 2011), these mechanisms explain why wages tend to be higher in cities and not merely compensate for higher urban living costs. These agglomeration economies may be experienced immediately upon moving into a city (static) or result in higher wage growth (dynamic).

The static advantages associated with dense urban areas refer to the benefits only enjoyed while working in cities (Rosenthal and Strange, 2004; Puga, 2010; Holmes, 2010). In this context, worker sorting describes the possibility that more productive workers may choose to locate in cities. In Denmark, higher educated and wealthier workers are more likely to work in bigger cities and closer to agglomerations (Hybel and Mulalic, 2022; Mulalic and Rouwendal, 2020; Gutiérrez-i Puigarnau et al., 2016). Combes et al. (2008) suggests that sorting may explain half of the urban wage premium, with estimates generally within the range of 30 to 70%. However, identification is difficult because of endogenous sorting. If, e.g., high-skilled workers prefer urban amenities, higher productivity workers could concentrate in cities, impacting the productivity variation, (Combes et al., 2011). To address this, econometricians use panel data methods, including individual fixed effects, to control for worker sorting on unobservable factors and historical instruments to manage feedback effects from worker migration to high-wage areas (Combes et al., 2010). Additionally, identification is improved by the use of commuting zones to account for labor market spillovers and experienced density, which takes into account the number of jobs in the surrounding instead of the population density (as in Duranton and Puga (2020)). Our contribution to this literature includes estimating the agglomeration elasticity for Denmark using the latest methodological advancements. We find a positive yet modest agglomeration effect.

Dynamic advantages suggest that the benefits of agglomeration, including enhanced human capital accumulation, are not immediate but accrue over time. This means that workers in cities thus accumulate higher-quality experience, i.e., through learning from superior interactions with other economic agents (Davis and Dingel, 2019; Eeckhout et al., 2014). The literature focuses on the role of learning as a potential mechanism in agglomeration economies (Glaeser, 1999; Duranton and Puga, 2001; Glaeser and Mare, 2001). Specifically, this literature examines how the location of experience accumulation and usage affects wage returns. To quantify the transferable agglomeration effects Combes and Gobillon (2015) exploit wage variations across geographic areas. Roca and Puga (2017) further argues that this accumulated human capital remains beneficial even when a worker relocates. However, estimating the effects is difficult, and the results are mixed. Roca and Puga

(2017) use a panel of Spanish male workers from 2004-2009 and find that experience gained in larger cities is highly rewarded, with a significant portion of this wage growth being portable, indicating the learning mechanism. In contrast, a study on British workers from 1998-2008 by D'Costa and Overman (2014) does not find evidence of an urban wage growth premium. Other studies have shown that in Italy, unskilled workers benefit more from a wage premium that accumulates over time, while skilled workers enjoy a premium when migrating to cities, with additional wage growth over time (Matano and Naticchioni, 2016). In Germany, both portable and non-portable agglomeration effects are significant (Frings and Kamb, 2021), and in Norway, college-educated workers experience a higher return on labor market experience acquired in cities (Carlsen et al., 2016). Our paper contributes to these empirical findings by improving the data quality and documenting the returns to experience accumulated for the universe of workers across Danish cities.

Importantly, this paper extends the literature by documenting the differences in the gender-specific benefits of urbanization. The economic literature offers numerous explanations for why the gender wage gap is also linked to agglomeration mechanisms sorting, matching, and learning. The mechanism and empirical results of sorting are well documented as women are known to sort into labor markets, e.g., along the lines of education and occupations, labor market participation, and work experience (for a discussion, see Blau and Kahn (2017)). For matching and learning in the context of agglomeration, the literature offers fewer explanations. While women do not have smaller networks, they are less likely to change jobs, leading to smaller wage gains and worse matches (Caldwell and Harmon, 2019). Additionally, beliefs about outside options, especially under constrained search (e.g., from higher commuting costs because of childcare, see Borghorst et al. (2021)) could also impact wages through bargaining and worse matches (Jäger et al., 2022; Caldwell and Danieli, 2020) which both are known to contribute to the urban wage premium (Hirsch et al., 2022; Dauth et al., 2016; Card et al., 2023) and the gender wage gap (Caldwell and Oehlsen, 2022; Manning, 2003b). Consequently, the learning benefits from agglomeration might be different for men and women. In the context of gender-specific benefits from agglomeration, studies exist that have explored the static urban wage premium, but the dynamic part of the gender gap is understudied. For instance, Hirsch et al. (2013) find variations in the gender wage gap in different areas of Germany, and Phimister (2005) estimate the static wage premium in the UK, relating it to the cognitive and social skills of different genders. Additionally, Rosenthal and Strange (2012) shows that women-owned businesses benefit less from network effects and agglomeration due to higher household burdens and commuting costs. To our knowledge, this paper is the first to contribute to this literature by identifying the role of learning in gender-specific urban wage premiums.

The remainder of the paper is organized as follows. In section 2, we describe the data, provide descriptive statistics, and present several empirical observations that suggest the gender-specific differences in the urban (growth) wage premium. Section 3 describes and discusses the empirical model and the estimation strategy. Section 4 presents the empirical results, and section 5 discusses the gender-specific urban wage premium. Section 6 concludes.

2.2 Data

The data used in the empirical analysis are derived from annual register data from Statistics Denmark for the years 2008–2016. For each year, we have information on the full population of workers, including the workers' workplace location at the municipality level, worker hourly wages, worker experience, job tenure, and a range of explanatory variables such as age, gender, and education.[4] experience for each worker is computed as the cumulative sum of the worker's work activity starting from the year 2003. The work activity is measured as the number of days worked during the year in primary employment. We also observe job tenure – the length of time a worker has been with an employer – for all workers. For each establishment, we observe the location of the establishment both on a municipality level and also more spatially detailed as the cell of location in the National Hectare Grid, which we use for the construction of travel to work areas (TTAs). We also observe the activity of the establishment measured as i) the number of workers employed at the establishment during the year, ii) the number of workers employed in November, and finally, iii) the number of full-time equivalents employed at the establishment during the year.

We focus on a sample of employed workers aged between 17 and 65. We exclude observations for workers who work in the public sector (health, education, and administration) because of fixed government wages and mining and agriculture because of dependence on location productivity. We also exclude immigrants because we do not always observe education and have sufficient information to calculate the work experience correctly. [5]

Table 2.1 shows the descriptive statistics, and we see no stark differences between

[4]To protect the identity of the companies for which data exist and to provide sufficient confidentiality protection, Statistics Denmark does not provide the exact workplace addresses for companies, but it does provide the municipality code for each establishment.

[5]We are then left with 7,246,703 observations (1,155,612 workers); this contains 704,008 men (4,441,873 observations) and 451,604 women (2,636,868 observations).

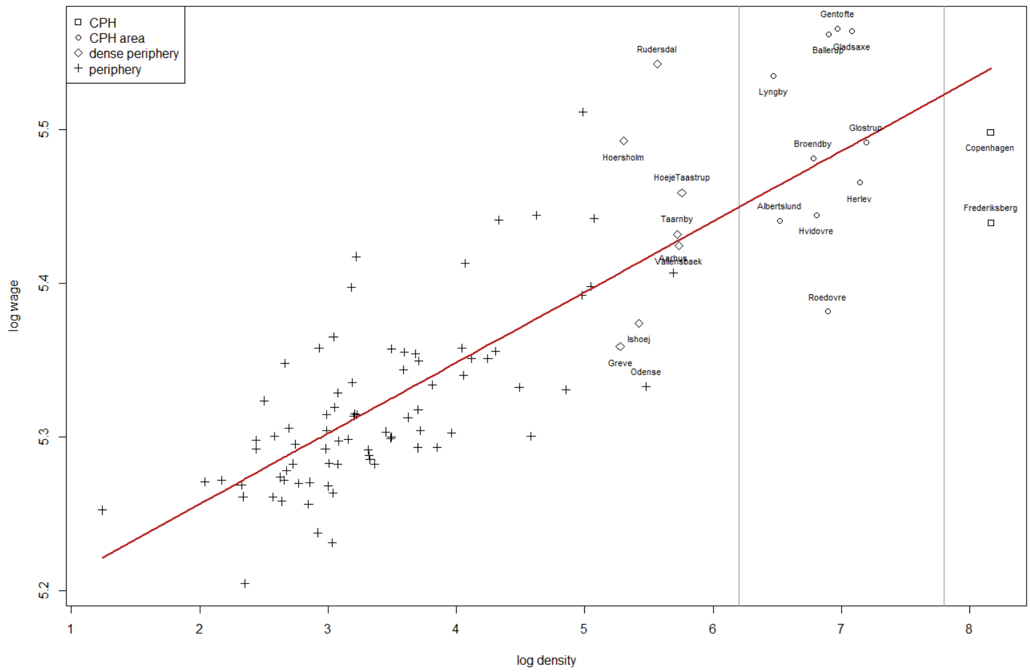
Table 2.1: Descriptive statistics

	Men		Women	
	mean	std.dev.	mean	std.dev.
Hourly wage (DKK)	230.59	79.34	200.38	67.60
Age	43.23	10.99	42.41	10.80
Activity (p.a., share)	0.91	0.21	0.91	0.21
Experience (based on activity)	10.48	3.32	10.35	3.37
Job-tenure (year)	4.97	4.52	4.88	4.48
Full-time (share)	0.92	0.28	0.82	0.39
Education (share)				
Primary	0.24		0.20	
Secondary	0.52		0.46	
Tertiary	0.23		0.33	
Occupational skill (share)				
Basic skill	0.65		0.62	
High skill	0.29		0.36	
Leading position	0.06		0.02	
Number of workers	704,008		451,604	
Number of observations	4,441,873		2,636,868	

men and women, except for hourly wage, which is about 15 % higher for men; see also Figure A.1 in the Appendix. Mean hourly wages are 231 and 200 DKK for men and women, respectively.[6] men are slightly older, have slightly longer job tenure, and have more often full-time jobs. Women obtain tertiary education more frequently, while men hold leading positions more often.

It appears that there is a significant positive correlation between log hourly wages and log job density; see Figure 2.1. Moreover, the Figure also shows the wage increase for the considered four density groups.

Figure 2.1: Wage against density anno 2016

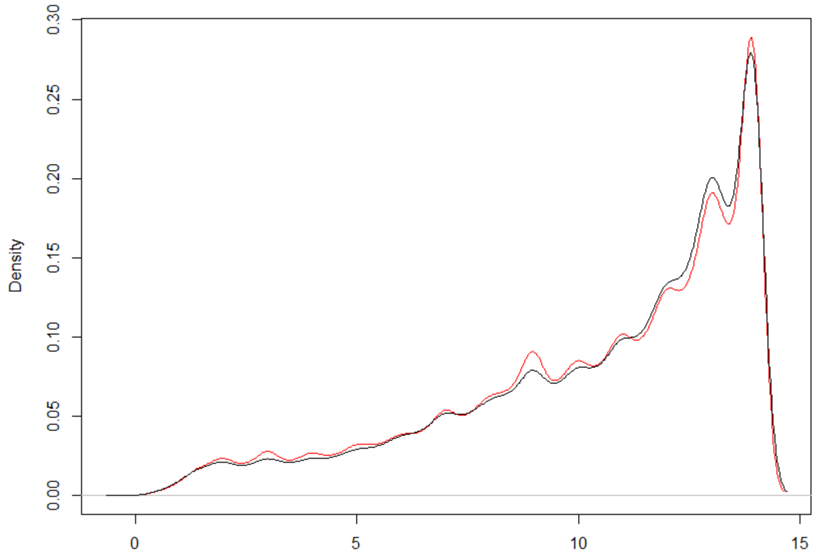


The mean worker experience is about ten years for men and women. The distributions of experience are also remarkably similar for both genders; see Figure 2.2. Table 2.2 suggests, however, that women use the accumulated experience, conditional on the area where the experience has been accrued, more intensively in high-density areas, compared to men. For example, of all the collected experience in the highest density area, women use 77% of the experience in the same area while men use 75%. The same is true for all the other considered areas, i.e., Women use the accumulated experience more intensively in high-density areas compared to men. We also find that the share of women increases with the job density; see Figure 2.3.

As we have seen, several systematic patterns emerge between men and women across the density areas in our sample. We summarize three facts about the gender-specific urban wage premium:

[6] 1 DKK ≈ 0.13 EUR.

Figure 2.2: Distributions of experience for men (red) and women (black)



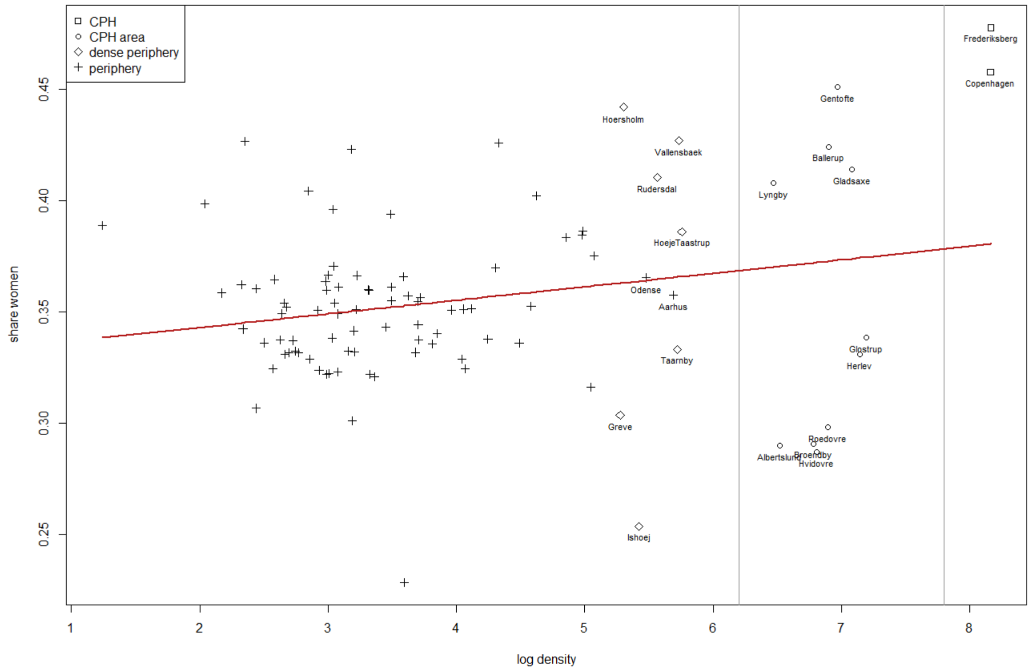
Notes: This Figure depicts the Gaussian kernel density distribution of the accumulated experience in the year 2016

Table 2.2: Areas of the used experience in shares, by gender

Origin	CPH	CPH area	Dense periph.	Periphery	Sum
Used in			Men		
CPH	0.75	0.11	0.04	0.09	1.00
CPH area	0.12	0.71	0.05	0.12	1.00
Dense periph.	0.11	0.13	0.63	0.13	1.00
Periphery	0.01	0.02	0.01	0.96	1.00
			Women		
CPH	0.77	0.10	0.04	0.08	1.00
CPH area	0.14	0.71	0.05	0.10	1.00
Dense periph.	0.15	0.12	0.62	0.11	1.00
periphery	0.02	0.02	0.01	0.96	1.00

Notes: CPH (highest density) includes municipalities Copenhagen and Frederiksberg, CPH area (second highest density) includes the rest of the Copenhagen metropolitan area, dense periphery (third highest density) includes among other municipalities Aarhus, the second largest city in Denmark, and periphery (lowest density) includes smaller towns and rural areas.

Figure 2.3: Share of women against density



1. wages and the share of women increase with job density;
2. distributions of the experience are similar for men and women, and
3. Women use the accumulated experience more intensively in high-density areas compared to men.

To measure employment density at the municipality level, we use the experienced density instead of the naive density (Duranton and Puga, 2020). The naive density is defined as the aggregated employment activity of the municipality, usually measured as the number of workers employed in the municipality divided by the geographical area of the municipality. The naive density measure assumes a uniform distribution of workers across space within a municipality. For municipalities with large work-wise uninhabited areas, this measure will underestimate the density experienced by most workers.

To measure the experienced density for municipality a, we calculate the average of the experienced density of all the workers employed in the municipality. The experienced density for municipality a is therefore given as

$$\bar{d}_{ad}^{k} := \frac{1}{N_{at}} \sum_{i \in a} d_{it}^{k}, \tag{2.1}$$

where N_{at} is the number of workers employed in municipality a at time t and d_{it}^{k} is the experience density for individual worker i at time t using distance k. The variable d_{it}^k is calculated using spatial information on the level of the National Hectare Grid. Specifically, d_{it}^k is the sum of employment activities for the establishments within distance k of the establishment where worker i is currently employed divided by the geographic size of the area.

2.3 The econometric model

In this section, we introduce a reduced-form wage model that includes the dynamic effects of experience. We introduce the stylized model in subsection 2.3.1 and describe how the dynamic effects of experience include a portable benefit, that we refer to as a *city premium* and a non-portable benefit, that we refer to as a *city use premium*. We then, in subsection 2.3.2, explain how we specify the learning effects. In subsection 2.3.3, we study the simplified model using linear wage paths. Finally, in subsection 2.3.4, we derive the bias in the estimates of the area fixed effects when the dynamic effects of experience or unobserved worker heterogeneity are ignored.

2.3.1 The wage equation

We use w_{ait} to denote the log wage of worker i in year t employed in area a and assume that the log wage is given by the equation

$$w_{ait} = \sigma_a + \mu_i + l_{it} + \mathbf{x}_{it}^\top \beta + \epsilon_{it}, \tag{2.2}$$

where σ_a is the unobserved area a fixed effect, μ_i is the unobserved individual fixed effect, l_{it} are the learning effects to be specified later, $\mathbf{x}_{it}$ is the vector of observable worker characteristics, β is a vector of parameters and ϵ_{it} is unobservable error term.

We use the function $a(i,t)$ to specify the area a in which worker i at time t is employed and let $1[a(i,t) = j]$ be the indicator function that the area of employment is area j. The indicators for the areas $j = 1,...,J$ are collected in the $J \times 1$ vector $\iota_{it} := (1[a(i,t) = 1],...,1[a(i,t) = J])^\top$ just as the area fixed effects $\{\sigma_j\}_{j=1}^J$ are collected in the vector $\sigma := (\sigma_1,...,\sigma_J)^\top$. We assume that the worker's choice of area of employment $1[a(i,t) = j]$ is uncorrelated with the individual and time-specific error terms ϵ_{it}.[7]

The *static* advantages of working in high-density areas are the advantages gained while working there but lost immediately upon being employed elsewhere. A worker changing area of employment from $a = a(i,t)$ to $a' = a(i,t')$ will immediately experience a change in wage due to the difference in the area fixed effects $\sigma_{a'} - \sigma_a$. This change is immediately lost again should the worker change her area of employment

[7] See Appendix B in Combes et al. (2008) for a detailed discussion of this assumption in a dynamic framework.

back to the area a. The wage equation (2.2) therefore allows for a static earnings premium of being employed in a high-density area if area fixed effects $\{\sigma_j\}_{j=1}^J$ are positively correlated with the employment density.

We refer to workers with an above-average value of μ_i as initial high-wage earners. The inclusion of the unobserved individual fixed effect μ_i also allows for sorting, where initial high-wage earners are predominantly employed in areas of high density. When this is the case, the covariance $Cov(1[a(i,t)=a],\mu_i)$ will be positive for all high-density areas a. For areas with some employment, this is equivalent to $\mathbb{E}[\mu_i|a(i,t)=a] > \mathbb{E}[\mu_i]$, such that the workers of the area a have a higher expected value of μ_i than the expected value $\mathbb{E}[\mu_i]$ for the population in general. Such sorting effects imply that high-density areas offer certain amenities favored by the high-wage earners. The initial high-wage earners are, therefore, willing to pay higher housing prices in high-density areas as suggested by Glaeser and Mare (2001).

Finally, the model allows for learning effects l_{it}, which capture part of the value of the workers' experience distinguished by where the experience is accumulated and where it is used. The value of worker i's experience at time t is given by

$$V(\{e_{ait}\}) = \sum_{a=1}^{J} \phi_{a'a} 1[a(i,t)=a'] e_{ait}, \tag{2.3}$$

where e_{ait} is the years of experience accumulated in area a at time t by worker i. The coefficient $\phi_{aa'}$ measures the value of a year of experience accumulated in area a when used in area a'.

To estimate a model that allows for these wage effects of experience, we arrange the areas according to their level of employment density into groups $g = \mathcal{G}(a(i,t))$, with a set g_0 consisting of areas with low employment density serving as reference group. We then specify the learning effects as

$$l_{it} = \sum_{g \neq g_0} \lambda_g e_{git} + \sum_{g} \delta_g \tilde{e}_{git}, \tag{2.4}$$

where $\lambda_g e_{git}$ is the value of experience e_{git} accumulated in any area belonging to the group g and $\delta_g \tilde{e}_{git}$ the value of experience accumulated in group g when not used in group g_0, achieved by defining $\tilde{e}_{git} := 1[a(i,t) \notin g_0] e_{git}$.[8]

This allows for a *city premium* where experience accumulated in the high-density areas is worth more used anywhere in a case of $\lambda_g > 0$. It also allows for a *city use premium* where the experience of different origin g is rewarded higher when used in high-density areas $a(i,t) \notin g_0$ in which case $\delta_g > 0$. Importantly, the *city premium*

[8]The estimated specification also allows for non-linear effects, but we ignore these for now for ease of presentation.

λ_g is the portable part of the experience in comparison to the *city use premium* δ_g, which is not portable, hence lost when the worker is employed in areas belonging to the low-density group of areas g_0.

2.3.2 The specification of learning effects

In this section, we derive a specification for the learning effects of the form given in equation (2.4) when there are only two areas: a highly urbanized area c (city) and a less urbanized area r (rural area).

The experience accumulated by worker i at time t by working in the city is denoted by e_{cit}, and the experience accumulated by working in the rural area is denoted by e_{rit}. The total experience accumulated is simply denoted by e_{it}. It equals the sum of the accumulated experiences in the city and the rural area, i.e., $e_{it} = e_{cit} + e_{rit}$. The wage benefit of experience is distinguished by the area where the experience is used. We use ϕ_{hj} as the benefit for an extra year of experience used in area h accumulated in area j. The log-wage w_{rit} for a worker currently employed in the rural area is now given by

$$w_{rit} = \sigma_r + \phi_{rr}e_{rit} + \phi_{rc}e_{cit} + u_{it}, \tag{2.5}$$

where σ_r is the static wage effect of the rural area and u_{it} is the unobserved error term. We define $\lambda_{rc} := (\phi_{rc} - \phi_{rr})$ as the measure of the wage premium of the city experience relative to rural experience when used in the rural area. Adding and subtracting $\phi_{rr}e_{cit}$ to the wage equation for the worker employed in the rural area, we get

$$w_{rit} = \sigma_r + \phi_{rr}e_{it} + \lambda_{rc}e_{cit} + u_{it}, \tag{2.6}$$

that includes only the total experience e_{it} of the worker and the city experience.

We then consider the wage of a worker employed in the city, which is given by

$$w_{cit} = \sigma_c + \phi_{cr}e_{rit} + \phi_{cc}e_{cit} + u_{it}, \tag{2.7}$$

consisting of a static city-specific wage effect σ_c and the wage benefits from experience accumulated in the rural area and the city. Adding and subtract the wage benefit $\phi_{rr}e_{rit} + \phi_{rc}e_{cit}$ we get

$$w_{rit} = \sigma_c + \phi_{rr}e_{rit} + \phi_{rc}e_{cit} + (\phi_{cr} - \phi_{rr})e_{rit} + (\phi_{cc} - \phi_{cr})e_{cit} + u_{it}, \tag{2.8}$$

where $\delta_{cr} := (\phi_{cr} - \phi_{rr})$ measures the extra benefit a worker receives from experience accumulated in the rural area when this experience is used in the city and similarly

$\delta_{cc} := (\phi_{cc} - \phi_{cr})$ measures the extra benefit a worker receives from experience accumulated in the city area when used in the city. Importantly a worker changing employment from the rural area to the city, having the stock of experience (e_{rit}, e_{cit}), would immediately receive a wage change $\sum_{a\in\{r,c\}} \delta_{ca} e_{ait} = \delta_{cr} e_{rit} + \delta_{cc} e_{cit}$ in addition to the change $\sigma_c - \sigma_r$ from the static area effects.

Finally we rewrite $\phi_{rr} e_{rit} + \phi_{rc} e_{cit}$ for the worker employed in the city to $\phi_{rr} e_{it} + \lambda_{rc} e_{cit}$ using the same approach as for the worker employed in the rural area and combine equation (2.6) and (2.8) to get

$$w_{a(i,t),it} = \iota_{it}^{\top}\sigma + \phi_{rr} e_{it} + \lambda_{rc} e_{cit} + \delta_{cr}\tilde{e}_{rit} + \delta_{cc}\tilde{e}_{cit} + u_{it}, \tag{2.9}$$

where $\iota_{it} = (1[a(i,t) = r], 1[a(i,t) = c])^{\top}$, $\tilde{e}_{rit} := 1[a(i,t) = c] e_{rit}$, $\tilde{e}_{cit} := 1[a(i,t) = c] e_{cit}$. In this specification, the learning effect

$$l_{it} = \lambda_{rc} e_{cit} + \delta_{cr}\tilde{e}_{rit} + \delta_{cc}\tilde{e}_{cit} \tag{2.10}$$

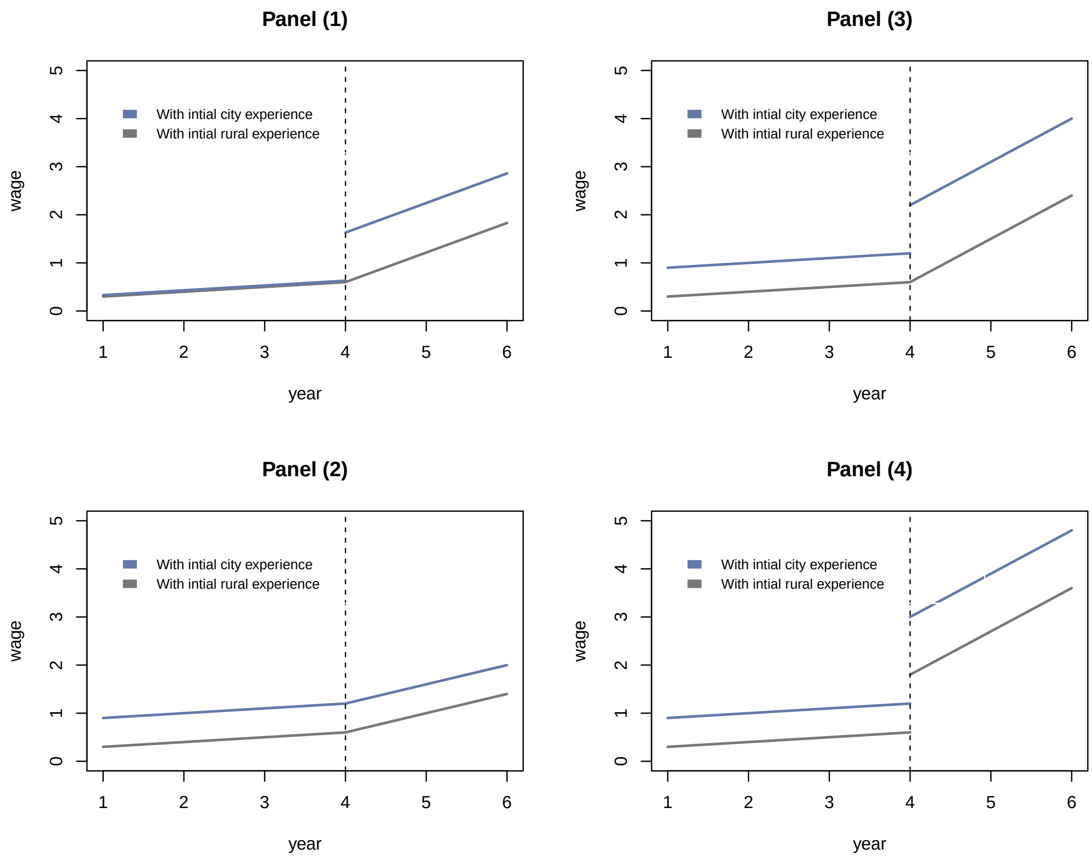
consists of the *city premium* λ_{rc} – the wage benefit of the city experience gained irrespective of where it is used – in addition to the *city use premiums* for the rural experience used in the city δ_{rc} and the city experience used in the city δ_{cc}. In the next section, we illustrate the learning effects of this stylized model using linear wage paths.

2.3.3 Wage paths in the stylized model

Figure (2.4) illustrates the wage paths of four different scenarios using the stylized model given in equation (2.9). The scenarios differ only in terms of the pay-off of experience; see Table 2.3. In all the scenarios, we consider two workers for $T = 6$ periods. Both workers are employed in rural areas for the first four years and then change the area of employment to the city for the last two years. We assume that workers accumulate one year of experience with the passing of every period. We also assume that the workers differ in their initial experience. The first worker is assumed to have two years of initial rural experience, while the second worker is assumed to have two years of initial city experience.

Figure 2.4 shows that in three scenarios - panel (1), (3), and (4) - there is a jump in at least one of the workers' wage paths. At the time of change in job location, the worker with the initial city experience has also accumulated rural experience, and the worker, therefore, has both types of experience. For this worker, the jump is therefore present if either the city use the premium of city experience δ_{cc} or the city use the premium of rural experience δ_{cr} are positive. In comparison, the worker with the initial rural experience has only accumulated rural experience at the time

Figure 2.4: Wage curve scenarios



of change in job location. Therefore, there is only a jump in the wage path of this worker in panel (4) where $\delta_{cr} > 0$.

Table 2.3: Scenarios

	Exp. ϕ_{rr}	City exp. λ_{cr}	Rural exp. used in city δ_{cr}	City exp. used in city δ_{cc}	Static effect $\sigma_r = \sigma_c$
Panel (1)	0.1	0	0	0.5	0
Panel (2)	0.1	0.3	0	0	0
Panel (3)	0.1	0.3	0	0.5	0
Panel(4)	0.1	0.3	0.2	0.5	0

We also explore the wage growth rate, which is determined by the value of experience in the current area of employment. This is trivial because the current area of employment determines what type of experience the worker accumulates. This implies that the wage growth is determined solely by $\phi_{rr} = 0.1$ while the workers are employed in the rural area. When employed in the city, the growth rate increases to $\lambda_{cr} + \delta_{cc} + \phi_{rr}$ adding the sum of the city premium and the city use premium of city experience. The higher growth rate of wage during the city employment, due to $\lambda_{cr} > 0$ and $\delta_{cc} > 0$, implies that the value of the experience for both workers is above the average while they are employed in the city. The city use premium of rural experience δ_{cr} cannot affect any wage growth rate because the rural experience is never accumulated when the worker is employed in the city.

2.3.4 The bias of static wage effects

In this section, we expand the simple wage model of equation (2.9) to include an individual fixed effect μ_i, observed worker characteristics $\mathbf{x}_{it}$ and the learning effects l_{it}. The model is, therefore, equivalent to our full model given in (2.2). However, we still assume that there are only two areas of employment. The wage of the individual worker is therefore given as

$$w_{a(i,t),it} = \iota_{it}^{\top}\sigma + \mathbf{x}_{it}^{\top}\beta + \mu_i + l_{it} + \epsilon_{it} \tag{2.11}$$

$$l_{it} = \lambda_{rc} e_{cit} + \delta_{cr}\tilde{e}_{rit} + \delta_{cr}\tilde{e}_{cit}$$

where $\mathbf{x}_{it}^{\top}\beta$ includes the wage pay-off of the experience $\phi_{rr}e_{it}$ not distinguished by the origin or the place of use.

We first consider the bias of the city fixed effects estimates if the econometrician fails to control both for the unobserved individual effects μ_i and the learning effects l_{it}. The workers employed in the city have higher unobserved fixed effects μ_i than workers in general. This induces the $bias(\hat{\sigma}_c) := \hat{\sigma}_c - \sigma_c$ in a positive direction. The same is the case for the learning effects under the assumption that the combined gain of the city experience relative to the rural areas $\lambda_{cr}+\delta_{cc}$ is positive. Moreover, workers with city experience are not uniformly distributed across the urban landscape; hence, city workers, on average, have more city experience. When this experience is valued higher, the city area becomes more productive. Failing to control for this explicitly, therefore, affects the $bias(\hat{\sigma}_c) := \hat{\sigma}_c - \sigma_c$ positively. See Appendix B for a detailed derivation.

We now consider the bias of the city fixed effect estimates if the econometrician fails to control for the leaning effect but uses the within estimator to control for the unobserved individual fixed effects μ_i. In this case, it is only workers who change their area of employment who identify the area-fixed effects. Under the assumption that the city premiums of experience λ_{cr}, δ_{cc} and δ_{cr} are all positive, the workers who change their area of employment from the rural area to the city will - while working in the city - have a higher value of experience than their average value of experience and will, therefore, affect the $bias(\hat{\sigma}_c) := \hat{\sigma}_c - \sigma_c$ positively. The workers who migrate away from the city will, on the other hand, affect the $bias(\hat{\sigma}_c) := \hat{\sigma}_c - \sigma_c$ negatively, unless it is the case that the value of the city experience to a very high degree is not portable. For a detailed derivation, see Appendix C.

2.4 Empirical results

This section presents the empirical results. We first estimate a specification ignoring both the learning effect and the unobserved worker heterogeneity. The results of this estimation are given in subsection 2.4.1. We then, in subsection 2.4.2, include individual fixed effects but still do not control for learning effects. Finally, in subsection 2.4.3, we also consider the learning effects.

2.4.1 The static wage gains of density

In this section, we estimate the static wage gains of density, not controlling for the individual fixed effects or the learning effect. The wage equation is therefore given by

$$w_{a(i,t),it} = \iota_{it}^{\top}\sigma + \mathbf{x}_{it}^{\top}\beta + \gamma_s + \eta_t + v_{it}, \tag{2.12}$$

where γ_s are sector fixed effects and η_t are year fixed effects.

The results of the estimation are reported in Table 2.4 column (1). The wage value of the experience is concave. The first year of experience increases wages by 4.5% and the fifth year of experience by 3.1%.[9] Job tenure is also concave. The first year of employment at a workplace implies a wage increase of 1.1%. The wage increase by the fifth year of employment at the same workplace is reduced to 0.5%. As expected, the wages increase with the level of occupational skill and with the level of education.

Column (2) in Table 2.4 shows the results of regressing the estimates of the area fixed effects against the log density. We instrument employment density with the population densities for the years 1801 and 1834. The elasticity of wages with respect to employment density is estimated to be 0.0226.[10].

The estimate of the elasticity of wages with respect to employment density is likely biased due to the bias in the estimates of the area fixed effects. This bias was

[9] These increases as calculated as $(\exp(0.0453 - 0.0016) - 1) \cdot 100$ and as $(\exp(0.0453 \cdot 5 - 0.0016 \cdot 5^2 - (0.0453 \cdot 4 - 0.0016 \cdot 4^2)) - 1) \cdot 100$.

[10] We have also estimated the elasticity of wages with respect to employment density using the area-year fixed effects. It is then 0.0176. We have finally estimated the elasticity of wages with respect to effective density using area-year fixed effects. This elasticity is 0.0182 (see Table 5 in Hybel (2020)). The effective density is a transportation time-weighted employment density. This method has been used by Combes et al. (2010).

found in the two-area example to be positive for the high-density areas, assuming that workers in these areas have higher values of the individual fixed effects μ_i and that the experience gained and used in the high-density areas has a higher value $(\lambda_{cr} + \delta_{cc}) > 0$ (see equation B.4). It was also assumed that workers did not change their areas of employment. This assumption is justified by the fact that when individual fixed effects are not included, the area fixed effects are also identified by the workers who do not change the employment area. These workers make up the majority of workers in our sample.

Table 2.4: Estimation of the static wage gains of density

	Dependent variable:			
	Log wage	Area indicator	Log wage	Area indicator
	(1)	(2)	(3)	(4)
Log density		0.0226^{***}		0.0107^{***}
		(0.0033)		(0.0017)
Area fixed effects	+		+	
Individual fixed effects	-		+	
Experience	0.0453^{***}		0.0408^{***}	
	(0.0002)		(0.0004)	
Experience^2	-0.0016^{***}		-0.0014^{***}	
	(0.00001)		(0.00001)	
Job tenure	0.0121^{***}		0.0097^{***}	
	(0.0001)		(0.0001)	
Job tenure^2	-0.0008^{***}		-0.0006^{***}	
	(0.00001)		(0.00001)	
Medium skilled occupation	0.1583^{***}		0.0209^{***}	
	(0.0005)		(0.0005)	
High skilled occupation	0.3351^{***}		0.0644^{***}	
	(0.0011)		(0.0008)	
Secondary Education	0.0487^{***}		0.1906^{***}	
	(0.0005)		(0.0017)	
University Education	0.2040^{***}		0.3202^{***}	
	(0.0008)		(0.0023)	
Male	0.1412^{***}			
	(0.0005)			
Sector fixed effects	+		+	
Year fixed effects	+		+	
Observations	7,246,703	98	7,246,703	98

Notes: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

2.4.2 The static wage gains of density and the individual fixed effects

In this section, we estimate the static wage gains of density, controlling for the individual fixed effects but still not for the learning effects. The wage equation is now given by

$$w_{a(i,t),it} = \iota_{it}^{\top}\sigma + \mu_i + \mathbf{x}_{it}^{\top}\beta + \gamma_s + \eta_t + u_{it}. \tag{2.13}$$

The results of the estimation are reported in Table 2.4 column (3). The coefficients for experience and job-tenure are robust to the inclusion of the individual fixed effects. The wage value of the first year of experience is slightly reduced to 4.0% (versus 4.5%) and the fifth year of experience to 2.9% (versus 3.1%). The wage benefit for job tenure is also slightly reduced, with a wage increase for the first year of employment of 0.9% (versus 1.1%), while the wage increase of the fifth year of employment is reduced to 0.4% (versus 0.5%). Also, in this specification, the wage increases with the level of occupational skill and the level of education. The estimates are, however, less robust to the addition of the individual fixed effects and are associated with the significantly higher wage benefit.

Column (4) in Table 2.4 shows the results of regressing the estimates of the area fixed effects against the log density of the area using historical instruments of population density for the years 1801 and 1834. The elasticity of wage with respect to employment density is reduced by more than 50% (from 0.0226 to 0.0107). This is close to the 47% reduction reported by Roca and Puga (2017).[11]

Ignoring the learning effects likely results in a biased elasticity of wages with respect to employment density. When individual fixed effects are included in the estimation equation, only workers who change employment areas identify the area fixed effects. Workers changing their area of employment to the high-density areas bias the estimates of area fixed effects positively if the city premium and the city use premium are positive (see equation C.4). This happens because workers who relocate are more likely to have a value of experience greater than their average value of experience while working in a high-density area. On the other hand, workers moving to a low-density area can bias the estimates for the area fixed effects either positively or negatively, depending on how portable is the value of the accumulated experience (see equation C.5). Suppose the value of experience is extremely portable. In that case, these workers will have a value of experience that is lower than their average value of experience. At the same time, they work in the high-density area, inducing a negative bias of the high-density area fixed effects.

[11] Using data for Denmark, Knudsen et al. (2019) and De Borger et al. (2019) estimate the elasticity of wage with respect to job accessibility of a similar magnitude.

2.4.3 Dynamic wage benefits of density

In this section, we estimate both the static and dynamic wage gains of density. The wage equation we estimate is now

$$w_{a(i,t),it} = \iota_{it}^{\top}\sigma + \mu_i + \mathbf{x}_{it}^{\top}\beta + l_{it} + \gamma_s + \eta_t + \epsilon_{it} \tag{2.14}$$

$$l_{it} = \sum_{g \neq g_0} \lambda_g e_{git} + \sum_{g \neq g_0} \alpha_g e_{git} e_{it} + \sum_{g} \delta_g \tilde{e}_{git} + \sum_{g} \psi_g \tilde{e}_{git} e_{it},$$

where the learning effects l_{it} include both linear and non-linear effects. Specifically we allow for the city premiums $\lambda_g e_{git} + \alpha_g e_{git} e_{it}$ to depend on the workers total experience e_{it} and similarly for the city use premiums $\delta_g \tilde{e}_{git} + \psi_g \tilde{e}_{git} e_{it}$.

The areas are divided into four different groups $g \in \{1, \dots, 4\}$ depending on their level of employment density: three groups with the high employment density collective referred to as *the top* and the final group of comparatively low employment density (the reference group). The group with the highest level of employment density consists of municipalities Copenhagen and Frederiksberg, which together make up the center of the Greater Copenhagen area, the largest urban area in Denmark. The group with the second-highest density consists of municipalities located in the proximity of this center and are thus all part of the Greater Copenhagen area. The third group includes dense municipalities in mostly rural areas, including the second-largest city in Denmark, Aarhus. The low employment density group includes the periphery, i.e., smaller towns and rural areas.

The estimation results are reported in Table 2.5 column (1). The wage gain from experience accumulated outside the top is 3.4% (versus 4.0%) for the first year and 2.4% (versus 2.9%) for the fifth year when the worker uses the experience outside the top region. This is a reduction compared to similar estimates in the model given in equation 2.13 without learning effects. The wage gain from experience accumulated in the low-density areas is higher if it is used in the top areas. The coefficient of experience from the areas of lowest density, when used at the top, is 0.0067, and the second-order term is -0.0004. The first year of experience accumulated in the low-density area thus gives a 4.0% wage increase, while the fifth year gives a 2.7% wage increase. The experience gained outside the high-density areas is rewarded more when used in the high-density areas. How much a worker gains from working in a high-density area depends, therefore, on the level of the accumulated experience. This result is in line with the results of Roca and Puga (2017).

The experience accumulated in the areas of the highest and the second-highest density is rewarded higher independent of where it is used. This is apparent from the coefficients of 0.0147 and 0.0078 for experience gained in the areas of highest density

Table 2.5: Estimation of the dynamic and static wage gains of density

	Dependent variable:		
	Log wage	Area indicator	Area indicator + City premium (7 years of experience)
	(1)	(2)	(3)
Log density		0.0052***	0.0228***
Experience	0.0344***		
Experience^2	−0.0012***		
Experience (highest density)	0.0147***		
Experience (second highest density)	0.0078***		
Experience (third highest density)	0.0067***		
used in top	(0.0007)		
Experience used in top (highest density)	0.0107***		
Experience used in top (second highest density)	0.0079***		
Observations	7,246,703		

Notes: * p<0.1; ** p<0.05; *** p<0.01*

and for the experience gained in the areas of second-highest density, respectively. The first year of experience from the highest-density areas is 1.5% more rewarding than the first year of experience from the low-density areas and 0.8% for the areas of the second-highest density. The experience from the areas of the third-highest density is insignificantly different from the experience accumulated in the low-density areas.

The value of the experience accumulated in the high-density areas is only partly portable. Experience gained in the top - the areas of highest, second highest, and third highest density - is rewarded higher when used in the top. This implies, for example, that the wage gain from the experience accumulated and used in the areas of the highest density is 5.9% for the first year and 4.0% for the fifth year, while the portable gain is slightly lower at 4.9% for the first year and 3.6% for the fifth year.

Finally, column (2) in Table 5 shows that the estimate of the elasticity of wage with respect to employment density, conditional on the learning effects, is reduced again by about 50%. Notice, however, that for workers with seven years of experience accumulated in the high-density area (city), this elasticity increases more than four times (from 0.0052 to 0.0228); see column (3) in Table 5.

Earnings Profiles

In this section, we illustrate the results from Table 2.5 using earning profiles. We assume that workers initially have zero years of experience and accumulate one year of experience each year for every year of the period under consideration. We also assume that workers change their employment area after five years.

Figure 2.5 illustrates the case of a worker who relocates from the low-density area to either the highest-density area (blue wage-path) or the second-highest-density area (red wage-path). The left panel shows the wage benefit of the learning effects

Figure 2.5: Migration to high-density area

as defined in equation (2.14). These wage benefits do not include benefits of the experience accumulated and used in the low-density areas. Because the areas of low employment density are used as the reference, these benefits are part of the wage benefits earned by the experience accumulated and used anywhere. For the first five years, the worker accumulates and uses experience in an area with low employment density. The wage benefits of the learning effects are, therefore, zero. However, as illustrated in the right panel of Figure 2.5, the wage still increases due to the benefit of experience accumulated and used in a low-density area.[12]

When a worker relocates, the additional wage benefits of experience increase to approximately 4%. This increase consists of two parts. The first part is due to the additional wage benefit of using the already acquired experience in a new employment area. The size of this part does not depend on whether the worker migrates to an area of the highest employment density or second-highest employment density. The size does, however, depend on how much experience the worker has accumulated before the relocation. The second part shows one additional year of experience accumulated from year 5 to year 6 in the new area of employment. It depends on (i) the wage benefit of experience accumulated in the new employment area and (ii) the wage benefit of the experience accumulated in the new area of employment when used in the top-density area. These two components combined are larger for the areas of the highest employment density compared to the second-highest employment density area (see Table 2.5). This difference explains why the jump from year 5 to year 6 is slightly larger when a worker migrates to the top-density area rather than to the area of the second-highest employment density. Furthermore, it also explains why the growth of the additional benefits is larger for the high-density areas. After five

[12]The wage benefit of the experience accumulated and used in the low-density area is determined by the coefficients of the experience and experience squared in Table 2.5.

years of working in the highest density area, the additional wage benefit reaches a level of approximately 6-8%.[13]

Figure 2.6 illustrates the reverse case, i.e., when a worker relocates from a high-density to a low-density area. The left panel of Figure 2.6 shows how the additional wage benefit increases the longer the worker is employed in the high-density area. After the first year, the gain of the experience accumulated and used in one of the areas of the highest density relative to the experience accumulated and used in an area of the low density is only 1.5%. In contrast, after only five years, it reaches a level of about 10%. Due to the additional wage benefits of experience accumulated and used in the high-density areas, the additional wage benefits decrease when the worker leaves the city.

The right panel of Figure 2.6 pictures how a reduction in the additional wage benefit of the accumulated experience results in a decrease in the total benefit of the accumulated experience. Having accumulated as much as five years of experience in high-density areas, workers who migrate to low-density areas thus have to accept a decrease in the total wage benefits of experience. Hence, the additional wage benefits of experience are not completely portable and are potentially functioning as a strong disincentive for workers with many years of experience to leave the high-density areas, v.i.z. Cities.

Figure 2.6: Migration from high-density area

[13]It is easy to see from Figure 2.5, that during the period of employment in the area with high employment density, the additional wage benefits are above average. Using equation (C.4), This implies that workers who migrate to an area of the highest or the second-highest density will bias the estimates of the area fixed effects positively.

2.5 Gender and the dynamic wage gains of density

In this section, we focus on the gender-specific dynamic gains of density. We estimate the full wage model as specified in equation (2.14) separately for women and men. The results are given in columns (1) and (2) in Table 2.6.

We first focus on the rewards of experience gained outside the top-density areas. Here, we find that the wage gains of experience are lower for women than for men when the experience is used outside the top-density areas. For example, men's wages increase by 3.6% in the first year, compared to a lower 3.3% wage increase for women. For both genders, the top-density areas offer opportunities to increase the returns to experience accumulated outside the top-density areas. However, the additional increase in wages when using the experience in the top-density areas is slightly larger for women. For the first year of experience, men get an additional 0.5% wage increase when using the experience in the top area, while women get a higher 0.7% additional wage increase. However, despite this additional gain, men still get a higher wage increase for the first year of the experience accumulated in the low-density areas and used in the top-density areas, 4.2%, compared to women who get 4.0%. The use of the experience in top-density areas by women reduces the gender wage gap but does not close it.

We now focus on the rewards of the experience gained in the top-density areas when used outside the top-density areas. We find that for both men and women, the experience accumulated in the highest and the second-highest density areas is rewarded higher than the experience gained elsewhere when used outside the top-density areas. This additional gain from the accumulated experience is lower for women than for men.[14] This implies that the first year of experience accumulated in the highest-density area awards men with an additional 1.8% wage increase relative to the experience accumulated in the low-density areas. Women receive the lower 1.1% additional wage increase. The comparative percentages for the experience accumulated in the areas of second-highest density are 0.9% and 0.7% for men and women, respectively. The gender gap in the additional reward for the experience accumulated in high-density areas is, therefore, more pronounced in cities.

The experience accumulated in the high-density areas for both genders is rewarded extra when used in these areas. Moreover, only for the experience gained in the area of the highest density, the additional wage increase of using the experience in the top density areas is higher for men than for women. The coefficient of the experience for the highest density areas, when used in the top, is 0.0107 for men

[14]For men, the coefficient of the experience accumulated in the highest density area is 0.0179 compared to 0.0110 for women. It is 0.0086 for the areas of the second-highest density for men, while only 0.0072 for women (see Table 2.6).

Table 2.6: Estimation of the dynamic wage gains of density for men and women

	Dependent variable:	
	Log wage (male)	Log wage (women)
	(1)	(2)
Experience	0.0368***	0.0332***
	(0.0005)	(0.0007)
Experience^2	−0.0012***	−0.0012***
	(0.00002)	(0.00002)
Experience (highest density)	0.0179***	0.0110***
	(0.0014)	(0.0015)
Experience (second highest density)	0.0086***	0.0072***
	(0.0012)	(0.0015)
Experience (third highest density)	0.0014	−0.0015
	(0.0017)	(0.0021)
Experience (highest density)	−0.0004***	−0.0004***
× experience	(0.0001)	(0.0001)
Experience (second highest density)	−0.0002*	0.00003
× experience	(0.0001)	(0.0001)
Experience (third highest density)	0.00005	0.0001
× experience	(0.0001)	(0.0002)
Experience (lowest density)	0.0057***	0.0078***
used in top	(0.0009)	(0.0011)
Experience (highest density)	0.0107***	0.0099***
used in top	(0.0014)	(0.0015)
Experience (second highest density)	0.0058***	0.0107***
used in top	(0.0012)	(0.0015)
Experience (third highest density)	0.0082***	0.0118***
used in top	(0.0017)	(0.0022)
Experience (lowest density)	−0.0004***	−0.0004***
× experience used in top	(0.0001)	(0.0001)
Experience (highest density)	−0.0009***	−0.0006***
× experience used in top	(0.0001)	(0.0001)
Experience (second highest density)	−0.0004***	−0.0007***
× experience used in top	(0.0001)	(0.0001)
Experience (third highest density)	−0.0006***	−0.0006***
× experience used in top	(0.0001)	(0.0002)
Observations	4,441,873	2,636,868

Notes: * p<0.1; ** p<0.05; *** p<0.01* and 0.0099 for women (see Table 2.6). This is equivalent to 1.1% and 1.0% increases in the wages for men and women, respectively. For the experience accumulated in the areas of the second-highest density when used in the top, women get a larger 1.1% wage gain while men get a lower 0.6% wage gain. Hence, the first year of experience, when used in the top-density areas, awards men a 2.8% additional wage increase compared to the experience accumulated in the low-density areas and used in the low-density areas. For women, this additional wage increase is only 2.0%. The wage growth rate of working in the areas of highest density thus appears larger than the wage growth rate of working outside the top density areas, with a higher growth rate for men than for women. However, for the experience accumulated in the second-highest areas and used in the top-density areas, both genders get almost the same benefit, i.e., 1.36% and 1.40% for men and women, respectively. For the experience accumulated in the third-highest area and used in the top, men get an additional 0.8%. In contrast, women get 1.1%.[15]. High-density areas thus offer men and women similar opportunities, i.e., faster wage growth.

Finally, we compare the share of the additional value of portable experience for men and women accumulated in the high-density areas. For the areas with the highest density, the portable share for men is 0.63, while for women, it is 0.53.[16] For the areas of the second-highest density, the portable share for men is 0.60, while for women, it is 0.40. Finally, for the areas of the third-highest density, the portable share is 0 for both genders.[17] For women, the portable part of the value of experience is thus lower than for men, partly due to the lower additional gains from experience accumulated in high-density areas but also due to the larger additional benefits from using the accumulated experience in the top-density areas relative to men. Therefore, women continually have to work in high-density areas to have similar benefits as men.

[15]This is computed using only the coefficients of experience accumulated in the areas of third-highest density when used in the top density areas because the coefficients on the experience accumulated in the areas of the third-highest density are statistically insignificant.

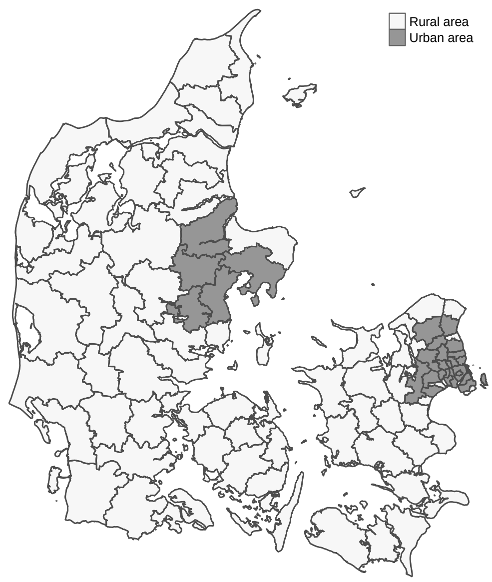
[16]For men, the share is calculated using the coefficient of experience accumulated in the highest density areas and the coefficient for experience accumulated in the highest density areas when used in the top, as given in Table 2.6. Hence, the portable share for men is $0.063 = 0.0179/(0.0179+0.0107)$, ignoring the second-order effects from the non-linear terms. The calculation of the other shares is performed similarly.

[17]We assume that the coefficients on experience accumulated in the areas of third highest density are 0 because they are not statistically significant (see Table 2.6).

2.5.1 Commuting zones

Using the density classification does not take labor market spillovers into account (e.g., see Borghorst et al. (2021) for a discussion of gender-specific commuting behavior and labor market access). To address this, we create commuting zones for the largest labor markets in Denmark, Copenhagen, and Aarhus. The densest zone is associated with the capital Copenhagen. It employs about one-third of the labor force, while the second densest zone is associated with the city of Aarhus and employs ten percent of the labor force. The commuting zones are constructed as Travel to Work Areas, with 80% of the workers employed in the zone also being resident in the zone and with 80% of the workers living in the zone also being employed within the zone. The two urban zones are illustrated in Figure (2.7), and we will henceforth simply refer to the two commuting zones as 'Copenhagen' and 'Aarhus.' In contrast, the rest of the municipalities will be referred to as the 'rural' area.

Figure 2.7: Urban areas



Notes: Commuting zones around Aarhus (west) and Copenhagen (east)

The results in table 2.7 show that taking into account labor market spillovers from commuting behavior matters when estimating the urban wage premium. The overall return to the experience of working in the two largest labor markets is similar at 1%, which is in line with the results in Table 2.5. Similarly, if workers with experience from rural areas work in the top labor markets, they would gain a premium of 0.4% over the rural labor markets. For workers with experience from Copenhagen, it is equivalent. Workers who gained experience in Aarhus would benefit by 0.4% if they worked in the rural labor markets. Estimated separately, we find for women, the return to experience from the largest labor markets is only half of that of men (in Copenhagen 0.6% vs. 1.2% and Aarhus 0.5% and 1%). These results are even stronger than in Table 2.6 and stress once again the importance of labor market spillovers when assessing the gender wage gap. However, women only benefit from working in the top markets when they have accumulated experience in Copenhagen before (.01%). The benefit from having gained experience from Aarhus for working in the top labor markets is zero for women and negative for men (which might be explained by the dominance of the Copenhagen labor market in the Danish economy).

2.5.2 Gender and earnings profiles

In this subsection, we illustrate the results shown in Table 2.6 using earning profiles. We do this because using only percentages for the value of the first year of experience understates the economic significance of the results since the difference becomes larger as workers accumulate more experience.

Migration to a low-density area

Figure 2.8 illustrates the additional wage gains from the experience accumulated in a high-density area for men and women. These gains are the learning effects defined in equation 2.14. We assume again that all workers initially have zero years of experience and each year accumulate one year of experience from the area of current employment. For the first five years, the workers are assumed to work in the high-density area, and then they migrate to one of the low-density areas.

Figure 2.8: Accumulated additional gains from experience when migrating from the city

The left panel in Figure 2.8 shows the additional wage gains from experience accumulated in the high-density area. The gains are larger for men. While employed

Table 2.7: Results for commuting zones

	Dependent variable:		
	log wage		
	(1)	(2)	(3)
	All	Men	Women
Copenhagen	0.040***	0.039***	0.038***
	(0.001)	(0.002)	(0.002)
Aarhus	0.005***	0.005***	0.004
	(0.002)	(0.002)	(0.003)
Experience	0.020***	0.021***	0.017***
	(0.0002)	(0.0002)	(0.0003)
Experience^2	−0.0003***	−0.0004***	−0.0002***
	(0.00000)	(0.00000)	(0.00001)
Experience (Copenhagen)	0.010***	0.012***	0.006***
	(0.0003)	(0.0004)	(0.001)
Experience (Aarhus)	0.009***	0.010***	0.005***
	(0.0005)	(0.001)	(0.001)
Experience (Copenhagen)	−0.0004***	−0.0004***	−0.0003***
× experience	(0.00002)	(0.00002)	(0.00003)
Experience (Aarhus)	−0.0004***	−0.0004***	−0.0002***
× experience	(0.00002)	(0.00002)	(0.00004)
Experience (rural)	0.004***	0.005***	0.002***
used in top	(0.0003)	(0.0003)	(0.0005)
Experience (Copenhagen)	0.00003	0.0002	0.001*
used in top	(0.0003)	(0.0004)	(0.001)
Experience (Aarhus)	−0.004***	−0.004***	−0.001
used in top	(0.0005)	(0.001)	(0.001)
Experience (rural)	−0.0002***	−0.0003***	−0.0001***
× experience used in top	(0.00001)	(0.00002)	(0.00002)
Experience (Copenhagen)	0.0001***	0.00004**	0.00001
× experience used in top	(0.00002)	(0.00002)	(0.00003)
Experience (Aarhus)	0.0002***	0.0002***	0.0001
× experience used in top	(0.00002)	(0.00003)	(0.00004)
Observations	7,246,703	4,441,873	2,636,868

Notes: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

in the high-density area, the additional wage gains from experience include the additional wage benefit of experience accumulated in the area of the highest density, as well as the additional wage benefit of the experience accumulated in the area of the highest density when used at the top. After five years of work in the high-density area, the benefit from an accumulated additional wage increase for men is approximately 12%, compared to only 8% for women.

Both genders experience a drop in the additional wage gains when relocating. This happens because they lose the additional wage benefit of the experience accumulated and used in the high-density area. The size of this reduction is approximately 4%. Furthermore, relative to the additional wage increase, at the time of migration, the reduction is larger for women, who lose 4% out of the 8% gained, while men lose 4% out of the 12% gained. The smaller relative loss for men reflects how their additional wage returns are more portable than the additional wage return for women. After migration, the additional wage returns to experience a slight decrease as the workers accumulate further experience.

The right panel of Figure 2.8 shows the outcome of relocation from the area of the second-highest density. Now, the additional wage gains from experience are larger for women than for men. After five years of work in the high-density area, women have reached an additional wage increase of approximately 7.5% while men have reached approximately 6.5%. When migrating, the additional wage gains of experience are reduced to approximately 4%. Therefore, the reduction is the largest for women.

Figure 2.9: Total accumulated gains from experience when migrating from the city

Figure 2.9 illustrates the total wage gains of the experience. Here, the gains of experience earned and used anywhere are added to the additional gains illustrated in Figure 2.8. The left panel in Figure 2.9 illustrates migration from a high to a low-density area. For both genders, the wages increase faster for the first five

years of employment in the high-density area compared to the latter five years of employment in the low-density area. The reduction in the wage growth rate is due to workers no longer receiving the benefits of using the accumulated experience in the top, as well as due to accumulating low-density experience rather than high-density experience[18]. In conclusion, the wage increases faster for men than for women both for the first five years of employment in the high-density area as well as for the latter five years of employment in the low-density area.

[18]There is also a third effect due to the nonlinearity of the learning effects. This effect is because the worker's experience accumulated in the high-density area becomes less worthwhile for every year of experience accumulated in the low-density area. This effect is very little, as shown in Figure 2.8 and is therefore ignored.

2.6 Conclusions

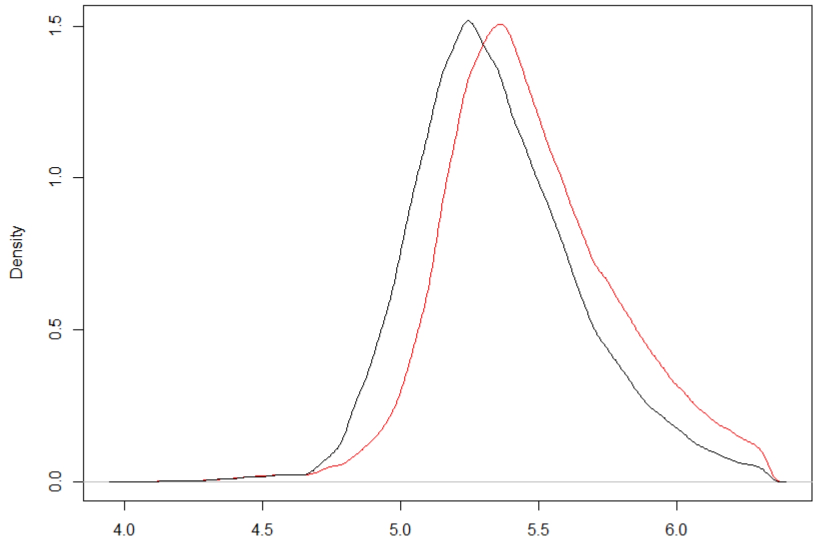
The gender pay gap is a long-standing inequality, and its origins are multidimensional. Empirically, we know very little about the impact of the agglomeration economies on this inequality. This article seeks to identify the gender-specific urban wage premium. Using register data for workers in Denmark, we first prove the existence of an urban premium for wage levels and a city-size premium on wage growth. The estimated effects imply individual-level compensating differentials for agglomeration economies as predicted by urban economic models that allow for productivity advantages emerging from improved sharing, matching, or learning in dense labor markets (Roca and Puga, 2017; Duranton and Puga, 2004). We also identify three empirical facts about the gender-specific urban wage premium: i) wages and the share of women increase with job density, ii) distributions of the work experience are similar for both genders and iii) women use the accumulated experience more intensively in cities. Finally, our empirical findings suggest that the value of the portable part of the accumulated work experience for women is below that of men. This explains, at least partially, the puzzle of simultaneously positive correlations between wages and the share of women on the one hand and between job density and significantly higher mean hourly wages for men on the other.

Policymakers and academics who are interested in the gender wage gap, agglomeration economies, and urbanization may be interested in our results. We emphasize that our results do not say anything explicitly about the gender pay gap because it might reflect broader inequalities in society. Still, they do indicate that women gain less from agglomeration. It is plausible that our results do not hold for other countries with different labor market structures. It would be interesting to apply the methodology introduced to countries with a larger number of cities that vary in size to examine the underlying mechanisms in more detail.

Appendix

2.A Data

Figure A.1: Distributions of the log hourly wages for men (red) and women (black)



Notes: This Figure depicts the Gaussian kernel density distribution of the hourly wages by gender for the period 2008 until 2016

2.B Ignoring individual fixed effects and learning effects

In the case where both individual fixed effects and learning effects are ignored, the econometrician uses the estimation equation

$$w_{a(i,t),it} = \iota_{it}^{\top}\sigma + \mathbf{x}_{it}^{\top}\beta + v_{it}, \tag{B.1}$$

implying that the estimates $(\hat{\sigma}_c, \hat{\sigma}_r)$ of the static area effects can be written as

$$\begin{pmatrix}\hat{\sigma}_c \\ \hat{\sigma}_r\end{pmatrix} = \begin{pmatrix}\sigma_c \\ \sigma_r\end{pmatrix} + (\mathbf{D}^{\top}\mathbf{D})^{-1}\mathbf{D}^{\top}\mathbf{X}(\beta - \hat{\beta}) + (\mathbf{D}^{\top}\mathbf{D})^{-1}\mathbf{D}^{\top}\mathbf{v}, \tag{B.2}$$

where $\mathbf{D}^{\top} = [\mathbf{D}_1^{\top}, ..., \mathbf{D}_N^{\top}]$, $\mathbf{D}_i^{\top} = [\iota_{i1}, ..., \iota_{iT}]$ and with $\mathbf{X}$ and $\mathbf{v}$ being defined analogously to $\mathbf{D}^{\top}$ as stacked matrices of $\mathbf{x}_{it}^{\top}$ and v_{it} respectively. The error term $v_{it} = l_{it} + \mu_i + \epsilon_{it}$ contains both the left-out learning effects and the left-out unobserved fixed effects. To derive an approximate formula of the bias, we assume that $\mathbf{D}$ is orthogonal to $\mathbf{X}$ such that the only source of bias stems from $(\mathbf{D}^{\top}\mathbf{D})^{-1}\mathbf{D}^{\top}\mathbf{v}$. The matrix $\mathbf{D}^{\top}\mathbf{D}$ is diagonal with diagonal terms $\sum_i T(i,a)$ where $T(i,a) := \sum_{t=1}^{T} 1[a(i,t) = a]$ is the duration of time individual i is employed in area a. In this case, the bias can, therefore, be written as

$$bias(\hat{\sigma}_a) = \frac{\sum_{i=1}^{N} T(i,a)\mu_i}{\sum_{i=1}^{N} T(i,a)} + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} 1[a(i,t) = a] l_{it}}{\sum_{i=1}^{N} T(i,a)}, \tag{B.3}$$

so a sum of duration-weighted averages of the unobserved individual characteristic μ_i and duration-weighted averages of learning effects l_{it} (see the following subsection for a more detailed derivation).

In the case where no worker changes area of employment, where $\mu_i = \mu$ for all workers in the city and where $\mu_i = 0$ for workers in the rural area, it follows that

$$\frac{\sum_{i=1}^{N} T(i,a)\mu_i}{\sum_{i=1}^{N} T(i,a)} = \mu.$$

Because no worker changes the area of employment, the city workers have no rural experience such that the rural experience used in the city $\tilde{e}_{rit} = 0$. No migration also implies that their city experience e_{cit} is equal to their city experience used in the city $\tilde{e}_{cit} := 1[a(i,t) = c]e_{cit}$. It therefore follows that the learning effects for a city worker are given as

$$l_{it} = (\lambda_{cr} + \delta_{cc}) e_{cit},$$

in which case the total bias becomes

$$\mu + (\lambda_{cr} + \delta_{cc}) \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} 1[a(i,t) = c] e_{cit}}{\sum_{i=1}^{N} T(i,a)} = \mu + (\lambda_{cr} + \delta_{cc}) \left(\frac{T+1}{2} \right), \tag{B.4}$$

where $(T+1)/2$ is the average number of years of city experience of a worker employed in the city.

Technical note I

To derive the bias, we first consider the estimation of the panel data model with static area-specific effects $\sigma = (\sigma_1, ..., \sigma_J)^\top$ defined by

$$w_{a(it),it} = \mathbf{z}_{it}^\top \theta + u_{it} = \iota_{it}^\top \sigma + \mathbf{x}_{it}^\top \beta + u_{it} \tag{B.5}$$

where $\mathbf{z}_{it}^\top := (\iota_{it}^\top \ \mathbf{x}_{it}^\top)$, $\theta := (\sigma^\top, \beta^\top)^\top$ and $\iota_{it}^\top := (1[a(i,t) = a_1], ..., 1[a(i,t) = a_J])$ is the vector of dummy variables for the location of employment of individual i at time t. Stacking the model in the time index to get

$$\mathbf{w}_i = \mathbf{Z}_i \theta + \mathbf{u}_i = \mathbf{D}_i \sigma + \mathbf{X}_i^\top \beta + \mathbf{u}_i, \tag{B.6}$$

and defining the associated pooled OLS estimator $\hat{\theta}_N$ as the solution to the normal equations

$$\left(\sum_i \mathbf{Z}_i^\top \mathbf{Z}_i \right) \hat{\theta}_N = \left(\sum_i \mathbf{Z}_i^\top \mathbf{w}_i \right), \tag{B.7}$$

it follows by matrix partition that

$$\begin{pmatrix} \sum_i \mathbf{D}_i^\top \mathbf{D}_i & \sum_i \mathbf{D}_i^\top \mathbf{X}_i \\ \sum_i \mathbf{X}_i^\top \mathbf{D}_i & \sum_i \mathbf{X}_i^\top \mathbf{X}_i \end{pmatrix} \begin{pmatrix} \hat{\sigma}_N \\ \hat{\beta}_N \end{pmatrix} = \begin{pmatrix} \sum_i \mathbf{D}_i^\top \mathbf{w}_i \\ \sum_i \mathbf{X}_i^\top \mathbf{w}_i \end{pmatrix}, \tag{B.8}$$

using the equations associated with $\hat{\sigma}_N$ it follows that

$$\left(\sum_i \mathbf{D}_i^\top \mathbf{D}_i \right) \hat{\sigma}_N + \left(\sum_i \mathbf{D}_i^\top \mathbf{X}_i \right) \hat{\beta}_N = \sum_i \mathbf{D}_i^\top \mathbf{w}_i \tag{B.9}$$

where we then substitute $\mathbf{w}_i$ with the model equation to get

$$\hat{\sigma}_N = \sigma + \left(\sum_i \mathbf{D}_i^\top \mathbf{D}_i\right)^{-1} \left(\sum_i \mathbf{D}_i^\top \mathbf{X}_i\right)(\beta - \hat{\beta}_N) + \left(\sum_i \mathbf{D}_i^\top \mathbf{D}_i\right)^{-1} \left(\sum_i \mathbf{D}_i^\top \mathbf{u}_i\right), \tag{B.10}$$

multiplying and dividing with N and taken probability limits under the assumption that $\mathbb{E}[\mathbf{D}_i^\top \mathbf{X}_i] = \mathbf{0}$ it follows that

$$plim\ \hat{\sigma}_N = \sigma + plim \left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \mathbf{D}_i\right)^{-1} \left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \mathbf{u}_i\right). \tag{B.11}$$

The error term $u_{it} = \mu_i + l_{it} + \epsilon_{it}$ when the true model includes individual fixed effects and learning effects as in model (2.2) implying that

$$plim\ \hat{\sigma}_{N,a} = \sigma_a + plim\ \frac{\sum_{i=1}^N T(i,a)\mu_i}{\sum_{i=1}^N T(i,a)} + plim\ \frac{\sum_{i=1}^N \sum_{t=1}^T 1[a(i,t) = a] l_{it}}{\sum_{i=1}^N T(i,a)}, \tag{B.12}$$

under the assumption that $\mathbb{E}[\mathbf{D}_i^\top \epsilon_i] = \mathbf{0}$ where $T(i,a) := \sum_{t=1}^T 1[a(i,t) = a]$.

In the more general case where the assumption that $\mathbb{E}[\mathbf{D}_i^\top \mathbf{X}_i] = \mathbf{0}$ is not imposed, the expression for the probability limit includes further the term

$$-\left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \mathbf{D}_i\right)^{-1} \left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \mathbf{X}_i\right)(plim\ \hat{\beta}_N - \beta). \tag{B.13}$$

2.C Ignoring learning effects

In the second case, the econometrician estimates the equation

$$w_{a(i,t),it} = \iota_{it}^\top \sigma + \mu_i + \mathbf{x}_{it}^\top \beta + u_{it}, \tag{C.1}$$

controlling for the unobserved individual fixed effects while still ignoring the learning effects. In this case, the city-fixed effects are only identified by the workers who change their area of employment. This can be seen by using dot notation $\dot{z}_{it} := z_{it} - (1/T)\sum_t z_{it}$ and writing the wage equation time demeaned

$$\dot{w}_{a(i,t),it} = \sum_a \sigma_a(\iota_{ait} - \bar{\iota}_{ai}) + \dot{\mathbf{x}}_{it}^\top \beta + \dot{u}_{it}, \tag{C.2}$$

where $\iota_{ait} := 1[a(i,t) = a]$ such that for any worker not changing area of employment $(\iota_{ait} - \bar{\iota}_{ai}) = 0$. Because the time de-meaned learning effects are contained in the error term $\dot{u}_{it} = \dot{l}_{it} - \dot{\epsilon}_{it}$ the bias will in general depend on the time de-meaned value of experience $\dot{l}_{it}$. Specifically, it can be shown that the probability limit of the estimate of the city fixed effect is given as

$$plim\ \hat{\sigma}_c = \sigma_c + plim\left(\frac{1}{N}\sum_i T(i,c)\left(1-\frac{T(i,c)}{T}\right)\right)^{-1}\left(\frac{1}{N}\sum_i \mathbf{d}_i^\top \dot{\mathbf{l}}_i\right), \tag{C.3}$$

the derivation of which is given in the following subsection.

Consider the first case where migration is from the rural area to the city. For the first m periods, the workers are employed in the rural area, accumulating one year of experience each year. Their rural experience is, therefore, given as

$$e_{rit} = \begin{cases} t & \text{if } t \le m \\ m & \text{if } t > m \end{cases},$$

and their city experience is given as

$$e_{cit} = \begin{cases} 0 & \text{if } t \le m \\ t-m & \text{if } t > m \end{cases}.$$

Since the workers have no city experience prior to migration, the city experience used in the city $\tilde{e}_{cit}$ is equal to the city experience e_{cit} such that the learning effects become

$$l_{it} = (\lambda_{cr} + \delta_{cc})e_{cit} + \delta_{cr}\tilde{e}_{rit}.$$

The sum of the time-demeaned learning effects while working in the city $\mathbf{d}_i^\top \dot{\mathbf{l}}_i$ can therefore be written as the sum of the two components of experience

$$(\lambda_{cr} + \delta_{cc})\sum_{t=1}^{T} 1[a(i,t)=c]\dot{e}_{cit} + \delta_{cr}\sum_{t=1}^{T} 1[a(i,t)=c]\dot{\tilde{e}}_{rit}.$$

The first component $\sum_{t=1}^{T} 1[a(i,t)=c]\dot{e}_{cit}$ is equal to $\sum_{t=1}^{T} 1[a(i,t)=c]e_{cit} - ((T-m)/T)\sum_{t=1}^{T} 1[a(i,t)=c]e_{cit}$. Isolating the factor $\sum_{t=1}^{T} 1[a(i,t)=c]e_{cit} = (T-m)(T-m+1)/2$ it follows that

$$\sum_{t=1}^{T} 1[a(i,t)=c]\dot{e}_{cit} = \frac{(T-m+1)}{2}(T-m)\left(1-\frac{T-m}{T}\right).$$

For the rural experience used in the city the component $\sum_{t=1}^{T} 1[a(i,t)=c]\dot{\tilde{e}}_{rit} = \sum_{t=1}^{T} 1[a(i,t)=c]\tilde{e}_{rit} - ((T-m)/T)\sum_{t=1}^{T} 1[a(i,t)=c]\tilde{e}_{rit}$. Now isolating the factor $\sum_{t=1}^{T} 1[a(i,t)=c]\tilde{e}_{rit} = m(T-m)$ and use that in all the $(T-m)$ periods where rural experience is used in the city, it is m years of experience determined as the number of periods of previous employment in the rural area. We, therefore, find that.

$$\sum_{t=1}^{T} 1[a(i,t)=c]\dot{e}_{rit} = m(T-m)\Big(1-\frac{T-m}{T}\Big).$$

Dividing this with the factor $T(i,c)\Big(1-\frac{T(i,c)}{T}\Big)=(T-m)\Big(1-\frac{T-m}{T}\Big)$ it follows that the bias is given as

$$bias(\hat{\sigma}_c) = (\lambda_{cr}+\delta_{cc})\frac{(T-m+1)}{2} + \delta_{cr}m, \tag{C.4}$$

showing that the workers who migrate to the city will have a value of experience greater than their average value of experience under the assumption that $(\lambda_{cr}+\delta_{cc})$ and δ_{cr} are both positive.

Next, we consider the case where workers are employed for the first m periods in the city and the latter $T-m$ periods in the rural area. Because the workers have no rural experience while working in the city, the learning effects while employed in the city can be written as

$$\dot{l}_{it} = \lambda_{cr}\dot{e}_{cit} + \delta_{cc}\dot{\tilde{e}}_{cit}.$$

It follows that

$$\sum_{t=1}^{T} 1[a(i,t)=c]\dot{l}_{it} = \lambda_{cr}\sum_{t=1}^{T} 1[a(i,t)=c]\dot{e}_{cit} + \delta_{cc}\sum_{t=1}^{T} 1[a(i,t)=c]\dot{\tilde{e}}_{cit}.$$

Dividing both sums $\sum_{t=1}^{T} 1[a(i,t)=c]\dot{\tilde{e}}_{cit} = \frac{m(m+1)}{2}(1-m/T)$ and $\sum_{t=1}^{T} 1[a(i,t)=c]\dot{e}_{cit} = \frac{m(m+1)}{2}(1-m/T)-\frac{m^2(T-m)}{T}$ with the factor $T(i,c)\Big(1-\frac{T(i,c)}{T}\Big)=m(1-m/T)$ it follows that the total bias is

$$bias(\hat{\sigma}_c) = (\lambda_{cr}+\delta_{cc})\frac{m+1}{2} - \lambda_{cr}m, \tag{C.5}$$

which is negative if and only if $\theta > 1/2+1/2m$ where $\theta := \lambda_{cr}/(\lambda_{cr}+\delta_{cc})$ is the share of value of experience that is portable. If the value of experience is highly portable, it is possible that workers, while working in the city, have a lower value of experience than their average value of experience. The portability of the value of experience implies that when workers change areas of employment, they do not experience a large drop in the value of their experience. On the other hand, if the experience is not portable, then workers who initially work in the city experience a fast growth in wages and may be reaching a peak level. They then migrate to the rural area and experience a large drop in their wages never again reaching the peak level. This implies that these workers, while in the city, have a value of experience above their average value of experience and hence they contribute positively to the bias of the city fixed effect.

Technical note II

For the case where the estimation is done with individual fixed effects the derivations of the bias are very similar. We start by defining the symmetric, idempotent, rank $T-1$ demeaning matrix $Q = I_T - \iota_T(\iota_T^\top \iota_T)^{-1}\iota_t^\top$ and premultiplying the stacked version of the model with Q to get

$$Q\mathbf{w}_i = Q\mathbf{Z}_i\theta + Q\mathbf{u}_i = Q\mathbf{D}_i\sigma + Q\mathbf{X}_i\theta + Q\mathbf{u}_i, \tag{C.6}$$

using dot notation $\dot{\mathbf{A}} = Q\mathbf{A}$ the model can then be written as

$$\dot{\mathbf{w}}_i = \dot{\mathbf{Z}}_i\theta + \dot{\mathbf{u}}_i = \dot{\mathbf{D}}_i\sigma + \dot{\mathbf{X}}_i\theta + \dot{\mathbf{u}}_i. \tag{C.7}$$

Using derivations similar to the case without individual fixed effects, it can be shown that

$$plim\ \hat{\sigma}_N = \sigma + plim\left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \dot{\mathbf{D}}_i\right)^{-1}\left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \dot{\mathbf{u}}_i\right), \tag{C.8}$$

under the assumption that $\mathbb{E}[\dot{\mathbf{D}}_i^\top \mathbf{X}_i] = \mathbb{E}[\mathbf{D}_i^\top \dot{\mathbf{X}}_i] = \mathbf{0}$. The matrix $\sum_i \mathbf{D}_i^\top \dot{\mathbf{D}}_i$ is not invertible if all areas are included; however, this is easily fixed by dropping a row of $\mathbf{D}_i$ using one area as a reference level.

Before simplifying the expression for the bias further, we note that in the more general case where the assumption that $\mathbb{E}[\dot{\mathbf{D}}_i^\top \mathbf{X}_i] = \mathbf{0}$ is not imposed, the expression for the probability limit includes further the term

$$-\left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \dot{\mathbf{D}}_i\right)^{-1}\left(\frac{1}{N}\sum_i \mathbf{D}_i^\top \dot{\mathbf{X}}_i\right)(plim\ \hat{\beta}_N - \beta). \tag{C.9}$$

Simplification of the bias under the assumption $\mathbb{E}[\mathbf{D}_i^\top \dot{\mathbf{X}}_i] = \mathbf{0}$ is still more challenging than for the case without individual fixed effects because the matrix $\mathbf{D}_i^\top \dot{\mathbf{D}}_i$ is not diagonal. However, in the two-area case, the matrix reduces to a scalar because a row corresponding to one area is removed. In this case, the probability limit of the area fixed effect for the area not used as a reference level is given as

$$plim\ \hat{\sigma}_{N,a} = \sigma_a + plim\ \left(\frac{1}{N}\sum_i \mathbf{d}_i^\top \dot{\mathbf{d}}_i\right)^{-1}\left(\frac{1}{N}\sum_i \mathbf{d}_i^\top \dot{\mathbf{u}}_i\right), \tag{C.10}$$

where $\mathbf{d}_i = (1[a(i,1)=a],...,1[a(i,T)=a])^\top$. The scalar $\mathbf{d}_i^\top \dot{\mathbf{d}}_i$ can also be written as

$$\mathbf{d}_i^\top \dot{\mathbf{d}}_i = \sum_{t=1}^{T} 1[a(i,t)=a]\left(1[a(i,t)=a] - (1/T)\sum_{t=1}^{T} 1[a(i,t)=a]\right) \tag{C.11}$$

$$= T(i,a)\left(1 - \frac{T(i,a)}{T}\right), \tag{C.12}$$

from which it follows that if we consider a worker who works in the city for the first m periods and changes employment location to the rural area for the last $T-m$ periods, then

$$\mathbf{d}_i^\top \dot{\mathbf{d}}_i = m\left(1 - \frac{m}{T}\right), \tag{C.13}$$

when the rural area is used as a reference. Analogously, if the worker is employed for the first m periods in the rural area and then changes job location to the city for the last $T-m$ periods, then

$$\mathbf{d}_i^\top \dot{\mathbf{d}}_i = (T-m)\left(1 - \frac{(T-m)}{T}\right). \tag{C.14}$$

still assuming the rural area is used as a reference.

CHAPTER 3

Transfer taxes and housing affordability: Insights for markets with a substantial rental sector

Malte Borghorst

Abstract: This study examines the impact of real estate transfer taxes on housing affordability in a market with a substantial rental share. Using data from the whole of Germany, I exploit spatial and temporal variation in the tax rates across states and find that for each one-percentage-point increase in the tax rate, house prices decrease by 4.4%, and rents drop by 1.7%. The effects vary, with no discernible impact in urban or rapidly growing markets, but more than double the average effects in rural or declining markets.

3.1 Introduction

The real estate transfer tax is an ad valorem tax, levied on the buyer side. It has been the subject of frequent policy debate as policymakers grapple with its impact on housing affordability in particular lock-in effects for house owners, and increased costs for first-time buyers (Mirrlees et al., 2010). It also constitutes a significant source of revenue for local governments. For instance, in Germany, it comprised 60% (€15.8 billion) of state (Bundesländer) taxes in 2019[1]. Despite the tax's economic relevance, empirical studies on the effect of the transfer tax on housing affordability are relatively rare, and most research focuses on countries with predominantly owner-occupancy markets, overlooking the existence of rental markets. The rental market is important because here investors are better at mitigating the costs of transfer taxes (see i.e. Han et al. (2022)). Investors have deeper pockets for financing, can adjust holding periods, and theoretically could pass on tax increases into the rental market. In the Literature empirical evidence for the effects of the transfer tax on housing affordability taking into account rental markets is scarce and evidence of the role of the transfer tax in rental markets is just emerging (Schindlbauer, 2020).

This study fills this gap by estimating the impact of transaction taxes on housing affordability (house prices and rents) in a nationwide market with a substantial rental share. For this purpose, the German housing market stands out with a rental market share of 50.5%[2] and investors can use tax loopholes share deals to avoid paying the tax[3]. The real estate transfer tax rates in Germany vary across states and over time, offering important variation for the identification of the impact on the housing market. In 2006 a reform granted states the authority to set their own tax rate, many of which subsequently raised these taxes. Tax rates in Germany, ranging from 3.5% to a maximum of 6.5%, increased periodically in our sample period from 2007 to 2019. Importantly, these rate hikes were contingent on state-level decisions, ensuring their exogeneity to local housing markets. This contrasts with settings where municipal financial constraints can affect local taxes, such as the Mansion tax in New York (Kopczuk and Munroe, 2015) and the Land Transfer Tax in Toronto (Dachis et al., 2012). Notably, the tax increase did not impose additional administrative burdens on taxpayers or government agencies, as the tax predated the reform.

[1] Bundesfinanzministerium Steuereinnahmen nach Steuerarten 2000 – 2022.

[2] See EU-SILC survey data, EUROSTAT (2022)

[3] The design of the German transfer tax system allows investors to exploit a loophole to entirely avoid the tax by consolidating properties in legal entities and acquiring a share of the entity instead of the asset itself.

To estimate the effect of the transfer tax on house prices and rents, I first construct a house price index which keeps monthly municipal level variation, using data from the largest German internet platform for housing listings, Immoscout[4]. I define movers as municipalities in states that changed the tax rates and stayers as municipalities in the two states that kept it at 3.5% over the whole period of observation, Bavaria and Saxony. The estimation thus compares movers to stayers in a Two-Way-Fixed-Effects (TWFE) estimator on the continuous tax rates.

I find that house prices respond significantly to variations in the transfer tax rate, experiencing a substantial 4.4% decline for each one-percentage-point increase in the tax rate[5]. In a novel contribution, I find rents decrease by 1.7%. Importantly, these effects show no anticipation and vary considerably, with Urban and growing markets showing minimal (zero) effect and rural or declining areas more than double the average marginal effect.

This study contributes to the empirical literature which documents a substantial incidence of transfer taxes on sellers, consistently finding significant negative effects on asking prices by confirming these findings and extending it by providing new evidence, that these results hold even in the presence of a substantial rental market (Van Ommeren and Van Leuvensteijn, 2005; Dachis et al., 2012; Best and Kleven, 2018; Besley et al., 2014; Kopczuk and Munroe, 2015; Davidoff and Leigh, 2013). Most empirical studies focus on the owner-occupier market, neglecting the rental market. In the case of owner-occupier market markets, a mover is searching for a new house while offering their old house simultaneously, thus facing both sides of the market at the same time. To address this, Wheaton (1990) and Lundborg and Skedinger (1999) developed a search model for the housing market[6]. Lundborg and Skedinger (1999) predict, that the effect of the tax depends on whether it is levied on the buyer or the seller side. The intuition of their model is quite simple. If the ad valorem tax is raised on the buyer side (as is the case in Germany), buyers need to be compensated by lower ask prices. However, many markets differ from the

[4]DOI 10.7807/immo:red:wk:suf:v7 www.Immoscout24.de

[5]The published paper that I am aware of is Fritzsche and Vandrei (2019) who use a smaller sample of transaction survey data on selected cities across some German states and document that the tax led to bunching and decreased transaction numbers, but find no effect on prices. However, working papers exist: Petkova and Weichenrieder (2017) show price and quantity effects using state-level data, Frenzel Baudisch and Dresselhaus (2018) on the commercial real estate market, Dolls et al. (2021) price effects and tax burden, Christofzik et al. (2020) on heterogeneity. The size of the effect is consistent with the international literature (Van Ommeren and Van Leuvensteijn, 2005; Dachis et al., 2012; Best and Kleven, 2018; Besley et al., 2014; Kopczuk and Munroe, 2015; Davidoff and Leigh, 2013).

[6]These models face the same criticism as Mortensen and Pissarides (1994) as pointed out by Van Ommeren and Van Leuvensteijn (2005)

owner-occupier case because of a substantial rental sector. To address this, Han et al. (2022) extent the search models by Wheaton (1990); Lundborg and Skedinger (1999) and show that the prediction of lower housing prices as a response to a transfer tax, remains valid if investors and a rental market are included. My results confirm this hypothesis.

This study is the first to document the effect of the transfer tax on rents in a housing market with a substantial rental share, like Germany. Han et al. (2022) model the decision of owner-occupiers who move to a new house to rent out their old house instead of selling and becoming investors. They point out, that in rental markets, the role of investors is important because investors are generally better at mitigating the tax and profit from the lower ask prices. In their model, investors do not pay the tax when a tenant moves, resulting in longer holding periods, and for each transaction, they pay only one market side of the burden. Investors also have deeper pockets for financing, can adjust holding periods, and theoretically could pass the prices on into the rental market, depending on the regulation. However, thick rental markets, like Germany's, are mainly comprised of institutional investors (investors with more than one rental property) who face different incentives and constraints than owner-occupiers decide to become an investor. Furthermore, the existence of loopholes for institutional investors in the German tax design implies that the transfer tax falls more heavily on owner-occupiers than on buy-to-rent investors, placing them on uneven footing, which could further increase the share of the rental market. All these factors could increase the supply of rental units and lead to congestion and market thickness externalities that arise with search frictions. When a seller posts a vacancy, she does not internalize that by doing so she is making it harder for other sellers to find a buyer (congestion externality) while making it easier for buyers to find a home (thick market externality), (Hosios, 1990; Pissarides, 2000; Gabrovski and Ortego-Marti, 2021). Although Han et al. (2022) predict higher rents in a mainly owner-occupier market, in a thick rental market with institutional investors, these externalities and the strong incentives for an increase of rental supply means, that direction of the effect is less clear. Even though identifying the channel is out of the scope of this project, this paper is the first to exploit timing and spatial variation in the transfer tax rates to measure the impact on municipality-level rents characterized by a high rental share. The results show a negative effect on rents of 2% with substantial heterogeneity across market thickness and structure.

This study further extends the empirical literature by estimating heterogeneous effects across multiple labor and housing markets. The intuition of search models predicts that the size of the effect of the transfer tax on house prices and rents depends on the search costs depending on market thickness and demand structure, see i.e. for an analysis of the Beveridge curve in the housing marketGabrovski and Ortego-Marti (2019). The comparison of heterogeneity in regional dynamics to capture demand effects, and heterogeneity across urban and rural areas as measures of market thickness adds a novel dimension to the evolving empirical literature on heterogeneous effects of the transfer tax, see i.e. Poulhès et al. (2020) and Christofzik et al. (2020). Estimating heterogeneous effects for house prices and rents, I find that thicker markets like cities and stable-demand regions with strong growth show zero effects, while in rural and shrinking markets prices and rents react very strongly (twofold).

The exploration of the heterogeneous effects also improves identification compared to similar studies using Difference-in-Difference (DiD) or Two-Way-Fixed-Effects (TWFE) (Fritzsche and Vandrei, 2019; Petkova and Weichenrieder, 2017; Frenzel Baudisch and Dresselhaus, 2018; Dolls et al., 2021; Christofzik et al., 2020). In the context of DiD or TWFE, the econometric literature discusses concerns about the violations of the homogeneous treatment assumption in the presence of staggered adoption and differnces in treatment intensity[7]. The German real estate transfer tax is additionally characterized by continuous and multiple treatments[8]. This study addresses concerns regarding the homogeneous treatment assumption by documenting robustness in the total marginal effects, taking into account heterogeneity by density and regional growth dynamics.

This study contributes to the existing literature on the timing of the reactions to the tax (Fritzsche and Vandrei, 2019; Petkova and Weichenrieder, 2017; Frenzel Baudisch and Dresselhaus, 2018; Dolls et al., 2021; Christofzik et al., 2020) all present timing effects, by estimating the dynamic effects of the tax changes on house prices. For this purpose, I estimate an event study over the period 6 months before and 12 months after individual tax increases, interacting the effects with indicators for multiple treatments. I show that the effects are immediate and persist for at least 5 months, while no anticipation effects in prices or rents exist. These findings also verify the exogeneity and common trend assumption.

Finally, I document further heterogeneity by studying the effect of the transaction tax on single houses. This is a market segment most likely reflecting the owner-

[7]See i.e., Callaway and Sant'Anna (2021); De Chaisemartin and d'Haultfoeuille (2020); Borusyak et al. (2021); Goodman-Bacon (2021); Sun and Abraham (2021); Athey and Imbens (2006) and for a survey see, i.e., de Chaisemartin et al. (2022)

[8]To address these concerns, the literature and estimation procedures are just evolvingde Chaisemartin et al. (2022) suggest a solution in the case of continuous and multiple treatment, which only relies on the parallel trends assumption and that the treatment is never lower than in period one, which might offer a solution in the future.

occupier case and the only published study on the German housing market uses transaction survey data on single houses (Fritzsche and Vandrei, 2019). In line with their results, I find no effect of the tax on single houses but substantial heterogeneity across density and regional dynamics. As for the whole market, the transfer tax shows no effect on house prices of single single houses in cities and the fastest growing regions but in rural and shrinking regions the effect is up to -6% per percentage point increase in the tax rate. While this confirms the pattern from the whole market, the effects are substantially smaller and are more comparable to the effect on rents.

The paper proceeds as follows. Section 3.2 presents the institutional context, section 3.3 describes the data, section 3.4 describes the empirical strategy, section 3.5 discusses the results and section 3.6 concludes.

3.2 Institutional context and identification

In Germany, the real estate transfer tax dates back to the Danube monarchies of the 19th century (MacGregor Pelikánová and Jánošíková, 2017) and is an ad valorem tax, levied on the buyer side. It is the most important independent source of revenue for the German states and throughout this study, it increased from around 30% (6.9 bn.€, 2007) to 60% (15.8 bn.€, 2019)[9]. The real estate transfer tax is often criticized for its impact on housing affordability and mobility (Mirrlees et al., 2010). In the German context, ongoing parliamentary debates and studies highlighted concerns about the tax's design, particularly its investor loopholes, lock-in effects, and increased costs for first-time buyers (Bechtoldth et al., 2014; Zander and Faller, 2006; Voigtländer et al., 2013).

This paper exploits temporal and spatial variation following a policy shift that occurred in 2006 when German states gained the authority to adjust the residential transfer tax rate, which was initially set at 3.5%. Over the following years, many states independently raised the transfer tax rate by varying amounts leading to variations in timing and extent across states, see Fig. 3.1. The maximum 3-percentage point tax hike (i.e., from 3.5% to 6.5%) may seem small relative to the total purchase price. Still, in Germany, as in most countries, mortgage rates significantly improve with a 20% deposit. This means that for a buyer aiming for this 20% threshold and facing liquidity constraints, the house price is effectively capped at five times their savings net transfer tax payments (since transfer taxes cannot be mortgaged). Consequently, the house price responds directly and proportionally, by a factor of five, to changes in the transfer tax. Additionally banks might give different financing conditions depending on the market conditions (e.g. constraints for financing are more constraint in declining areas). For more insights into finance constraints and timing regarding the transfer tax, see Best and Kleven (2018).

The variability in tax rates resulting from this policy change offers a unique opportunity to estimate the effects of the transfer tax on the housing and rental markets, because of its temporal and spatial variation in the tax rate. In the German design, these tax increases are contingent on state-level decisions, rather than municipal financial constraints, and are therefore plausibly exogenous to local housing markets. This improves identification and sets the German context apart from settings, where the tax levied at the municipal level (or school district level as in some parts of the US), thus can be impacted by municipal financial constraints and, worse, dynamics on the housing market[10]. Furthermore, this policy change didn't introduce added

[9] Bundesfinanzministerium Steuereinnahmen nach Steuerarten 2000 – 2022

[10] For analysis and identification in these contexts see i.e. New Yorks Mansion tax (Kopczuk and Munroe, 2015) and the Land Transfer Tax in Toronto (Dachis et al., 2012).

Figure 3.1: Transfer tax rate by state and year

States/year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Schleswig-Holstein	3.5%						5.0%		6.5%							
Hamburg	3.5%			4.5%												
Lower Saxony	3.5%					4.5%			5.0%							
Bremen	3.5%					4.5%			5.0%							
North Rhine-Westphalia	3.5%						5.0%			6.5%						
Hesse	3.5%							5.0%	6.0%							
Rhineland-Palatinate	3.5%						5.0%									
Baden-Wurttemberg	3.5%						5.0%									
Bavaria	3.5%															
Saarland	3.5%					4.0%		5.5%		6.5%						
Berlin	3.5%	4.5%					5.0%		6.0%							
Brandenburg	3.5%					5.0%				6.5%						
Mecklenburg-Vorpommern	3.5%						5.0%							6.0%		
Saxony	3.5%															
Saxony-Anhalt	3.5%				4.5%		5.0%									
Thuringia	3.5%					5.0%						6.5%				

Note: Illustration purpose only - the analysis uses monthly data.

administrative burdens for taxpayers or government agencies, as the tax existed before the reform.

Unlike countries with a predominant owner-occupier market characterized by frequent housing upgrades, the primary reasons for purchasing a house in Germany mainly revolve around long commutes and starting a family, resulting in a predominantly family-oriented and long-tenured owner-occupier demographic (Bechtoldth et al., 2014; Zander and Faller, 2006). Estimating the impact of transfer tax changes on long-distance residential moves across labor markets is challenging. This is due to the simultaneous changes in house prices, labor market conditions, and amenities, making it difficult to isolate tax-related relocations. Additionally, in Germany, residential mobility data is limited and relies on sporadic census data or household surveys with a small sample of movers, further complicating the analysis. However, Hilber and Lyytikäinen (2017) found that even in markets with a high share of owner-occupancy, like the UK, there were no identifiable reactions to transfer tax changes for long-distance moves. For this study, I will assume that the tax changes will not prompt moves across states. The share of first-time buyers of total transactions in the housing market is likely on the lower side, with investors being responsible for the main share of the market. While transaction data for the whole market is not available, there are an estimated 1 million total transfers per year, with about 500,000 of these being owner-occupier transfers, including first-time buyers. This figure also includes subsidies for new construction (Voigtländer et al., 2013)[11].

[11] In the Immoscout sample, I observe around 1,000,000 objects for sale and 1,000,000 objects for rent per year.

The low number of real owner-occupier transfers underscores the importance of the rental market in Germany.

3.3 Data

Data on house prices are obtained from the online platform Immoscout24[12], as provided by the FDZ of the RWI[13]. The dataset contains ask prices and characteristics of houses or apartments as posted by the seller[14]. The average price of an object is 274,900 €(sd: 266,448 €) with an average floor space of 131 m^2 (sd 70m^2) resulting in a price of 2,127 €per m^2 (sd: 1,414 €per m^2). The average rent is 7.7 €per m^2 (sd 3 €per m^2). Han and Strange (2016) show that ask prices posted on websites close proxies for the realized transfer prices[15]. Unlike other (administrative) data sources the structure of this particular type of data consists of repeated cross-sections of listed units. The duration of ownership for these units is usually quite long, using a panel structure on the objects would lead to a skewed sample because regularly listed units do not accurately represent the current market. Additionally, there's a possibility that object IDs might have been reused for similar apartments within the same buildings or different houses by the same agency, and characteristics are not entered consistently and often missing. To improve comparability while preserving regional variation, I restructure the data in the form of a quality-adjusted, monthly municipal House Price Index (HPI), using the following equation,

$$\ln \mathbf{P_{hit}} = \alpha \ln d_{it} + \mathbf{X}'_{hit}\beta + \delta_{it} + \varepsilon_{hit}. \tag{3.1}$$

where P_{hit} represents the price of the house (€ per m^2), X_{hit} includes a set of house characteristics[16], and δ_{it} denotes municipality-year-month fixed effects which are used for the creation of the index. Consequently, the log hedonic price index in €per m^2 is adjusted for inflation using the German consumer price index (CPI).[17] For the regional controls, I use labor market regions (following Kosfeld and Werner (2012)) and classify them based on their settlement structure dynamics and urbanization[18].

[12] www.immobilienscout24.de

[13] DOI 10.7807/immo:red:wk:suf:v7

[14] See Schaffner (2020) for a description and the initial preparation (i.e. removing duplicates) and it can be prepared using the code provided by Beze and Gutschlhofer (https://github.com/eyayaw/cleaning-RWI-GEO-RED)

[15] For a discussion in the context of the Immoscout data see Ahlfeldt et al. (2023).

[16] Housing characteristics include the distance of the object to the city center, floor space, number of rooms, the age of the house, categories for the number of bedrooms, bathrooms, and floors, the housing type, holiday houses, indicators for heating type, basement, guest washroom, construction phase, equipment, condition, balcony, garden, kitchen, floor

[17] https://www-genesis.destatis.de/genesis/online?sequenz=statistikTabellen&selectionname=61111&language=en

[18] As in Bundesinstitut für Bau-, Stadt und Raumforschung (BBSR), 2023) https://www.inkar.de/

Categorization into growing and shrinking regions follows an index by Häußermann and Siebel (2004) and consists of five categories: strong growth, growing, stagnating, shrinking and strongly shrinking. Because of the regional fixed effects, it is important to keep this categorization constant over time, such that the effects are not identified by category movers. I therefore choose to keep it constant at the pre-treatment level (2006). The labor market region density is categorized into three types: cities, rural, and countryside[19] and the average disposable household income is calculated on the district level (Kreis). Distances to the closest border and illustration are based on geoinformation from the Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie, BKG) and are calculated as the distance from the centroid to the closest border.

3.4 Empirical strategy

The institutional framework of the German real estate transfer tax divides states into two categories: "stayers" (Saxony and Bavaria) maintaining a tax rate of 3.5%, and "movers" (all other states) with tax rate increases ranging from 0.5 to 2.5 percentage points. These changes occur at staggered adoption times, often leading to multiple treatment scenarios with varying treatment intensity. Because these tax increases are determined at the state level, rather than being influenced by municipal financial constraints, they are plausibly exogenous to local housing markets. Moreover, these tax increases did not introduce additional administrative burdens for taxpayers or government agencies, as the tax existed before the reform, making the treatment well-defined by the changes in tax rates.

This spatial and temporal variation in tax rates allows for a comparison between movers and stayers using a Two-Way-Fixed-Effects (TWFE) approach. However, the presence of staggered adoption and variations in treatment intensity raises concerns about the violation of the homogeneous treatment assumption. While the econometric literature provides solutions for testing treatment heterogeneity and staggered adoptions, the German real estate transfer tax presents additional challenges due to its continuous treatment intensity variation and multiple treatments[20]. The German real estate transfer tax is additionally characterized by variation in *continuous* treatment intensity and multiple treatments. To address these concerns, the literature and estimation procedures are just evolving. de Chaisemartin et al. (2022) suggests

[19] As in BBSR (2023) `https://www.inkar.de/`

[20] For a discussion and solutions see Callaway and Sant'Anna (2021); De Chaisemartin and d'Haultfoeuille (2020); Borusyak et al. (2021); Goodman-Bacon (2021); Sun and Abraham (2021); Athey and Imbens (2006), for a survey see i.e. de Chaisemartin et al. (2022).

a solution in the case of continuous and multiple treatments, which only relies on the parallel trends assumption and that the treatment is never lower than in period one, which might offer a solution in the future. For this project, I document the robustness of total marginal effects while accounting for heterogeneity in density and regional growth dynamics.

To study the effect of the transfer tax (τ) on the house price index from Eq. 3.1: $\hat{\delta_{it}}$ I employ following regression

$$\hat{\delta_{it}} = \gamma \cdot \tau + \mathbf{X'}_{it}\kappa + \theta_i + \omega_t + \nu_{it} \tag{3.2}$$

I control for state fixed effects θ_i and time fixed effects, ω_t (year and calendar month), such that the coefficient γ is the effect of a 1 percentage point increase in the tax rate of a mover, compared to the tax rate staying at 3.5%. The time-fixed effects further control for different price levels at the national level (i.e. the financial crisis of 2008) and the state-fixed effects reflect unobserved differences across states such as amenities and purchasing power. Standard errors are clustered at the state level, which is the level of treatment. Additional controls X_{it} include the average disposable household income at the districts (Kreis) level and labor markets specific time trends categorized by five growth dynamic sub-classifications.

However, it is essential to discuss the common trends and stable unit treatment value assumption (SUTVA). While the SUTVA would hold for properties themselves since they are immobile and changing i.e. the size or other features to adapt to price changes are costly and take time, the market spillover effects close to borders are more concerning. A person planning to buy property in a treated labor market close to the border could choose to commute across the state border. I therefore validate the estimates with a buffer zone, where I exclude the houses within 10km if a state border[21].

To test for common trends and no anticipation, I employ a dynamic event study approach, as in Kleven et al. (2019b). I narrowed down the sample to a window of 6 months before the tax change and one year afterward ($e = [-6, 12]$), ensuring that there is no overlap between multiple treatment periods. I use a regression model similar to the one outlined in Equation 3.2, as follows:

$$\hat{\delta_{it}} = \sum_{e \neq -1}^{e} \gamma_e \tau_{it} + \mathbf{X}_{it}\kappa + \omega_t + \nu_{it} \tag{3.3}$$

In this Event study, I estimate event time indicators (γ_e) for each month in the sample

[21] Dachis et al. (2012) find a setup to exploit this effect by employing a difference in difference setup in combination with a regression discontinuity design at the border, but for this paper I will address the market spillovers by creating buffer zones around the borders.

window, except for the month immediately preceding the tax change ($\gamma_e \neq -1$). Compared to the previous estimation Eq. 3.2, I omit the state fixed effects (θ_i), which allows the results to be interpreted as changes in levels compared to the month just before the tax change. Additional controls X_{it} are as in the Eq. 3.2.

3.5 Results

Most empirical studies focus on the owner-occupier market, neglecting the rental market. In the case of owner-occupier market markets, a mover is searching for a new house while offering their old house simultaneously, thus facing both sides of the market at the same time. To address this, Wheaton (1990) and Lundborg and Skedinger (1999) developed a search model for the housing market[22]. In the prediction of the price reaction to a transfer tax, Lundborg and Skedinger (1999) distinguishes between a tax on the buyer and the seller side. The intuition of their model is quite simple. If the ad valorem tax is raised on the buyer side (as is the case in Germany), buyers need to be compensated by lower ask prices. However, many markets differ from the owner-occupier case because of a substantial rental sector of around. To address this, Han et al. (2022) extent the search models by Wheaton (1990); Lundborg and Skedinger (1999) and show that the prediction of lower housing prices as a response to a transfer tax, remains valid if investors and a rental market are included.

The results in Table 3.1 col. 1 and 4 confirm this hypothesis and show a very stable quasi-elasticity of -.04. This implies that an increase of the real estate transfer tax by 1 percentage point leads to a reduction in asking prices of approximately 4%. This strong negative effect is in line with estimates in the literature[23]. An average increase of 2.5% in the property tax would then decrease the house prices by around 10% or 27k at an average price of 276,000€. While this sounds like a large effect, note that the tax directly translates into the costs associated with the transfer. If the tax increases the price on the buyer side, the buyer has to increase their savings. If we estimate a savings amount of 20-30% to cover downpayment, agent fees, and taxes (which cannot be borrowed) a 3% increase would translate to a 10% increase in savings required for the transaction. When a 10% increase in savings only translates into a 10% decrease in asking prices, these estimates are on the lower side. In section 3.5.3, I show that this amount varies by market structure and real estate type.

In this analysis, I normalized square meter prices using the municipal-level House

[22] These models face the same criticism as Mortensen and Pissarides (1994) as pointed out by Van Ommeren and Van Leuvensteijn (2005).

[23] See e.g. Van Ommeren and Van Leuvensteijn (2005); Dachis et al. (2012); Best and Kleven (2018); Besley et al. (2014); Kopczuk and Munroe (2015); Davidoff and Leigh (2013).

Table 3.1: Effect of the transfer tax on prices - heterogeneity and rents

	(1)	(2)	(3)	(4)	(5)	(6)
	baseline	effect heterogeneity			single	
dep. var.: ln(HPI)		growth	density	both	houses	rents
total marginal	-.041***	-.034***	-.051***	-.044***	-.019	-.017*
effect of tax	(.009)	(.010)	(.015)	(.012)	(.011)	(.007)
transfer tax (in %)						
× fast shrinking		-.145***		-.108***	-.064**	-.031
		(.013)		(.014)	(.023)	(.019)
× shrinking		-.114***		-.101***	-.049**	-.046**
		(.014)		(.016)	(.018)	(.016)
× stagnant		-.060**		-.064**	-.027*	-.027
		(.018)		(.020)	(.012)	(.016)
× growing		-.043**		-.054***	-.025*	-.020**
		(.012)		(.013)	(.011)	(.006)
× fast growing		.005		-.010	.007	-.002
		(.010)		(.010)	(.011)	(.013)
transfer tax (in %)						
× cities			-.033*	-.030**	-.004	-.012
			(.014)	(.010)	(.011)	(.007)
× rural			-.088***	-.072***	-.038**	-.033**
			(.019)	(.017)	(.013)	(.012)
× countryside			-.116***	-.095***	-.055***	-.057***
			(.026)	(.020)	(.010)	(.012)
year-month FE	+	+	+	+	+	+
state fixed effects	+	+	+	+	+	+
frequency weights	+	+	+	+	+	+
year interacted with	growth	-	-	-	-	-
disp. income (in €)	+	+	+	+	+	+
buffer	$> 10km$	$> 10km$	$> 10km$	$> 10km$	$> 10km$	$> 10km$
N(incl. weights)	10,095,674	10,095,674	10,095,674	10,095,674	3,079,125	10,179,671

Notes: Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, *clustered at state level*

Price Index (HPI) for consistency while maintaining regional heterogeneity. While the municipal heterogeneity ensures sufficient variation in the data, controls must be carefully chosen to address concerns about comparability and, more importantly, the common trends and SUTVA assumptions. One downside of using an HPI is that large markets with many observations would be weighted the same as a rural area with few observations. Furthermore, when the index is created using only a few observations, it potentially contains more measurement errors compared to a larger market. I incorporated weights based on posting frequency in a given municipality and month to account for these shortcomings. If I were not to weight the estimation by the number of postings, the results would be half of my baseline specification, see

Appendix A Table A.1 Col. 7. This suggests I need to address the spatial structure of my data.

One point to address here is the existence of regional income disparities, which raises concerns about differences in local purchasing power. Therefore, I control for the annual average disposable income in the region, measured at the district level. Appendix A Table A.1 Col. 5 shows that leaving out the control changes the coefficient marginally to 4.7% while increasing standard errors slightly.

While I am controlling for yearly and monthly fixed effects on the national level, time trends in the local labor markets could differ. To account for regional labor market dynamics, I use linear time trends for each labor market region as defined by Kosfeld and Werner (2012). Appendix A Table A.1 Col. 1-4 show that leaving out the time trend interacted with labor markets reduces our estimate to 3.8% (Col. 4). If I use individual labor market time trends, the standard errors increase (Col. 3). This would be a bad control since the individual labor market time trend would be influenced by the increase in the transfer tax and, therefore, are endogenous. To address this, I use two subclassifications along the dimensions urban-rural and growing-shrinking as provided by Häußermann and Siebel (2004). If I use urban-rural classification, the estimates would be similar at 4.3% (Col. 2), but the standard errors would be slightly larger.

As discussed in the previous section, the SUTVA assumption still needs to hold. It implies no tax spillovers across state borders affecting house prices. This may seem like a strong assumption, especially since the spillovers have been exploited for identification, see Dachis et al. (2012); Han et al. (2022). To address this, I calculate distances from municipalities' centroids to the nearest border and create buffer zones around state borders (Appendix A Table A.2 Col. 1-4). In the baseline specification (Col.1), I employ a 10km buffer around the border, resulting in the exclusion of municipalities within this zone. In the baseline specification (Col.1), a 10km buffer around the border is used, resulting in the exclusion of municipalities within this zone. The estimates show consistency across various buffer widths, including none (Col.2), 5km (Col. 3), or 15km (Col. 4). These estimates exhibit minimal changes, with the 10km buffer zone demonstrating the lowest standard errors.

While a theoretical argument can be made that amenity differences between regions do not play a major role since it is shown that the tax only impacts short-distance moves within the same labor market (Hilber and Lyytikäinen, 2017), amenity differences within a labor market do matter. By creating a municipality price index, I implicitly keep intra-state variation across municipalities. Consequently, I control for small area differences in amenities. Potential controls for amenities such as municipal spending are likely "bad controls" because while the size of the tax itself is exogenous to the municipality, spending is endogenous since the revenue generated by the tax directly impacts public expenses on local amenities such as schools and parks. Municipal fixed effects are nested in the state fixed effects as treatment occurs at the state level and the effect remains stable (with smaller standard errors), see Table A.1 Col. 4.

3.5.1 Rental market

The exact size of the tax on the rental market is understudied. At the time of writing, the only study that I am aware of is Han et al. (2022) and they restrict their analysis to one City - Toronto. Han et al. (2022) extend a search model by including a rental market featuring investors. They point out that investors do not pay the tax when a tenant moves, resulting in longer holding periods, and for each transaction, they pay only one market side of the burden. This means investors are generally better at mitigating the tax and profit from the lower ask prices.

I find that the tax increase of one percentage point is *decreasing* rents by 1.7%. While this contradicts the results of Han et al. (2022), who find higher rents in the particular case of the Toronto market, a negative effect is still plausible in the presence of a thick rental market. Since 50% of the German population is renting, institutional investors, who own multiple rental properties are common. These institutional investors face even less constraints compared to owner-occupiers becoming investors, particularly to the German transfer tax design tax loopholes that exists for institutional investors[24]. In a large rental market, these factors mean that institutional investors can fully profit from the lower asking prices and could potentially increase the supply of rental units substantially, leading to downward pressure on the rental prices.

Also in the rental market, substantial heterogeneity exists. The effect is relatively homogeneous across most growing categories at -2% to -3% (with p-values below 10%), except the fast-growing category where rents seem to be unaffected, again issuing the importance of the market structure. Similarly, cities with their thick rental markets show no effect of the transfer tax on rents but in more rural regions the effect is a decrease of 5.7% per percentage point increase in the transfer tax. Note that while the rent level is generally only slowly adjusting to housing market trends due to old contracts and rent regulation, the present sample considers only new rental postings and renegotiating rents is rare, such that I expect little measurement error. This effect could hint at the presence of large institutional investors, who can benefit from the lower prices because of fewer constraints and the institutional context in Germany, which creates loopholes for the investors. Rather than single

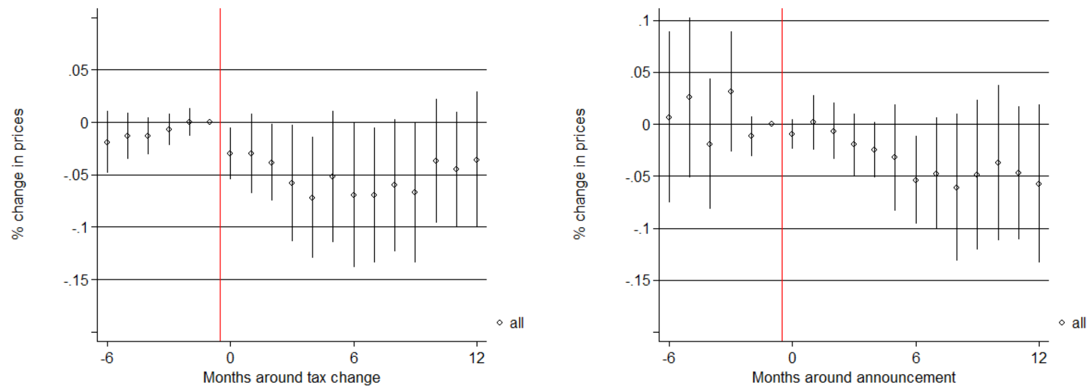
[24]The investor is bundling properties in a company and trades shares of the company, avoiding the real estate transfer tax.Schindlbauer (2020)

house owners deciding to rent out instead of selling, big companies could pass on the lower prices to the rental market in all regions besides the most stable demand (urban centers and fast-growing).

3.5.2 Common trend and no anticipation

To test for anticipation effects, I use the months when the first draft of the tax change was published in the state parliament bulletin. Thankfully, Dolls et al. (2021) collected the data on the date of announcement and the date of taking into effect, such that I can use their timeline here. Figure 3.2(a) shows the timing of the price

Figure 3.2: Analysis of the dynamics around announcement and tax change



(a) Tax change: prices **(b)** Tax announcement: prices

Note: The reference period in panel 2a is the month before the tax change and in panel 2b the month before the announcement of the tax change in the parliamentary bulletin. The sample has been restricted to half a year before the event and one year after. Due to these sample restrictions confidence intervals are at 90% and standard errors are clustered at the state level.

effect, where we see no anticipation in the prices and a sharp and clear drop in prices in the months after the introduction of the tax. The effect from the tax change in the first five months after the tax change is estimated with relatively large standard errors at around -5% compared to the year before the tax change.

Similarly, the prices show no anticipatory effects when we look at the announcement of the tax, see Fig. 3.2(b), and the estimates show large variation. Furthermore, we see no immediate reaction in the prices to the announcement of the tax change. The lower point estimates after five months of the tax announcement are likely the result of the tax change following the announcement.

3.5.3 Heterogeneity across density and labor market dynamics

As discussed in section 3.4 a large literature is discussing concerns about the homogenous treatment assumption in the case of TWFE[25]. Unfortunately at the point of writing, the literature has not yet developed feasible solutions for continuous treatment as in this setup, although promising work exists (de Chaisemartin et al., 2022). Instead, I am addressing concerns about treatment heterogeneity by estimating the effects for heterogeneously treated regions and comparing the total marginal effect of the tax rate. In the process, I expand the literature by exploring heterogeneity across markets with different dynamics and densities. Table 3.1 Col. 2 shows a specification where the tax rate interacted with 5 levels of labor market region dynamics[26] We see that accounting for heterogeneous effects across labor market region dynamics changes the marginal effect of the tax rates from -.041 to -.034. While this is a 0.5 pct. point smaller effect, this difference is not statistically significant. More interesting is the heterogeneity itself. When comparing the effect of a 1 percentage point increase in shrinking and growing regions I find, that the effect ranges from zero in the growing regions to -14% in the fast-shrinking regions. This is a substantial gap and shows that the market structure matters could change bargaining power or financing constraints. When regions grow fast, demand for real estate seems to be more inelastic and under bargaining the transfer tax is not passed through into ask prices. Alternatively financing constraints could be looser in growing markets leading to less liquidity constraints for buyers. This means the burden is completely passed onto the buyer. On the contrary in shrinking labor market regions, sellers could have lower bargaining power and face and buyers could face more rigid financing constraints from their bank for a house in a lower tier market. This means the seller carries a larger share of the burden. This is an important insight because this means that the total marginal effect varies with the growth dynamics in a region.

Alternatively, this heterogeneity could also reflect the thickness of urban housing markets. Col. 3 shows the heterogeneous effects of the tax rate interacting with rural and urban classifications. Here we can see that the tax rate bites more in rural areas (-11.6%) than in cities (-3.3%). This is also reflected in the total marginal effect of -5.1%.

Comparing these heterogeneous effects by combining them into one model, Col. 4, confirms the results from Col. 2 and 3, even though now the effects range is more moderate with fast shrinking regions at -10.8% and countryside interaction at -9.5%. The zero effect of the fastest-growing regions remains and cities still experience a

[25]see e.g. De Chaisemartin and d'Haultfoeuille (2020) for an overview.

[26]The classification follows Häußermann and Siebel (2004), see section 1.2 for a description of their index.

3.0% drop in prices. The consistency in the pattern with Col. 2 and 3 underlines the importance of including both classifications when analyzing the heterogeneity. However, the total effect is very similar to the baseline result of 4.4%.

3.5.4 Single houses

The author is aware of one similar paper published on the topic of the German real estate transfer tax, but it restricts their sample to single houses (Fritzsche and Vandrei, 2019), even though they do not find any effect. Comparing estimates for single houses of this sample to theirs is therefore a relevant exercise. Table 3.1 Col. 5 documents an effect of -1.9% per 1 percentage point increase of the transfer tax at a p-value of 8.1%, but again substantial heterogeneity depending on the market structure, ranging from -5.4% in fast shrinking regions to zero effects in fast-growing regions. For cities, I find a marginal effect of zero and in the most rural areas, the market reacts stronger with -5.5%.

3.6 Conclusion

This study examines the impact of real estate transfer taxes on housing affordability in a market with a substantial rental share. Using data from the whole of Germany, I exploit spatial and temporal variation in the tax rates across states and find that for each one-percentage-point increase in the tax rate, house prices decrease by 4.4%, and rents drop by 1.7%. The effects vary, from no observable impact in urban or rapidly growing markets, but more than doubled effects in rural or declining markets. Our findings provide valuable insights into the effects of transfer taxes on housing affordability and market dynamics, aiding policymakers and researchers. These findings hold significant implications for policy decisions in a predominantly rental-oriented market like Germany. Given the prevailing concerns about housing affordability in Germany, policymakers may consider adjusting tax policies to achieve specific objectives. If the aim is to reduce rents, the tax structure should encourage institutional investing while making owner-occupied housing more affordable. On the owner-occupier market, transfer taxes introduce distortions in housing tenure choices and create lock-in effects, where potential movers are discouraged from relocating to a more suitable home due to tax implications. In Germany, the existence of tax loopholes for institutional investors exacerbates the unequal tax burden between owner-occupiers and buy-to-rent investors, potentially leading to a misallocation of properties between rental and ownership markets.

Appendix

3.A Tables

Table A.1: Effect of the transfer tax on prices - specification

dep. var.: ln(HPI)	(1) baseline	(2) urban	(4) growth	(5) income	(6) both	(7) no weights	(8) amenities
	time trend		leaving controls out				
transfer tax (in %)	-0.041***	-0.043**	-0.038***	-0.047**	-0.044**	-0.023*	-0.049***
	(0.009)	(0.011)	(0.009)	(0.012)	(0.010)	(0.009)	(0.009)
year & month FE	+	+	+	+	+	+	+
state FE	+	+	+	+	+	+	municipality
frequency weights	+	+	+	+	+	-	+
time trend cat.	growth	urban	-	growth	-	growth	growth
disp. income (€)	+	+	+	-	-	+	+
buffer	$> 10km$	$> 10km$	$> 10km$	$> 10km$	$> 10km$	$> 10km$	$> 10km$
N	10,095,674	10,095,674	10,095,674	10,095,674	10,095,674	805,479	10,095,647

Notes: Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, *clustered at state level*

Table A.2: Effect of the transfer tax on prices - SUTVA

dep. var.: ln(HPI)	(1) baseline	(2) no buffer	(3) $> 5km$	(4) $> 15km$
transfer tax (in %)	-0.041***	-0.040**	-0.041***	-0.041**
	(0.009)	(0.010)	(0.009)	(0.012)
year-month FE	+	+	+	+
state fixed effects	+	+	+	+
frequency weights	+	+	+	+
year interacted with	growth	growth	growth	growth
disp. income (in €)	+	+	+	+
buffer	$> 10km$	no buffer	$> 5km$	$> 15km$
N	10,095,674	14,064,429	12,608,094	8,906,433

Notes: Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, *clustered at state level*

Concluding Remarks

In conclusion, the complex dynamics of spatial disparities in income, gender-gaps, and housing affordability remain a important and interesting issue. The research projects I've discussed here provide valuable insights into these disparities in the job and housing markets. They shine a light on specific aspects like gender differences, economic factors, balancing work and life, location dynamics, and policy implications.

The first project reveals a noticeable difference between men and women in commuting behavior and how it affects their job mobility around childbirth. This highlights the importance of recognizing the roles and responsibilities that come with gender and family obligations in the context of regional disparities.

The second project emphasizes the advantage of living in cities for higher wages and how this advantage may not benefit women as much as men. This draws attention to how gender plays a role in the differences between regions in the job market.

The third project focuses on the housing market and how taxes on property transfers affect affordability. It shows that these effects vary depending on the regional dynamics, which has implications for policymakers aiming to address housing disparities.

Together, these research projects give us a more detailed understanding of the differences between regions in terms of jobs and housing, especially when it comes to ongoing regional disparities. They highlight the importance of considering gender-related factors, family responsibilities, economic costs and benefits, and geographic variations when we try to tackle the effects of these regional differences. These insights can help guide policymakers, city planners, and researchers as they work to create more inclusive and fair urban environments while grappling with the persistent challenge of regional disparities.

Bibliography

Ahlfeldt, G. M., Heblich, S., and Seidel, T. (2023). Micro-geographic property price and rent indices. *Regional Science and Urban Economics*, 98:103836.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189.

Aizawa, N. and Fang, H. (2020). Equilibrium labor market search and health insurance reform. *Journal of Political Economy*, 128(11):4258–4336.

Altonji, J. and Paxson, C. (1992). Labor supply, hours constraints, and job mobility. *Journal of Human Resources*, 27(2):256–278.

Altonji, J. G. and Blank, R. M. (1999). Chapter 48 race and gender in the labor market. In *Handbook of Labor Economics*, volume 3, pages 3143–3259. Elsevier.

Angrist, J. D. and Keueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106(4):979–1014.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics*. Princeton university press.

Athey, S. and Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2):431–497.

Balli, H. O. and Sørensen, B. E. (2013). Interaction effects in econometrics. *Empirical Economics*, 45(1):583–603.

Barlevy, G. (2008). Identification of Search Models using Record Statistics. *The Review of Economic Studies*, 75(1):29–64.

Barth, E. and Dale-Olsen, H. (2009). Monopsonistic discrimination, worker turnover, and the gender wage gap. *Labour Economics*, 16(5):589–597.

Baum-Snow, N. (2010). Changes in transportation infrastructure and commuting patterns in us metropolitan areas, 1960-2000. *American Economic Review*, 100(2):378–382.

Bechtoldth, S., Freier, R., Geyer, J., and Kühn, F. (2014). Acht jahre nach der reform der grunderwerbsteuer: Bundesländer nutzen ihre spielräume für steuererhöhungen. *DIW Wochenbericht*, 81(50):1283–1290.

Besley, T., Meads, N., and Surico, P. (2014). The incidence of transaction taxes: Evidence from a stamp duty holiday. *Journal of Public Economics*, 119:61–70.

Best, M. C. and Kleven, H. J. (2018). Housing market responses to transaction taxes: Evidence from notches and stimulus in the uk. *The Review of Economic Studies*, 85(1):157–193.

Böheim, R. and Taylor, M. P. (2004). Actual and preferred working hours. *British Journal of Industrial Relations*, 42(1):149–166.

Biasi, B. and Sarsons, H. (2021). Flexible Wages, Bargaining, and the Gender Gap*. *The Quarterly Journal of Economics*, 137(1):215–266.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

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

Bonhomme, S. and Jolivet, G. (2009). The pervasive absence of compensating differentials. *Journal of Applied Econometrics*, 24(5):763–795.

Borghorst, M., Mulalic, I., and Van Ommeren, J. (2021). Commuting, children and the gender wage gap. Technical report, Tinbergen Institute Discussion Paper.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American statistical association*, 90(430):443–450.

Börjesson, M., Isacsson, G., Andersson, M., and Anderstig, C. (2019). Agglomeration, productivity and the role of transport system improvements. *Economics of Transportation*, 18:27 – 39.

Caldwell, S. and Danieli, O. (2020). *Outside options in the labor market*. Pinhas Sapir Center for Development, Tel Aviv University.

Caldwell, S. and Harmon, N. (2019). Outside options, bargaining, and wages: Evidence from coworker networks. *Unpublished manuscript, Univ. Copenhagen*, pages 203–207.

Caldwell, S. and Oehlsen, E. (2022). Gender differences in labor supply: Experimental evidence from the gig economy. *University of California, Berkeley Working Paper*.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Card, D. (2022). Who set your wage? *American Economic Review*, 112(4):1075–90.

Card, D. and Hyslop, D. R. (2021). Female earnings inequality: The changing role of family characteristics and its effect on the extensive and intensive margins. *Journal of Labor Economics*, 39(S1):S59–S106.

Card, D., Rothstein, J., and Yi, M. (2023). Location, location, location. Technical report, National Bureau of Economic Research.

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

Christofzik, D. I., Feld, L. P., and Yeter, M. (2020). Heterogeneous price and quantity effects of the real estate transfer tax in germany. Technical report, Freiburger Diskussionspapiere zur Ordnungsökonomik.

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

Combes, P.-P., Duranton, G., and Gobillon, L. (2011). The identification of agglomeration economies. *Journal of economic geography*, 11(2):253–266.

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

Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In Glaeser, E. L., editor, *Agglomeration Economics*. University of Chicago Press.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of Regional and Urban Economics*, volume 5, pages 247–348. Elsevier.

Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly Exogenous. *The Review of Economics and Statistics*, 94(1):260–272.

Cortés, P. and Pan, J. (2020). Children and the remaining gender gaps in the labor market. Technical report, National Bureau of Economic Research.

Couture, V., Duranton, G., and Turner, M. A. (2018). Speed. *The Review of Economics and Statistics*, 100(4):725–739.

Dachis, B., Duranton, G., and Turner, M. A. (2012). The effects of land transfer taxes on real estate markets: evidence from a natural experiment in toronto. *Journal of economic Geography*, 12(2):327–354.

Daly, M. and Groes, F. (2017). Who takes the child to the doctor? mom, pretty much all of the time. *Applied Economics Letters*, 24(17):1267–1276.

Dauth, W., Findeisen, S., Moretti, E., and Suedekum, J. (2016). Spatial wage disparities–workers, firms, and assortative matching. Technical report, Mimeo, Heinrich Heine University Düsseldorf.

Davidoff, I. and Leigh, A. (2013). How do stamp duties affect the housing market? *Economic Record*, 89(286):396–410.

Davis, D. R. and Dingel, J. I. (2019). A spatial knowledge economy. *American Economic Review*, 109(1):153–170.

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

De Borger, B., Mulalic, I., and Rouwendal, J. (2019). Productivity effects of an exogenous improvement in transport infrastructure: Accessibility and the great belt bridge. Technical report, Tech. rep. 065/VIII.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

de Chaisemartin, C., d'Haultfoeuille, X., Pasquier, F., and Vazquez-Bare, G. (2022). Difference-in-differences estimators for treatments continuously distributed at every period. *arXiv preprint arXiv:2201.06898*.

Dolls, M., Fuest, C., Krolage, C., and Neumeier, F. (2021). Who bears the burden of real estate transfer taxes? evidence from the german housing market. *working paper*.

Duranton, G. and Puga, D. (2001). Nursery cities: urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91(5):1454–1477.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of Regional and Urban Economics*, volume 4, pages 2063–2117. Elsevier.

Duranton, G. and Puga, D. (2020). The economics of urban density. *Journal of economic perspectives*, 34(3):3–26.

Eeckhout, J., Pinheiro, R., and Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, 122(3):554–620.

Farré, L., Jofre-Monseny, J., and Torrecillas, J. (2020). Commuting Time and the Gender Gap in Labor Market Participation. IZA Discussion Papers 13213, Institute of Labor Economics (IZA).

Frenzel Baudisch, C. and Dresselhaus, C. (2018). Impact of the german real estate transfer: Tax on the commercial real estate market. Technical report, Finanzwissenschaftliche Arbeitspapiere.

Frings, H. and Kamb, R. (2021). *What explains the urban wage premium? Sorting, non-portable or portable agglomeration effects?* Number 916. Ruhr Economic Papers.

Fritzsche, C. and Vandrei, L. (2019). The german real estate transfer tax: Evidence for single-family home transactions. *Regional Science and Urban Economics*, 74:131–143.

Fujita, M. (1989). *Urban Economic Theory*. Cambridge University Press.

Gabrovski, M. and Ortego-Marti, V. (2019). The cyclical behavior of the beveridge curve in the housing market. *Journal of Economic Theory*, 181:361–381.

Gabrovski, M. and Ortego-Marti, V. (2021). Efficiency in the housing market with search frictions. Technical report.

Galanaki, E. and Papalexandris, N. (2012). A profile of the organization investing in fringe benefits: Drawing on data from 15 European countries across 10 years. *International Federation of Scholarly Associations of Management*, pages 26–29.

Gallen, Y., Lesner, R. V., and Vejlin, R. (2019). The labor market gender gap in denmark: Sorting out the past 30 years. *Labour Economics*, 56:58 – 67.

Gaubert, C. (2018). Firm sorting and agglomeration. *American Economic Review*, 108(11):3117–3153.

Glaeser, E. L. (1999). Learning in cities. *Journal of urban Economics*, 46(2):254–277.

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

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Gronberg, T. and Reed, W. (1994). Estimating workers' marginal willingness to pay for job attributes using duration data. *Journal of Human Resources*, 29(3).

Gruber, J. and Lettau, M. (2004). How elastic is the firm's demand for health insurance? *Journal of Public Economics*, 88(7-8):1273–1293.

Gruber, J. and Madrian, B. C. (2004). Health insurance, labor supply, and job mobility: A critical review of the literature. In McLaughlin, C., editor, *Health Policy and the Uninsured*, chapter 11. ington, D.C.: Urban Institute Press.

Gutiérrez-i Puigarnau, E., Mulalic, I., and van Ommeren, J. N. (2016). Do rich households live farther away from their workplaces? *Journal of Economic Geography*, 16(1):177–201.

Gutiérrez-i Puigarnau, E. and Van Ommeren, J. N. (2011). Welfare effects of distortionary fringe benefits taxation: The case of employer-provided cars. *International Economic Review*, 52(4):1105–1122.

Han, L., Ngai, L. R., and Sheedy, K. D. (2022). *To Own Or to Rent?: The Effects of Transaction Taxes on Housing Markets*. Centre for Economic Policy Research.

Han, L. and Strange, W. C. (2016). What is the role of the asking price for a house? *Journal of Urban Economics*, 93:115–130.

Harding, M. (2014). Personal tax treatment of company cars and commuting expenses: Estimating the fiscal and environmental costs.

Häußermann, H. and Siebel, W. (2004). Schrumpfende städte-schrumpfende phantasie. *Merkur*, 58(664):682–692.

Heblich, S., Redding, S. J., and Sturm, D. M. (2020). The Making of the Modern Metropolis: Evidence from London*. *The Quarterly Journal of Economics*.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics*, 53(1):1–28.

Hilber, C. A. and Lyytikäinen, T. (2017). Transfer taxes and household mobility: Distortion on the housing or labor market? *Journal of Urban Economics*, 101:57–73.

Hirsch, B., Jahn, E. J., Manning, A., and Oberfichtner, M. (2022). The urban wage premium in imperfect labor markets. *Journal of Human Resources*, 57(S):S111–S136.

Hirsch, B., König, M., and Möller, J. (2013). Is there a gap in the gap? regional differences in the gender pay gap. *Scottish Journal of Political Economy*, 60(4):412–439.

Holmes, T. J. (2010). Structural, experimentalist and descriptive approaches to empirical work in regional economics. *Journal of Regional Science*, 50(1):5–22.

Hosios, A. J. (1990). On the efficiency of matching and related models of search and unemployment. *The Review of Economic Studies*, 57(2):279–298.

Hwang, H.-s., Reed, W., and Hubbard, C. (1992). Compensating wage differentials and unobserved productivity. *Journal of Political Economy*, 100(4):835–58.

Hybel, J. and Mulalic, I. (2022). Transportation and quality of life: Evidence from denmark. Technical report.

Jäger, S., Roth, C., Roussille, N., and Schoefer, B. (2022). Worker beliefs about outside options. Technical report, National Bureau of Economic Research.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*, 109:122–26.

Kleven, H., Landais, C., and Søgaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Knudsen, E., Hjorth, K., and Pilegaard, N. (2019). Wages and accessibility – evidence from denmark. Technical report, Tech. rep. DTU.

Kopczuk, W. and Munroe, D. (2015). Mansion tax: The effect of transfer taxes on the residential real estate market. *American economic Journal: economic policy*, 7(2):214–57.

Kosfeld, R. and Werner, A. (2012). German labour markets—new delineation after the reforms of german district boundaries 2007–2011. *Raumforschung und Raumordnung*, 70:49–64.

Koslowsky, M., Kluger, A. N., and Reich, M. (2013). *Commuting stress: Causes, effects, and methods of coping*. Springer Science & Business Media.

Künn-Nelen, A. (2016). Does commuting affect health? *Health economics*, 25(8):984–1004.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, 136(1):381–426.

Levkovich, O., Rouwendal, J., and Van Ommeren, J. (2020). The impact of highways on population redistribution: The role of land development restrictions. *Journal of Economic Geography*, 20(3):783–808.

Lucas, R. E. and Rossi–Hansberg, E. (2002). On the internal structure of cities. *Econometrica*, 70(4):1445–1476.

Lundborg, P. and Skedinger, P. (1999). Transaction taxes in a search model of the housing market. *Journal of Urban Economics*, 45(2):385–399.

Maasoumi, E. and Wang, L. (2019). The gender gap between earnings distributions. *Journal of Political Economy*, 127(5):2438–2504.

MacGregor Pelikánová, R. and Jánošíková, P. (2017). The danube dynamics of the real estate transfer tax. *DANUBE: Law, Economics and Social Issues Review*, 8(4):191–206.

Madden, J. F. (1985). Urban wage gradients: Empirical evidence. *Journal of Urban Economics*, 18(3):291 – 301.

Manning, A. (2003a). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

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

Manning, A. and Petrongolo, B. (2008). The part-time pay penalty for women in britain. *The Economic Journal*, 118(526):F28–F51.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–59.

Matano, A. and Naticchioni, P. (2016). What drives the urban wage premium? evidence along the wage distribution. *Journal of Regional Science*, 56(2):191–209.

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

Mirrlees, J., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., and Poterba, J. (2010). Reforming the tax system for the 21st century.

Moretti, E. (2011). Local labor markets. In *Handbook of Labor Economics*, volume 4, pages 1237–1313. Elsevier.

Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415.

Mulalic, I. and Rouwendal, J. (2020). Does improving public transport decrease car ownership? evidence from a residential sorting model for the copenhagen metropolitan area. *Regional Science and Urban Economics*, 83:103543.

Mulalic, I., Van Ommeren, J. N., and Pilegaard, N. (2013). Wages and commuting: Quasi-natural experiments' evidence from firms that relocate. *The Economic Journal*, 124(579):1086–1105.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, pages 693–709.

Oi, W. Y. and Idson, T. L. (1999). Firm size and wages. *Handbook of labor economics*, 3:2165–2214.

Ophem, H. V. (1991). Wages, nonwage job characteristics and the search behavior of employees. *The Review of Economics and Statistics*, 73(1):145–151.

Paetzold, J. (2019). Do commuting subsidies increase commuting distances? evidence from a regression kink design. *Regional Science and Urban Economics*, 75:136–147.

Petkova, K. and Weichenrieder, A. J. (2017). Price and quantity effects of the german real estate transfer tax.

Petrongolo, B. and Ronchi, M. (2020). Gender gaps and the structure of local labor markets. *Labour Economics*, 64:101819.

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

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory, 2nd Edition*, volume 1 of *MIT Press Books*. The MIT Press.

Potter, S., Enoch, M., Rye, T., Black, C., and Ubbels, B. (2006). Tax treatment of employer commuting support: An international review. *Transport Reviews*, 26(2):221–237.

Poulhès, M. et al. (2020). Housing market responses to transfer taxes: Evidence from a french reform. Technical report, Center for Research in Economics and Statistics.

Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, 50(1):203–219.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9(1):21–58.

Roca, J. D. L. and Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, 84(1):106–142.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of Regional and Urban Economics*, volume 4, pages 2119–2171. Elsevier.

Rosenthal, S. S. and Strange, W. C. (2012). Female entrepreneurship, agglomeration, and a new spatial mismatch. *Review of Economics and Statistics*, 94(3):764–788.

Ross, S. L. and Zenou, Y. (2008). Are shirking and leisure substitutable? an empirical test of efficiency wages based on urban economic theory. *Regional Science and Urban Economics*, 38(5):498–517.

Schaffner, S. (2020). *FDZ data description: Real-estate data for Germany (RWI-GEO-RED v3) - Advertisements on the internet platform ImmobilienScout24. 2007-06/2020*. Number 242996 in RWI Projektberichte. RWI - Leibniz-Institut für Wirtschaftsforschung.

Schindlbauer, M. (2020). New real estate transfer tax (rett) rules for share deals in germany. *Int'l. In-House Counsel J.*, 13:1.

Skora, T., Rüger, H., and Stawarz, N. (2020). Commuting and the motherhood wage gap: evidence from Germany. *Sustainability*, 12(14):5692.

Sullivan, P. and To, T. (2014). Search and nonwage job characteristics. *The Journal of Human Resources*, 49(2):472–507.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Timothy, D. and Wheaton, W. C. (2001). Intra-urban wage variation, employment location, and commuting times. *Journal of Urban Economics*, 50(2):338–366.

Turner, R. W. (1987). Taxes and the number of fringe benefits received. *Journal of Public Economics*, 33(1):41–57.

Van Ommeren, J. and Dargay, J. (2006). The optimal choice of commuting speed: consequences for commuting time, distance and costs. *Journal of Transport Economics and Policy (JTEP)*, 40(2):279–296.

Van Ommeren, J. and Fosgerau, M. (2009). Workers' marginal costs of commuting. *Journal of Urban Economics*, 65(1):38–47.

Van Ommeren, J. and Rietveld, P. (2005). The commuting time paradox. *Journal of Urban Economics*, 58(3):437–454.

Van Ommeren, J., van den Berg, G., and Gorter, C. (2000). Estimating the marginal willingness to pay for commuting. *Journal of Regional Science*, 40(3):541–563.

Van Ommeren, J. and Van Leuvensteijn, M. (2005). New evidence of the effect of transaction costs on residential mobility. *Journal of regional Science*, 45(4):681–702.

Van Ommeren, J. and Wentink, D. (2012). The (hidden) cost of employer parking policies. *International Economic Review*, 53(3):965–978.

Van Ommeren, J. N. and Gutiérrez-i Puigarnau, E. (2011). Are workers with a long commute less productive? an empirical analysis of absenteeism. *Regional Science and Urban Economics*, 41(1):1–8.

Voigtländer, P., Henger, D. R., Haas, H., Schier, M., Jaroszek, L., Kröncke, D., Just, P., Bienert, P., Geiger, P., Hesse, M., et al. (2013). Gesamtwirtschaftliche bedeutung der immobilienwirtschaft. *Deutscher Verband für Wohnungswesen, Städtebau und Raumordnung eV (DV) und Gesellschaft für Immobilienwirtschaftliche Forschung eV*.

Wheaton, W. C. (1974). A comparative static analysis of urban spatial structure. *Journal of Economic Theory*, 9(2):223–237.

Wheaton, W. C. (1990). Vacancy, search, and prices in a housing market matching model. *Journal of political Economy*, 98(6):1270–1292.

Zander, C. and Faller, B. (2006). Internationaler vergleich von kosten und dienstleistungseffizienz bei der transaktion von wohneigentum. *Optionen für Deutschland. Bonn: Selbstverlag des Bundesamtes für Bauwesen und Raumordnung (BBR)*.

Zax, J. (1991). The substitution between moves and quits. *Economic Journal*, 101(409):1510–21.

Zenou, Y. (2002). How do firms redline workers? *Journal of urban Economics*, 52(3):391–408.

Zenou, Y. (2009a). *Urban Labor Economics*. Cambridge University Press.

Zenou, Y. (2009b). Urban search models under high-relocation costs. theory and application to spatial mismatch. *Labour Economics*, 16(5):534–546.

Erklärung

gemäß §10 Abs. 6 der Promotionsordnung der Mercator School of Management, Fakultät für Betriebswirtschaftslehre der Universität Duisburg-Essen, vom 11. Juni 2012.

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und ohne unerlaubte Hilfe angefertigt und andere als die in der Dissertation und angegebenen Hilfsmittel nicht benutzt habe. Alle Personen, die an den Gruppenarbeiten mitgewirkt haben sind benannt. Alle Stellen, die wörtlich oder sinngemäß aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht.

Duisburg, 13. Oktober 2023

Malte Janosch Borghorst

DOI: 10.17185/duepublico/81709
URN: urn:nbn:de:hbz:465-20240313-152018-4