UNIVERSITY OF COPENHAGEN
DEPARTMENT OF ECONOMICS

PhD Dissertation

Maria Juul Hansen

Topics in Urban Economics:

Non-Market Valuation and Location Choice

MAIN SUPERVISOR: BERTEL SCHJERNING

CO-SUPERVISOR: NIKOLAJ HARMON

SUBMITTED: NOVEMBER 18, 2019

Contents

Acknowledgements

In sixth grade I filled out my first mandatory study plan that should follow me for the many coming years. I planned to become either a radio host or riding master. The former was more of an exotic idea, while the latter made better sense given I pursued my passion for horseback riding for another ten years after that. Around the same time, however, my childhood friend's mother claimed I would become a researcher one day. Not knowing what that or the academic world involved, this was just a term in the back of my head with no real substance. Nevertheless, in high school I was lucky to have a teacher who had studied economics himself and who encouraged me to apply for the economics program. The rest is history.

During my time as a master's student, I was very luck to meet Bertel Schjerning who has been my supervisor during my PhD. Besides teaching me microeconometrics, he also hired me as a research assistant which provided me the first glimpse of what it would feel like becoming a researcher. Bertel, I owe you deep gratitude for believing in me and for paving my way into academia. I have enjoyed working with you for the past five and a half years and have benefitted greatly from your supervision during my time as a PhD student. Our research camp trips to Hornbæk and the U.S. togehter with the rest of the co-authors at the URBAN project have always been very instructive and fun. My co-supervisor Nikolaj Harmon also deserves my thanks.

I am also very thankful to my co-author Christopher Timmins who hosted me during my one-year research visit at Duke University. Thank you for setting aside time every week to discuss my research and for making me feel very welcome in Durham and at Duke. In that connection I also thank Augustinus Fonden, Økonomisk Instituts Rejsefond, Oticon Fonden, Konsul Axel Nielsens Mindelegat, Knud Højgaards Fond and Danmark Amerika-Fondet for financial support of my stay abroad and Ernst og Karen Marie Andersens Fond for financial support during eight years of my studies. I also thank Department of Economics, University of Copenhagen, for believing in my research ideas and financing my PhD.

Thank you to my friends and family for bearing with me, especially during the last six months of finishing my dissertation. It has been tough at times and I appreciate how you have taken your time to listen, have backed me up and in general provided me a really good reason not to spend all my time behind the desk. Special mention is devoted to my grandmother Kirsten Marie Mortensen. She has always been a great inspiration with her

curiosity and eager for obtaining new knowledge, teaching me everything she knows and never give up on her dreams. As a young woman her's was to get a high school exam, but she ended up working in the mail services instead. At the age of almost 59 she got her diploma after all, something I really admire her for. I dedicate this dissertation to her.

Maria Juul Hansen

Copenhagen, November 2019

Sammenfatning

Denne ph.d.-afhandling består af tre selvstændige kapitler, der alle handler om emner i byøkonomi. De komplementerer hinanden ved at anskue spørgsmålet om, hvad der driver beslutningen om bopælsvalg, og i kapitel 2 og 3 også arbejdspladsvalg, fra tre forskellige vinkler.

Det første kapitel anvender en reduceret form-tilgang til at estimere marginale betalingsvilligheder for non-marginale ændringer i voldelig kriminlaitet under tidsvarierende præferencer. Kapitlet bidrager til litteraturen om værdisætning af goder, der ikke handles på markedsvilkår, ved at udvikle en metode, der identificerer heterogene kurver for marginal betaillingsvillighed. Det gøres ved at udnytte data på individer, som køber en bolig to gange i løbet af dataperioden. Individidet afslører derfor sin efterspørgsel efter det tilhørende niveau for voldelig kriminalitet i nabolagene to gange, og disse punkter forbindes til en lineær kurve for marginal betalingsvillighed. Det vigtigste resultat i kapitlet er, at det forårsager et signifikant bias i estimatet for marginal betalingsvillighed, hvis man ignorerer den individspecifikke og tidsvarierende heterogenitet, som der har været tradition for i litteraturen. Således er betalingsvilligheden for et 80 procent fald i voldelig kriminalitet som andel af individets indkomst overvurderet med 1,4-2,8 procentpoint (svarende til 13,2-18,7 procent) i de traditionelle metoder.

Det andet kapitel modellerer derimod individers placeringsbeslutninger eksplicit. Det anerkender, at beslutningerne om, hvor man skal bo og arbejde er dynamiske i sig selv og opstiller derfor en dynamisk strukturel model for boligstørrelse og bopæls- og arbejdspladsvalg. Kapitlet fokuserer på at estimere effekterne på (kortsigtede) ligevægtspriser, urbanisering og pendling af ændringer i infrastruktur og placering af jobs. Disse effekter undersøges i to kontrafaktiske eksperimenter, hvor i) boligudbuddet øges eksogent med 5 procent i København og Frederiksberg og ii) pendlingsomkostninger øges med 50 procent. i) resulterer i lavere priser i ligevægt og højere grad af urbanisering. ii) medfører lavere gennemsnitlige pendlingstider, men også en højere andel, som ikke er i beskæftigelse. Ligevægtspriserne falder i udkantsregioner med lavere jobtæthed. Dette kapitel bygger på og gentager tekst fra mit kandidatspeciale “A Dynamic Structural Approach to Individual Home and Job Location Decisions”, som blev indleveret til bedømmelse 16 måneder efter starten på ph.d.-programmet. I dette arbejde udviklede jeg en dynamisk ligevægtsmodel for bopæls- og arbejdspladsbeslutninger, men fokuserede på modeludvikling og -simulationer og efterlod derfor estimationen til fremtidig forskning. Mens det nuværende kapitel skal anses som en udvidelse af mit speciale, er der tale om væsentlige udvidelser og forbedringer i adskillige dimensioner, inkl. estimation af modellen.

Det tredje kapitel tager udgangspunkt i kapitel 2. Det ser bort fra ligevægten på boligmarkedet, men udvider modellen i en anden vigtig dimension: nemlig ved at opstille en dynamisk kollektiv model for beslutningen om, hvor man skal bo og arbejde, når man er en husholdning med to partnere, der begge arbejder. Denne slags husholdninger står overfor en særskilt udfordring, da de skal enes om, hvor de skal bo, og samtidig finde arbejdspladser for hvert medlem af husholdningen. Dette kapitel analyserer effekten af at udflytte jobs fra København centrum, men i modsætning til det forrige kapitel, koncentrerer det sig om effekten på fordelingen af pendling inden for husholdningen og på løngabet mellem mænd og kvinder. Af beregninsmæssige årsager er det en statisk version af modellen, der indtil videre er blevet estimeret. Et af de foreløbige resultater er, at udflytningen af jobs påvirker bopæls- og arbejdsbeslutninger, forskellen mellem mænd og kvinders pendling mindskes en smule og tilsvarende for løngabet.

Summary

This dissertation consists of three self-contained chapters on topics in urban economics. They complement each other by taking three different angles on what drives the decisions on where to live, and in chapter 2 and 3 also where to work.

The first chapter uses a reduced-form approach to estimate marginal willingness to pay (MWTP) for non-marginal changes in violent crime under time-varying preferences. It contributes to the literature on non-market valuation by developing a method that identifies heterogeneous linear MWTP curves. It does so by exploiting data on individuals who purchase two homes with associated levels of violent crime during the sample period. This allows the researcher to observe two points on the individual demand curve for violent crime and then 'connect the dots'. The key finding is that ignoring individual time-varying heterogeneity in MWTP, as has been the tradition in the literature, induces a significant upward bias in the MWTP for crime reductions and downward bias for increases in crime. Hence, the willingness to pay for an 80 percent reduction in violent crime as a share of individual income is overstated by 1.4-2.8 percentage points (corresponding to 13.2-18.6 percent) using the traditional methods.

The second chapter rather models the individuals' location decisions explicitly. It recognizes that the decions on where to live and work are inherently dynamic and therefore sets up a dynamic structural model for housing size and home and work location choices. The focus of this chapter is on estimating the effects on (short-run) equilibrium house prices, urbanization and commuting from changes in housing supply and infrastructure. These effects are explored in two counterfactuals where i) the housing supply is increased exogenously in Copenhagen and Frederiksberg by 5 percent and ii) commute costs are increased by 50 percent. i) results in lower prices in equilibrium and more urbanization. ii) implies lower average commute times, but also a higher share of people in non-employment. The equilibrium prices dropped in peripheral regions with low job density. This chapter builds on and repeats text from my master's thesis "A Dynamic Structural Approach to Individual Home and Job Location Decisions", which after 16 months in the PhD program was handed in for assesment. In that work I developed a dynamic equilibrium model of residential and work locations, but focused on the model development and simulations and hence left estimation for future research. While the current chapter should be considered an extension of the master's thesis, those are considerable extensions and improvements in several dimensions, including the estimation part.

The third chapter uses chapter 2 as a starting point. It abstracts from the equilibrium on the housing market, but instead extends the model in another important dimension by modelleing the decision on where to live and work for dual-earner households in a collective dynamic model. These households face a co-location problem, where they must agree where to live and then allocate the commute across the two spouses by choosing separate work locations. This chapter analyses the effect of relocating jobs away from the Copenhagen center, but unlike the previous chapter it concentrates on the effects on intra-household allocation of commuting and the gender wage gap. For computational reasons it estimates a static version of the model for now and finds that the relocation of jobs does affect location decisions for both home and work, the difference between male and female commute time is slightly lowered and so is the gender wage gap.

Chapter 1

Connect-The-Dots: Identification of Heterogeneous Marginal Willingness to Pay Functions under Time-Varying Preferences

Maria Juul Hansen, University of Copenhagen

Christopher Timmins, Duke University and NBER

Abstract

Models based on residential location choice have become commonplace in the non-market valuation literature. Rosen (1974) provided a utility-theoretic basis for hedonic models to be used to measure the welfare consequences of changes in local public goods and amenities. However, his proposed two-stage estimation procedure embodied a number of difficult econometric problems that became the focus of research for decades. Our paper builds upon the "inversion" approach suggested by Bajari and Benkard (2005) and the buyer-panel extension of that work proposed by Bishop and Timmins (2018). The latter paper shows how data on repeat purchases can be used to flexibly recover preferences with rich individual heterogeneity but is unable to deal well with time-varying individual attributes that might prompt residential location changes. We expand that approach to deal with any number of time-varying individual attributes including income, family structure and other drivers of housing choice. We apply that method to detailed longitudinal data from the Danish census, and use our estimates to value non-marginal changes in violent crime rates. We demonstrate a significant and policy-relevant bias from failing to properly account for the endogeneity problems in Rosen (1974).

1 Introduction

The valuation of neighborhood amenities and local public goods is important for the allocation of public funds and the measurement of the benefits of regulation and other policies. The value of these goods cannot generally be measured from market prices, but because many of them are, as their name suggests, local (i.e., consumption varies with geography), their values can be recovered from residential choices. Rosen (1974) provides the theory that connects those decisions to utility-theoretic measures of welfare, setting the stage for hedonic theory to be used in a variety of policy contexts.

While it is the basis for an entire literature, Rosen's procedure for recovering preferences from housing decisions is problematic. By the 1980s it was realized that there were important endogeneity problems inherent in his approach, (Epple, 1987; Bartik, 1987), and the literature began to suggest alternatives. This paper builds upon an "inversion" approach suggested by Bajari and Benkard (2005), in which preferences are not estimated in a traditional sense, but rather they are recovered at an individual level from the conditions imposed by optimizing behavior. This avoids the need for an unobserved (to the econometrician) preference shock, which is the source of the econometric problems mentioned above.

Specifically, our approach extends the idea in Bajari and Benkard (2005) to include the information available in a buyer-panel – individuals who are observed to buy more than one housing unit over the span of many years. This repeat-buyer information allows one to more richly describe individual preferences than was feasible using the cross-sectional individual information about consumers described by Bajari and Benkard (2005). Bishop and Timmins (2018) make this point and demonstrate the power of a buyer-panel in this context. However, their paper also reveals the weakness inherent in buyer-panel data. Specifically, it takes many years to observe individuals buying multiple houses, and during that time horizon, we would expect many of their circumstances to change. Indeed, it is often a change in household circumstances that prompts a move. Changing circumstances implies a time-varying form of individual heterogeneity that may be hard to observe. Indeed, panel data sets in the U.S. generally either (i) are available at a level of spatial disaggregation that is insufficient for modeling exposure to local public goods and amenities (e.g., the PSID or NLSY), or (ii) do not contain information about salient household attributes (e.g., data constructed from housing transaction information linked by buyers' names)[1]. We overcome this data constraint by employing restricted-access census data from Denmark. These data provide detailed information about numerous static and time-varying household attributes, in addition to specific information about the location and value of purchased homes. These data allow the researcher to control for

[1] Bishop and Timmins (2018). use data on housing transactions from a real estate data services provider, linked by buyer name and sale/purchase dates. These data can be linked to information about race and income collected under the Home Mortgage Disclosure Act, but that is the extent of household heterogeneity that can be included.

family structure (e.g., marriage, divorce, death, or the birth of a new child), changes in income or wealth, along with race, education, and other important characteristics. We can control for all the things mentioned here, but do not control for race as race has usually not been debated to the extent it has in the U.S. Bishop and Timmins (2018) ignored these time-varying drivers of demand for amenities because the data needed to address them were unavailable. Moreover, with the theory in that paper, it is difficult to address the problem even with the requisite data, as adding more time-varying attributes requires adding more repeat purchases, meaning a longer panel is required which generally means that more household attributes are likely to change. In this paper, we break that cycle with an alternative technique that uses the large size of the Danish census to adjust age, education, marital status, number of children and income such that we can argue that we observe the individual on the same demand curve twice when she buys a new property. With two points on the same demand curve we are able to identify her marginal willingness to pay (MWTP) for local amenities in a flexible way that avoids the well-known problems with Rosen's procedure.

We apply our method to a wide array of local public goods and amenities that might determine residential location choice. We focus our attention on violent crime, showing how the value of an inframarginal change differs dramatically depending upon whether one recovers an unbiased measure of preferences in the form of a MWTP function or not. Our results suggest that ignoring the endogeneity problems of Rosen's procedure means overstating the willingness to pay for an 80 percent reduction in violent crime by 16-20 percent and understating the costs of an 80 percent increase in violent crime by 13-16 percent. We also show there is a high degree of heterogeneity in the willingness to pay for reductions or to avoid increases in violent crime across the violent crime distribution.

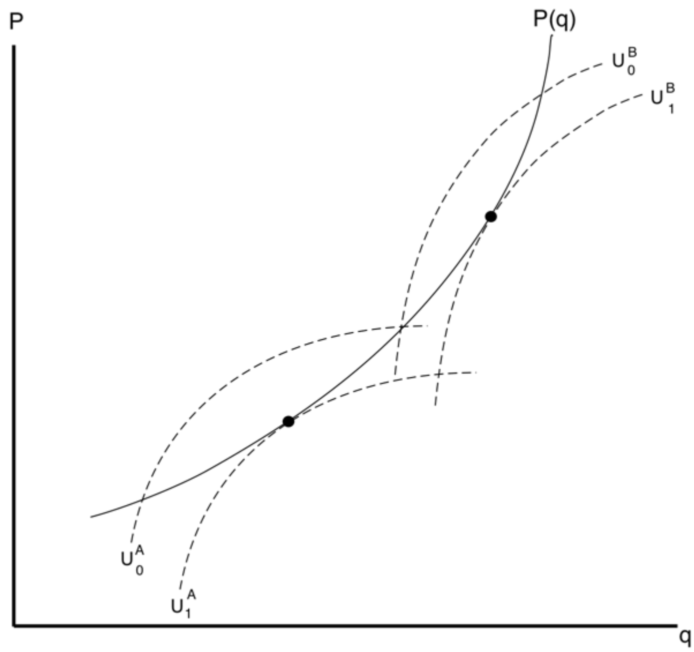
This paper proceeds as follows. Section 2 describes in more detail the source of the difficulty in recovering an unbiased estimate of the MWTP function. Section 3 discusses our data, which are particularly well-suited to modeling panel variation in house purchase decisions. Section 4 lays out the theory of our proposed estimation strategy and Section 5 reports results, with a particular emphasis on the differences in estimated value for a large change in violent crime using the alternative estimation strategies. Section 6 concludes.

2 The Difficulty with Estimating MWTP

In his 1974 article, Rosen proposed a two-step estimator for the MWTP function. His insight was the following: The slope of an individual's indifference curves in (q, P) space, where q represents some amenity and P is the price of the house associated with that amenity, reflects the willingness to give up additional units of other consumption (in the form of paying more for a house) in exchange for more q. Conveniently, individuals will sort into the housing unit that maximizes their utility, and that point on the slope of the hedonic price function will reveal the slope of their indifference curve. Figure 1.1

illustrates how agents A and B optimize where their indifference curves U^A and U^B are tangents to the hedonic price function P.

Figure 1.1: Picking q to optimize utility



Rosen proposed a two-step approach to recovering preferences from the hedonic price function. In the first step, the hedonic relationship between price (P), the amenity (q) and other housing characteristics (X) is estimated. For each observed house purchase, the implicit price of q, ($\frac{\partial P_i}{\partial q_i}$) is calculated, and is then used as the dependent variable in an estimation of the demand for q:

$$P_i^q = \frac{\partial P_i}{\partial q_i} = \gamma_0 + \gamma_1 q_i + \gamma_2 W_i + \epsilon_i \tag{1.1}$$

where W_i is a vector of individual attributes[2]. Rosen suggested estimating this equation by OLS and using the result to measure the value of a large change in the amenity q by integrating to yield measures of consumer surplus[3].

The literature eventually pointed out two potential problems with this approach. The first is that if P_i^q is just a function of q_i, there is no additional information introduced by the hedonic gradient. This exercise then amounts to regressing a linear function of q_i on q_i, and Brown and Rosen (1982) show that this will just reproduce the hedonic gradient. Brown and Rosen (1982) and Mendelsohn (1985) show how this problem can be addressed by imposing functional form restrictions on the hedonic price and MWTP functions, or

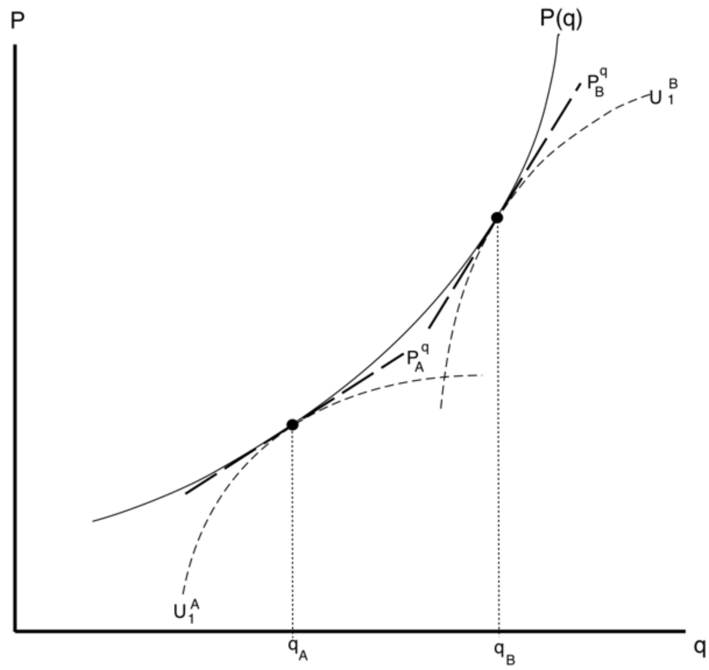
[2]Note that it is relatively easy to observe detailed information about households' characteristics if one is able to use cross-sectional data for estimation, as is required for Rosen's method.

[3]Willig (1976) discusses the role of income effects and the difference between compensated and uncompensated demand for welfare analysis.

by exploiting data from multiple markets. More recently, Ekeland et al. (2004) show that the type of linearity that causes the problems noted by Brown and Rosen (1982) and Mendelsohn (1985) is a special case and need not be a concern.

A second form of endogeneity was noted by Epple (1987) and Bartik (1987). They pointed out that, when an individual sorts along the hedonic price function, she both chooses the level of the amenity (q) and the implicit price that she pays for it (P_i^q). If the hedonic price function is non-linear, in the notation of the previous equation, individuals with large values of ϵ_i (i.e., strong preferences for q) will choose a high value of q (q_B in Figure 1.2) and, if the hedonic price function is convex, a high value for the implicit price of q (P_B^q in Figure 1.2). ϵ_i is thus correlated with both q_i and P_i^q. Epple (1987) notes that the traditional approach to using instruments in a system of supply and demand equations will not work here because buyers and sellers are systematically matched with one another by the sorting process. The literature struggled with this problem for more than a decade.

Figure 1.2: Picking q to optimize utility with non-linear hedonic price function



Bajari and Benkard (2005) provide a solution to this problem by replacing the estimation approach suggested by Rosen (1974) with an inversion approach. In particular, they propose writing down a utility function with heterogeneity embodied in utility function parameters at the individual level. For the sake of simplicity, we use a linear utility function for exposition:

$$U(q, x; \kappa) = \kappa_{1,i} q + \kappa_{2,i} x + c \tag{1.2}$$

which is maximized subject to a budget constraint:

$$c + P(q, x) = I, \tag{1.3}$$

where q is an amenity, x is other house characteristics and c is numeraire consumption and I is total income. The hedonic price function $P(q, x)$ represents the equilibrium of interactions between housing buyers and sellers and is assumed to be continuous and dense in attribute space. Solving for indirect utility (V) and taking first-order conditions with respect to q and x, one gets

$$\frac{\partial V_i}{\partial q} : \kappa_{1,i} - \frac{\partial P}{\partial q} = 0 \tag{1.4}$$

$$\frac{\partial V_i}{\partial x} : \kappa_{2,i} - \frac{\partial P}{\partial x} = 0. \tag{1.5}$$

These two equations can then be easily solved for values of $(\kappa_{1,i}, \kappa_{2,i})$ for each individual. The attractiveness of this approach comes in that individual heterogeneity is captured directly by utility function parameters, rather than in an econometric error term. It therefore avoids the endogeneity problems that accompany the latter. The downside to this approach is that it puts strong constraints on the shape of the MWTP function. In the simple example described above, those MWTP functions are necessarily horizontal lines (i.e., elasticity of zero). By assuming a Cobb-Douglas utility function, one assumes a MWTP function with elasticity of -1. If the goal is to let the data reveal the elasticity of demand for q, this is a severe limitation.

Bishop and Timmins (2018) show how the set of MWTP parameters one can recover from this inversion procedure can be expanded to include both intercept and slope parameters if one is able to observe two purchases by each individual. Importantly, that approach requires that both purchases lie on the same demand curve (i.e., preferences do not shift between house purchases). That paper also shows that identifying the coefficients on time-varying determinants of preferences requires additional repeat sales data (e.g., identifying a linear MWTP with one time-varying preference shifter would require data on three house purchases during which time no other individual attributes could vary). This creates a vicious cycle, whereby more housing transactions are required to identify the effects of time-varying preference shifters, but by including more housing transactions one increases the time dimension of the panel, and number of other time-varying attributes that might change. In this paper, we demonstrate a way to break out of this cycle by using a flexible function relating consumption of q to individual attributes to adjust that consumption to a counterfactual that holds time-varying individual attributes fixed. This approach relies on having rich data describing individual characteristics, which we get from the Danish census.

3 Data

The data come from Statistics Denmark's confidential registers[4]. Overall, we use three types of data: housing data, individual demographic data and neighborhood data[5], all of which are observed on an annual basis. The housing and individual demographic data cover the period 1992-2015 while the neighborhood data cover a shorter time span such that we end up using data for 2008-2014.

3.1 Individual demographic data

The datasets describing demographic information like home address, age, marital status, number of kids (all from the population register BEF), education (from the register UDDA), home ownership (from the register EJER), and income, wealth and debt (from the register IND) are merged based on the unique personal identifier PNR. While information from BEF is posted on January 1st of the year, EJER and UDDA are posted in the beginning of October each year and IND by the end of the year. To ensure that the observations are as close in time as possible, we merge BEF from year t together with UDDA, EJER and IND from $t-1$.

From the population register we get information on all individuals living in Denmark for 1992-2015. We restrict attention to all home owners in the Copenhagen local labor market defined according to Statistics Denmark's definition from 2014, cf. Figure 1.3, from 2008-2014[6]. The address information we get from this register is an anonymized address for the street, number, floor and door such that apartment complexes consist of several unique addresses.

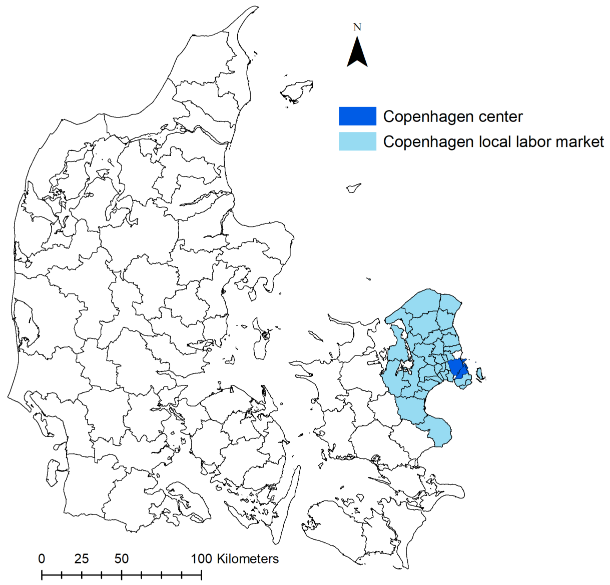
We only include home owners in the analysis because the decisions to own versus rent a house cannot be directly compared. Contrary to renters, home owners are making an actual investment in the property and may therefore consider how local amenities may evolve in the future. The hedonics literature typically ignores the dynamic component of the decision process (one notable exception is Bishop and Murphy (2011), but we do focus on buyers to avoid confounding different objective functions). The information on home ownership comes from the ownership register EJER. Every property in Denmark has a record indicating which individuals own it and at what date they took over the ownership. For apartment complexes, each apartment is considered a separate property. It also tells how large a fraction of the property each individual owns, so if e.g. a couple buys a house and they own 50% each, this information will be available. We define home owners to be every person who owns more than 0% of an address.

[4]See Table A1 for an overview of the registers we use.

[5]The authors thank Jørgen Brandt (Department of Environmental Science, Aarhus University) for providing access to data on air pollution.

[6]Incorporating additional housing markets would complicate the model by requiring that we consider tradeoffs in both labor and property markets, see Roback (1982).

Figure 1.3: Overview of municipalities in Denmark



Note: Statistics Denmark's definition on local labor markets is based on municipalities, cf. Statistics Denmark (2016). Each municipality consists of several parishes.

3.2 Housing and transactions data

Data on housing characteristics come from BOL which holds a description of every housing unit in Denmark such as number of rooms, square meter living space, construction year, number of bathrooms, whether the building is historically preserved and if there is access to a kitchen. Sales prices come from the EJSA register and have been deflated to 2011 prices using the consumer price index. EJSA contains an observation for every housing unit sold including the transaction price, the type of sale (e.g. single-family house, commercial or farm property), number of square meters sold and the type of post-sale ownership (e.g. private, association, company or state). Lastly, we have data on valuations of all properties in Denmark from the register EJVK. These valuations are made by the tax authorities for property tax purposes every other year and consist of, among others, an assesment of the land value and the property value. EJER and EJVK have a unique housing unit identifier which can also be found in the housing characteristics dataset BOL. We use this variable to merge EJSA and EJVK on to BOL which can then be merged on the personal dataset by using the address and PNR from BOL.

3.3 Neighborhood data

The amenity data we have access to include recordings of air pollution and crime. Data on air quality come from the Danish Center for Environment and Energy who model the air pollution in Denmark for a number of pollutants. These are, most importantly, NO_x, O_3, $PM_{2.5}, PM_{1.0}$ and CO, all computed in $\mu g/m^3$ for each 1×1 km square cell of Denmark for 1979-2017 on a monthly basis. To be able to compare this dataset with the other registers, we have computed annual averages for each pollutant and square cell. We then matched the square cells to parishes by use of mapping software. To ensure that all parishes do get a measurement for the pollutants, we include all square cells within 300 m from the boundaries of the parishes since a few parishes do not have a square cell within its boundaries.

From Statistics Denmark we have got access to information on number of victims by type of crime by year and parishes for 2005-2017. The types of crime are violent crime, sexual crime and property crime. More detailed crime types are available. For example, we can distinguish between burglary and theft in property crime. To avoid inclusion of types of crime that are unlikely to be reported or may not have anything do with the area itself (e.g. incesteous crime), we define violent crime to include serious violent crime, rape, crime against life and body, murder and attempted murder and violence against public authorities. The excluded groups are simple violence, threats and crime against personal freedom. We define property crime to include thefts and robberies and exclude blackmailing.

Despite having data on population, housing units and sales since 1992, we focus on the years 2008-2014 because we only have data on school districts[7] from Statistics Denmark's register SKOL for that subperiod. The school district data contain a link between all addresses in Denmark and a code for the school district that any home address belongs to in a given year. The school district boundaries can change over time and determine which public school parents are guaranteed to have their children accepted to. In theory there is free school choice implying parents are not forced to choose the local school, but to get their children into another school they must first apply and only if there are available seats will the parents' request be accepted. In the analysis, we exploit the school district boundaries to construct school fixed effects to account for the possibility that households sort based on school quality which then may influence house prices.

3.4 Sample selection

In the analysis, we only include sales that fulfill a number of criteria: the valuation of the property of the sale must exceed the value of the entire lot as the lot value is the value without any buildings and should therefore represent only a fraction of the property value. Further criteria are to only include sales that are not flagged as problematic by Statistics Denmark, the home is owned by a private individual or private housing co-operatives[8], the type of the sale belongs to one of the groups: single-family houses on private land, two-apartment houses or double houses on private land, three-apartment houses on private land, residential-only property with 4-8 apartments on private land,

[7] In Denmark each shcool has an associated district, thus there is no distinction between school districts and catchment zones.

[8] Denmark has a tradition for housing cooperatives which are associations whose purpose is to buy, own and manage residential properties for the members of the association. Each member does not own his residence, but does own a share of the association and thereby the right to use one of its residences.

residential-only properties with 9 or more apartments on private land, mixed residential and business properties on private land excluding owner-occupied flats, developed farms, owner-occupied flats for residential use on private land, lots below 2000 square meter and other developed land. Excluded categories are: business-only properties, factories and warehouses, summer houses and other properties not belonging to any of the beforementioned groups. Moreover, we only include adresses that have been sold once during the year, where only one household (family unit) lives, and where the parish code of the home is known as we use parishes to define neighborhoods for which we have amenity data. We also delete observations where the area sold is zero and for apartments where the area sold is above 500 square meters to avoid interpreting sales of whole apartment blocks as sales of single apartments. In general, we remove observations with sold area above the 99th percentile of the area distribution or if the number of rooms exceeds the 99th percentile of the rooms distribution[9].

Table 1.1 shows summary statistics for the property transactions we use, while Table 1.2 shows summary statistics for the buyers of properties in the Copenhagen local market during the period. In the estimation of the hedonic price function we use everyone with one or two purchases, while we only use those with two purchases when we identify individual-specific demand curves for violent crime. Compared to the one-purchase individuals, the individuals who bought two homes during the period are 7.5 percent less likely to be in a couple, earn 12.4 percent more on average, are 11.1 percent more likely to have children but those who have children are about as likely to have children in school age. They are a bit more likely to have a medium-length higher education[10] or more (59 percent against 54 percent for first-time buyers) and 61 percent live in a big city compared to 56 percent of the one-purchase households. The two-purchase individuals move to a place which is closer to their previous home measured in travel time, but have stayed 1.7 years in their previous home on average while the first-time buyers have stayed 5.7 years before they buy and move to their first owneroccupied dwelling. Both types of purchasers have been living with the same household for four to five years, i.e. they have stayed in the same couple or stayed single for that time period on average.

There is a significant amount of variation in the violent crime rates across houses in different parishes, cf. Figure 1.4. The distributions do not move much across years. Figure 1.5a depicts the average number of victims of violent crime for 2008-2014 by parish and Figure 1.6a zooms in on the Copenhagen local labor market. The higher crime rates tend to be within a 10-15 km radius of the Copenhagen center. The same goes for property crime, cf. Figure 1.5b and Figure 1.6b.

[9]See Table A2 for an overview of the sample selection process

[10]Short-length higher education corresponds to 1.5-2.5 years of study after high school, medium-length higher education to 3-4 years and long-length higher education to 5-6 years.

Table 1.1: Summary statistics of property transactions

	Mean	S.d	Median	N
Violent crime	10.18	16.51	6.00	58,920
Property crime	150.28	503.49	49.00	58,920
$PM_{2.5}$, μ_g/m^3	9.66	0.85	9.54	58,920
# sqm sold	475.51	453.00	347.00	58,920
I[apartment]	0.37	0.48	0.00	58,920
I[bath]	0.99	0.09	1.00	58,920
I[preserved]	0.02	0.13	0.00	58,920
Build year	1956	33.02	1963	58,542
# rooms	4.02	1.39	4.00	58,920
Km to Copenhagen center	17.23	13.47	12.61	58,920
Inhabs. pr. sqm	3,748	5,456	1,720	58,920

Sample criteria: Only using one property observation within the household in the year.

Figure 1.4: Probability density function of violent crime

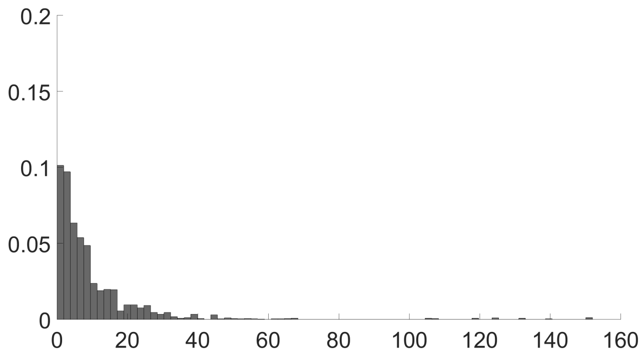


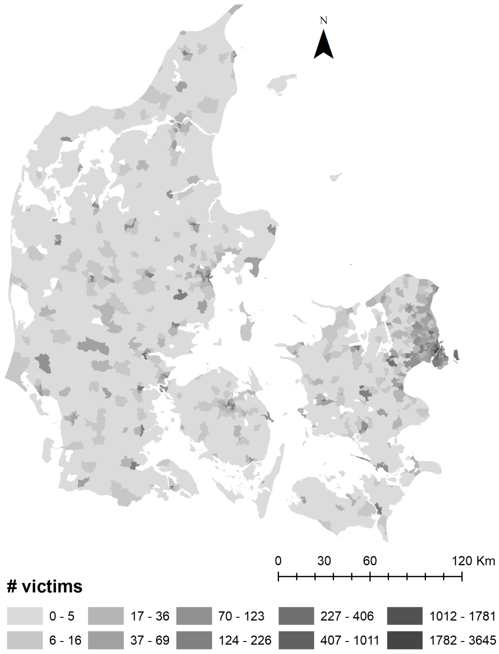
Table 1.2: Summary statistics of buyers by number of purchases 2008-2014

	Mean	S.d	N
1st purchase			
I[couple]	0.82	0.38	95,844
I[male]	0.50	0.50	95,844
I[has children]	0.55	0.50	95,844
I[has school age child]	0.23	0.42	95,844
Education			
Unskilled	0.04	0.19	95,844
High school	0.18	0.38	95,844
Vocational/Short-length	0.25	0.43	95,844
Medium-length	0.29	0.45	95,844
Long-length	0.25	0.43	95,844
Years household existed	4.87	4.69	95,844
I[divorce]	0.97	0.18	78,939
I[new couple]	0.08	0.28	16,905
Household total inc. (10,000 real DKK)	70.73	25.29	81,700
Household assets (10,000 real DKK)	258.26	114.70	88,510
Household debt (10,000 real DKK)	247.47	160.28	95,824
I[new job municipality]	0.37	0.48	95,844
I[live in big city]	0.55	0.50	95,844
Home move distance (minutes)	17.32	30.83	93,074
Years in previous home	5.68	8.49	76,410
2nd purchase			
I[couple]	0.79	0.41	3,935
I[male]	0.52	0.50	3,935
I[has children]	0.63	0.48	3,935
I[has school age child]	0.23	0.42	3,935
Education			
Unskilled	0.03	0.16	3,935
High school	0.14	0.35	3,935
Vocational/Short-length	0.24	0.43	3,935
Medium-length	0.30	0.46	3,935
Long-length	0.30	0.46	3,935
Years household existed	4.23	4.22	3,935
I[divorce]	0.97	0.16	3,121
I[new couple]	0.09	0.29	814
Household total inc. (10,000 real DKK)	74.33	26.00	3,112
Household assets (10,000 real DKK)	273.84	118.39	3,514
Household debt (10,000 real DKK)	290.15	184.63	3,935
I[new job municipality]	0.34	0.47	3,935
I[live in big city]	0.58	0.49	3,935
Home move distance (minutes)	12.03	11.82	3,906
Years in previous home	1.84	1.58	3,935

Note: I[new job municipality]$= 1$ if either or both of the household members gets a job in t in another municipality than where they had a job in $t-1$.

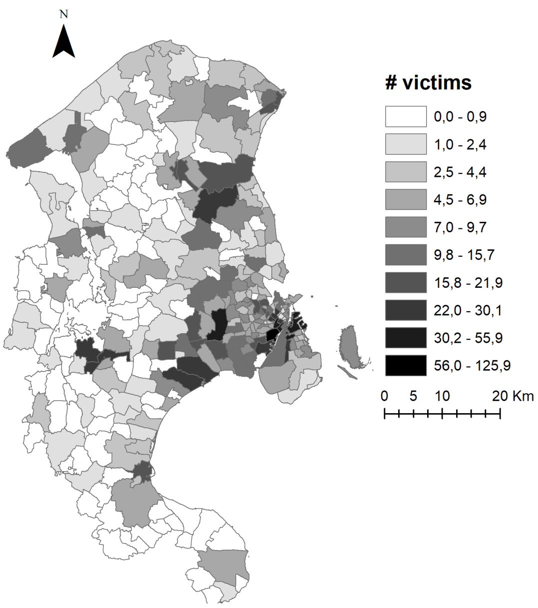
Figure 1.5: Average number of victims of crime 2008-2014 by parish in Denmark

(a) Violent

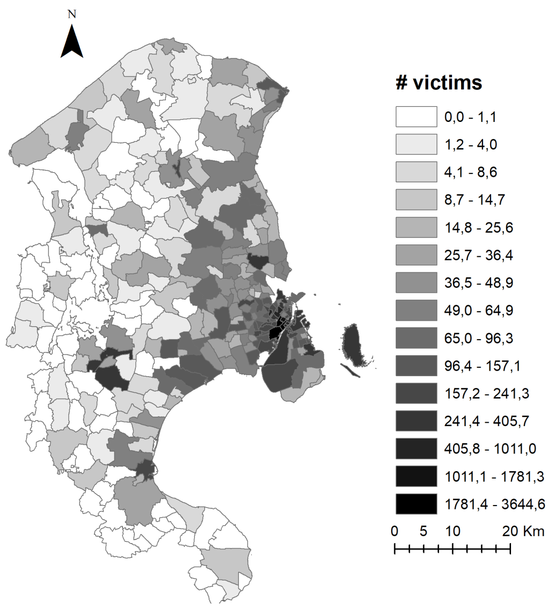


(b) Property

Figure 1.6: Average number of victims of crime 2008-2014 by parish in Copenhagen local labor market



(a) Violent



(b) Property

4 Estimation Strategy

We estimate heterogeneous linear marginal willingness to pay curves for an amenity q with parameters $\mu_i = (\mu_{i0}, \mu_{i1})$.

$$MWTP_{q_{it}} = \mu_{i0} + \mu_{i1} q_{it} \tag{1.6}$$

Linear functions are identified whenever we observe two points on the line. We therefore exploit that some individuals buy a home twice and therefore have revealed their demand for q in two different markets (time periods) under, potentially, two different price schedules. However, these two points would only identify μ_i as long as the individual's preferences are unchanged between the first and second purchases. Since individuals' preferences may change throughout their lives and it may be these changing preferences that make them decide on buying a new home, we are not guaranteed that these individuals' quantity of q chosen are actually observed along the same demand curve.

To circumvent this problem we exploit the richness of our data and predict the level for q that the individual would have chosen had her attributes not changed. We argue that due to the richness of our data, we are able to control for changing preferences by modeling the demand as a flexible function of a large set of individual attributes.

4.1 Step 1: Hedonic gradient

We first estimate a linear model for $\log(P_{it})$, where P_{it} is the total real property price that individual i paid at time t:

$$\log(P_{it}) = \beta_{0,t} + q_{it}\beta_{1,t} + q_{it}^2\beta_{2,t} + x_{it}'\beta_{3,t} + (x_{it}^2)'\beta_{4,t} + d_{it}'\beta_{5,t} + \epsilon_{it}, \tag{1.7}$$

where $\beta_{0,t}$ is a time-specific constant, q_{it} is a scalar describing the amenity of interest (i.e. violent crime), x_{it} is a vector of continuous covariates, namely number of victims of property crime, $PM_{2.5}$ in $\mu g/m^3$, square meters sold, year of construction of the housing unit, number of rooms, distance to Copenhagen city centre and population density. d_{it} is a vector of dummy variables for whether the housing unit sold has a bathroom, has a kitchen and if it is an apartment. Our model does not rely on any particular assumptions about the shape of the hedonic price function, but it can be allowed to vary across time. The demand for x_{it} is not the focus of this paper, but x_{it} is included to avoid bias in our estimates of $\beta_{1,t}$ and $\beta_{2,t}$.

The implicit price of q_{it}, P_{it}^q, is then given by:

$$\begin{aligned}
\frac{d \log P_{it}}{dq_{it}} &= \beta_{1,t} + 2\beta_{2,t}q_{it} \equiv \log(P_{it})^q \Rightarrow \\
\frac{dP_{it}/P_{it}}{dq_{it}} &= \beta_{1,t} + 2\beta_{2,t}q_{it} \Leftrightarrow
\end{aligned}$$

$$\frac{dP_{it}}{dq_{it}} = (\beta_{1,t} + 2\beta_{2,t}q_{it})P_{it} \equiv P_{it}^q \tag{1.8}$$

Denoting predicted quantities by "hat", the estimated implicit price for q_{it} is:

$$\hat{P}_{it}^{q_{it}} = (\hat{\beta}_{1,t} + 2\hat{\beta}_{2,t}q_{it})P_{it}(q_{it}), \tag{1.9}$$

while the estimated implicit log price is

$$\log(\hat{P}_{it})^q = \hat{\beta}_{1,t} + 2\hat{\beta}_{2,t}q_{it}. \tag{1.10}$$

4.2 Step 2: Segmentation equations

To predict the quantity of violent crime at the time of the second purchase, we estimate the relationship between individual demographic characteristics and the violent crime rate associated with the parish in which the individual buys a property. The relatonship between the quantitiy of q consumed and the attributes of the individuals doing that consumption describe how the market is segmented based on individual attributes and arise from sorting in hedonic equilibrium (Mendelsohn, 1985). Again letting q_{it} denote the number of victims of violent crime the individual demands, we use a linear model of the form[11],

$$q_{it} = w_{it}'\delta_t + u_{it}, \tag{1.11}$$

where w_{it} is a vector of individual demographics that we expect to affect the preferences for violent crime (constant term, age, educational level, marital status, number of children, log of real household income, and interactions between marital status and age), and u_{it} is an random regression error. In an effort to make this function as flexible as possible, we allow the parameter vector δ_t to vary over time as well.

We then adjust the individual's characteristics back to the values she had at the time of the first purchase and call these adjusted characteristics $\tilde{w}_{i2}$. By using the segmentation euqation we can predict adjusted demand at the second purchase denoted $\tilde{q}_{i2}$:

$$\tilde{q}_{i2} = \tilde{w}_{i2}\hat{\delta}_2. \tag{1.12}$$

Demand at the first purchase, q_{i1}, is directly observed in the data together with the price P_{i1}.

We can now compute the implicit price that i would have had to pay for $\tilde{q}_{i2}$ in period

[11] Using a Tobit model to account for the zero lower bound does not change the results. Given the functional form of the price function it is unlikely that the segmentation equation would actually be linear. Equation 1.11 therefore works as a linear approximation.

two (i.e., if her attributes had stayed at her period one values):

$$\tilde{P}^{\tilde{q}}_{i2} = (\hat{\beta}_{1,2} + 2\hat{\beta}_{2,2}\tilde{q}_{i2})\tilde{P}_{i2}(\tilde{q}_{i2}) \tag{1.13}$$

where

$$\tilde{P}_{i2}(\tilde{q}_{i2}) = \exp[\log(\tilde{P}_{i2}(\tilde{q}_{i2}))] = \exp[\hat{\beta}_{0,2} + \tilde{q}_{i2}\hat{\beta}_{1,2} + \tilde{q}^2_{i2}\hat{\beta}_{2,2} + x'_{i2}\hat{\beta}_{3,2} + (x^2_{i2})'\hat{\beta}_{4,2} + d'_{i2}\hat{\beta}_{5,2}] \tag{1.14}$$

The estimated implicit price for q_{i1} at time 1 is:

$$P^q_{i1} = (\hat{\beta}_{1,1} + 2\hat{\beta}_{2,1}q_{i1})P_{i1}. \tag{1.15}$$

4.3 Step 3: MWTP function inversion

We now have two observations of marginal price and chosen level of q along the same MWTP curve for each individual (i.e. holding the individual's time-varying attributes fixed). In equilibrium, the implict price will be equal to the MWTP function and this relationship allows us to write the problem of estimating linear MWTP curves as two equations with two unknowns (μ_{i0}, μ_{i1}):

$$P^q_{i1} = \mu_{i0} + \mu_{i1}q_{i1} \tag{1.16}$$
$$\tilde{P}^{\tilde{q}}_{i2} = \mu_{i0} + \mu_{i1}\tilde{q}_{i2} \tag{1.17}$$

This system of two equations can be used to solve for the two unknowns (μ_{i0}, μ_{i1}) for each individual. Because individual heterogeneity is embodied in the preference parameters rather than in an additive regression error we avoid the endogeneity problems described by Epple (1987) and Bartik (1987). This solution process amounts to finding the parameters of the MWTP function that "connect the dots" (CTD) for each individual.

$$\mu_{i1} = \frac{\tilde{P}^{\tilde{q}}_{i2} - P^q_{i1}}{\tilde{q}_{i2} - q_{i1}} \tag{1.18}$$

$$\mu_{i0} = P^q_{i1} - \mu_{i1}q_{i1} \tag{1.19}$$

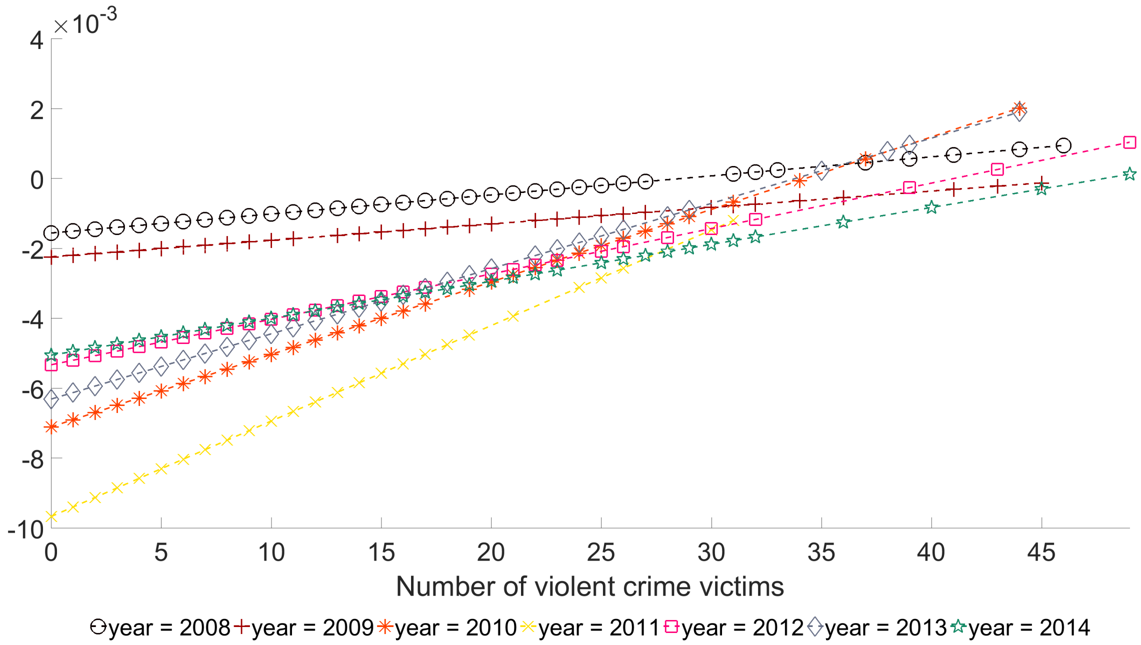
5 Results

In the following section we present the results for the three steps.

5.1 Step 1: Hedonic gradient

We start by estimating the hedonic price function according to Equation 1.7. The results are presented in Table 1.3. While the hedonic price function does show variability over time, it is remarkably stable with respect to the signs of the parameters associated with each variable. We focus our attention on violent crime which has a negative, but upward-sloping gradient (implicit price) in each year, as illustrated in Figure 1.7. Property crime and $PM_{2.5}$ pollution exhibit counter-intuitive signs, suggesting that they are likely correlated with some unobservable determinants of the housing prices. For instance, property crimes may be more likely to be reported in higher income neighborhoods. Higher levels of $PM_{2.5}$ may be correlated with desirable downtown locations. As these variables are not the focus of our analysis, we include them as valuable proxies to soak up variation in unobservables and focus our attention on violent crime.

Figure 1.7: Hedonic gradient of violent crime by year



Note: The hedonic gradient (corresponding to the implicit price) plotted is the estimated $\log(\hat{P}_{it})^q$. Restricting to number of victims of violent crime ≤ 50.

5.2 Step 2: Segmentation equations

Next, we estimate Equation 1.11 which is described in Table 1.4. Segmentation equation estimates vary over time but the signs on coefficients of different variables are very stable. A few trends can be observed. In general, consumption of violent crime tends to drop as one's education increases. Younger individuals choose neighborhoods with higher levels of violent crime. Violent crime falls with the introduction of more children, but does so

Table 1.3: 1st stage OLS regression of log(real property price) by year

	(1) 2008	(2) 2009	(3) 2010	(4) 2011	(5) 2012	(6) 2013	(7) 2014
Violent crime (100s)	-0.1555*	-0.2235	-0.7103***	-0.9666***	-0.5324***	-0.6306***	-0.5050***
	(0.0921)	(0.1361)	(0.1195)	(0.1540)	(0.1099)	(0.1118)	(0.0719)
Violent crime (100s)^2	0.2724**	0.2349	1.0372***	1.3662***	0.6497***	0.9333***	0.5284***
	(0.1156)	(0.1491)	(0.1800)	(0.2154)	(0.1356)	(0.1287)	(0.0850)
Property crime (100s)	0.0105	0.0353***	0.0537***	0.0325***	0.0412***	0.0262***	0.0317***
	(0.0079)	(0.0124)	(0.0088)	(0.0075)	(0.0068)	(0.0056)	(0.0054)
Property crime (100s)^2	-0.0012**	-0.0015**	-0.0020***	-0.0010***	-0.0010***	-0.0008***	-0.0009***
	(0.0005)	(0.0006)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0002)
$PM_{2,5}$	1.9272***	1.3856**	0.7631*	0.5260	2.5365**	1.9631**	0.5667
	(0.5287)	(0.6050)	(0.4453)	(1.0679)	(1.1709)	(0.7773)	(1.4103)
$PM^2_{2,5}$	-0.0935***	-0.0716**	-0.0387*	-0.0279	-0.1461**	-0.1102***	-0.0486
	(0.0248)	(0.0294)	(0.0212)	(0.0485)	(0.0656)	(0.0422)	(0.0753)
# 100 sqm sold	0.0328***	0.0283***	0.0369***	0.0375***	0.0352***	0.0349***	0.0313***
	(0.0027)	(0.0023)	(0.0017)	(0.0019)	(0.0026)	(0.0019)	(0.0032)
# 100 sqm sold^2	-0.0005***	-0.0003***	-0.0006***	-0.0006***	-0.0006***	-0.0005***	-0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
I[apartment]	-0.2250***	-0.2879***	-0.2465***	-0.2458***	-0.2528***	-0.2245***	-0.2307***
	(0.0129)	(0.0139)	(0.0116)	(0.0138)	(0.0133)	(0.0124)	(0.0136)
I[bath]	0.0606*	0.1103***	0.0566*	-0.0058	0.0588	0.0767**	0.0412
	(0.0335)	(0.0351)	(0.0324)	(0.0414)	(0.0393)	(0.0369)	(0.0409)
I[preserved]	-0.0227	0.0110	-0.0529*	0.0306	-0.0601	-0.0034	0.0400
	(0.0410)	(0.0393)	(0.0311)	(0.0319)	(0.0380)	(0.0336)	(0.0317)
Build year	-0.1983***	-0.2103***	-0.1636***	-0.1872***	-0.2179***	-0.2246***	-0.1876***
	(0.0121)	(0.0137)	(0.0127)	(0.0116)	(0.0136)	(0.0117)	(0.0113)
Build year^2	0.0001***	0.0001***	0.0000***	0.0000***	0.0001***	0.0001***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
# rooms	0.3916***	0.3320***	0.3286***	0.3593***	0.3641***	0.3619***	0.4124***
	(0.0104)	(0.0127)	(0.0122)	(0.0145)	(0.0121)	(0.0135)	(0.0113)
# rooms^2	-0.0284***	-0.0210***	-0.0210***	-0.0234***	-0.0241***	-0.0237***	-0.0291***
	(0.0012)	(0.0013)	(0.0013)	(0.0015)	(0.0012)	(0.0015)	(0.0012)
Dist. Cph.	-0.2195**	-0.6606***	-0.4158***	-0.8002***	-0.4061***	-0.6390***	-0.4529***
	(0.0939)	(0.1119)	(0.1059)	(0.1100)	(0.0992)	(0.0856)	(0.0758)
Dist. Cph.^{2}	0.0462***	0.0847***	0.0336*	0.1094***	0.0373**	0.0817***	0.0384***
	(0.0152)	(0.0201)	(0.0203)	(0.0212)	(0.0178)	(0.0142)	(0.0133)
Pop. dens.	-0.0166	-0.2379***	-0.1882***	-0.3234***	-0.1866***	-0.1945***	-0.0969***
	(0.0386)	(0.0386)	(0.0399)	(0.0410)	(0.0375)	(0.0297)	(0.0284)
Pop. dens.^{2}	-0.0057	0.0513***	0.0442***	0.0724***	0.0455***	0.0420***	0.0232***
	(0.0112)	(0.0104)	(0.0110)	(0.0110)	(0.0100)	(0.0073)	(0.0071)
Constant	195.1696***	209.7520***	167.2577***	191.5577***	212.4795***	221.8604***	193.3556***
	(12.1407)	(13.6503)	(12.6725)	(12.6889)	(13.7981)	(11.7014)	(12.6978)
School district FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log lik.	740.09	97.12	-80.83	-343.65	-419.78	-619.33	-569.76
r2	0.7529	0.7265	0.6994	0.7060	0.7063	0.7014	0.7119
Observations	7,889	6,937	8,845	7,316	8,222	8,745	10,588

Sample criteria: Individuals buying a property in the year. *Note:* Robust standard errors in parentheses. Violent and property crime measured as number of victims in units of 100 by parish code. $PM_{2,5}$ measured in μ_g/m^3. Distance to Copenhagen center (Dist. Cph.) measured in 10 km. Population density (Pop. dens.) measured as 10,000 inhabitants per km^2.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

more slowly if the parents are living in a couple. Violent crime also falls with increased income.

To predict the individual's quantity of violent crime at the time of the second purchase were she to have the attributes she had at the time of the first purchase, we first adjust the individual back to her characteristics at the time of the first purchase. This is how we control for potentially changing preferences between the two purchases.

Table 1.4: 2nd stage segmentation equation for number of victims of violent crime by year

	(1) 2008	(2) 2009	(3) 2010	(4) 2011	(5) 2012	(6) 2013	(7) 2014
Age	-0.098***	-0.060***	-0.045***	-0.035***	-0.048***	-0.061***	-0.050***
	(0.0134)	(0.0146)	(0.0134)	(0.0107)	(0.0124)	(0.0123)	(0.0115)
Education (ref. unskilled)							
High School	-0.901	0.123	-0.492	-0.717*	-1.187***	-1.180***	-2.236***
	(0.5856)	(0.5368)	(0.4623)	(0.4074)	(0.4385)	(0.4414)	(0.4817)
Vocational/Short Cycle	-1.903***	-0.250	-0.803*	-0.729*	-1.469***	-1.644***	-2.456***
	(0.5738)	(0.5283)	(0.4497)	(0.3948)	(0.4234)	(0.4249)	(0.4661)
Medium Cycle	-1.541***	-0.126	-0.677	-0.328	-0.932**	-1.388***	-2.485***
	(0.5751)	(0.5278)	(0.4454)	(0.3924)	(0.4221)	(0.4145)	(0.4599)
Long cycle	-1.923***	-0.547	-1.020**	-0.282	-0.975**	-1.613***	-2.350***
	(0.5818)	(0.5339)	(0.4492)	(0.3948)	(0.4254)	(0.4159)	(0.4653)
I[couple]	-3.281***	-1.535**	-1.544**	-1.054*	-1.014	-1.180*	-1.018
	(0.7430)	(0.7568)	(0.7273)	(0.6060)	(0.7172)	(0.6906)	(0.6601)
Kids (ref. 0)							
Kids=1	-1.944***	-1.808***	-1.602***	-0.652	-1.162**	-1.530***	-1.377***
	(0.6075)	(0.6585)	(0.5319)	(0.5044)	(0.4960)	(0.5189)	(0.4904)
Kids=2	-2.770***	-2.912***	-2.379***	-1.827***	-1.683***	-2.296***	-1.641**
	(0.6708)	(0.6950)	(0.6377)	(0.6727)	(0.5847)	(0.6736)	(0.7421)
Kids=3	-2.992**	-3.721***	-3.262***	-0.501	-2.147	-3.436***	-1.589
	(1.3476)	(1.2713)	(0.9384)	(1.4792)	(1.8014)	(1.2408)	(1.6073)
Kids=1 × I[couple]	0.754	1.110	0.614	-0.426	-0.109	0.849	0.321
	(0.6483)	(0.6949)	(0.5640)	(0.5376)	(0.5316)	(0.5589)	(0.5350)
Kids=2 × I[couple]	0.673	1.796**	1.330**	0.757	0.197	1.088	0.054
	(0.7044)	(0.7293)	(0.6647)	(0.6965)	(0.6121)	(0.7040)	(0.7684)
Kids=3 × I[couple]	1.771	3.633***	2.262**	0.145	1.256	2.482*	0.271
	(1.3991)	(1.3347)	(0.9814)	(1.5108)	(1.8264)	(1.2861)	(1.6475)
log(real hh. income)	-1.395***	-0.561**	-0.843***	-0.614***	-0.584***	-0.689***	-0.661***
	(0.2231)	(0.2196)	(0.1946)	(0.1731)	(0.1988)	(0.1925)	(0.1991)
I[couple] × Age	0.082***	0.011	0.027*	0.018	0.014	0.027*	0.023*
	(0.0162)	(0.0170)	(0.0153)	(0.0127)	(0.0142)	(0.0143)	(0.0137)
Constant	34.147***	19.835***	22.676***	18.229***	19.474***	21.921***	22.604***
	(2.9515)	(2.8939)	(2.6021)	(2.3160)	(2.6496)	(2.5467)	(2.6475)
R^2	0.028	0.015	0.012	0.010	0.015	0.011	0.011
N	12,775	11,335	14,773	12,032	13,564	14,331	17,160

Sample criteria: Individuals buying at least one property during 2008-2014. *Note:* Estimated by OLS. Standard errors in parentheses clustered at the individual level. Violent crime measured as number of victims by parish code.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3 Step 3: MWTP function inversion

Having obtained the predicted demand for violent crime at the second purchase, we use the estimates from Table 1.3 to get the implicit price that the individual would have had to pay for that level of violent crime in the year when the second purchase took place. To get the individual's MWTP function for violent crime, we evaluate Equation 1.18 and Equation 1.19 using the observations of predicted demand for violent crime and the associated implicit price at the second purchase together with the observed demand for violent crime and the price paid at the first purchase.

Interpretation of results

In order to summarize the results of the many heterogeneous MWTP functions that we recover, we run a regression of the slope (μ_1) and the intercept (μ_0) on a number of demographic controls, cf. Table 1.5 and Table 1.6. As expected from the theory, μ_1 is negative on average for almost everyone. Individuals above the age of 50 start to exhibit positive μ_1 according to specification 3 in Table 1.5, but they only account for 9.8 percent of the sample. Though most regressors are statistically insignificant in specification 4 of Table 1.5, individuals with medium-length higher education tend to increase their MWTP for reductions in violent crime as crime increases more than any of the other groups. The opposite holds for individuals with children. Moving on to Table 1.6, the intercept is negative for all subgroups as expected and especially so for individuals with at least a college degree and those who have children.

Table 1.5: OLS of MWTP slope parameters using CTD

	Spec. 1	Spec. 2	Spec. 3	Spec. 4
Educ. (ref. unskilled)				
High school	-99.9			-22.7
	(285.2)			(288.3)
Vocational or short-length	-34.5			-109.8
	(276.7)			(278.9)
Medium-length	-319.8			-375.1
	(278.9)			(281.7)
Long-length	134.3			60.1
	(286.4)			(287.6)
Age		22.8^{**}		22.0^{**}
		(9.0)		(9.2)
I[kids]			424.1^{**}	384.1^{*}
			(197.2)	(200.1)
Constant	-238.6	-1115.0^{***}	-511.7^{***}	-1132.5^{***}
	(200.2)	(325.7)	(123.1)	(366.5)
R^2	0.001	0.002	0.002	0.005
N	5,082	5,082	5,082	5,082

Note: Standard errors in parentheses clustered at the individual level. Violent crime measured as number of victims of violent crime by parish code. Removing observations with $\mu_{i1} > p99^{\mu_1}$ or $\mu_{i1} < p1^{\mu_1}$ or $\mu_{i0} > p99^{\mu_0}$ or $\mu_{i0} < p1^{\mu_0}$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Analysis using Rosen (1974)

We use the hedonic price function from Table 1.3, combined with information about individual homeowners using the theory described in Section 2. Estimates reported in Table 1.7 come from Rosen's second-stage regression, using the implicit price of violent crime derived from the hedonic gradient as the dependent variable. The endogeneity

Table 1.6: OLS of MWTP intercept parameters using CTD

	Spec. 1	Spec. 2	Spec. 3	Spec. 4
I[college]	-5202.6***			-4515.0**
	(1832.4)			(1837.2)
Age		-222.9***		-194.0***
		(67.3)		(66.8)
I[kids]			-6165.7***	-5438.8***
			(1560.1)	(1565.7)
Constant	-8550.2***	-2085.9	-7100.9***	379.3
	(884.5)	(2513.6)	(1030.0)	(2609.6)
R^2	0.003	0.003	0.006	0.011
N	5,082	5,082	5,082	5,082

Note: Standard errors in parentheses clustered at the individual level. Violent crime measured as number of victims of violent crime by parish code. Removing observations with $\mu_{i1} > p99^{\mu_1}$ or $\mu_{i1} < p1^{\mu_1}$ or $\mu_{i0} > p99^{\mu_0}$ or $\mu_{i0} < p1^{\mu_0}$.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

concern raised by Epple (1987) arises because the individual's choice of violent crime is determined by the same unobserved preference shock that determines the implicit price of violent crime because of the process of sorting along a non-linear budget constraint. The concern in that setting is that the coefficient on violent crime will have a positive bias. That is indeed what we find. That bias is severe enough that in every specification, the coefficient on violent crime (i.e., the slope of the MWTP function) is positive and statistically significant. In the following section, we demonstrate how the counterintuitive slope affects our welfare measure associated with non-marginal changes in violent crime.

5.4 Welfare analysis

In this section we demonstrate the consequences of mis-measuring the slope of the MWTP function by comparing the value of a large change in violent crime derived using Rosen's two-stage approach to that using our CTD procedure. In practice, we calculate the willingness to pay (WTP) for a large reduction and large increase in crime using CTD and Rosen's method. The WTP is the area between the MWTP curve and the horizontal axis between the current and new level of crime. Let q_0 denote the current level of crime and q_{low} the level of crime after the 80 percent reduction, i.e. $q_{low} = q_0 * 0.2$. Figure 1.8 illustrates the concept: the WTP of a reduction in violent crime from q_0 to q_{low} using the CTD procedure corresponds to area (1), while it corresponds to area (1) + (2) if we just assumed a horizontal MWTP curve[12] and area (1) + (2) + (3) if we used Rosen's

[12] That is the implication of Bajari and Benkard (2005) unless the researcher assumes an explicit function form for preferences that ensures downwards-sloping demand curves, e.g. a Cobb-Douglas utility function.

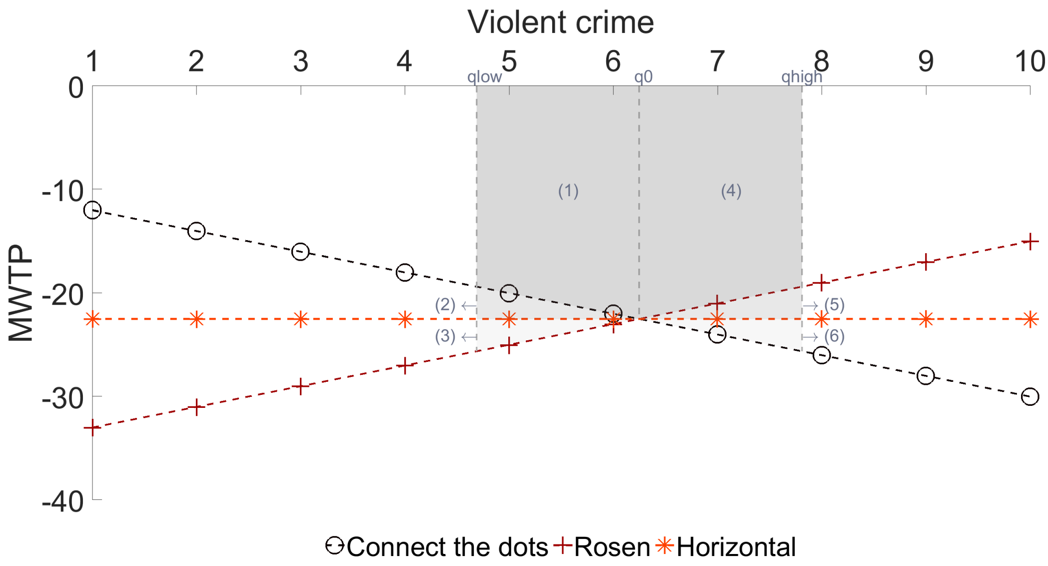
Table 1.7: Rosen 2nd stage: OLS of MWTP

	Spec. 1	Spec. 2	Spec. 3	Spec. 4
Violent crime	278.6***	193.7***	263.8***	79.2***
	(7.6)	(9.0)	(3.4)	(8.0)
Educ. (ref. unskilled)				
High School × Violent crime	47.8***			40.1***
	(9.1)			(6.3)
Vocational/Short-length × Violent crime	58.7***			43.8***
	(8.3)			(6.1)
Medium-length × Violent crime	-9.6			-14.9***
	(8.0)			(5.6)
Long-length × Violent crime	141.9***			115.8***
	(11.3)			(8.6)
I[college grad.]	-4958.8***			-4541.9***
	(99.6)			(89.6)
Violent crime × Age		3.8***		4.0***
		(0.2)		(0.2)
Age		-82.9***		-110.3***
		(2.9)		(2.6)
I[kids] × Violent crime			150.0***	140.8***
			(5.6)	(5.2)
I[kids]			-4076.3***	-3942.4***
			(71.9)	(68.9)
Constant	-12619.6***	-10677.0***	-11636.0***	-6356.2***
	(41.0)	(118.4)	(48.5)	(110.4)
R^2	0.252	0.210	0.239	0.296
N	97,966	97,966	97,966	97,966

Sample criteria: Individuals buying a property during 2008-2014. *Note:* The dependent variable is the estimated implicit price of violent crime using estimates from Table 1.3. Standard errors in parentheses clustered at the individual level. Removing observations with $MWTP > p99^{MWTP}$ or $MWTP < p1^{MWTP}$. Violent crime measured as number of victims by parish code.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

method. Likewiese, if we considered an increase in violent crime from q_0 to q_{high}, we would get a WTP of area (4) + (5) + (6) if we used the CTD approach, area (4) + (5) if we used the horizontal MWTP function and area (4) if we relied on Rosen's method. I.e. because the Rosen model predicts upward-sloping MWTP functions for violent crime, the MWTP for additional crime reductions rises as the level of crime drops, meaning that the Rosen model predcits a larger benefit from a reduction than the CTD approach does. The opposite argument can be used to show that the Rosen model predicts a smaller cost from a violent crime increase.

Figure 1.8: Example: computing WTP using different methods



Algebraically, the MWTP function for individual i using the CTD approach is given as

$$MWTP_{it}^{CTD} = \mu_{i0} + \mu_{i1} q_{i,t} \tag{1.20}$$

The WTP for a reduction in violent crime for i is then:

$$\begin{aligned} WTP_{it}^{CTD} &= \int_{q_{low}}^{q_0} (\mu_{i0} + \mu_{i1} q) dq \\ &= \mu_{i0} \cdot (q_{it,low} - q_{it,0}) + 0.5 \cdot \mu_{i1} (q_{it,low}^2 - q_{it,0}^2). \end{aligned} \tag{1.21}$$

For Rosen's method we use the estimate from the Rosen second stage where the MWTP, i.e. the estimated implicit price, has been regressed on the level of crime:

$$MWTP_{it}^{R} = \alpha_0 + \alpha_1 \cdot q_{it} + \epsilon_{it}, \tag{1.22}$$

where ϵ_{it} is a regression error. To get the WTP for Rosen's method, we integrate Equation 1.22 between q and q_{low}:

$$WTP_{it}^{R} = \alpha_0 \cdot (q_{it,low} - q_{it,0}) + 0.5 \cdot \alpha_1 (q_{it,low}^2 - q_{it,0}^2). \tag{1.23}$$

Table 1.8 shows summary statistics of the WTP for an 80 percent reduction in violent crime and Table 1.9 for an 80 percent increase using CTD and Rosen's method, respectively. Each row in the tables corresponds to different specifications of the regressions of μ_1 and μ_0 for the CTD approach and for different specifications of the 2nd stage in Rosen's method. Hence, each specification number in Table 1.8 corresponds to the same specification in Table 1.5 for μ_1 and Table 1.6 for μ_0 when we compute WTP using the CTD. For Rosen's method, the specification number corresponds to the specifications in Table 1.7.

As Table 1.8 shows, Rosen's method overstates the median WTP by 6,863-11,073 DKK dependent on the specification used. This corresponds to an overstatement in the range 16-20 percent. For the case of an increase in violent crime, Table 1.9 reveals that Rosen's method understates the median negative WTP (i.e. the compensation required in order accept an increase in violent crime) by 7,028-11,074 DKK, equivalent to approximately 13-16 percent. Figure D1 and Figure D2 illustrate the median difference in WTP by either method and for all specifications.

To put the magnitudes of the WTP numbers into context, we compute the WTP as the share of total annual income for each specification and method. Results are presented in Table D1 for the reduction in crime and in Table D2 for the increase. The median WTP for an 80 percent reduction in violent crime out of the individual's total income is 10.6 percent for specification 4 and rising to 15.0 percent for specification 1. Using Rosen's method, the numbers are 12.0 percent for specification 4 and 17.8 percent for specification 1. Rosen's method overstates the median WTP as a percent of total income by 1.4 to 2.8 percentage points dependent on the specification. Looking at the WTP for avoiding an 80 percent increase in violent crime, the median as percent of total income is 17.9 percent for specification 1 using CTD and 12.1 percent for specification 4. For Rosen's method it is 15.2 and 10.7 percent for specification 1 and 4, respectively. His method understates that WTP by 1.4 to 2.7 percentage points.

We choose to focus on specification 4 since that allows for more heterogeneity in the MWTP. Figure 1.9 displays the distribution of WTP for an 80 percent reduction in violent crime by either method and Figure 1.10 for an 80 percent increase. Both figures reveal a large amount of heterogeneity in WTP across the violent crime distribution[13].

[13]See Appendix B for the Rosen method where we control for individual-specific fixed effects in Rosen's second stage and Appendix C for the case of a horizontal MWTP function (i.e. no individual heterogeneity)

Table 1.8: Summary statistics of WTP (DKK) for 80 pct. reduction in violent crime

Spec.	Model	Mean	S.d.	Median	N
1	CTD	55,003	9,455	60,648	1,927
	Rosen	65,075	11,347	71,721	1,927
2	CTD	49,152	10,569	49,192	2,469
	Rosen	56,331	9,915	57,241	2,469
3	CTD	46,487	5,531	41,623	2,535
	Rosen	52,937	3,215	50,110	2,535
4	CTD	44,486	20,110	35,122	2,524
	Rosen	49,763	19,824	41,985	2,524

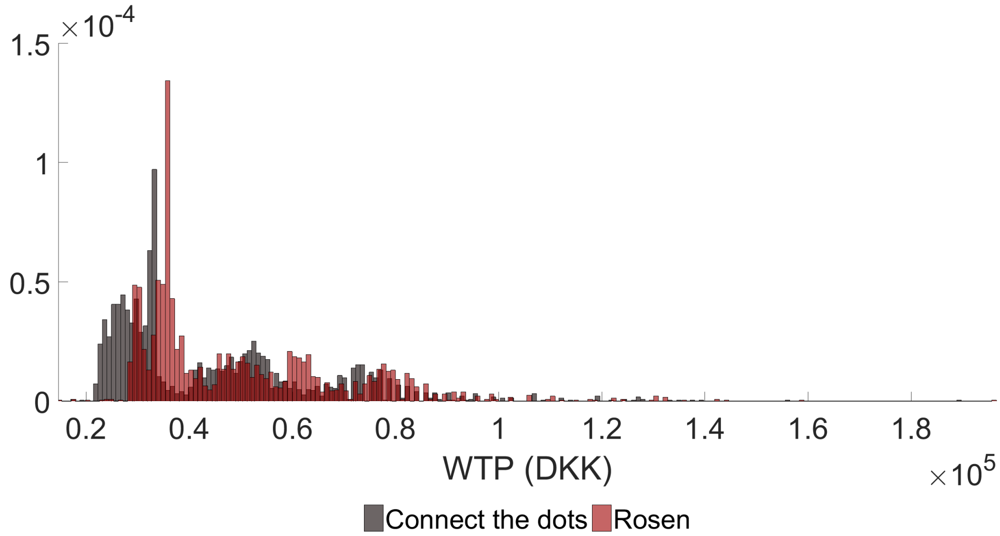
Note: Violent crime measured as number of violent crime victims by parish code. Only computing summary statistics based on observations who fulfill: $q_0 > 0, q_0 \leq 50, q_0 \cdot 0.2 \leq 50,\ q_0 \cdot 0.8 \leq 50, \mu_1 > p1^{\mu_1}, \mu_1 < p99^{\mu_1}, \mu_0 > p1^{\mu_0}, \mu_0 < p99^{\mu_0}$ and who bought two homes during the period.

Table 1.9: Summary statistics of WTP (DKK) for 80 pct. increase in violent crime

Spec.	Model	Mean	S.d.	Median	N
1	CTD	-65,935	10,354	-72,906	1,927
	Rosen	-55,862	8,471	-61,832	1,927
2	CTD	-56,397	6,593	-58,194	2,469
	Rosen	-49,218	4,922	-50,338	2,469
3	CTD	-53,383	637	-52,823	2,535
	Rosen	-46,933	2,953	-44,336	2,535
4	CTD	-48,845	15,141	-44,109	2,524
	Rosen	-43,568	15,186	-37,081	2,524

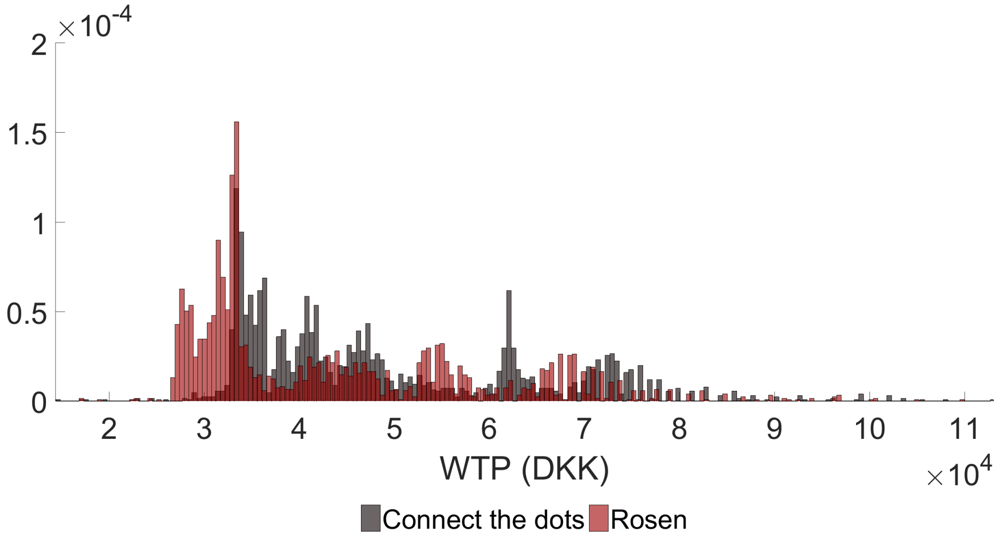
Note: Violent crime measured as number of violent crime victims by parish code. Only computing summary statistics based on observations who fulfill: $q_0 > 0, q_0 \leq 50, q_0 \cdot 0.2 \leq 50,\ q_0 \cdot 0.8 \leq 50, \mu_1 > p1^{\mu_1}, \mu_1 < p99^{\mu_1}, \mu_0 > p1^{\mu_0}, \mu_0 < p99^{\mu_0}$ and who bought two homes during the period.

Figure 1.9: Distribution of WTP (DKK) for 80 pct. reduction in violent crime (spec. 4)



Note: Using WTPs from specifications 4 in Table 1.8.

Figure 1.10: Distribution of WTP (DKK) for 80 pct. increase in violent crime (spec. 4)



Note: Using WTPs from specifications 4 in Table 1.9.

6 Conclusion

For many years, the hedonics literature has struggled with how to recover preferences underlying the choices observed in the housing market. Accurately recovering these preferences is necessary for measuring the value of non-marginal changes in (dis)amenities. Because of the econometrics problems described above, simple "first-stage" (in the parlance of Rosen (1974)) techniques have been used instead, but these methods only provide valid approximations for marignal changes in amenities or local public goods. Most policy-relevant changes tend to be non-mariginal. Over the last two decades, a number of techniques have been developed to address this problem. This paper contributes to that literature, extending the analysis in Bajari and Benkard (2005) to allow individual MWTP functions to have both heterogeneous intercepts *and* slopes. In so doing, we extend the method developed by Bishop and Timmins (2018) to incorporate rich infomation about time-varying individual attributes. We implement that method using detailed data from the Danish census. Applying our model to valuing large reductions in violent crime and comparing it to Rosen's procedure, we find that the Rosen model does indeed lead to biased estimates of the MWTP function. These biases lead to large overstatements of the value of crime reductions and understatements of the costs of increases in crime.

A Overview of data

Table A1: Overview of confidential registers from Statistics Denmark

Register	Description	Availability	Update
BEF	Main register listing all individuals with official address in Denmark. Information on basic information like social security number (SSN), age, home address, marital status, spouse's SSN, country of origin, and children and their SSN, parents' SSN and gender.	1992-current	January 1st
UDDA	Education register with information on highest obtained education of the individual including code for the educational institution and detailed fields and levels of study.	1992-current	October 1st
IND	Income register with information on annual total income, total wage income, assets, debt, public transfers and tax payments.	1992-current.	December 31st
EJER	Property ownership register with information on the SSN of owner, property identification code, type of ownership (e.g. public or private), ownership share of property and start date of ownership.	1992-current	October 1st
BOL	Property census register with information on every property unit in Denmark, e.g. number of rooms, living space , type of property, address and construction year. The register is based on The Central Register of Buildings and Dwellings (BBR) which is used for property assesments.	1992-current	January 1st
EJSA	Property transactions register with information on transactions of all real properties in Denmark such as the transactions prices, type of sale, land value, square meters sold and property identification code.	1992-current	January 1st
EJVK	Property assesment register based on BBR. As a general rule, the tax authorities assess the value of all owneroccupied dwellings in uneven years and other dwellings in even years.	1992-current	October 1st
KROF	Register of reports of victims of criminal offense by type of offense with information on e.g. address of the crime scene, victim's SSN and gender. We have not used individual-level data on crimes but instead got Statistics Denmark to deliver a dataset holding the number of victims by type of crime, parish and year based on data in KROF.	2001-current, but detailed address of crime scene incl. parish only 2005-current	January 1st
SKOL	Register of school districts with information on the home addresses that belong to the district. Data is based on CPR Vejregister which is a complete registry of all roads in Denmark including certain distric divisions such as school district. Municipalities decide the school districts and report these to CPR Vejregister.	2007-current	January 1st

Table A2: Overview of sample selection process

Selection criteria	N
Main dataset (all adult individuals in Denmark 2008-2014, sales and no sales)	29,460,516
Non-missing sales price (i.e. potential sales observation)	1,265,418
Property value > Lot value and private sale	664,928
Rooms < 99th percentile	662,776
Area sold < 99th percentile	656,439
Sale marked as OK by Statistics Denmark	656,439
Private owner post sale	591,492
Sales type either single-family houses on private land, two-apartment houses or double houses on private land, three-apartment houses on private land, residential- only property with 4-8 apartments on private land, residential-only properties with 9 or more apartments on private land, mixed residential and business properties on private land excluding owner-occupied flats, developed farms, owner-occupied flats for residential use on private land, lots below 2000 square meter and other developed land	582,920
Max number of distinct sales of the property: 1	580,395
Max number of sales of the property on the same date: 1	518,207
Property sold on open market terms	513,655
Property is sold in the current year	513,655
Max number of households (family units) on the address: 1	443,179
Parish of the address of the sale is known	443,179
Property sold to household who lives on address	314,199
Property was bought in Copenhagen local labor market	115,882
Individual’s education known	99,779

B Rosen 2nd stage with individual fixed effects

Table B1: Rosen 2nd stage: OLS of MWTP with individual fixed effects

	Spec. 1	Spec. 2	Spec. 3	Spec. 4
Violent crime	397.9***	110.1**	351.3***	66.3
	(33.5)	(47.4)	(17.5)	(53.6)
Educ. (ref. unskilled)				
High School × Violent crime	20.4			9.2
	(40.0)			(39.0)
Vocational/Short-length × Violent crime	1.6			48.6
	(39.1)			(40.9)
Medium-length × Violent crime	-47.3			34.4
	(44.8)			(46.2)
Long-length × Violent crime	88.3**			117.0***
	(40.5)			(41.0)
I[college grad.]	-8152.4***			-2418.5***
	(731.4)			(713.4)
Violent crime × Age		7.0***		6.1***
		(1.2)		(1.3)
Age		-1444.5***		-1387.7***
		(41.3)		(46.9)
I[kids]=1 × Violent crime			75.7**	52.3*
			(33.5)	(27.8)
I[kids]			-4722.3***	-944.0**
			(432.5)	(404.4)
Constant	-12504.8***	41569.1***	-11731.7***	40499.9***
	(237.9)	(1619.6)	(265.5)	(1723.2)
R^2	0.268	0.443	0.283	0.448
σ_e	6,860.2	5,981.6	6,782.0	5,959.0
σ_u	7,274.3	17,492.9	7,211.0	16,773.6
N	97,966	97,966	97,966	97,966

Sample criteria: Individuals buying a property during 2008-2014. *Note:* The dependent variable is the estimated implicit price of violent crime using estimates from Table 1.3. Standard errors in parentheses clustered at the individual level. Removing observations with $MWPT > p99^{MWTP}$ or $MWTP < p1^{MWTP}$. Violent crime measured as number of victims by parish code. σ_e: variance of random error component. σ_u: variance of fixed error component.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Summary statistics of WTP for 80 pct. reduction in violent crime using Rosen 2nd stage with FE (DKK)

Spec.	Model	Mean	S.d.	Median	N
1	CTD	64,700	59,927	46,558	1,583
	Rosen	83,832	87,228	54,551	1,583
2	CTD	75,758	70,061	52,887	1,563
	Rosen	96,805	98,664	61,350	1,563
3	CTD	70,059	64,615	49,061	1,701
	Rosen	91,733	93,781	60,594	1,701
4	CTD	68,264	70,850	42,465	1,631
	Rosen	86,606	95,603	51,159	1,631

Note: Violent crime measured as number of violent crime victims by parish code. Only computing summary statistics based on observations who fulfill: $q_0 > 0, q_0 \leq 50, q_0 \cdot 0.2 \leq 50, q_0 \cdot 0.8 \leq 50, \mu_1 > p1^{\mu_1}, \mu_1 < p99^{\mu_1}, \mu_0 > p1^{\mu_0}, \mu_0 < p99^{\mu_0}$ and who bought two homes during the period. Rosen specifications are the ones in Table B1 and individual u_i has been predicted and added to the constant.

Table B3: Summary statistics of WTP for 80 pct. increase in violent crime using Rosen 2nd stage with FE (DKK)

Spec.	Model	Mean	S.d.	Median	N
1	CTD	-82,902	86,446	-54,185	1,583
	Rosen	-63,770	59,015	-46,249	1,583
2	CTD	-94,576	96,243	-61,999	1,563
	Rosen	-73,529	67,346	-53,250	1,563
3	CTD	-91,722	93,259	-61,821	1,701
	Rosen	-70,048	63,192	-50,712	1,701
4	CTD	-83,287	89,592	-53,632	1,631
	Rosen	-64,945	63,745	-44,714	1,631

Note: Violent crime measured as number of violent crime victims by parish code. Only computing summary statistics based on observations who fulfill: $q_0 > 0, q_0 \leq 50, q_0 \cdot 0.2 \leq 50, q_0 \cdot 0.8 \leq 50, \mu_1 > p1^{\mu_1}, \mu_1 < p99^{\mu_1}, \mu_0 > p1^{\mu_0}, \mu_0 < p99^{\mu_0}$ and who bought two homes during the period. Rosen specifications are the ones in Table B1 and individual u_i has been predicted and added to the constant.

C Horizontal MWTP function

Table C1: 1st stage OLS regression of log(real property price) with only linear terms by year

	(1) 2008	(2) 2009	(3) 2010	(4) 2011	(5) 2012	(6) 2013	(7) 2014
Violent crime	-0.00126*	-0.00225**	-0.00376***	-0.00215**	-0.00231***	-0.00033	-0.00133***
	(0.0007)	(0.0011)	(0.0009)	(0.0010)	(0.0007)	(0.0006)	(0.0004)
Property crime	-0.00002	0.00014***	0.00013***	0.00006***	0.00008***	0.00004**	0.00007***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$PM_{2.5}$	-0.08973**	-0.22616***	-0.08176**	-0.30361***	-0.25292***	-0.20729***	-0.54059***
	(0.0422)	(0.0437)	(0.0390)	(0.0583)	(0.0622)	(0.0535)	(0.0687)
# sqm sold	0.00014***	0.00014***	0.00018***	0.00016***	0.00014***	0.00016***	0.00012***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
I[apartment]	-0.32804***	-0.36035***	-0.32271***	-0.33069***	-0.34356***	-0.30353***	-0.32359***
	(0.0133)	(0.0135)	(0.0114)	(0.0133)	(0.0131)	(0.0125)	(0.0111)
I[bath]	0.06597*	0.12222***	0.08163**	0.03809	0.07762**	0.08064**	0.05222
	(0.0358)	(0.0357)	(0.0354)	(0.0409)	(0.0392)	(0.0364)	(0.0430)
I[preserved]	0.03010	0.06202*	-0.00145	0.05077	0.00631	0.07861***	0.11206***
	(0.0393)	(0.0375)	(0.0340)	(0.0340)	(0.0344)	(0.0301)	(0.0314)
Build year	0.00193***	0.00224***	0.00219***	0.00222***	0.00250***	0.00202***	0.00205***
	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
# rooms	0.15403***	0.14992***	0.14052***	0.15478***	0.14953***	0.15176***	0.15972***
	(0.0034)	(0.0035)	(0.0031)	(0.0037)	(0.0034)	(0.0036)	(0.0032)
Dist. Cph.	-0.00395	-0.02550***	-0.02627***	-0.02243***	-0.02458***	-0.0267***	-0.02522***
	(0.0055)	(0.0059)	(0.0056)	(0.0062)	(0.0056)	(0.0052)	(0.0046)
Pop. dens.	-0.003**	-0.007***	-0.005***	-0.007***	-0.004***	-0.004***	-0.001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	12.0***	12.0***	11.0***	14.0***	12.0***	13.0***	16.0***
	(0.4867)	(0.5332)	(0.4631)	(0.7194)	(0.6249)	(0.5380)	(0.6761)
School district FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.697597	0.675861	0.655373	0.653081	0.645719	0.640704	0.651151
N	7,889	6,937	8,845	7,316	8,222	8,745	10,588

Sample criteria: Individuals buying a property in the year. *Note:* Standard errors in parentheses clustered at the individual level. Violent and property crime (Prop. crime) measured as number of victims by parish code. $PM_{2,5}$ measured in μ_g/m^3. Distance to Copenhagen center (Dist. Cph.) measured in km. Population density (Pop. dens.) measured as 1,000 inhabitants per km^2.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

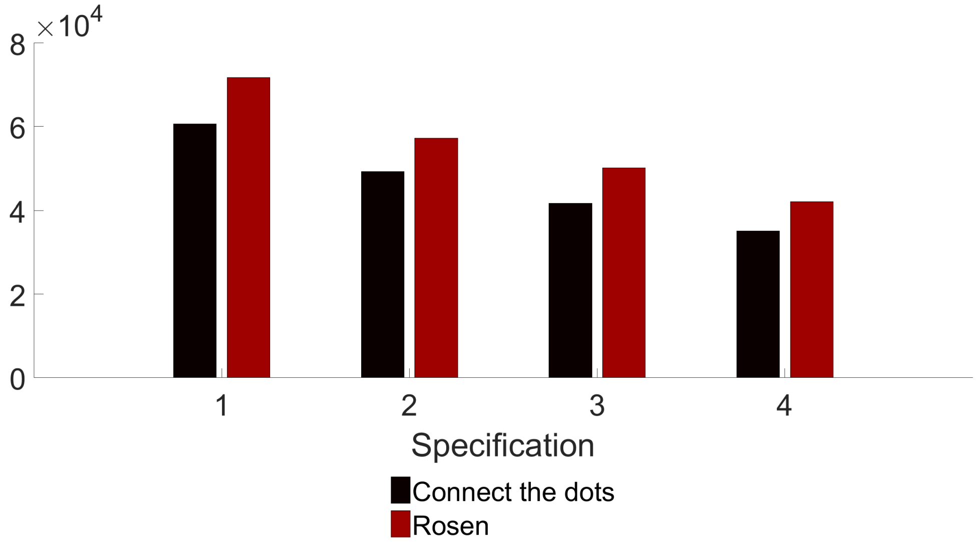
Table C2: Summary statistics of WTP (DKK) for violent crime changes using horizontal MWTP function

Policy change	Mean	S.d.	Median	N
80 pct. increase	-42,401	27,845	-35,061	2,535
80 pct. reduction	42,401	27,845	35,061	2,535

Note: Violent crime measured as number of violent crime victims by parish code.

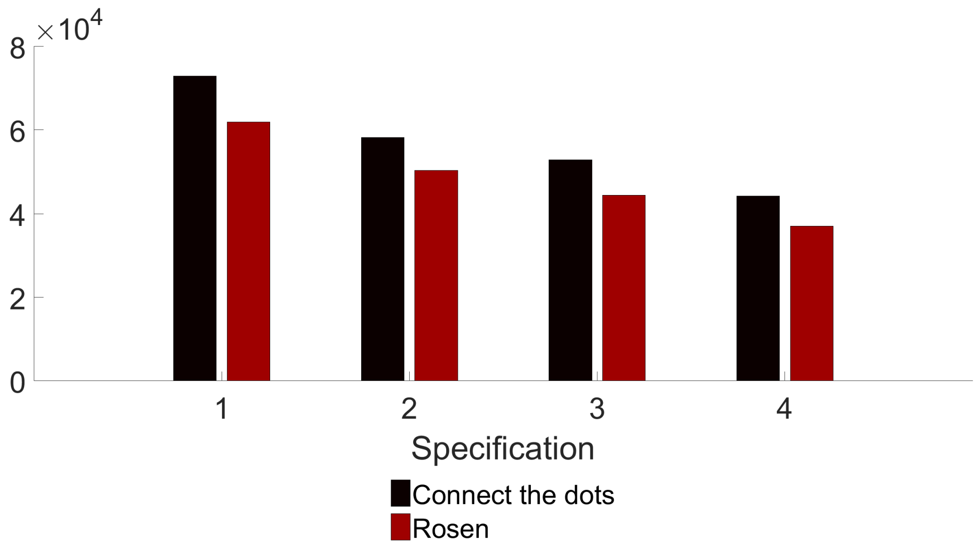
D WTP comparisons

Figure D1: Median WTP (DKK) for 80 pct. reduction in violent crime by method and specification



Note: Corresponding to median of Table 1.8.

Figure D2: Median WTP (DKK) for 80 pct. increase in violent crime by method and specification



Note: Corresponding to median of Table 1.9.

Table D1: Summary statistics of WTP as share of income for 80 pct. reduction in violent crime

Spec.	Model	Mean	S.d.	Median	N
1	CTD	0.187	0.163	0.150	1,891
	Rosen	0.222	0.193	0.178	1,891
2	CTD	0.149	0.120	0.124	2,431
	Rosen	0.172	0.141	0.143	2,431
3	CTD	0.142	0.114	0.120	2,493
	Rosen	0.162	0.130	0.136	2,493
4	CTD	0.129	0.133	0.106	2,493
	Rosen	0.146	0.143	0.120	2,493

Note: Violent crime measured as number of violent crime victims by parish code. Income is total annual real income. *Sample criteria:* Total real income below 99th percentile and above 1st percentile.

Table D2: Summary statistics of WTP as share of income for 80 pct. increase in violent crime

Spec.	Model	Mean	S.d.	Median	N
1	CTD	-0.225	0.194	-0.179	1,891
	Rosen	-0.190	0.164	-0.152	1,891
2	CTD	-0.175	0.144	-0.145	2,431
	Rosen	-0.151	0.122	-0.126	2,431
2	CTD	-0.164	0.131	-0.136	2,493
	Rosen	-0.144	0.115	-0.121	2,493
2	CTD	-0.145	0.128	-0.121	2,493
	Rosen	-0.128	0.115	-0.107	2,493

Note: Violent crime measured as number of violent crime victims by parish code. Income is total annual real income. *Sample criteria:* Total real income below 99th percentile and above 1st percentile.

Table D3: Summary statistics of absolute change in violent crime numbers following 80 pct. change

Spec.	Mean	S.d.	Median	N
1	5.47	1.03	6.02	1,927
2	4.59	0.97	4.70	2,469
3	4.34	0.38	4.68	2,535
4	4.14	1.13	4.35	2,524

Note: Violent crime measured as number of violent crime victims by parish code.

Bibliography

Bajari, P. and C. L. Benkard (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of political economy 113*(6), 1239–1276.

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. *Journal of political Economy 95*(1), 81–88.

Bishop, K. C. and A. D. Murphy (2011). Estimating the willingness to pay to avoid violent crime: A dynamic approach. *American Economic Review 101*(3), 625–29.

Bishop, K. C. and C. Timmins (2018). Using panel data to easily estimate hedonic demand functions. *Journal of the Association of Environmental and Resource Economists 5*(3), 517–543.

Bishop, K. C. and C. Timmins (2019). Estimating the marginal willingness to pay function without instrumental variables. *Journal of Urban Economics 109*, 66–83.

Brown, J. N. and H. S. Rosen (1982). On the estimation of structural hedonic price models.

Ekeland, I., J. J. Heckman, and L. Nesheim (2004). Identification and estimation of hedonic models. *Journal of political economy 112*(S1), S60–S109.

Epple, D. (1987). Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products. *Journal of political economy 95*(1), 59–80.

Mendelsohn, R. (1985). Identifying structural equations with single market data. *The Review of Economics and Statistics*, 525–529.

Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, 1257–1278.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy 82*(1), 34–55.

Statistics Denmark (2016). Pendlingsområder - metode. Technical report.

Willig, R. D. (1976). Consumer's surplus without apology.

Chapter 2

A Dynamic Equilibrium Model of Commuting, Residential and Work Location Choices

Christian Langholz Carstensen, University of Copenhagen

Maria Juul Hansen, University of Copenhagen

Fedor Iskhakov, Australian National University

John Rust, Georgetown University

Bertel Schjerning, University of Copenhagen

Abstract

In this chapter we develop and estimate a dynamic equilibrium life cycle model of residential and work location choices. In our model, commuting is endogenously determined by the distance between work and residence, and house prices are determined in equilibrium. We estimate the model using Danish register data for the entire population of households in the Greater Copenhagen area (GCA). Assuming a fixed supply of housing in the short run, we consider the effects on house prices, job mobility, residential sorting and commuting in two counterfactual equilibria with i) increased supply of housing in the center of the GCA and ii) increased cost of commuting between all residential and work location regions. We find that i) results in lower prices in equilibrium for all regions and a higher degree of urbanization. ii) implies lower average commute times, but also a higher share of people in non-employment, in particular for residents outside of the GCA. The equilibrium prices drop in peripheral regions with low job density.

Acknowledgements

The authors would like to acknowledge funding from the URBAN research project financed by Innovation Fund Denmark. It should be noted that this chapter builds on Maria Juul Hansen's master's thesis "A Dynamic Structural Approach To Individual Home And Job Location Decisions" (University of Copenhagen, 2017), which set up and simulated a dynamic equilibrium model of residential and work locations. The current chapter extends that work in several and important dimensions, including estimation of the model. As the curse of dimensionality was already a serious challenge in the master's thesis, it allowed only two time periods and a very limited choice set. In the current chapter we have made significant progress also in that regard.

1 Introduction

Denmark belongs to a large group of countries that are undergoing a process of strong urbanization and spatial concentration of economic activity. While this has led to increased productivity in the larger cities through agglomeration mechanisms (such as better accessibility of firms to both their markets and supply of specialized labor) it has also resulted in several major societal challenges, including the large and systematic flows of people and jobs with increased traffic congestion and large increases in house prices in urban areas. The result is a changed demographic composition of cities and increased regional inequality. The steady decoupling of urban and rural housing price trends is a clear testament to the latter effect and it has led to an increased inequality in wealth across regions.

A number of policies have been suggested to ameliorate some of the downsides of this development, including infrastructure investments and relocation of government jobs from Copenhagen to the rest of Denmark. However, the dynamic effects of such policies are not well understood due to the complexity of households' commute choices, job and residence mobility, and their interactions with the housing market. Dynamics are crucial for moving and job location decisions since these are made under uncertainty about future house prices and job opportunities, and due to the substantial fixed moving costs that are implied by partly irreversible investments in property and the cost of searching for a new job. These long-run implications of location decisions together with key life events imply that intertemporal incentives are likely to underlie much of the observed behavior. It would be unrealistic to assume that households are not forward-looking and that location choices and housing demand are stationary through the life cycle. Due to commute costs, we must keep track of these mechanisms simultaneously to credibly predict the quantitative consequences of such policies on local house prices, commuting patterns, residential sorting, and inequality. This poses a key challenge that we will address in this chapter.

We develop and estimate a dynamic equilibrium life cycle model of residential and work locations to investigate how individuals choose the location of their job and residence, and how urbanization affects house prices, commuting patterns, and demographic composition of cities. We use this model to consider the effects of a number of changes to the economic environment including the effects of an increase in the local housing supply and increased cost of commuting between all residential and work location regions. In doing so, we consider the implications for job mobility, residential sorting, commuting patterns, local housing demand, and house prices.

The modeling framework is a structural life cycle model formulated at the individual level, where people simultaneously choose residential and job locations by taking into account their need for housing, their wage potential at different job markets, commuting costs, amenities, and moving costs. Our model is inspired by the work of Buchinsky et al. (2014), which we extend to a dynamic discrete-continuous choice setting with endogenous house prices and equilibrium constraints, where households make a continuous choice of house size and discrete choice of location of work and residence. Commuting is endogenously determined by the distance between chosen work and residence, and house prices are determined in equilibrium.

By using a life cycle model, we account for individual heterogeneity. To avoid the complexity associated with a complete modeling of life cycle consumption, savings, and borrowing decisions, we assume individuals have quasi-linear utility functions and do not face any borrowing constraints. Instead, we approximate some of these effects by allowing marginal utility of money to be increasing in individual income to reflect that richer households have a higher demand for housing and sort into more expensive regions.

Given the complexity of the model, we also abstract from fully modeling the dynamics of the housing size decision, and we assume that households ignore any adjustment costs of changing house size. The abstraction from such adjustments costs as well as the stylized modeling of consumption, savings, and demand for house size, effectively means we do not distinguish between home owners and renters, but rather model everyone as renters who pay a fixed share of the total house price each period. This essentially amounts to assuming that renting a home or owning a home are nearly equivalent with the "rent" a homeowner pays consisting of the sum of mortgage payments and the opportunity cost of the equity capital the owner has in the home.

The described simplifications allow to keep the model computationally tractable while studying location decisions in more detail. With quadratic utility of house size we can derive a closed form for the optimal amount of housing in each residential location as the solution to a static housing subproblem that can be solved independently of the overall discrete dynamic programming problem governing residential location choices. Given the solution of optimal choice for house size for a given residential location, all dynamic decisions are discrete (job and residential locations) and all of the state variables of the model are also discrete.

Even with these simplifications a challenge of the model presented in this chapter is that both the number of states and choices are proportional to the number of combinations of work and residence locations. To ameliorate this curse of dimensionality and avoid solving the model for extremely many combinations of states and choices, we aggregate location choices to the municipal level and restrict attention to the island Zealand (which includes Copenhagen and its surroundings). Out of the 98 municipalities in Denmark, we consider the 16 municipalities located in the Greater Copenhagen Area in detail as well as the outside region (rest of Zealand).

We estimate this model using high-quality Danish administrative data that allows us to track the entire population of households, its members, their jobs, and residential locations for the period 1992-2016. In the estimation we focus on the subperiod 2005-2010. These data contain very detailed linked information about location and size of houses, individual employment, wages, and residential and work location dynamics for all individuals and households in Denmark. To estimate the model and compute the equilibrium house prices, our model is repeatedly resolved for many types of individuals as a subroutine of both i) a structural nested backward induction maximum likelihood estimation routine to estimate the preference parameters, and ii) an equilibrium solver that finds paths of housing prices that equate the demand for the available supply of houses in Denmark implied by a microaggregation and simulation of the model based on the parameters estimated in part i).

We assume a fixed supply of housing in the short run and thus abstract from the longer run dynamics where new houses are built in response to changes in house prices. Given that the supply of housing is quite inelastic compared to housing demand, we think this is a reasonable approximation in the shorter run. Using this model, we compute the effects of i) increased housing stock in the two most urbanized areas of Denmark, and ii) increase in marginal cost (per hour of) commuting between residential and work-location regions.

We find that i) results in increased degree of urbanization as households move from the peripheral regions towards the center. Equilibrium prices fall in all locations, especially in the two locations where the policy was implemented. ii) implies lower commuting on average among employed individuals, but also increases the share of non-employed individuals, especially for those residing in the remote regions, where the equilibrium prices also drop.

The rest of the chapter is organized as follows: Section 2 gives an overview of the existing literature and summarizes our contribution relative to existing studies. Section 3 presents the data sources and describes the institutional setting. It also provides descriptive evidence of house prices, residential and work location choices, and the resulting commuting and spatial sorting. Section 4 outlines the model. Section 5 introduces the algorithm that we use to solve and estimate the model. Section 6 describes how we solve for equilibrium prices in the short run. Section 7 presents the parameter estimates and model fit and makes a number of counterfactual simulations focusing on how changes in the local house stock, job density, and commuting costs affect house prices and optimal location decisions. Section 8 concludes and gives directions for future research.

2 Related Literature

This chapter builds on and contributes to several strands of the literature covering theoretical and empirical models of location choice in continuous and discrete settings. This section provides a short review of the literature, leading to the dynamic equilibrium model of simultaneous choice of both residence and work location that we develop in this chapter.

2.1 Monocentric city model

The literature on household location decisions is based on theory and methodology developed in industrial organization and labor economics. The literature deals with sorting, i.e. with the mechanism that market forces make people with similar preferences and personal characteristics self-select and cluster in certain locations. The urban economics literature is a separate research field dating back to seminal papers by Hicks (1932) and Sjaastad (1962). They made economists interested in understanding the driving forces and implications of how individuals locate. But there were also other contributions that led to this rising interest. Tiebout (1956) was the first to argue that when people sorted ("voted with their feet") in terms of residential location they implicitly revealed their demands for local public goods that were exclusively available in different locations. He focused on the effect that fiscal competition had on income sorting between jurisdictions. At the same time Alonso et al. (1964) developed the monocentric city model, which was enriched by Mills (1967) and Muth (1969). In contrast to the Tibeout model, this was a model for income sorting across geographical space and has become the foundation for many analyses of locations within a city. The main idea was to consider job locations to be exogenous at the center of the city which reduces the residential problem to a choice of how far to commute to one's job, thus ignoring any other travel time and distance that the individual might use to decide on his optimal location (e.g. travel to shops, family or daycare). In a strict sense, the model thereby took as given that people like to live in big cities, but does not explain why they wish to live close to the Central Business District (CBD), except the fact that commute is shorter since all jobs are located in the CBD. It therefore focused on the trade-off between living close to the CBD to get a shorter commute at the cost of more expensive housing. Overall, the consumer maximization problem is standard, except that consumers also choose a residential location (a distance from the CBD) on top of the optimal amount of housing and a composite good. The housing prices at each location respond to offset the marginal decrease in utility that stems from living further away from the CBD. Besides modelling the consumer behaviour, a construction sector which builds houses by use of land and capital is part of the set-up. Land prices are therefore endogenously determined in equilibrium and the model predicts that as land prices increase closer to the CBD, construction firms tend to build with a higher density, i.e. to build tall apartment blocks rather than one story houses.

Extensions of the monocentric city model framework include work by, among others, McMillen (2006) and Ahlfeldt and McMillen (2015) (who focus on the intensity of development) and the Ogawa and Fujita (1980) model that endogenises the location of firms as well to explicitly model agglomeration economics. These other branches of the urban economics literature are out of the scope of our short-run study.

The general scope of the monocentric city model is to predict how geographical space is divided between residential and land use and thus to predict the size of urban areas. The urban area increases until the marginal value of devoting more land to cities equals the marginal value of decreasing agricultural land use. The most prominent recent paper within this branch of the literature is Ahlfeldt et al. (2015). They set up a general equilibrium model of internal city structure where people select a combination of residence and job and where wages and prices of land adjust in response to moving patterns. Their focus is on estimating the extent of agglomeration on productivity, not on the location choice per se. Even though they do study how equilibrium land prices change in response to altered moving patterns, they can only estimate the long-run effects in a static modeling framework.

2.2 Reduced-form models

The discrete choice framework that has been increasingly popular in recent decades was initiated by McFadden (1974). He provided a methodology for analyzing choice behaviour when the agent optimizes with respect to choices from a discrete rather than continuous set as in the monocentric city approach. He was also the first to really contribute to the discrete sorting literature in McFadden (1978), where individuals choose a specific location rather than just a certain distance from the CBD.

Another distinction within the location choice literature is that it originally centered around two types of models: human capital models on the one hand and hedonic models on the other. Human capital models are, among others, motivated by Topel (1986). Hedonic models were introduced around the same time as McFadden came up with the discrete choice estimation methods, namely by Rosen (1974) and further extended by Roback (1982). Rosen (1974) set forth a method for estimating marginal willingness to pay (MWTP) for goods for which there were no formal markets such as air pollution, crime rates and scenic views. Introducing hedonic price functions, he explained how researchers could use data on observed location choices by households and housing prices for the different locations to compute implicit price indices of these non-traded amenities. Whereas the human capital models argue that migration occurs due to disequilibrium in the labor market (people move to a new location to earn a higher wage), the hedonic approach asserts that individuals might move even if housing and labor markets are in equilibrium because they might have changed demand for location-specific non-traded amenities. While the human capital literature sees earnings differentials as temporary circumstances that will mitigate when workers relocate in order to get the highest possible return on their human capital investments, hedonic models explain how wage and housing price differentials may not be completely eliminated, as they may fail to compensate individuals for location-specific (dis)amenities. With the emergence of the sorting literature, which has been thorougly surveyed by Kuminoff et al. (2013), these two approaches were combined into a unified framework.

2.3 Structural static models

To explicitly take this sorting into account, a growing number of papers structurally model location decisions, though until very recently mainly by using static models. One example is Borjas (2000) who looked at how immigrants affect the equilibrium in the local labor markets across several geographic areas. This paper points out that the possibility of moving to another location for work among natives is not sufficient to cancel all wage differentials across locations. This is because people face high moving costs that make them reluctant to move for the best wage offer. In contrast, immigrants from other countries do not to the same extent incur moving costs on top of those associated with leaving their home country and are therefore more prone to settle and work in the area characterized by the best wage offer. Bayer et al. (2009) is another example which stresses the importance of not only relying on the first-order conditions from the traditional hedonic model, but rather combine the frameworks originally presented in McFadden (1978) and Rosen (1974). This allows for explicitly accounting for moving costs which is necessary in order to get unbiased estimates of MWTP for non-traded amenities. Besides the change in amenities, the rising prices in cities across the world has also led researchers to study the effect of altering the housing supply, including Nathanson (2019) which estimates a static equilibrium model of residential location, while the job search behaviour across local labor markets has been studied in Manning and Petrongolo (2017). They find that there is a sharp decay in attractiveness of jobs as they get further away from the home, which also speaks in favor of modelling the decisions of home and job jointly as we do in this chapter.

However, the combined modelling of work and home locations has not been the focus of the literature until recently. Tsivanidis (2019) is a very recent exception. He estimates a static general equilibrium model of home and work locations as well as car ownership and housing size to quantify the effects on sorting of workers and their welfare from a large infrastructure investment. He documents that it is indeed essential to account for the spatial reallocation of workers and general equilibrium effects as we do in the model of this chapter. A related question is discussed in Albouy and Stuart (2019) which decomposes the determinants of residential location choices into a number of amenities using a neo-classical spatial equilibrium framework. They conclude that quality of life (i.e. factors related to the utility of residing in a region) are more important than trade productivity (i.e. determinants that affect people's taste for jobs there), but that both have an effect.

2.4 Structural dynamic models

The lack of appropriate data and computational difficulties are the main reasons why the literature has focused on static models for so many years. Dynamics are crucial, however, as outlined in Section 1. Kennan and Walker (2011) were the first to add dynamics to a structural model where individuals optimize over a set of residential locations each period.

There were a few predecessors in the dynamic location choice literature such as Holt (1997) and Tunali (2000). However, they both did not distinguish between alternative locations, but modeled only the move-stay decision. Dahl (2002) did do so by allowing individuals to choose between all U.S. states, but individuals only made one moving decision for their entire life. Gallin (2004) looked into how changes in expected future wages affected net migration in an area, but he used aggregates and thus did not model how the individual responded to this. Lastly, Gould (2007) studied how workers choose between residential locations and occupations, but only distinguishes between rural and urban locations.

Kennan and Walker (2011) was the first paper to broaden the application to a more detailed setting, where they allow for many different locations (U.S. states). However, they restrict people to live and work in the same location. They find that better income prospects associated with moving to another location is an important driver of migration decisions. Bishop (2012) also uses a dynamic model but has another focus: namely to set up an equilibrium model to estimate willingness to pay for air quality while controlling for moving costs and forward-looking behavior. In her model, individuals are forward-looking with respect to local amenities, and they can move in expectation of how they evolve over time. Her model can also result in a huge choice set, but she does not address the question of how work and residential locations are interconnected. Bayer et al. (2016) adopt her approach, but go one step further and estimate the willingness to pay for several non-traded amenities of a neighborhood. Additionally, their paper allows for household wealth to evolve endogenously with housing prices such that households' expectations about future housing prices can affect their decisions. Location choice hence becomes dynamic both due to moving costs and wealth accumulation.

Another recent contribution to the literature is Oswald (2019). He models the choice of consumption, home ownership, and residential location. Whereas there has been a tradition of ignoring the choice of home ownership, he integrates this decision into the model to account for the fact that home owners' wealth declines when house prices do, while renters may benefit from lower rents. He argues this is important as the option of moving is a way to self-insure against local shocks to the housing and labor market and estimates the value of this self-insurance mechanism. The question of owning or renting is also taken up in Favilukis et al. (2019) which calibrates a rich dynamic equilibrium overlapping generations model for commuting, consumption, housing, residential location, and own/rent decisions for New York City. They use the model to assess the implications of zoning and rent control policies. However, they assume away moving costs and assume a two-alternative choice set for residential locations. Halket et al. (2015) studied the allocation of properties to ownership and rental from a supply side perspective, while Attanasio et al. (2012) added a choice of housing size and consumption over the life cycle on top of the choice of owning or rentning. They abstract from the discrete location decisions though, but find that demand for housing does indeed react to prices and income shocks.

While the papers mentioned above do not model detailed labor markets, a number of papers concentrate on the dynamic aspects of migration decisions, employment status, and job search such as Ransom (2016), Schmutz and Sidibé (2015) and Mangum (2015). The latter two include equilibrium constraints on the labor market and real estate market, respectively, but none of them model the joint decisions of home and work. Guglielminetti et al. (2017) model the job search process for unemployed people and take home location into account, but for two possible locations.

Buchinsky et al. (2014) are the first to structurally estimate a dynamic model of residential location that also adds the choice of work location. The model mainly builds on Kennan and Walker (2011), but a very important extension is that they distinguish between home and work locations. Hence, individuals choose home and job locations as well as labor market status and sector each period. By relaxing the assumption of zero commute they are able to model commute costs. Moreover, they allow people to have expectations about the job offerings before they decide on their locations. This complicates the empirical implementation of the model since job offers are not observed, but also adds a more realistic aspect to the model. We employ a similar approach in the model of this chapter. Even though their model is very rich in terms of its choice set, it is a partial equilibrium model where both wages and prices are taken as exogenous. The authors provide arguments why this is not too important for their setting. Another limitation of the paper is that it restricts attention to the case of hihgly educated immigrants from the Soviet Union migrating to Israel, which is likely to be a very selected group. Their results therefore mainly regard immigrants instead of people who migrate within a country, which is the focus of the current chapter.

3 Data and Institutional Background

This section provides the description of the data we use to estimate the model. We also provide descriptive statistics on property prices, urbanization, overall life cycle patterns of moving home and job and demand for housing size, while we use Section 7 to go into more details of the location and sorting patterns when we present the model fit from the estimation.

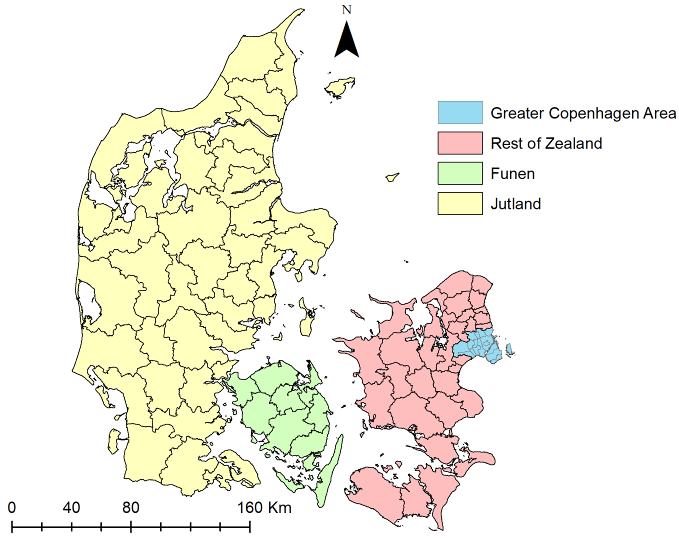
3.1 Danish register data

We use administrative data provided by Statistics Denmark which holds information on every individual living in Denmark in the period 1998-2010[1], although for estimation we use the years between 2005 and 2010. The personal registers contain information on individual background characteristics like home and work address, education and number of children, while other registers hold data on home sales and prices, home owners and dwelling characteristics. Since we focus on the Greater Copenhagen Area in the current

[1]But can be extended to 1992-2016.

chapter, we redefine regions according to Figure 2.1. We include everyone who has either work or home region within the Greater Copenhagen Area[2]. Below we go over each of the separate data sources.

Figure 2.1: Definition of regions



In 2007, a municipality reform took place in Denmark which reduced the number of municipalities from 271 to 98. We are able to track how municipalities were combined, and use the more coarse definition for all years.

Register data of individual background characteristics is recorded on January 1st each year and list all individuals who are officially registered with an address in Denmark. Each individual in the registers is associated with a family identifier[3] and an anonymized version of their official social security number which allows linking of different registers on the individual level. We use the age, gender, address identifier[4], whether she has children, how old the children are, and if she lives with a partner (we track both marriages and cohabitations) as background characteristics. The data on workplace and other workplace-related variables such as industry and occupation, along with the wage, are recorded in the end of November, and are linked to the individual data from the previous year. For each individual who has more than one wage-earning job, we use information from the main occupation which is determined by the largest source of income. Individuals who do not work are either classified as unemployed or outside the labor force.

Commute time data come from The Danish Traffic Model (LTM) which has been developed by researchers at The Technical University of Denmark (DTU). They divide

[2]Appendix A provides more details on the geographic units in Denmark. Including the entire Denmark is left for future research.

[3]Families are defined as everyone on an address who are related biologically, registered partners, or opposite gender couples. Singles are families of one. However, for each individual we also observe the identifier of their partner.

[4]Addresses are anonymized within a region, but associated with unique identifiers.

Denmark into 907 traffic zones (LTM zones) and modelled commute time between each pair of regions. Since our model is formulated in terms of municipalties, not LTM zones, we compute a commute time measure by each transport mode between any pair of LTM zones within a municipality pair. For a given pair of LTM zones in a municipality pair, we use the commute time from the mode with the shortest commute time. We then weigh the commute time of each observed LTM pair in the municipality pair with its estimated number of trips by that mode from the traffic model and thereby get a trip-weigted average commute time between any pair of municipalities.

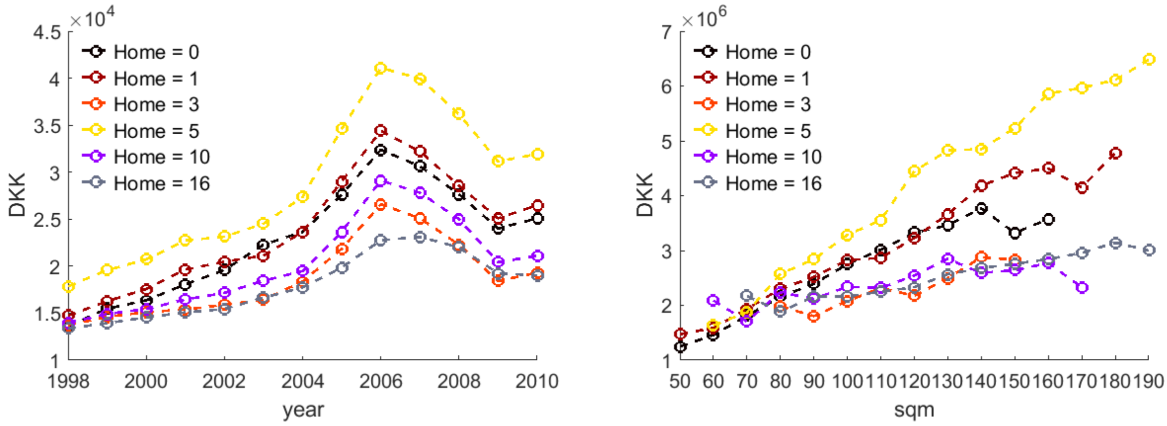
Information on property prices come from the sales transaction register. We deflate all sales prices by the consumer price index with 2011 as the reference year. We use only private sales and disregard properties with commercial-only purpose. A more detailed description of the Danish register data is available in Appendix F.

3.2 Property prices

Property prices in the Copenhagen area, and especially in the center and northern parts, have tended to be higher than prices in the remaining regions of Denmark for the last three decades. Since the begininning of the 1990s, however, the prices by regions have diverged. The hierarchy of regions from lowest to highest price per square meter is more or less unchanged over time. But the prices in central parts and north of Copenhagen and its nearest surroundings started to increase in the mid 1990s, the rest of the country experienced much more modest or stagnating prices. Figure 2.2a documents this evolution since 1998 to 2010. In the years just before the financial crisis, prices of all regions began to rise, but still steeper for areas close by Copenhagen. The gap between prices of the most urbanized areas compared to the suburbs and rural areas widened as a result. While prices in the center of the GCA (Copenhagen and Frederiksberg) continued to grow until the outset of the financial crisis in 2006, the flow of people to that same area showed the reverse trend until 2005, cf. Figure 2.3. As prices reached the highest level in decades in 2005, the net outflow from the center of the GCA topped at 3,000 people. From 2006, on the other hand, the net outflow rose towards 0 in 2010 while prices dropped significantly during the years 2006-2009. Historically, we have therefore seen increasing prices both in times of increasingly negative net inflows to the center of the GCA and in years where the negative net inflow was getting more modest.

Another stylized fact about prices is how they relate to the size of the homes. As Figure 2.2b depicts, prices are almost a linear function of square meter living space. However, the strength of the relationship differs across regions with the Copenhagen, Frederiksberg and Gentofte north of Copenhagen showing the steepest relationships. On Zealand, which is a much less dense area, the slope is almost five times lower than in Copenhagen.

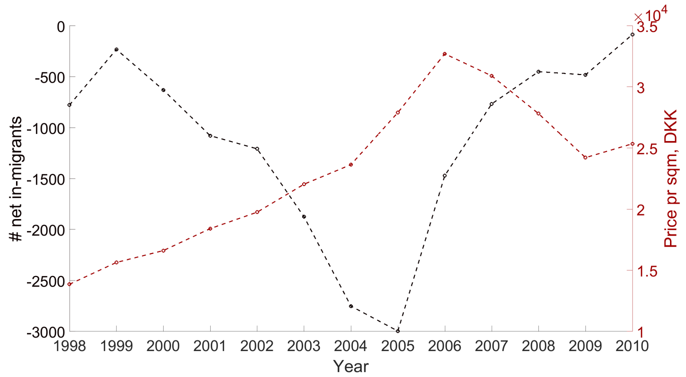
Figure 2.2: Prices by year and square meters for selected regions



(a) Price per sqm (b) Total price by sqm

Note: $Home = 0$: Copenhagen municipality, $Home = 1$, Frederiksberg, $Home = 3$: Broendby, $Home = 5$: Gentofte, $Home = 10$: Hvidovre, $Home = 16$: Rest of Zealand. The figure shows real sales prices deflated by the 2011 consumer price index.

Figure 2.3: Sales price per sqm and net in-migrants for Copenhagen and Frederiksberg regions over time



Note: Real sales prices deflated by the 2011 consumer price index.

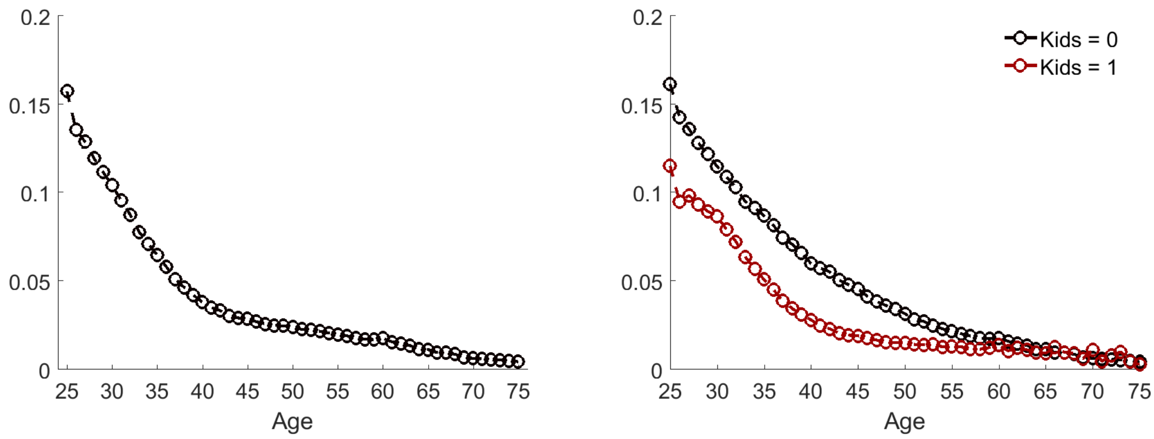
3.3 Life cycle patterns

In this section we summarize descriptive statistics on home and work location choices, commuting and house size demand. We show a very clear life cycle profile on all margins. This underlines the need for modelling these decisions as being affected by dynamic incentives.

Home decision

On average over the life cycle, 4.3 percent of people move to another home during a year. These moves only include moves across municipality boundaries. Intra-regional moves are thus disregarded, but both renters and home owners are included. The moving probability

Figure 2.4: Share moving home region by age



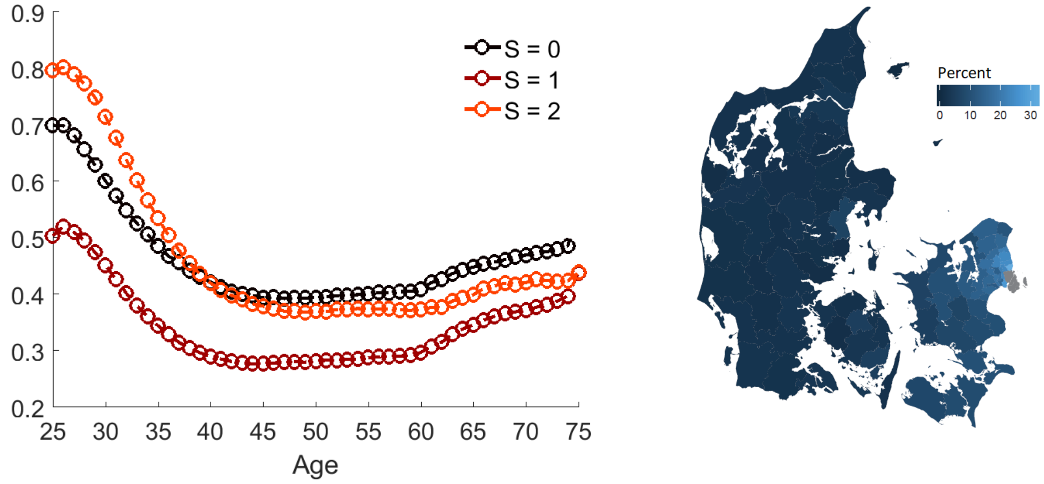
(a) Aggregate (b) By children

changes a lot over the life cycle as Figure 2.4a shows. While 25-year-olds have a 16 percent probability of moving to another home region, this drops almost linearly until the late 30s where the probability is about 4 percent. These numbers also depend significantly on the parental status of individuals. Those with children have a steeper decline in moving probability from the beginning of their 30s until the mid-40s compared to those without children. At age 35, 5 percent of people with children move their home to another region compared to 9 percent of those with no children. By the end of the 50s the shares are almost identical across the two groups, consistent with the fact that most individuals no longer have children living at home anymore. The clear life cycle perspective in moving behaviour speaks in favor of using a dynamic model.

There is furthermore a clear life-cycle profile in sorting patterns as we illustrate in Figure 2.5a. It shows the share of individuals living in Copenhagen municipality and Frederiksberg by age and schooling level. Clearly, young and highly-educated people are particularly more likely to live in these municipalities which make up the centre of the GCA. One attraction is that the GCA educational institutions and universities are primarily placed there, but it is also the place with most high-skilled jobs. This can be seen in Table B1 in appendix which shows the average job density by work region and education group. At the age of 25, 80 percent of the high-skilled people live in Copenhagen and Frederiksberg compared to 70 percent of medium-skilled and 50 percent of low-skilled individuals. The probability stagnates at around 40 percent for high- and medium-skilled by age 40 and at 30 percent for low-skilled people.

Turning to Figure 2.5b, it shows the distribution of home locations at time $t-1$ for people who move to Copenhagen and Frederiksberg at time t. Clearly, the main part of in-migrants come from municipalities located close by Copenhagen and Frederiksberg. The probability of originating from a region is thus decreasing in the distance to Copenhagen and Frederiksberg. Odense and Aarhus stand out as they comprise a relatively large share

Figure 2.5: Sorting into GCA centre, Copenhagen and Frederiksberg



(a) Share of GCA population living in Copenhagen and Frederiksberg

(b) Inflow by home $t-1$

Note: In (a): $S = 0$: low education, $S = 1$: medium education, $S = 2$: high education. In (b): shows the distribution of in-migrants to Copenhagen and Frederiksberg across home regions in $t-1$.

of the locations from which the in-migrants originate given their distance from Copenhagen. This indicates that people from other big cities of Denmark are attracted to Copenhagen despite the distance, but overall less than 22 percent of the in-migrants come from Funen or Jutland. This underlines that when we focus on Copenhagen in the estimation it is less important to model location decisions for people living on Funen or in Jutland in detail.

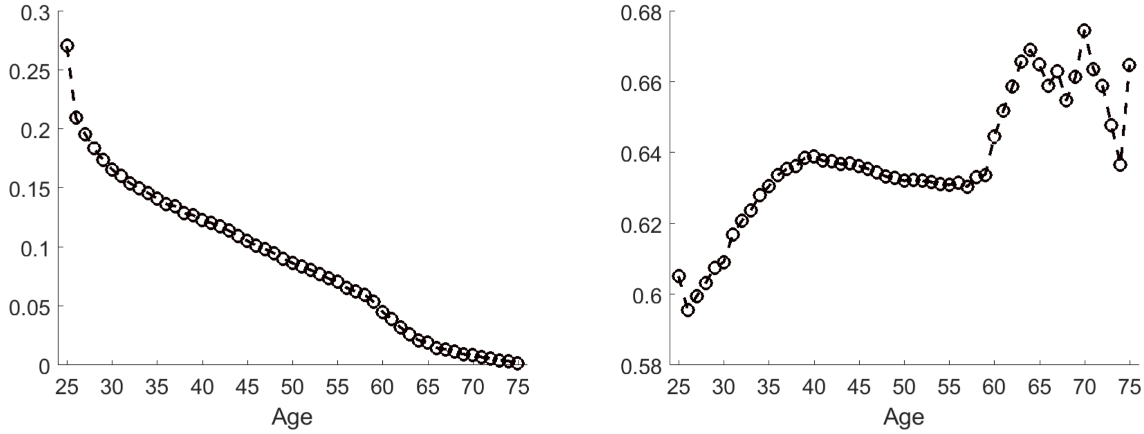
Work decision

The average share of people changing job locations from one municipality to another is 14.5 percent. As for the home moves, this share changes over the life course as Figure 2.6a pictures. While the 25-year-olds in the data have a 27 percent probability of moving to a job in another region (excluding transitions in and out of unemployment), the probability is 16 percent at age 30. From then it falls linearly to 6 percent at age 59 and then drops towards zero from there. The sharper drop from age 59 to 60 is due to people being eligible for early retirement benefits at that time.

By linking home and work locations we get the commute time. As Figure 2.6b shows, the average commute time is increasing from age 26 to 40 whereafter it starts to decline slightly. There is a sharp increase in commute time at age 60 until 65. This is explained by self-selection of workers who stay on the labor market even after they are eligible for retirement.

Having covered residential and work locations separately, Figure 2.7 shows the share of residential moves which are associated with a job move. For 58 percent of the inter-regional moves, there is no change in work region. However, 33 percent do change job either the

Figure 2.6: Share moving work region and average commute time

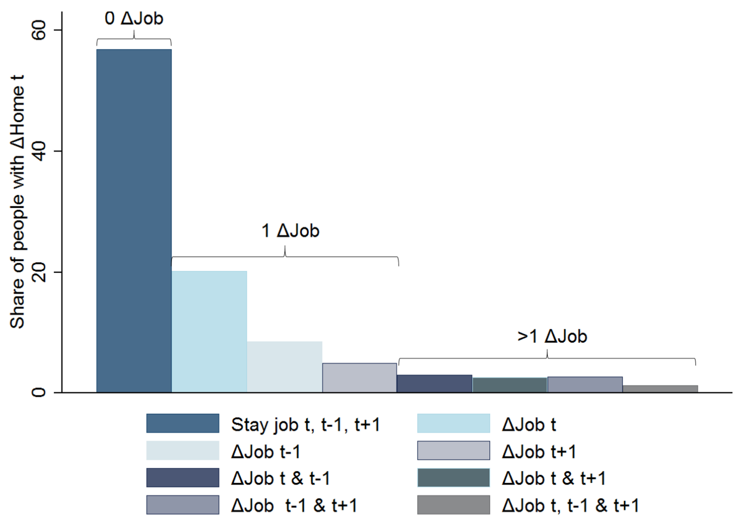


(a) Move work region

(b) Commute time (hours)

year before, in the same year or the year after they relocate their residence. The remaining 9 percent change jobs more than once during that time window. This finding underlines the importance of modelling residential and work location decisions as joint decisions.

Figure 2.7: Probability of moving home region at t by number job moves in $t-1, t$ and $t+1$

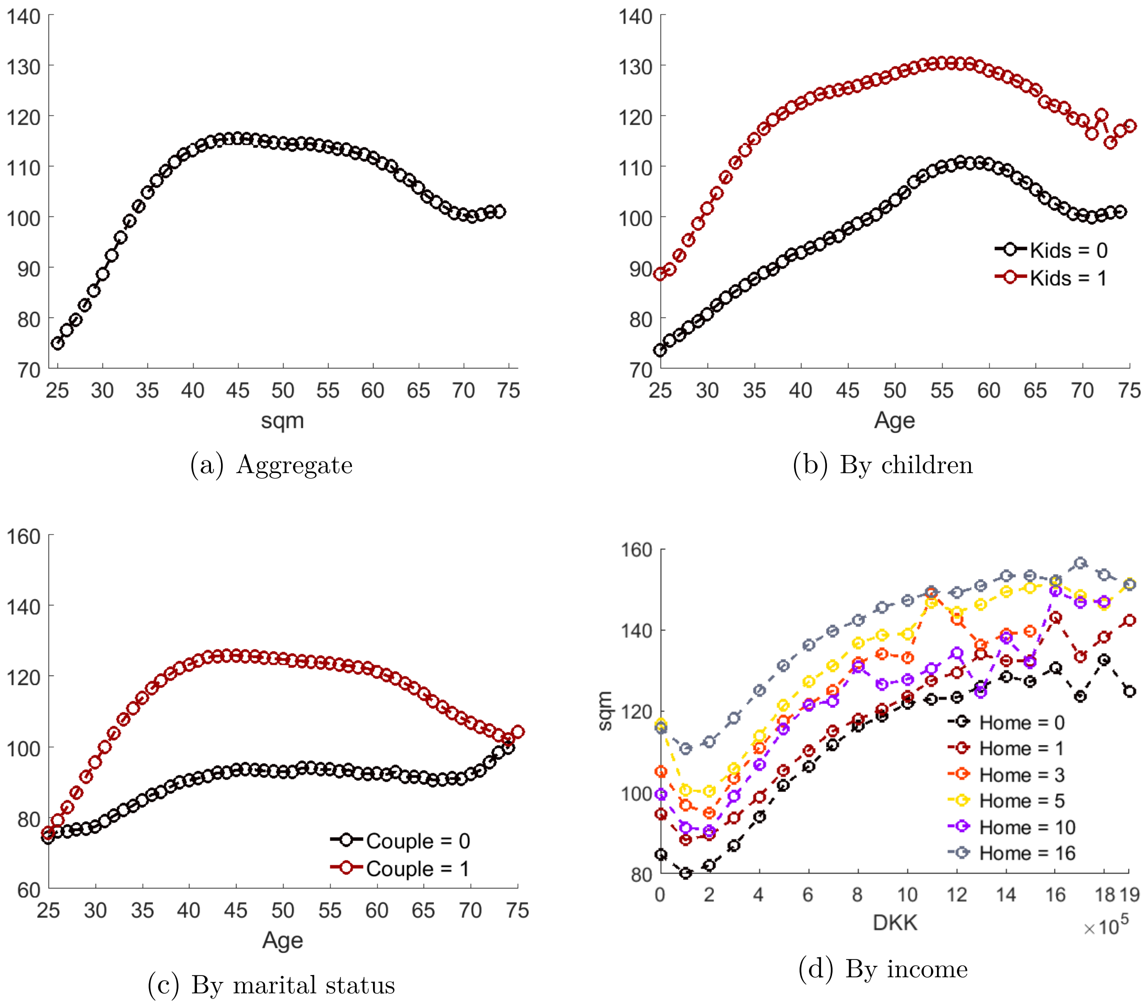


Note: ΔJob refers to "change in work region". t is the time of the home region move.

Housing demand

Another important aspect of home regions, besides square meter prices, is house sizes. Figure 2.8a illustrates how the average square meter demand evolves over the life cycle: it starts at 75 square meters at age 25, increases to 115 square meters in the beginning of the 40s, then levels off until the late 50s where it begins to decline. Multiple-member households generally demand more square meters throughout their lives as illustrated in Figure 2.8b, which distinguishes between individuals with and without children, and Figure 2.8c, which separates demand by marital status. Turning to Figure 2.8d, there clearly is a gradient in income too: higher-income people demand more square meters also when conditioning on their home region. People in Rest of Zealand therefore live in bigger houses for any given income bracket compared to those who live in e.g. Copenhagen. This is closely related to the spatial variation in house prices documented above; regions with a high square meter price are characterized by smaller houses, all else equal. Due to the very differentiated demand for square meters across regions, it is important to explicitly model this when modelling the home location choices.

Figure 2.8: Housing size demand by age and income



Note: Home $= 0$: Copenhagen, *Home* $= 1$, Frederiksberg, *Home* $= 3$: Broendby, *Home* $= 5$: Gentofte, *Home* $= 10$: Hvidovre, *Home* $= 16$: Rest of Zealand. The figure shows real sales prices deflated by the 2011 consumer price index.

4 A Dynamic Model of Residential and Work Locations

In this section we lay out the model of individual housing demand and location choice of work and residence, formulated *conditional* on prices and job opportunities in different regions. The individual choice model is embedded into the general equilibrium model in Section 6, where we derive the housing market equilibrium based on the solution to the individual problem.

4.1 Sequential choice of work and residence locations

Consider an individual decision maker who in each time t of her lifecycle, $t \in \{t_0, \dots, T\}$, chooses work and residence locations, as well as the size of residence. The decision maker in our model has unitary preferences, and we interpret t as the individual's age. Time horizon $T = 75$ is chosen to be sufficiently large such that very few changes of resident or work location occur after this age.

Modeling location choice is computationally burdensome when the number of regions is large. Denoting R the number of regions, even without the choice of house size, the number of discrete location choices for residence and work is R^2. Moreover, because the cost of moving is an unavoidable element of any realistic location choice model, the state space of the model necessarily includes the previously chosen locations, thus requiring R^2 states before any of additional heterogeneity is accounted for. Therefore, in order to keep the model tractable, we introduce a number of simplifying assumptions right from the outset.

First, we assume that there are no fixed costs associated with scaling the size of a house up or down. The decision maker only pays the rental value of a home, calculated by its size times local square meter prices, and she is in each period free to resize her house without moving. It follows that the first order conditions fully characterize the choice of the house size conditional on the individual's characteristics and the attributes of the region of residence, and it can thus be expressed analytically and substituted into the indirect utility function. We derive the corresponding static demand for home size in Section 5.2 below. In presenting the dynamic discrete choice set-up in this section we therefore abstract from the continuous choice of house size, which is subsumed by a general indirect utility specification.

Second, we assume that work and residence location choices are made sequentially, namely that work location choice is made first followed by the residence location choice. Even though this assumption does not immediately decrease the number of alternatives in the resulting nested choice model (R nests by R alternatives each), it allows us to introduce a sensible job matching process. Namely, we differentiate between the job transition *choice* that denotes intention, from the job *outcome* that becomes the next period work location. In other words, our model allows for unsuccessful attempts to change work location and involuntary unemployment.

In our computational approach we recognize the fact that the expected future value of the current period choices only depends on the work and residence locations *realized by the end of the period*. We therefore formulate the dynamic programming problem in terms of expected value functions, keeping its dimensionality on the order of R^2 rather than R^4 (R^2 states by R^2 choices) as would be required by the traditional solution in the space of choice-specific value functions.

Because "work location" appears in the model in three different forms (existing, intended, and realized work location), we use the following explicit notation to distinguish between them. We denote wl_t the beginning of the period existing work location (*state*) and d_t^w the period t choice of intended work location (*choice*). This may or may not be the same as wl_t. Finally, to denote the *outcome* of the job match process during period t we simply use wl_{t+1}, as the realized in period t work location becomes the existing one in period $t+1$[5].

To maintain uniform notation, rl_t denotes the existing period t residence location and, correspondingly, d_t^r denotes the choice of new residence location. We assume perfect control over the location of the residence (subject to the equilibrium house prices), and therefore the location of residence in period $t+1$ is given by the choice at period t, i.e. $rl_{t+1} = d_t^r$.

The precise timing convention we use is as follows. Each period t starts off with a given work and residence location, and other variables x_t to be described below, forming the vector of state variables $s_t = (wl_t, rl_t, x_t)$. We assume that individuals make their work and residential location choices sequentially but instantaneously at the start of each period t, with the intended work location decision made first, followed by the residential location decision made *conditional* on the realization of the employment search, i.e. realized work location wl_{t+1}. Once the intended work location is chosen, the job search outcome is realized, and the residence location choice is made, the household determines the optimal house size depending on their own characteristics and the chosen region of residence. Thereafter, the housing consumption is enjoyed for the rest of the period, and the process transitions to the next period. We describe the transition rules and list all the components of the state vector s_t in Section 4.4 below.

Following the tradition of the discrete choice literature, we assume that the choices of both work and residence locations depend on the *IID* extreme value idiosyncratic shocks $\epsilon_t = (\epsilon_t^w, \epsilon_t^r)$ that can be interpreted as the components of the utility that the econometrician does not observe. We assume that these stochastic components are revealed to the individual sequentially: at the time of the work location decision d_t^w only the "work location shocks" ϵ_t^w are known, whereas the "residential location shocks" ϵ_t^r are only revealed after the individual learns the outcome of their employment search. In other words we assume that the households find out the idiosyncratic attributes of the residence

[5]Using notation wl_{t+1} as the realized work location in period t involves a degree of confusion with the time subscripts, but we opt to bear this cost to avoid having an additional outcome variable.

locations only once they know where their job takes them. These assumptions lead to the standard nested choice structure of the work and residence location decisions, with standard analytic expressions for choice probabilities at each level, and inclusive values of the residential choice given by the McFadden's social surplus (*logsum*) functions, as we discuss below.

4.2 Specification of the job search process

Before deriving the recursive formulation of the model, we specify the possible transitions in the job search process. A spatial model with fixed wages could lead to the outcome where far more people want to move into a high wage region than there are available jobs. Introduction of the labor market into the model allows us to avoid this unrealistic scenario.

Let the spatial work region $wl_t = \emptyset$ denote the state of non-employment, which can naturally be combined with any residence region rl_t. We assume that unemployment can be chosen voluntarily, but also allow for involuntary job separations with a certain probability, including the cases when no job transition is intended ($d_t^w = wl_t$).

Let $\pi_t^n(d_t^w, wl_t, x_t)$, $d_t^w \neq wl_t$, denote the probability of successfully finding a new job in the region d_t^w, conditional on the household characteristics and other variables in state vector x_t. For the case $d_t^w = wl_t$, $\pi_t^n(wl_t, wl_t, x_t) = \pi_t^k(wl_t, x_t)$ simply denotes the probability of keeping the existing job in location wl_t[6]. With the complementary probability $1 - \pi_t^k(wl_t, x_t)$ a transition to non-employment $wl_t = \emptyset$ occurs.

If the individual chooses to stop working, $d_t^w = \emptyset$, then $\pi_t^n(\emptyset, wl_t, x_t) = 1$, i.e. there is "perfect control" over the decision to stop working. However, for an individual who is searching for a new job in the region $d_t^w \neq wl_t$ the work location transition probabilities are given by

$$wl_{t+1} = \begin{cases} d_t^w & \text{with probability} \;\; \pi_t^n(d_t^w, wl_t, x_t), \\ wl_t & \text{with probability} \;\; (1 - \pi_t^n(d_t^w, wl_t, x_t))\pi_t^k(wl_t, x_t), \\ \emptyset & \text{with probability} \;\; (1 - \pi_t^n(d_t^w, wl_t, x_t))(1 - \pi_t^k(wl_t, x_t)). \end{cases} \tag{2.1}$$

In words, the transition probability in equation (2.1) says that if an individual chooses to search for a job in some new location $d_t^w \neq wl_t$, then there are three possible outcomes: i) the individual could be successful and receive a job offer in this location; ii) the individual does not get a job offer in the new location but is able to keep her existing job; or iii) the individual's job search is unsuccessful and she is laid off from her current job.

Note the "on the job search" assumption which means that if a worker applies for a job in a different location and does not get that job, she still has the option of staying in her current job, unless she is laid off. The alternative assumption would be to assume that if a worker applies for a job in some other work location, she has to quit her current job first. We think the on the job search assumption is a better approximation to reality:

[6] We abstract from job transitions within the same region.

the assumption that workers must first quit their jobs (i.e. have no recourse of staying at their current job if they apply for another job and are unsuccessful) would likely make it artificially risky to change job locations, and such an assumption may result in underprediction of job mobility across different regions.

If the individual does not search for a job in a new location, $d_t^w = wl_t$, we assume that they will continue working in the same location as before, and set $\pi_t^n(d_t^w, wl_t, x_t) = 0$. In this case (2.1) takes the form

$$wl_{t+1} = \begin{cases} wl_t & \text{with probability } \ \pi_t^k(wl_t, x_t), \\ \varnothing & \text{with probability } \ 1 - \pi_t^k(wl_t, x_t). \end{cases} \tag{2.2}$$

Specification (2.1) can be applied for unemployed individuals as well, in which case the last two rows collapse into one, and it takes the form

$$wl_{t+1} = \begin{cases} d_t^w & \text{with probability } \ \pi_t^n(d_t^w, \varnothing, x_t), \\ \varnothing & \text{with probability } \ 1 - \pi_t^n(d_t^w, \varnothing, x_t). \end{cases} \tag{2.3}$$

This specification allows for the possibility that there may be a lower chance of getting a job if a person is currently unemployed, compared to an individual who is currently employed in this location or in some other location. That is, one possible ordering of the employment transition probabilities that might be supported empirically is

$$\pi_t^n(d_t^w, \varnothing, x_t) < \pi_t^n(d_t^w, wl_t, x_t) < \pi_t^k(wl_t, x_t) < \pi_t^n(\varnothing, wl_t, x_t) = 1, \tag{2.4}$$

so an individual who chooses to stay in their current work location has the highest probability of being employed in this location, $\pi_t^k(wl_t, x_t)$, apart from choosing to stop working $\pi_t^n(\varnothing, wl_t, x_t) = 1$. An individual who is applying for jobs in the region from the outside ($d_t^w \neq wl_t$), has a lower employment probability, while the lowest probability of employment corresponds to the unemployed individuals, $\pi_t^n(d_t^w, \varnothing, x_t)$.

Let $\pi_t(d_t^w, wl_t, x_t, wl_{t+1})$ denote the probability of transition from work in region wl_t to work in region wl_{t+1} (including non-employment $wl_{t+1} = \varnothing$), which encompasses all the transition probabilities described in this section in equations (2.1)-(2.3). We present the functional forms assumptions of $\pi_t^n(d_t^w, wl_t, x_t)$ and $\pi_t^k(wl_t, x_t)$ in Section 4.5.

4.3 Recursive formulation and Bellman equations

At each period $\{t = t_0, \ldots, T\}$ the individuals in the model maximize the expected discounted utility over the remainder of their life by making sequential work and home location decisions, as well as choosing the house of the optimal size. For every t the attainable maximum is given by the value function $V_t(s_t, \epsilon_t)$ which is a function of state variables $s_t = (wl_t, rl_t, x_t)$ and the stochastic taste shocks ϵ_t. As mentioned in Section 4.1, instead of the value function $V_t(s_t, \epsilon_t)$ we focus on the *expected value function* $EV_t(wl_{t+1}, rl_{t+1}, x_t)$, as a function of the realized work and home locations (after all relocations have been completed). Note that the expected value function at period t depends on the work and residence locations at period $t+1$. Even though this may appear as a type of "clairvoyance" of the decision makers, it is merely the consequence of our timing assumptions. The "next period" location (wl_{t+1}, rl_{t+1}) in fact just denotes the location outcome after the decisions and relocation stage is completed in the beginning of the period. According to the timing convention, during period t, the individual lives at location rl_{t+1} and works at location wl_{t+1}.

Unlike the expected value function $EV_t(wl_{t+1}, rl_{t+1}, x_t)$, the period t (deterministic) flow utility has to account for the switching costs of relocations, and therefore has to depend on both initial locations and the realized location. To allow for a flexible way that switching costs enter the model (both for changing the home and work locations and differentiated for heterogeneous households) in this section we use the generic form of the utility function given by $u(wl_t, rl_t, wl_{t+1}, rl_{t+1}, x_t)$[7]. Note that the choice variables enter into the utility function indirectly: choice of work location d_t^w governs the job search process described in previous section, and under assumed perfect control the choice of residence location, we have $d_t^r = rl_{t+1}$.

Given the nested discrete choice structure in the model described in Section 4.1, the extreme value shocks $\epsilon_t = (\epsilon_t^w, \epsilon_t^r)$ enter the Bellman equation in a non-trivial way. We build the Bellman equation in stages following the backward induction over the events within the time period.

Let β denote the discount factor of the individual. For simplicity we assume it is independent of individual survival rates. I.e. we do not take into account that the discounting of future expected values may lower as the individual ages.

Recall that $\epsilon_t^r \in \mathbb{R}^R$ are the stochastic components of the utility corresponding to the choice of residence location, once the outcome of the job search process is revealed, and the new work location wl_{t+1} is known. Let $\epsilon_t^r(d_t^r)$ be the idiosyncratic utility costs/benefits of choosing to move to location d_t^r. We assume it is extreme value with scale parameter σ_r. Let $EV_t^r(wl_t, rl_t, wl_{t+1}, x_t)$ be the *ex ante* expected value for an individual who knows her employment location outcome wl_{t+1} but has not learned the residential location shocks $\{\epsilon_t^r(d_t^r)\}$ yet. This is given by the usual log-sum formula

$$EV_t^r(wl_t, rl_t, wl_{t+1}, x_t) =$$

[7]Additional assumptions on the utility function could have drastically reduced the computational burden of the model. For example, assuming that the moving costs for residence (c^r) and work (c^w) are additively separable and only depend on the destination, i.e. $u(wl_t, rl_t, wl_{t+1}, rl_{t+1}, x_t) = u'(wl_{t+1}, rl_{t+1}, x_t) - c^w(wl_{t+1}, x_t) - c^r(rl_{t+1}, x_t)$, leads to much simplified expressions of the value functions that can be expressed with the values of households that do not change locations, modified by a collection of constant moving costs, and thus drastically simplifying the computation of the log-sum function and choice probabilities when solving the dynamic programs. However, in order to be able to match the data we have, we prefer to keep the model specification flexible at this stage.

$$\sigma_r \log \left(\sum_{d^r} \exp\{[u(wl_t, rl_t, wl_{t+1}, d^r, x_t) + \beta EV_t(wl_{t+1}, d^r, x_t)]/\sigma_r\} \right). \tag{2.5}$$

The implied residence location choice probabilities are given by the multinomial logit formulas

$$P_t^r(d_t^r | wl_t, rl_t, wl_{t+1}, x_t) = \frac{\exp\{[u(wl_t, rl_t, wl_{t+1}, d_t^r, x_t) + \beta EV_t(wl_{t+1}, d_t^r, x_t)]/\sigma_r\}}{\sum_{d^r} \exp\{[u_t(wl_t, rl_t, wl_{t+1}, d^r, x_t) + \beta EV_t(wl_{t+1}, d^r, x_t)]/\sigma_r\}}. \tag{2.6}$$

Now consider the choice of the work location at the beginning of period t, d_t^w. Because this choice is moderated by the job search process, we have to take into account the probabilities $\pi_t(d_t^w, wl_t, x_t, wl_{t+1})$ that govern how the intended job location d_t^w translates into the realized one wl_{t+1}. Let $v^w(wl_t, rl_t, x_t, d_t^w)$ denote the expected choice-specific value corresponding to the particular choice of job location d_t^w. We have

$$v_t^w(wl_t, rl_t, x_t, d_t^w) = \sum_{wl} \pi_t(d_t^w, wl_t, x_t, wl) EV_t^r(wl_t, rl_t, wl, x_t). \tag{2.7}$$

Now recall that $\epsilon_t^w \in \mathbb{R}^{R+1}$ are the stochastic components corresponding to the choice of work location, with additional voluntary choice of non-employment. Similar to the residential location choice, let $EV_t^w(wl_t, rl_t, x_t)$ be the ex ante expected value for an individual who has not learned the work location shocks $\{\epsilon_t^w(d_t^w)\}$ yet. Under the assumption that the shocks have an extreme value distribution with scale parameter σ_w, $EV_t^w(wl_t, rl_t, x_t)$ is given by the log-sum formula

$$EV_t^w(wl_t, rl_t, x_t) = \sigma_w \log \left(\sum_{d^w} \exp \left\{ \sum_{wl} \pi_t(d^w, wl_t, x_t, wl) EV_t^r(wl_t, rl_t, wl, x_t)/\sigma_w \right\} \right). \tag{2.8}$$

Similarly, we have the usual multinomial logit choice probability for the choice of work location

$$P_t^w(d_t^w | wl_t, rl_t, x_t) = \frac{\exp\{v_t^w(wl_t, rl_t, x_t, d_t^w)/\sigma_w\}}{\sum_{d^w} \exp\{v_t^w(wl_t, rl_t, x_t, d^w)/\sigma_w\}}. \tag{2.9}$$

After accounting for the transition probabilities $\pi^x(x_t, x_{t+1})$ of the non-location state variables, which we assume are independent of both the stochastic shocks $\epsilon_t = (\epsilon_t^w, \epsilon_t^r)$ and the labor market probabilities $\pi_t^n(d_t^w, wl_t, x_t)$ and $\pi_t^k(wl_t, x_t)$, we have by the definition of the expected value function

$$EV_t(wl_{t+1}, rl_{t+1}, x_t) = \sum_{x_{t+1}} \pi^x(x_t, x_{t+1}) EV_{t+1}^w(wl_{t+1}, rl_{t+1}, x_{t+1}). \tag{2.10}$$

Combining equations (2.5), (2.7) and (2.10), we obtain a Bellman operator in expected value functions that maps $EV_{t+1}(wl_{t+2}, rl_{t+2}, x_{t+1})$, that enters in the shifted one period forward equation (2.5), into $EV_t(wl_{t+1}, rl_{t+1}, x_t)$[8].

[8]Writing down the complete Bellman operator is straightforward, but we do not do that here for space

The computational algorithm for solving the model is straightforward. Because the model is formulated in finite horizon, this reduces to a backward induction calculation starting at the maximum possible age T. For each period t we compute the expected value functions $EV_t(wl_{t+1}, rl_{t+1}, x_t)$, and the corresponding choice probabilities $P_t^w(d_t^w|wl_t, rl_t, x_t)$, and $P_t^r(d_t^r|wl_t, rl_t, wl_{t+1}, x_t)$ that serve as the basis for formulating the likelihood function.

Suppose we have already computed the expected value function $EV_{t+1}(wl_{t+2}, rl_{t+2}, x_{t+1})$ for all possible values of the states $(wl_{t+2}, rl_{t+2}, x_{t+1})$ at age $t+1$. On iteration t we loop over all possible end-of-period combinations of locations (wl_{t+1}, rl_{t+1}) and over all non-location states x_{t+1}. In each such point we then use equation (2.5) to compute the inclusive values of the different work locations wl_{t+2}, $EV_{t+1}^r(wl_{t+1}, rl_{t+1}, wl_{t+2}, x_{t+1})$, and the probabilities of location choices $P_t^r(d_t^r|wl_t, rl_t, wl_{t+1}, x_t)$ given by equation (2.6). Then we compute the d_{t+1}^w choice-specific values $v_{t+1}^w(wl_{t+1}, rl_{t+1}, x_{t+1}, d_{t+1}^w)$ using equation (2.7), and the accompanying work location choice probabilities $P_t^w(d_t^w|wl_t, rl_t, x_t)$ using equation (2.9). After that, using the equations (2.8) and (2.10), we compute the period t expected value function $EV_t(wl_{t+1}, rl_{t+1}, x_t)$. Once $EV_t(wl_{t+1}, rl_{t+1}, x_t)$ is computed for all states $(wl_{t+1}, rl_{t+1}, x_t)$, the period t iteration is complete, and the algorithm moves to period $t-1$.

4.4 State space dynamics

There is always an awkwardness about formulating a discrete time model with actual data where transitions occur in continuous time. The discrete time model assumes decisions are made at specific instants in time: i.e. at the start of each period where the period in our setup be one year. We will defer a discussion of how to best match the actual data to the model when the precise state change date is not clear. But for our discussion suppose we have data on the state of an individual at the start of each year, i.e. on January 1st.

Table 2.1 lists all the state variables in the model that we include to control for the heterogeneity among the households. These variables enter the non-spatial part of the state vector x_t, and together with the two location variables form the full state vector.

As mentioned above, the transitions of the non-spatial state variables are governed by the transition probability $\pi^x(x_t, x_{t+1})$ which we describe in details below. However, a part of the state space is non-time-varying and constitute the *types* of households. We assume a finite number of types, and note that because these do not change during the backwards induction, the solution algorithm can in principle solve separate dynamic programming problems for separate types in parallel, thus assuming availability of the appropriate number of computing cores, without increasing the overall computational load. The time-invariant *type* variable is education (schooling) type edu_t and marital status ms_t[9], while children status cs_t evolve as an independent first order Markov processes

considerations.

[9]This could could be extended by a permanent income type $perminc_t$, and the propensity to move type uh_t to reflect the unobserved heterogeneity in the population. This is left for future work.

Table 2.1: Non-location state variables including household types that enter x_t.

Symbol	Description	Possible Values
cs_t	Number of children at home	0 no children 1 1 or more children
ms_t	Marital status	0 single 1 married/cohabitating
edu_t	Education (school) type	0 Less than medium cycle education 1 Medium cycle education (BA) 2 Long cycle education (master/PhD)

Notes: The table lists the non-location state variables entering x_t. In addition the value function depends on the beginning of the period work and residence locations (wl_t, rl_t), and the expected value function depends on the realized end-of-period work and residence locations (wl_{t+1}, rl_{t+1}).

with transition probabilities defined below. The potential outcomes of these non-spatial (exogenous) state variables are listed in Table 2.1

The evolution of children status depends on age. A trick to reduce computation is to assume that the number of children can maximally change by one every year. Obviously, this fails in case of twin births and couple formation where the spouse has more than one child, but in the data we do not distinguish between having one or more children. In sum, cs_t follows

$$cs_{it+1} \sim \mu_{cs}(\cdot|cs_t, age_t). \tag{2.11}$$

The transitions of children are estimated separately in a first step, and the transition probability of the exogenous part of the state space vector $\pi^x(x_t, x_{t+1})$ is given by (2.11).

4.5 Specification of the utility function

In Section 4.3 we used the general form of the current period utility $u(wl_t, rl_t, wl_{t+1}, rl_{t+1}, x_t)$, only specifying its dependence on both current and realized work and residence locations, and the non-spatial variables. In this and next sections we give a complete specification of the utility function and job probabilities, starting with the "direct utility" specification which depends on the size of the house that the household occupies. To arrive on the final specification of $u(wl_t, rl_t, wl_{t+1}, rl_{t+1}, x_t)$ which is independent of the house size, in the next section we express the demand for housing using the first order condition of the static choice of house size, which is then plugged back into the utility function.

To help the exposition, we first describe the parts of the utility function, and specify them fully one by one afterwards. The utility of any location choice can generally be written as the sum of the following components (suppressing arguments and indices)

$$u = u_m - u_w + u_h + \underbrace{amenities - swcost_r^p - ttimecost}_{u_o}, \tag{2.12}$$

where u_m is the *monetary utility* (income net of housing expenditures), u_w is *disutility of work* which is equal to zero when $wl_{t+1} = \emptyset$, u_h is the *housing utility* obtained from the utilization of a chosen home size, *amenities* reflects the *regional-specific* attractiveness of housing options, $swcost_r^p$ is the *psychological costs* of changing the location of residence, and *ttimecost* is the *cost of commuting* between the chosen locations of work and residence. According to our timing convention (described in Section 4.1), all the house and regional characteristics correspond to the chosen location rl_{t+1}, because it is the location enjoyed during period t, after the instantaneous moving phase in the beginning of the period has taken place.

First, consider the u_m component. It can be expressed as a product of the marginal utility of money $\kappa(inc_t)$ (which depends on household income), and the consumable earnings, which are given by the difference between the household income and the cost of maintaining the house ($hcost_t$). We have

$$u_m = \kappa(inc_t)(inc_t - hcost_t), \tag{2.13}$$

where inc_t denotes household income in period t.

We assume the following functional form for the marginal utility of money

$$\kappa(inc_t) = \kappa_0 + \kappa_y inc_t. \tag{2.14}$$

Assuming that $\kappa_y < 0$, we have linearly decreasing marginal utility of money, implying that richer households will be less sensitive to housing prices and moving costs. In the absence of a wealth state variable and a consumption/savings choice in the model, marginal utility subsumes all effects of the credit constraint or availability of mortgage.

Household income $inc_t = inc_t(wl_t, wl_{t+1}, x_t)$ is modeled by a set of Mincer-type equations that include age as a personal characteristic, and are estimated separately for all regions and education groups to reflect the regional and skill-specific variation. In addition, we introduce a wage penalty to being non-employed in previous period ($wl_t = \emptyset$), and we entitle the currently unemployed ($wl_{t+1} = \emptyset$) with unemployment benefits or pension income. Conditional on being employed ($wl_{t+1} \neq \emptyset$), household income for individual i in period t is modelled as

$$\log(inc_{it}) = \delta_0 + \delta_{age} age_{it} + \delta_{age^2} age_{it}^2 + \delta_u \mathbb{1}_{\{wl_{it}=\emptyset\}} + \xi_{it}, \tag{2.15}$$

where ξ_{it} is the idiosyncratic error component and all parameters vary by work region and education group. For non-employed persons we implement the following specification of non-employment income on age dummies:

$$\log(inc_{it}) = \sum_{t=t_0}^{T} b_t + \nu_{it}, \tag{2.16}$$

for each education group, where ν_{it} is a random error term.

The probabilities of getting a new job or keeping the existing job were introduced in Section 4.2. Below we present the functional forms, starting with the probability of getting a new job which is defined as

$$\begin{aligned}\pi_t^n(d_t^w, wl_t, x_t) = \Bigg[1 + exp\Big(- \Big(\beta_0^{\pi(new)} + \beta_a^{\pi(new)} age_t + \beta_{unemp}^{\pi(new)} \mathbb{1}_{wl_t=\varnothing} \\ +\beta_{jobdensity}^{\pi(new)} jobdensity(d_t^w) \\ + \sum_{k=1}^{2} (\beta_s^{\pi(new)}(k) \mathbb{1}_{edu_t=k})\Big)\Big)\Bigg]^{-1},\end{aligned} \tag{2.17}$$

where $jobdensity(d_t^w)$ is an index of type edu_t jobs in region d_t^w. By allowing the probability of landing a new job to depend on job density, then when $\beta_{jobdensity}^{\pi(new)} > 0$ the individual is more likely to receive a job offer from a region with more jobs. This helps the model predict how attractive each work region is. Admittedly, there is something awkward about this specification as long as we do not model the equilibrium on the labor market, since we attribute a high job density to a high fixed supply of jobs while in reality it is an interplay between supply and demand. We assume the individuals only care about and searches for jobs of their own skill type to capture the heterogeneity in job moving behaviour. The probability of keeping one's current job is defined by

$$\pi_t^k(wl_t, x_t) = \Bigg[1 + exp\Big(- \Big(\beta_0^{\pi(keep)} + \beta_a^{\pi(keep)} age_t + \sum_{k=1}^{2} (\beta_s^{\pi(keep)}(k) \mathbb{1}_{edu_t=k})\Big)\Big)\Bigg]^{-1}. \tag{2.18}$$

Finally, we allow for disutility of work u_w through the fixed constant, c_{work}, which is relevant when $wl_{t+1} \neq \varnothing$.

Amenities of home regions are modelled as region-specific constants. Another approach would be to model amenities as a function of region-specific observables such as crime rates, nature, restaurants etc., but given that most of these variables are regional-specific and time constant we use a fixed effects approach:

$$amenities(rl_{t+1}) = \sum_{rl=1}^{R} \boldsymbol{\alpha}_{rl} \mathbb{1}_{\{rl_{t+1}=rl\}}, \tag{2.19}$$

where α^{rl} is a vector of coefficients for each region.

The psychological moving cost $swcost_r^p$ is a function of the family characteristics, age and education. We use the following specification

$$swcost_r^p(x_t) = \mathbb{1}_{\{rl_t \neq rl_{t+1}\}}[\gamma_0 + \gamma_{age} age_t + \gamma_{ms} ms_t + \gamma_{cs} cs_t + \sum_{k=1}^{2} \phi_{s,k} \mathbb{1}_{\{edu_t = k\}}], \tag{2.20}$$

which reflects the fact that the propensity to move changes with the family situation and is different at different stages of life.

The costs of commuting between rl_{t+1} and wl_{t+1} are assumed to be proportional to the exogenous travel time between the work and home locations. Hence, we have

$$ttimecost = \eta_{ttime} ttime(rl_{t+1}, wl_{t+1}) \tag{2.21}$$

where the function $ttime(rl_{t+1}, wl_{t+1})$ denotes the travel time between work location wl_{t+1} and residence location rl_{t+1}.

When specifying the demand and utility of housing, we note that regional-specific price of housing is approximately linear in home size measured in square meters of floor space. It is therefore natural to specify housing demand (size of home) $h(rl_{t+1}, x_{it}; P^h(rl_{t+1}))$ in residential region rl_{t+1} as a function of the regional-specific housing price $P^h(rl_{t+1})$ (expressed as an equivalent annual rental price) per square meter. Housing costs $hcost_t$ are thus assumed to be proportional to the equilibrium per square meter prices $P(rl_{t+1})$ through the parameter ψ_{uc}. This translates housing prices into an annual user cost, and also reflects mortgage expenses and housing taxes. Hence housing costs are given by

$$hcost_t(rl_{t+1}, h_{t+1}) = \psi_{uc} P(rl_{t+1}) h_{t+1}, \tag{2.22}$$

where according to our timing convention rl_{t+1} denotes the house occupied during period t.

The demand for housing also depends on individual characteristics such as household size and income. This for example reflects that richer people can buy relatively more square meters and others less, and that larger families may substitute space for location. We define the utility u_h of living in a house as a quadratic polynomial of its size h_t with heterogeneous coefficients

$$u_h = \Phi(x_t) h_{t+1} + \frac{1}{2} \phi_{h2}\, h_{t+1}^2, \tag{2.23}$$

where $\phi_{h2} < 0$ governs the degree of diminishing returns to house size and $\Phi(x_t)$ is a heterogeneous parameter, which affects the baseline marginal utility of housing

$$\Phi(x_t) = \phi_0 + \phi_{age} age_t + \phi_{ms} ms_t + \phi_{cs} cs_t. \tag{2.24}$$

Given the form of the utility function specified by equations (2.12)-(2.18), the part of utility that is dependent on home size is equal to

$$\tilde{u}_h = \Phi(x_t)h_{t+1} + \frac{1}{2}\phi_{h2}\, h_{t+1}^2 - \kappa(inc_t)hcost_t(rl_{t+1}, h_{t+1}). \tag{2.25}$$

From the first order conditions for (2.25) and the specification of housing cost in (2.22) it follows that the optimal choice of the house size[10] is given by

$$h_{t+1} = \frac{\kappa(inc_t)P(rl_{t+1})\psi_{uc} - \Phi(x_t)}{\phi_{h2}}. \tag{2.26}$$

Substituting expression (2.26) back into the utility function defined in equations (2.12)-(2.25), we obtain the final specification of the indirect utility function $u(wl_t, rl_t, wl_{t+1}, rl_{t+1}, x_t)$.

5 Structural Estimation

This section describes the estimation strategy applied to the theoretical model. We estimate the model sequentially in three steps: i) we estimate the parameters governing the wage equations and transition probabilities of the children state; ii) we estimate a reduced form housing demand equation, and iii) we estimate the remaining structural parameters by maximum likelihood applying the parameters obtained in i) and ii). Below we go through each step in detail and discuss identification of parameters.

First a note on our data sampling. In Section 3 we provide descriptive statistics using full population register data for the period 1998-2010, yet for the estimation we focus on years 2005-2010. We do this to work with a relatively homogeneous subsample while also retaining relevance in terms of policy guidance. Further, even though we do observe each individual's choice $d_{i,t} \equiv \{rl_{i,t+1}, wl_{i,t+1}, h_{i,t+1}\}$ and state $s_{i,t}$ on an annual basis, we pool the data over all the years in estimation in the current version of the estimation.

For time-varying variables such as house prices, we average them over the same period. Hence, we assume that all choices made by households are made considering only the average housing prices and amenities over this period. Since local amenities are fixed throughout the estimation period, we cannot separately identify the effects of observed regional amenities and regional fixed effects. A more subtle note here is that since both local house prices and local amenities are regional-specific in the pooled sample, joint identification of the utility coefficients of the time-invariant local amenities and marginal utility of money can only be accomplished through individual-level taste variation for housing and individual-level variation in income. However, since both income and the demand for housing varies within each region we are able to simultaneously pin down amenities and marginal utility of money along with the remaining parameters of the model.

[10]At the time immediately after all moving has finished since, as is clear from the Section 6, the equilibrium of the model is Markov perfect and higher demand in a region drives up prices.

5.1 Wage equations and transition probabilities

The estimation of transition probabilities for children status, μ_{cs}, is performed non-parametrically on the pooled data as the share of individuals within each age-children cell who is observed in each possible transition. Since in the current implementation, cs only takes values 0 and 1 (at least one child living at home), this comes down to four possible transitions at each age.

In order to capture regional differences in both wage level and its age gradient, we estimate the coefficients of the wage offer equation in (2.15) separately for each combination of region and education level. We use the equivalent full-time income for each individual and condition on observed employment. The estimates are presented in Appendix C.

Similarly, we estimate the education level-specific equation for non-employment income in (2.16). Until retirement age, non-employment income will mainly consist of unemployment benefits. After the usual retirement age, income sources are more mixed as one will receive both public pension, private pension savings and possibly labor income from part time employment.[11]

5.2 Housing demand

As mentioned in Section 4.1, we assume that within each region and in each time period households can freely adjust the size of their home. This is equivalent to having no cost of moving within the region to the house of optimal size. Moreover, we abstract away from any savings, including in home equity, and let households consider only the "square meter rental costs" that pertains to homes in each region through local prices. Both of these assumptions allow for the optimal amount of housing to be separable from the dynamic choice of location and expressed as the solution to a static subproblem that enters into the indirect instantaneous utility described in Section 4.5. This greatly reduces the computational burden, effectively allowing the structural estimation of the model to be carried out.

The fact that the solution to the static housing size problem specified in (2.26) is detached from the dynamic location choice allows us to estimate a scaled version of it in a separate step before turning to the dynamic model. Hence, using the pooled micro data we estimate the following demand equation for housing by OLS

$$\begin{aligned} h_{it+1} = \tilde{\phi}_0 + \tilde{\phi}_{age} age_{it} + \tilde{\phi}_{ms} ms_{it} + \tilde{\phi}_{cs} cs_{it} \\ - \tilde{\kappa}_0[\psi_{uc} P(rl_{it+1})] + \tilde{\kappa}_y[inc_{it} \times \psi_{uc} P(rl_{it+1})] + \varrho_{it}, \end{aligned} \tag{2.27}$$

[11] We do not allow for regional differences in non-employment income, even though these are indeed observed due to differences in savings. Yet, since an individual would not be able to change her savings by moving, we abstract away from differences in average regional savings. The result is of course that the returns to income received while working is downward biased in rich areas and upward biased in poor areas. Estimates of the non-employment income regressions are not presented but available upon request.

where ϱ_{it} is a random error (see also equation (2.26)).

Note that the parameters $\tilde{\phi}$ and $\tilde{\kappa}$ in the reduced form demand equation in (2.27) are proportional to the structural parameters that index marginal utility of money $\kappa_{(\cdot)}$ and heterogeneous housing utility parameters in $\Phi_{(\cdot)}$, but scaled by $-1/\phi_{h2} > 0$ and $-\psi_{uc}/\phi_{h2} > 0$ respectively. We identify ϕ_{h2} and ψ_{uc} in conjunction with the remaining structural parameters using the cross-equation restrictions implied by the housing demand equation and the location choice model. When estimating these structural parameters, the reduced form estimates of parameters are then kept fixed during the structural estimation of the location choice model, and only *rescaled* using the values of the structural parameters ψ_{uc} and ϕ_{h2}. This two-step procedure significantly reduces the dimensionality of the maximum likelihood problem when estimating the full model.

5.3 Maximum likelihood estimation of structural parameters

Having obtained estimates for children transitions, wage equations and scaled housing demand we estimate the remaining structural parameters, θ, by maximum likelihood. To recount, θ includes parameters indexing probability of getting a new job, (2.17), probability of keeping current job, (2.18), marginal utility of money, (2.14), housing costs, (2.22), utility values of the amenities, (2.19), psychological costs of moving residence, (2.20), travel time costs, (2.21), the disutility of work, c_{work}, and the degree of diminishing returns to house size, ϕ_{h_2}. We fix the discount factor to $\beta = 0.95$.

The likelihood function is derived from the choice probabilities for work and home location decisions given in (2.6) and (2.9). Because we assume perfect control for residential location, the latter can be directly evaluated at the data, giving the likelihood of the observed location of residence. To calculate the likelihood of the observed work location, however, we have to integrate out the likelihood over the possible choices and only write the likelihood in terms of observed *work location transitions*, i.e. as transition probabilities from state wl_t to wl_{t+1}.

Observing a transition wl_t to $wl_{t+1} = wl_t$ could have resulted from both an individual deciding to keep their job, and being successful (with probability $\pi_t^k(wl_t, x_t)$), and an individual trying to find a new job d_t^w and being unsuccessful (with probability $(1 - \pi_t^n(d_t^w, wl_t, x_t))\pi_t^k(wl_t, x_t)$). Observing a transition wl_t to $wl_{t+1} \neq wl_t$ could have resulted only from an individual deciding to move jobs and being successful (with probability $\pi_t^n(wl_{t+1}, wl_t, x_t)$).

The above two cases also apply for $wl_t = \emptyset$, but $wl_{t+1} = \emptyset$ and $wl_t \neq \emptyset$, i.e. transition to unemployment, may happen in three different scenarios. First, with probability $(1 - \pi_t^n(d_t^w, wl_t, x_t))(1 - \pi_t^k(wl_t, x_t))$ an individual could have unsuccessfully tried to transition to a job d_t^w, and at the same time has been displaced. Or, with probability $1 - \pi_t^k(wl_t, x_t)$ an individual could have tried to keep job wl_t, yet being unsuccessful and displaced. Or finally, an individual could have voluntarily chosen to quit working, $d_t^w = \emptyset$.

Recall that $\pi_t(d_t^w, wl_t, x_t, wl_{t+1})$ summarizes the work transition probabilities as a function of the intended work location. The contribution to the likelihood for an individual who is in *observed* work location wl_t and residential location rl_t at time t and in *observed* work location wl_{t+1} and residential location rl_{t+1} at time $t+1$ is

$$L_t(wl_t, rl_t, wl_{t+1}, rl_{t+1}, x_t) = \\ P_t^r(rl_{t+1}|wl_t, rl_t, wl_{t+1}, x_t) \cdot \sum_{d^w} P_t^w(d^w|wl_t, rl_t, x_t)\pi_t(d^w, wl_t, x_t, wl_{t+1}). \tag{2.28}$$

The full log-likelihood is constructed from individual likelihoods in the standard way by collecting the individual likelihood contributions and the objective of the maximum likelihood estimation is thus

$$\text{argmax}_\theta \frac{1}{N}\sum_i \sum_t \{ \log P_t^r(rl_{it+1}|wl_{it}, rl_{it}, wl_{it+1}, x_{it}; \theta) + \\ \log \sum_{d^w} P_t^w(d^w|wl_{it}, rl_{it}, x_{it}; \theta)\pi_t(d^w, wl_{it}, x_{it}, wl_{it+1}; \theta)\}, \tag{2.29}$$

where N is the number of individuals. To estimate the structural parameters we proceed in the spirit of the Nested Fixed Point (NFXP) algorithm by Rust (1987, 1988) and solve the model via backwards induction for each evaluation of the likelihood function.

5.4 Identification

We now consider how the parameters of the model are identified from data. Starting with the first-step estimation of housing demand, recall that Section 3.3 clearly demonstrates variation in housing demand by age, number of children and marital status. As this variation is present within regions as well we have identification of the parameters in $\tilde{\Phi}$ which were scaled by $-\phi_{h2}$. The same goes for the parameters governing marginal utility of money in $\tilde{\kappa}$. The baseline marginal utility of money, κ_0, is identified from the spatial variation in house prices together with individual-level variation in housing demand and income. Further, the income dependence in marginal utility of money, κ_y, is identified from the clear sorting of high-income households into larger homes and more expensive regions.

It follows from (2.26) and (2.27) that we can only identify the parameter for diminishing utility of housing, ϕ_{h2}, within the dynamic model of location decision. The dynamic location choice involves a trade-off between home size, value of amenities and commuting time so the fact that we observe households substituting between locating in regions of high amenity levels in return for smaller homes than in low amenity regions allows us to determine ϕ_{h2}.

The user cost of housing, ψ_{uc}, and marginal utility of money, κ, can be separately identified using the variation in location decisions, house prices and wages across regions. For a given marginal utility of income, the sensitivity of individuals' location decisions to the spatial variation in house prices provides identification of ψ_{uc}. In that sense ψ_{uc} can also be thought of as a factor that distinguishes individuals' marginal utility of wage income from marginal (dis)utility of house prices.

The disutility of work, c_{work}, is identified through both the variation in labor market participation across education groups and through the participation over the life cycle. High skill workers have both higher wages and higher participation rates, implying that the opportunity cost of not working must be positive, thus $c_{work} > 0$. The same effect occurs within education groups as the wage offer declines after a certain age which coincides with increasing propensity to not work (i.e. retire).

The parameters that index moving costs, $\{\gamma_0, \gamma_a, \gamma_c, \gamma_{ms}, \gamma_s\}$, are easily identified since the propensity to relocate home differs substantially along the age, children, marital and schooling dimensions. The variation in moving propensity along age, children and education is evident in the graphs of Section 3.3 and corresponding graphs are shown in the results in Section 7 for education groups. Higher age, presence of children and lower education all reduces the likelihood of moving.

The coefficients with local amenities in $\alpha_{rl}(\cdot)$ are identified by observing that at a given level of income and commute distance, households are willing to pay a higher square meter price in one region compared to another. The only justification for such behavior is a higher amenity level.

The uncertainty of the job search process implies that large work regions in terms of number of jobs are more attractive than small regions conditional on the wage offer. The parameter $\beta^{\pi(new)}_{jobdensity}$ is therefore determined by differences in transitions of a given household type into regions that offer similar wages but have different sizes. Since we also observe that there is a negative gradient in job relocations over the life cycle without a corresponding drop in wage differences, we get identification of the age component of job search. Similarly, less schooling and past unemployment will affect job transitions negatively as long as the relative differences in income across regions do not decrease one for one with these measures. We do not try to identify any local effects in the probability of keeping a job. Hence, observed transitions into non-employment across regions that offer similar wages for a given household type (in the age and education dimension) yield identification of keep probabilities.

Conditional on a choice of work location, wl, we observe a decreasing probability of choosing a home location rl as the distance to wl increases. This relationship pins down the cost of traveltime η_{ttime}.

6 Solving for Equilibrium House Prices

We take a short run perspective and assume a fixed supply of housing and thus abstract from the longer run dynamics where new houses are built in response to changes in house prices. We also abstract from equilibrium formation in the labor market, and ignore that firms in reality may change labor demand in their locations (and thus the number of jobs offered in different locations) in response to changes in local labor supply. Hence the equilibrium object we solve for is housing/rental prices, whereas wages, job arrival and dismissal rates are taken as given and housing supply is assumed fixed in the short run.

In equilibrium we assume that prices have adjusted so that the total demand for housing measured in square meters equals the supply in each residential region. Thus, when solving for the housing market equilibrium, the R-dimensional vector of regional square meter prices $P^h = (P^h(1), .., P^h(R))$ is set to equate the inelastic, exogenously fixed supply $S_t(rl)$ of total square meters of housing to the demand for the available square meters $D_t(rl, P^h)$ in each residential region $rl = \{1, \ldots R\}$. For the supply, we simply aggregate the individual-level demand for observed square meters of housing h_{it} for people who already live in region $rl_{it} = rl$ at the beginning of each period t

$$S_t(rl) = \sum_{i=1}^{N} h_{it} \mathbb{1}(rl_{it} = rl) \tag{2.30}$$

where $\mathbb{1}$ is the indicator function.

The regional demand for housing $D_t(rl, P^h)$ is calculated as the *expected demand* by taking a population average of housing demand weighted by choice probabilities of either staying or moving to region rl at the end of period t. To obtain demand, we start by simulating N individual states by drawing from observed states in the dataset with replacement. We then simulate a work location *outcome*, wl_{t+1}, using the decision rule P_t^w and job transition probabilities π_t such that we can condition on these in the computation of demand below:

$$D_t(rl, P^h) = \sum_{i=1}^{N} h(rl, x_{it}; P^h(rl)) \Pi_t(rl | wl_{it+1}, rl_{it}, x_{it}; P^h), \tag{2.31}$$

where $\Pi_t(rl|wl_{it+1}, rl_{it}, x_{it}; P^h)$ is the probability that an individual in state $s_{it} = (wl_{it+1}, rl_{it}, x_{it})$ chooses to live in region rl given the vector of regional house prices, P^h and simulated work location wl_{it+1}. Π_t is given by the right hand side of (2.6), but here we have added P^h as an argument to signify its dependence on house prices.

The resulting simulator for demand is in principle not smooth given that we have simulated a work location *outcome*, wl_{t+1} using a simple accept/reject simulator. However, since the conditional demand for residence, $\Pi_t(rl|wl_{it+1}, rl_{it}, x_{it}; P^h)$, is smooth in the vector of housing prices and employment probabilities, we still found it smooth enough to use gradient-based methods to calculate equilibrium. We calculate the house price equilibrium by arraying all the excess demand equations to have a system of R excess demand equations (for the housing market) in R unknowns and solve for the R-dimensional price vector P^h using Newton's method.

The short run equilibrium concept is imposed for simplicity. To work with a long run equilibrium notion that endogenizes the supply of housing, we would need data on zoning regulations and decisions by home builders and developers where to build more in different regions. Finally, commuting times/costs are potentially something to endogenize too, including in the short run. If the counterfactual equilibrium results in changed location patters, the resulting utilization of the road network will change as well and thereby affect congestion and commuting times. Future work will focus on these more involved specifications.

7 Results

We start by presenting model fits and parameter estimates from the first-stage income and housing demand equations and then move to the remaining parameter estimates from the structural location choice model. Using the estimated model to solve for equilibrium prices, we analyze the in-sample fit of computed equilibrium house prices compared to observed house prices. Finally, we conduct counterfactual policy experiments where we increase housing supply and commute time and compare the predicted responses in terms of residential sorting and job location.

7.1 Parameter estimates and model fit

The parameters estimates are provided in Tables 2.2 through 2.6[12]. Table 2.2 presents the estimates obtained from the housing demand regression in (2.27). Note that both annual income and housing price per square meter are measured in units of 100,000 DKK and we therefore use this unit in the following examples. A slight complication is that annual income is recorded before taxes while housing expenses obviously must be paid after taxes. Therefore, the implicit willingness to pay for housing, amenities and commuting will be measured in pre-tax income rather than actual disposable income. To avoid this issue, one would have to model the tax system on top of the wage equations, which we have deferred from in the current version.

The coefficients of the reducd form housing demand presented in Table 2.2 have reasonable magnitudes and expected signs. Recall from Section 5 that our estimation strategy only allowed for identification of scaled parameters in the first step housing

Table 2.2: First Stage Parameter Estimates, Housing Demand

Variable (parameters)	Coeff. Estimates	Standard Error	t-statistic
Const ($-\phi_0/\phi_{h2}$)	122.3154	0.05752	2126.3
Married ($-\phi_{ms}/\phi_{h2}$)	19.4172	0.01517	1279.7
Children ($-\phi_c/\phi_{h2}$)	13.6033	0.01615	842.2
Age ($-\phi_a/\phi_{h2}$)	0.5824	0.00059	983.6
Price pr. sqm ($\kappa_0\psi_{uc}/\phi_{h2}$)	-304.1712	0.21142	-1438.7
Price pr. sqm × income ($\kappa_y\psi_{uc}/\phi_{h2}$)	21.3753	0.02827	756.1

[12]Parameter estimates from income regressions for employment regions available in Appendix C. Parameter estimates for non-employment not shown but available upon request.

Table 2.3: Curvature Parameter of Housing Demand and User Cost

	Coeff. Estimates	Standard Error	t-statistic
Coef. on h^2, ϕ_{h2}	-0.0007	0.00000	-865.0
User cost housing, ψ_{uc}	0.2466	0.00139	177.9
κ_0	0.863		
κ_y	-0.061		

demand. We can however deduce that demand is increasing as a function of age and household size, and couples live in homes that, on average, measure 19.4 m^2 more than singles. Having children living at home is associated with a 13.6 m^2 larger dwelling. Housing demand is decreasing in prices as $\frac{\kappa_0 \psi_{uc}}{\phi_{h2}} < 0$, yet to a lesser extent for richer individuals since the term interacted with income is positive.

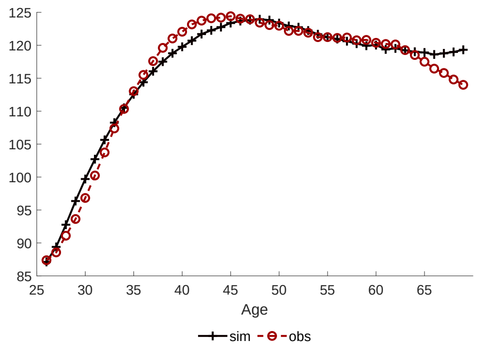
During the sample period, the average price per square meter was $26,661$ DKK in Copenhagen. Hence, individuals choosing to live in the Copenhagen municipality will on average demand $21.375 * 0.266 = 5.7$ more square meters of housing for each additional 100,000 DKK of individual annual income. Similarly, an individual with an income of 500,000 DKK living in Copenhagen demands 13.7 fewer square meters of housing compared to an individual with similar income living outside the capital area (Rest of Zealand) where square meter prices are $19,704$ DKK on average, i.e. around 7,000 DKK lower than in Copenhagen municipality[13].

The parameter estimates for ϕ_{h2} and ψ_{uc} are given in Table 2.3. We estimate the annual user costs of housing to $\hat{\psi}_{uc} = 0.247$, ie. 24.7% of the market value. This is definitely on the high side, but there are certain factors that may explain it. First of all, as noted above there is the complication that income is recorded before-tax. Since the tax burden lies in the interval 30-50%, the user cost measured in disposable income is correspondingly lower, around $12-17\%$. Furthermore, our estimation period 2005-2010 is mostly characterized by falling housing prices. In the standard user cost equation for housing, expected discounted capital gains reduce the user cost. If that equation truly lies in the back of people's mind when making housing purchases, then falling prices and pessimistic expectations work to increase user costs, and this might be what our estimate of ψ_{uc} is picking up.

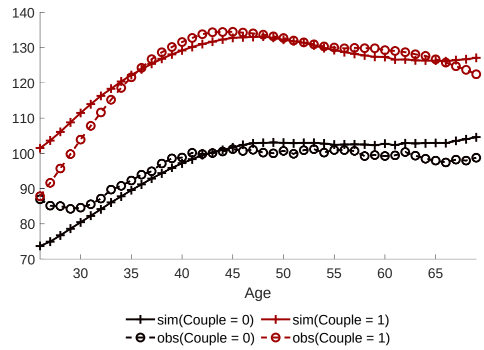
The model fit of chosen house size over the life cycle is shown in Figure 2.9a. The parameters capture the change in the demand over the life cycle closely. Separating by household size both in terms of having children and having a partner the model also provides a reasonable fit. There are some challenges capturing the demand at the beginning and end of the life cycle. The same goes for demand by education groups, where there is an underprediction of demand for the highly educated and overprediction for the medium- and low-skilled at the end of the life cycle. These obstacles are likely a result of the fact

[13]The difference in housing demand across these two regions is computed as $(-304.17 + 21.37 * 5) * (0.26661 - 0.19703) = -13.73$.

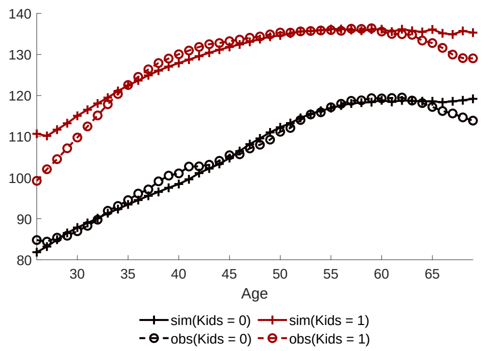
Figure 2.9: Model fit: housing size over the life cycle



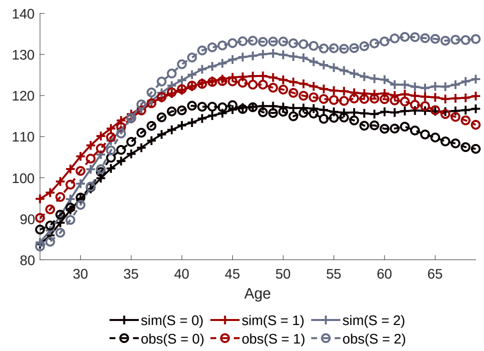
(a) Overall



(b) By marital status

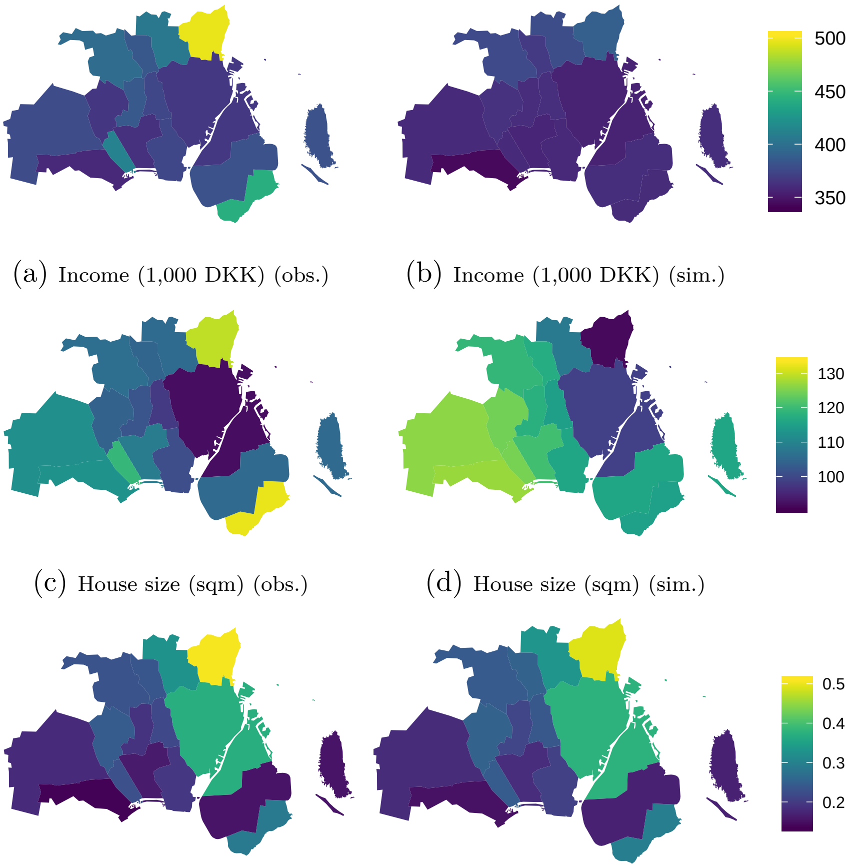


(c) By children



(d) By schooling

Figure 2.10: Model fit: income and housing size by home region



(a) Income (1,000 DKK) (obs.)

(b) Income (1,000 DKK) (sim.)

(c) House size (sqm) (obs.)

(d) House size (sqm) (sim.)

(e) Share highly educated (obs.)

(f) Share highly educated (sim.)

Note: Panels (a) and (b) show the average income in 1,000 DKK by home region. Panels (c) and (d) show the average size of homes in square meters by home region.

that we do not allow for adjustment costs and savings to affect housing size decisions.

Using the estimates of $\phi_{h2} = -0.0007$ and $\psi_{uc} = 0.2466$ together with the reduced form estimates in housing demand given in Table 2.2, we can back out the parameters that index marginal utility of money. We obtain $\kappa_0 = 0.86$ and $\kappa_y = -0.061$. Despite the strong negative gradient in income, these parameters result in relatively large estimates of marginal utility of money throughout most of the income distribution. Therefore the parameters imply a strong trade-off between home size and residential location and a clear sorting by richer individuals into more attractive and expensive regions and larger houses.

Figure 2.10 illustrates the ability of the model to fit sorting by highly educated individuals in conjunction with the variation in average income and housing demand across regions. The distribution of highly educated is captured very well because the income equation is specifically tied to the individual's education, and income predicts the home location through marginal utility of money. For example, the model is able to predict that the share of highly educated is high in Copenhagen, Frederiksberg and Gentofte where per square meter prices are high.

Although the model captures the educational sorting quite well, it has difficulty capturing the income levels in Gentofte, Dragoer and Vallensbaek. The latter would be improved if we included more heterogeneity in the income specifications such as the lagged dependent variable and persistent unobserved heterogeneity. Without sufficient variation in income, it is also hard to explain the spatial distribution of house sizes. The house sizes are especially underpredicted for Gentofte and Dragoer where individuals' incomes are the highest according to Figure 2.10a.

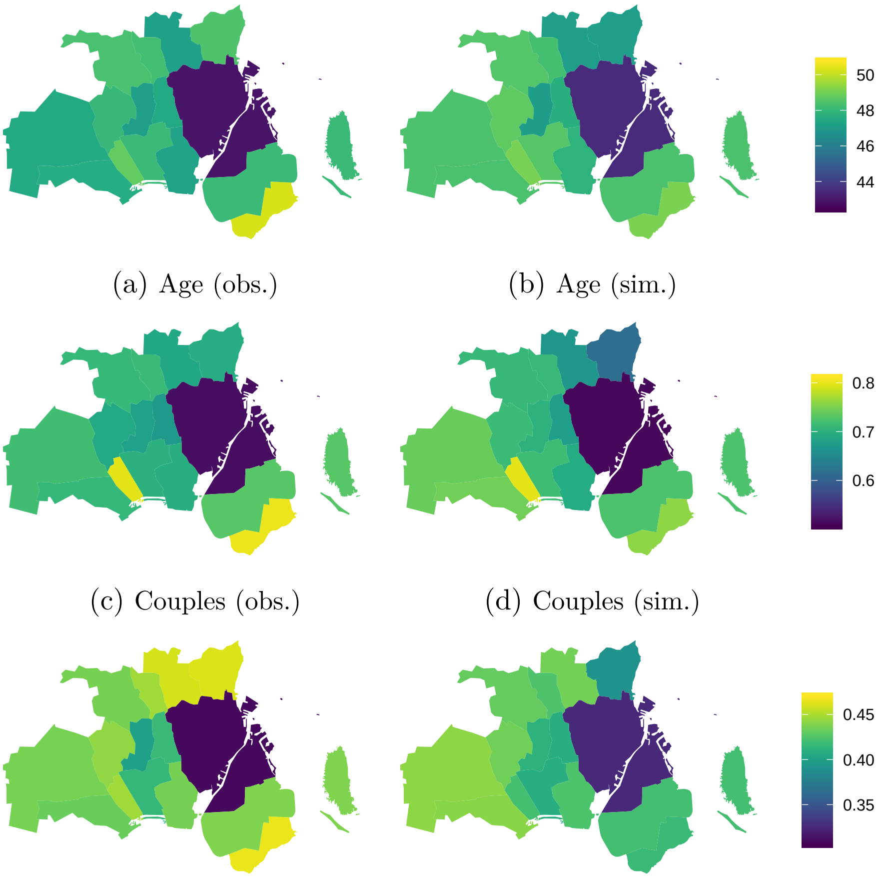
The residential sorting is driven mainly by four factors: i) regional variation in house prices and regional-specific amenities, ii) individual differences in housing demand, iii) individual differences in marginal utility of money and iv) distance to local labor markets. To flexibly capture regional-specific amenities, we include fixed effects for each residential region, $\alpha_{rl_{t+1}}$. The presence of local fixed amenities help rationalize why individuals prefer to live in regions where prices are high for reasons that are not explained by factors such as better access to local labor markets. The parameter estimates are presented in Table D1 in the appendix.

Gentofte (region 5) is associated with the highest amenity level and together with Frederiksberg (region 1) these are the only regions with better amenities than Copenhagen municipality (region 0) which is the outside category. The least attractive regions in terms of amenities are Albertslund (region 9), Hoeje-Taastrup (region 11), Ishoej (region 13) and Vallensbaek (region 15). They are all located in a cluster on the south-western border of the Greater Copenhagen Area. To exemplify the magnitudes, an individual with an annual income of $500,000$ DKK would need $1.7885/(0.86 - 0.061 \cdot 5) = 3.22$, i.e. 322,252 DKK, in compensation for living a year with the amenity level of Ishoej rather than that of Copenhagen municipality.

Figure 2.11 presents the model fit in terms of the residential sorting of household demographics. Starting with Figure 2.11a, the average age of the individual by home region is well captured. It is only slightly underpredicted in Dragoer, Gentofte and on the border between Rest of Zealand and the GCA. Looking at the share of couples in each region, the model fit also looks good. Again, Dragoer and Gentofte stand out as regions where the model underpredicts the shares the most. The distribution of families with children is less accurately captured, cf. Figure 2.11f. This could be improved by interacting the fixed effects by a dummy for children to pick up if the unobserved amenities of regions are valued differently by households with children.

Figure 2.12a shows corresponding fits for the probability of living in Copenhagen, but over the life cycle instead of the spatial allocation. The fit is very good in all respects. Only for the youngest cohorts is there a slight underprediction of the share living in Copenhagen. This is partly due to the fact that we do not model educational choice, and many higher educational institutions are located in Copenhagen. Figure 2.12c does indeed show that this problem is only evident for individuals with high education. It should be

Figure 2.11: Model fit: residential sorting



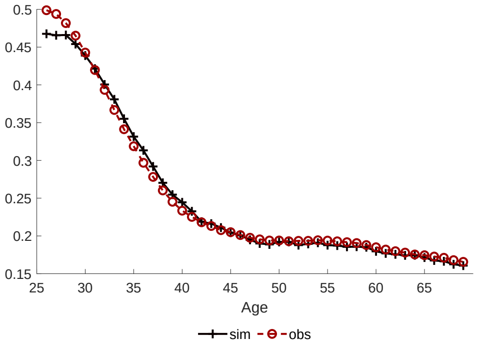
(a) Age (obs.) (b) Age (sim.)

(c) Couples (obs.) (d) Couples (sim.)

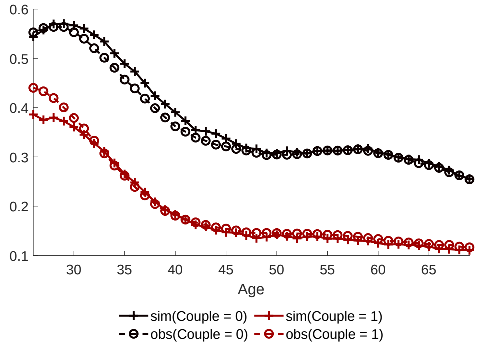
(e) Children (obs.) (f) Children (sim.)

Note: Panels (a) and (b) show average age by home region. Panels (c) and (d) show the share of couples by home region. Panels (e) and (f) show share of households with children by home region.

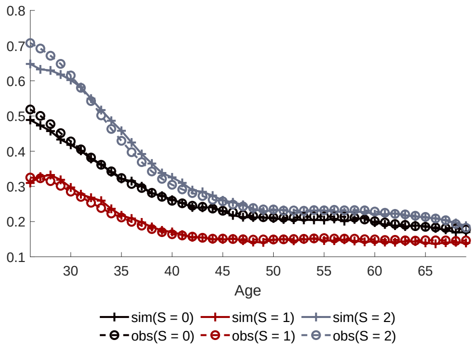
Figure 2.12: Model fit: share living in Copenhagen over the life cycle



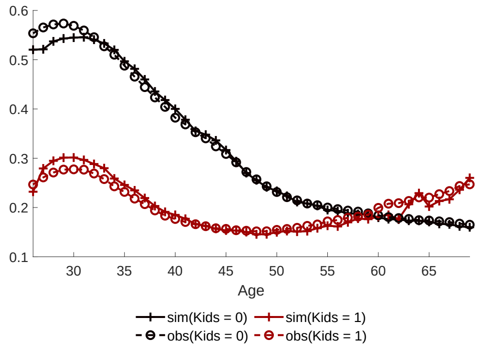
(a) Overall



(b) By marital status

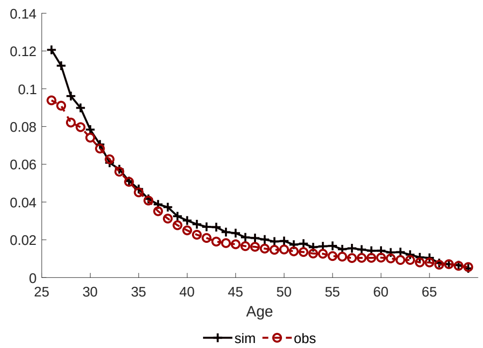


(c) By schooling

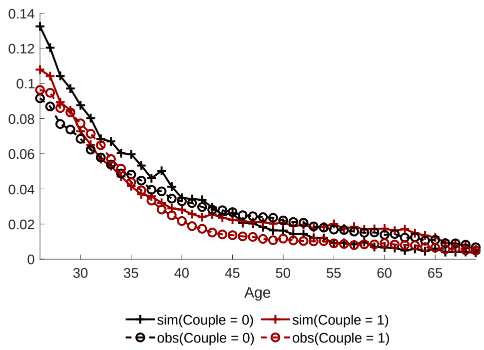


(d) By children

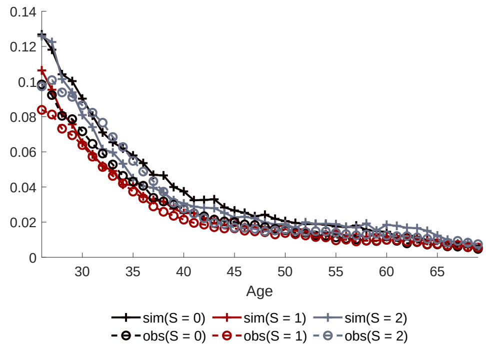
Figure 2.13: Model fit: share moving residential location over the life cycle



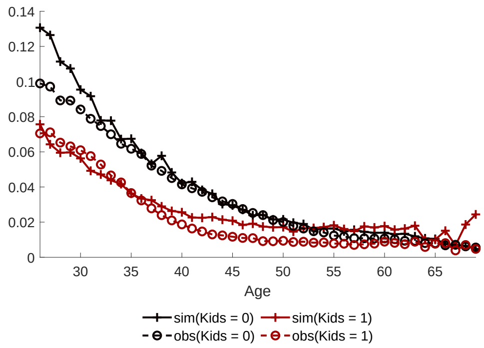
(a) Overall



(b) By marital status



(c) By schooling



(d) By children

noted that these moments are not only driven by the estimates of amenity values, but to a large extent by the moving costs that prevent people from moving away from their initial locations.

Table 2.4: Utility Cost of Moving Residence

	Coeff. Estimates	Standard Error	t-statistic
Const., γ_0	1.8363	0.00921	199.4
Age, γ_a	0.0881	0.00021	420.3
Married, γ_{ms}	0.0605	0.00485	12.5
Children, γ_c	0.8212	0.00523	156.9
Schooling, γ_s (1)	0.1797	0.00553	32.5
Schooling, γ_s (2)	-0.1470	0.00545	-27.0

Table 2.4 displays the estimates for the parameters γ that index the utility cost associated with moving residence. Married individuals and those with children are predicted to have higher moving costs and more so as they age. Medium-skilled individuals are less likely to move, all else equal, compared to low- and high-skilled types. Individuals with highest education are more mobile.

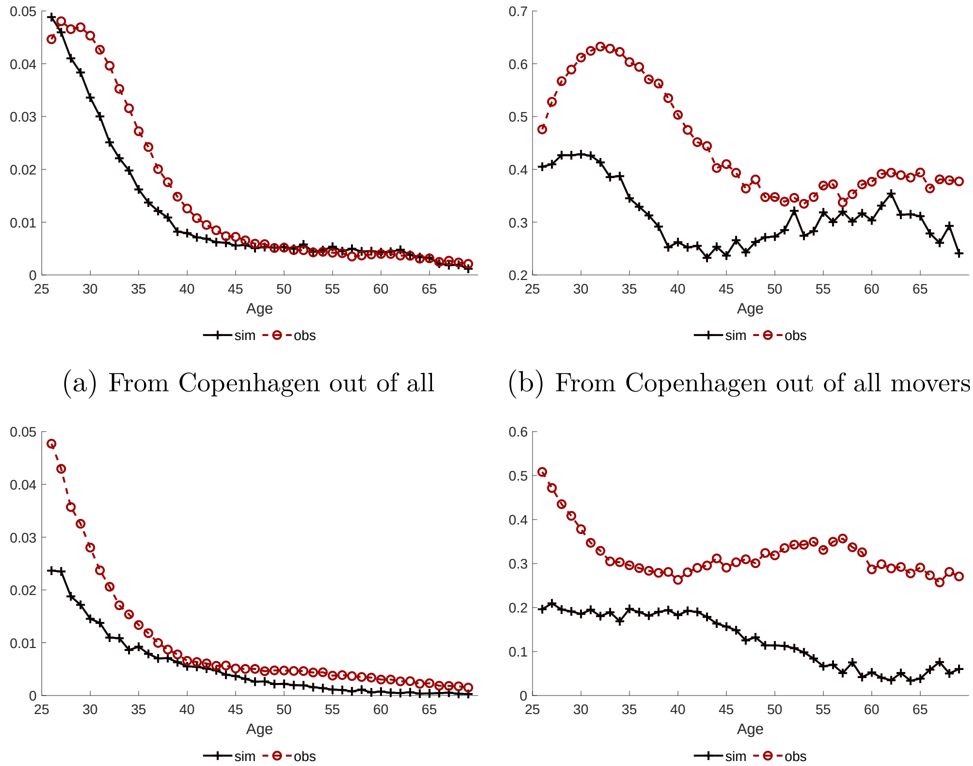
Table 2.5: Job Arrival and Dismissal

	Coeff. Estimates	Standard Error	t-statistic
Probability of keeping job: $\pi_t^k(wl_t, x_t; \beta^k)$			
Const., $\beta_0^{\pi(keep)}$	2.2226	0.04122	53.9
Age, $\beta_a^{\pi(keep)}$	0.0384	0.00098	39.0
Schooling, $\beta_s^{\pi(keep)}$ (1)	0.8267	0.02178	38.0
Schooling, $\beta_s^{\pi(keep)}$ (2)	0.5677	0.01633	34.8
Probability of new job: $\pi_t^n(d_t^w, wl_t, x_t : \beta^n)$			
Const., $\beta_0^{\pi(new)}$	-0.2453	0.00617	-39.7
Age, $\beta_a^{\pi(new)}$	-0.0624	0.00014	-457.6
Schooling, $\beta_s^{\pi(new)}$ (1)	0.1455	0.00347	41.9
Schooling, $\beta_s^{\pi(new)}$ (2)	0.2580	0.00375	68.8
Job density $\beta_{jobdensity}^{\pi(new)}$	2.9591	0.00700	422.7
Prev. unempl., $\beta_{unemp}^{\pi(new)}$	1.2326	0.00337	365.6

Overall, the model fit in terms of residential moving probabilities is good according to Figure 2.13a. There is a slight overprediction in the start of the life cycle, especially for individuals without children and singles. The largest prediction error is found for the probability of moving to and from Copenhagen. Figure 2.14 shows that the general shape of the probability of moving away from Copenhagen (as a share of all individuals in our data) is captured by the model, but it underpredicts the level until the age of 45. Conditioning on those who move, Figure 2.14b reveals that the same problem is observed among the movers, but the magnitude is larger. As the lower panel shows, the same can be said about the share migrating to Copenhagen. A key factor left out of the model is that we ignore the obvious fact that Copenhagen is a university city. Without explicitly modeling educational choice and the dynamics of occupational career choice it is hard to explain why younger individuals with low incomes choose to live in Copenhagen. Other omitted factors are individual taste variation for regional-specific amenities such as bars, restaurants, child care, and school quality which can readily be included into this model at a low computational cost. Also by modelling moving costs more carefully, e.g. including unobserved heterogeneity, we may be able to predict these shares better.

We now move to the ability of the model to predict work location outcomes. Table 2.5 displays estimates for the parameters for the job arrival and dismissal probabilities $\pi_t^n(d_t^w, wl_t, x_t : \beta^n)$ and $\pi_t^k(wl_t, x_t; \beta^k)$ that determines the work location transition probabilities. Starting with the probabilities of keeping the job, there is a positive effect of age and higher levels of schooling. An individual who is 40 years old and has a low-level education has a $(1 + \exp(-(2.435 + 0.014 \cdot 40)))^{-1} \cdot 100 = 97.7$ percent chance of keeping the job. A similar person, who was working in $t-1$ and searches for a new job has a 55.4 percent chance of being successful and ending in Copenhagen. In Hoeje-Taastrup, where

Figure 2.14: Model fit: share moving residential location from and to Copenhagen over the life cycle



(a) From Copenhagen out of all

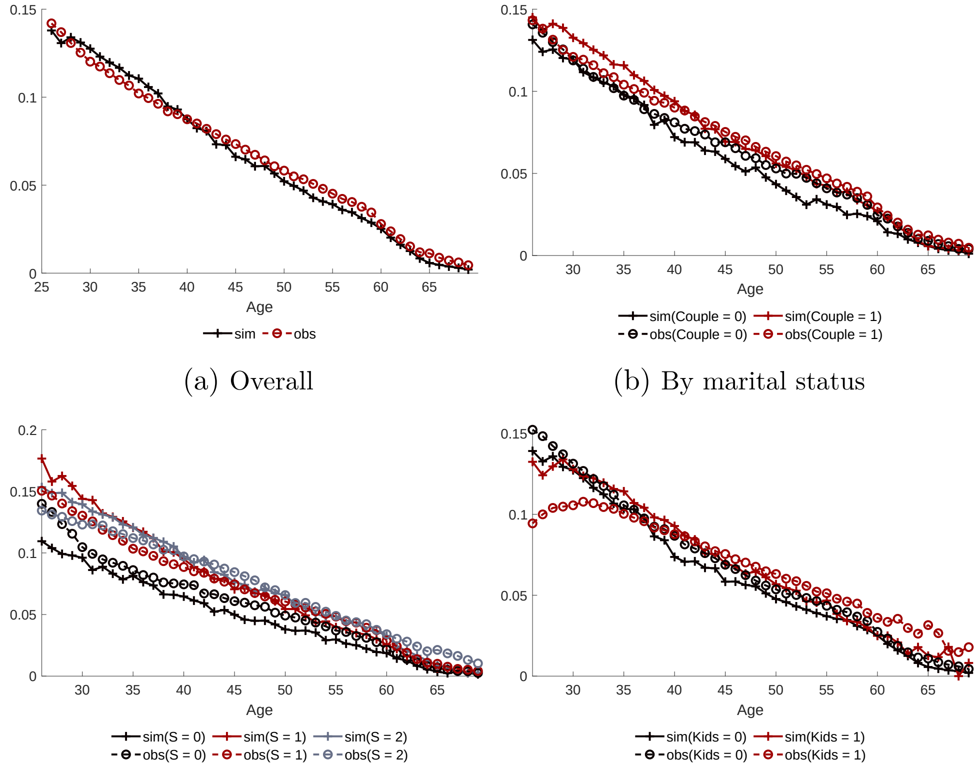
(b) From Copenhagen out of all movers

(c) To Copenhagen out of all

(d) To Copenhagen out of all movers

Note: Panels (a) and (c) show the share of all individuals in the data who move residential location from and to Copenhagen, respectively. Panel (b) and (d) show the share of all residential movers who move from and to Copenhagen, respectively.

Figure 2.15: Model fit: share moving work location over the life cycle



(a) Overall

(b) By marital status

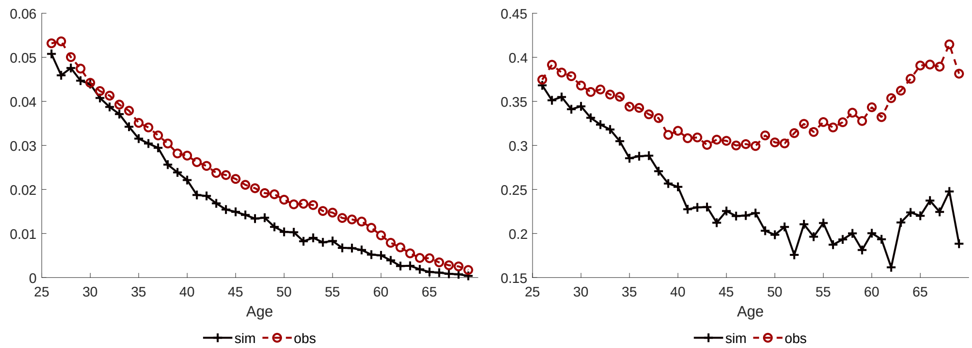
(c) By schooling

(d) By children

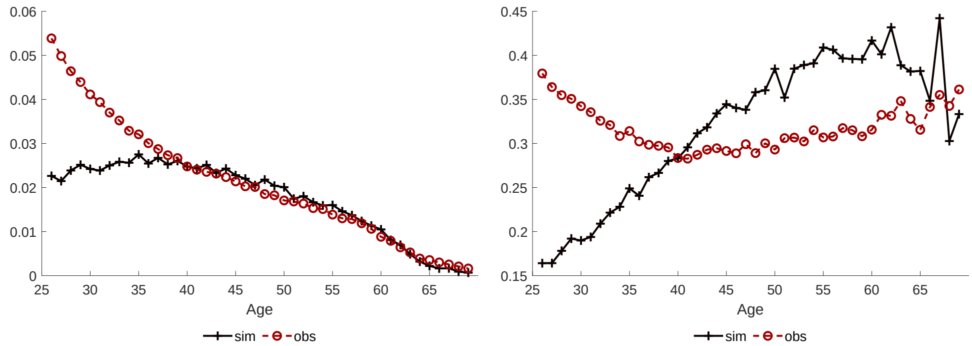
the job density for low-skilled jobs is 0.099 instead of 1 as in Copenhagen, the probability of ending up there would have been only 8.0 percent. The large regional differences in supply of jobs is therefore strongly reflected in the job probabilities.

Figure 2.15a also shows that the model can capture the share moving work location over the life cycle. There are some challenges of modelling the work transition probabilities for the younger individuals. Especially for those who have children where the model overpredicts the mobility while for low-skilled people it underprecits mobility. Looking specifically at the probability of moving work location to and from Copenhagen, Figure 2.16a illustrates that the share moving their job away from Copenhagen (out of all individuals in the data) shows a satisfactory fit though the model underpredits from age 40 onwards. The motivation for moving one's job conditional on the home location is shorter commute or higher wages. Commute distances are exogenous and thus independent of age, while wages have an age profile. We are aware wages may not exhibit enough variation across individuals as we do not allow for unobserved heterogeneity. Including this may improve on the fit since we would better capture whether the more mobile individuals are those who have a high unobserved fixed component of wages that they can bring with them when they move around. When zooming in on movers only in Figure 2.16b the fit is worse. Hence, the model cannot capture individuals who would like to move away from Copenhagen where there is such a high job density and thus high chance of being

Figure 2.16: Model fit: share moving work location from and to Copenhagen over the life cycle



(a) From Copenhagen out of all (b) From Copenhagen out of all movers



(c) To Copenhagen out of all (d) To Copenhagen out of all movers

Note: Panels (a) and (c) show the share of all individuals in the data who move work location away from and to Copenhagen, respectively. Panel (b) and (d) show the share of all work location movers who move from and to Copenhagen, respectively.

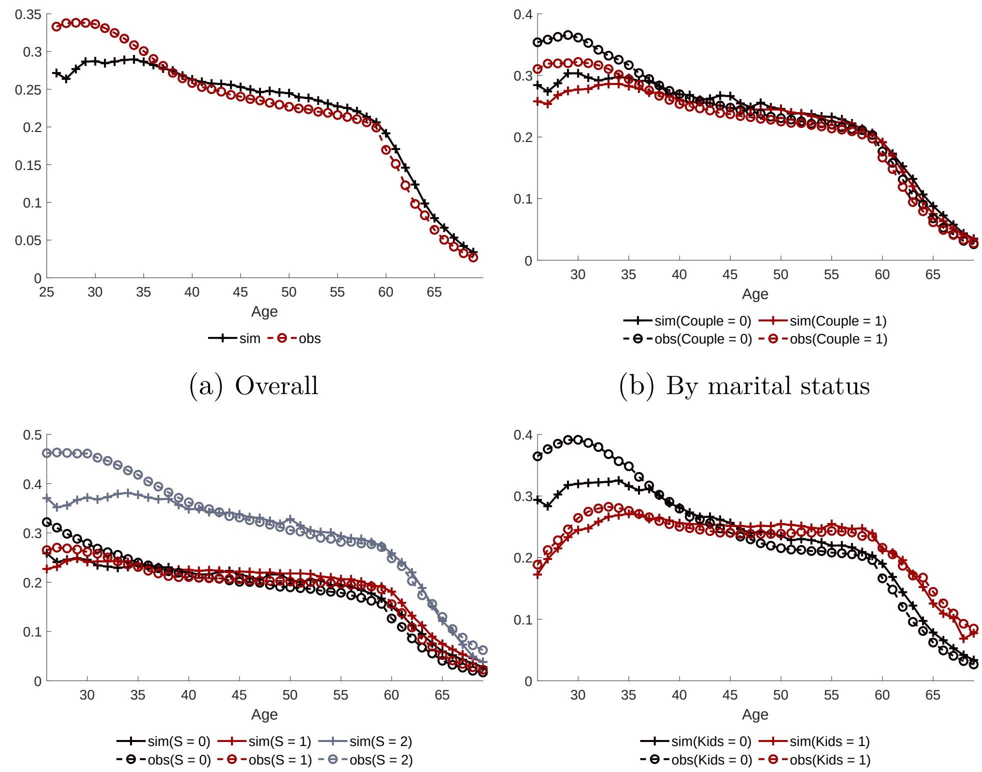
employed.

Considering instead the share working in Copenhagen in Figure 2.17a, the fit looks very good for the individuals older than 35. The heterogneity across individuals is also reflected in the model predictions. The work location decision is less well-captured for the young people because we do not model initial conditions or educational choice.

Looking at the share of individuals working in Copenhagen by their home municipality, the top panel of Figure 2.18 shows that the model captures the spatial distribution pretty well. It underpredicts the share somewhat for people also residing in Copenhagen. This can be improved by better capturing the share moving home location away from and to Copenhagen over the life cycle since that alone should make it more likely to also work in Copenhagen.

Table 2.6 provides estimates of the commute cost parameter, η_{ttime}, and disutility of working, c_{work}. The latter reflects the compensation one would require to take a job instead of being unemployed and corresponds to 280,393 DKK for a person with an annual (non-employment) income of 150,000 DKK. η_{ttime} indicates that an employed person with

Figure 2.17: Model fit: share working in Copenhagen over the life cycle



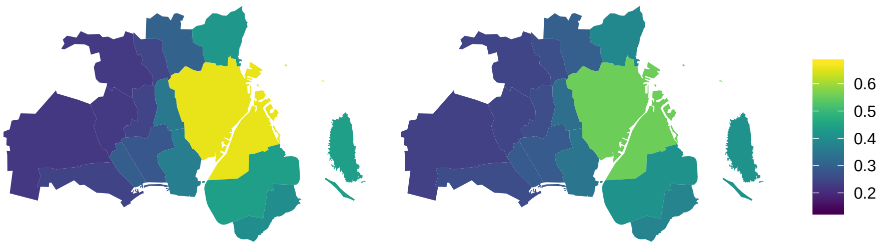
(a) Overall

(b) By marital status

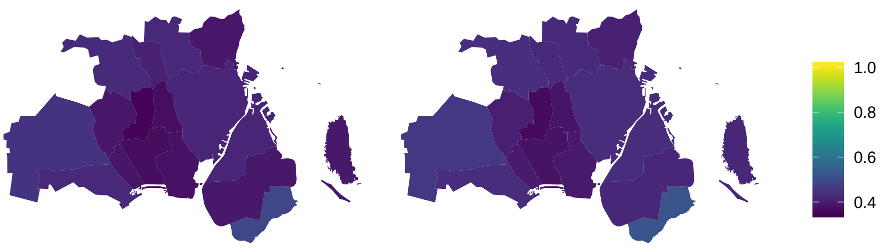
(c) By schooling

(d) By children

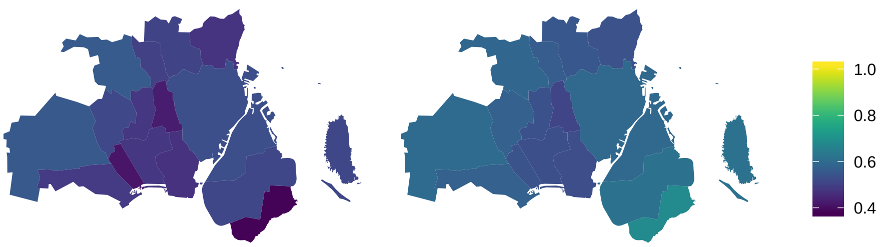
Figure 2.18: Model fit: working in Copenhagen and commute times (hours)



(a) Work in Cph. by home (obs.) (b) Work in Cph. by home (sim.)



(c) Commute time by home (obs.) (d) Commute time by home (sim.)



(e) Commute time by work (obs.) (f) Commute time by work (sim.)

Note: Panels (a) and (b) show the share of individuals in each home region who works in Copenhagen or Frederiksberg. Panels (c) and (d) show average commute time in hours by home region for employed individuals. Panels (e) and (f) show average commute time in hours by work region for employed individuals.

an income of 500,000 DKK would only be willing to commute one hour further if she earned an additional 42,685 DKK. The disutility of commuting in the data is therefore not overwhelming considering the fact that individual annual wage incomes typically are in the range 300,000-400,000 DKK.

Table 2.6: Commute Cost and Disutility of Work

	Coeff. Estimates	Standard Error	t-statistic
Cost of travel time, η_{ttime}	0.2369	0.00118	200.8
Disutil. of work, c_{work}	2.2163	0.00189	1175.6

As pictured in the middle and lower panels of Figure 2.18, the prediction error in the spatial allocation of commute times is low, especially by home locations. By work locations, the model predicts higher and a more uniform distribution of commute times than is observed in the data. The fact that average commute times are generally higher when splitting by work instead of home location is because people from Rest of Zealand also commute to the regions shown on the map. Figure 2.19a illustrates the commute time over the life cycle and across different types of individuals and it is predicted very accurately by the model. It is mainly for individuals above age 65 that the model starts to struggle, but there is also a strong selection among working individuals at that age. It is therefore not surprising that they cannot necessarily be compared to the younger working cohorts.

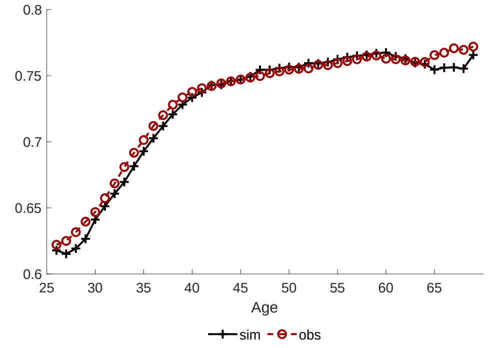
7.2 Baseline equilibrium

To assess whether the observed house prices in our data form an equilibrium at the housing market, we solve for the equilibrium prices following the procedure outlined in Section 6 and using the obtained structural estimates.

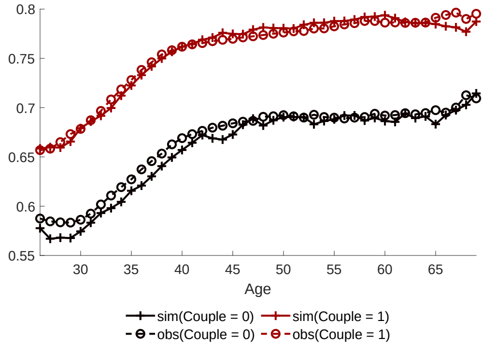
Figure 2.20 plots computed equilibrium prices against observed price data. The fit appears very good both in terms of the price ranking of the different regions as well as the overall levels. Here is it is important to emphasize that the model is estimated without explicitly imposing that the housing market is in equilibrium. The fact that the equilibrium prices predicted from our estimated model closely track the observed house prices in the different regions provides a good in-sample validation of the many cross-equation restrictions implied by our modeling of location choices and demand for house size.

With the overall fit being exceptionally good, there is a slight overprediction of prices in the cheapest regions and an underprediction in Gentofte, the most expensive region. Our parsimonious modeling of individual income and the lack of savings are again among the potential explanations as to why the model does not fully capture why people are willing to pay such high prices in Gentofte.

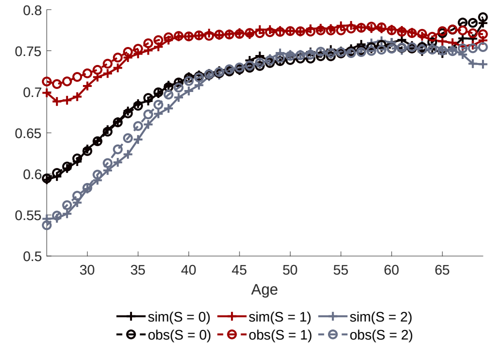
Figure 2.19: Model fit: commute time (hours) over the life cycle



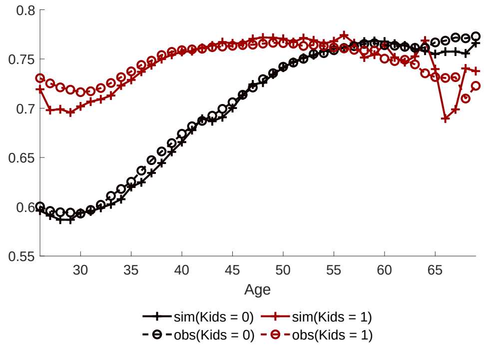
(a) Overall



(b) By marital status



(c) By schooling



(d) By children

Figure 2.20: Relation between observed and baseline equilibrium house prices per sqm (100,000 DKK)

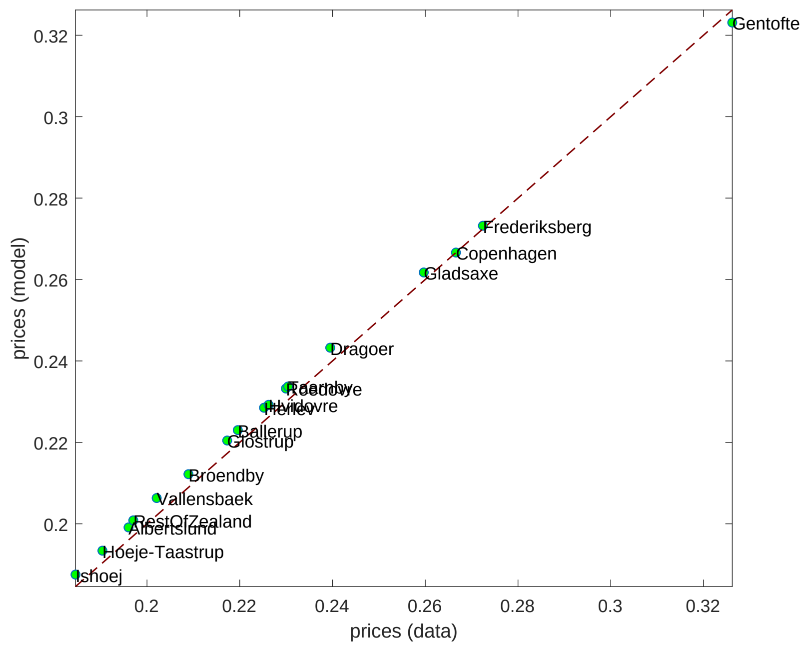
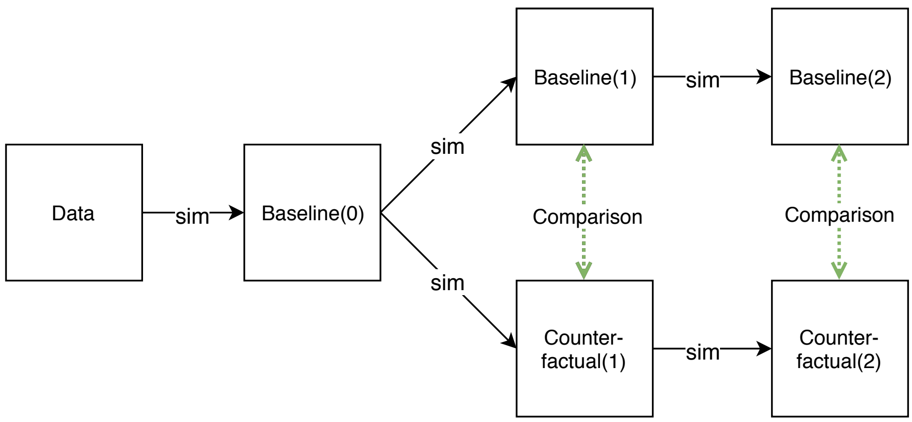


Figure 2.21: Structure of comparison between counterfactual and baseline



7.3 Counterfactual equilibrium

In order to make a valid comparison between the baseline of the model and a counterfactual simulation, we account for the implied relocations and price changes that were to occur even in the absence of any policy change (due to demographic trends) by simulating the model forward a number of periods. In doing so, we obtain a simulated household-level panel dataset with baseline outcomes. The baseline simulation starts at the empirical data on which the model was estimated. The outcome of simulating the model one period ahead from the empirical data yields the initial state for both the following baseline simulation steps as well as the counterfactual. This structure is illustrated in Figure 2.21, where the first baseline dataset is denoted Baseline(0). As Baseline(0) is the initial condition for the counterfactual, policy changes are imposed *at the beginning* of simulated period 1. At the end of period 1 it is therefore possible to identify all changes between baseline and counterfactual outcomes at the household level.

Counterfactual I: Increased housing supply

The first counterfactual experiment involves a 5 percent exogenous and permanent increase of the housing supply (square meters) in Copenhagen and Frederiksberg. Using the simulations for period 2, we study the implications for location choice, housing size, income sorting and equilibrium prices.

Table 2.7 summarizes the first four measures. As expected, the share living in Copenhagen and Frederiksberg increases, though just by 0.12 and 0.09 percentage points, respectively. Even though supply was constant in all other regions, the number of people living in Gentofte, Gladsaxe and Roedovre (which all share borders with Copenhagen) also show positive trends. As a result, the average housing size falls in these three regions. The demand for living in all the remaining regions drops, so in total the degree of centralization and urbanization has increased.

Individuals living in Copenhagen and Frederiksberg already can immediately adjust their housing demand upwards due to the lack of adjustment costs. However, some people from outside those regions, who did not prefer living there in baseline, are now inclined to relocate to those cities because there is a possibility of consuming more square meters. This starts the equilibrating process of some people moving out and substituting Copenhagen and Frederiksberg by Gentofte, Gladsaxe and Roedovre which are close substitutes in space. The increased demand for living in these nearby regions affects their equilibrium prices and prompts the original residents to consider moving too.

Looking at the resulting equilibrium prices, Figure 2.22 shows that all regions experience falling prices per square meter. This is especially true in Copenhagen, Frederiksberg, Gentofte and Roedovre despite the increased demand for living there. The average price per square meter in Copenhagen was 26,661 DKK and it falls by 750 DKK corresponding to 2.8 percent. Thus, locally in Copenhagen 56 percent of the supply shock is soaked up by falling prices. The lower prices in Copenhagen and Frederiksberg are caused by a more moderate increase in demand for living there than the increase in housing supply. In Gentofte and Roedovre there are two counteracting effects: increased demand for living there and spillovers from the generally lower price level in the GCA. The latter dominates. In the rest of the region the drop in prices is due to the substitution effect that induces people to move away and closer to the urban center.

Due to the reallocation of people across space, the sorting patterns have also changed in equilibrium. The second column of Table 2.7 shows the change in average income by home region and the third column the change in the within-region standard deviation in income. Copenhagen, Frederiksberg and Gentofte experience an increase in average income and incomes are more homoegeneous within those regions after the policy change, i.e. due to the resorting of individuals, the increased supply of housing does not induce many low-income households to live in Copenhagen and Frederiksberg in equilibrium despite the lower equilibrium prices. In Roedovre, on the other hand, incomes are lower and more heterogeneous. Dragoer stands out as the region with the highest increase in average income and the most significant fall in the standard deviation of income. This is consistent with the idea that the lower-income original residents of Dragoer move towards Copenhagen, where they can now consume a satisfying number of square meters at a more reasonable price than in baseline. The average income of residents in Dragoer is indeed higher than in Copenhagen, so it is likely that the lower-income outmigrants from Dragoer to Copenhagen have an income above the average for the original Copenhagen residents.

In conclusion, the effect on commute time is negligible and the effects on residential location and sorting are more complex. We will investigate this in more detail in future work.

Counterfactual II: Increased commute costs

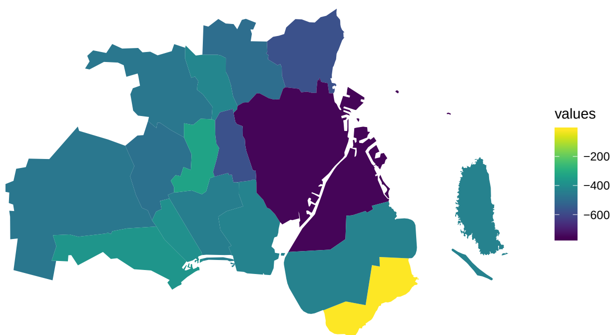
In the second counterfactual we increase the commute costs η_{ttime} by 50 percent. This might resemble an increase in monetary costs of traveling due to a removal of the mileage

Table 2.7: Simulated changes in $t = 2$ in Counterfactual I

	Population Share % points	E(inc) %	Std(inc) %
Copenhagen	0.12	0.05	-0.05
Frederiksberg	0.09	0.13	-0.47
Ballerup	-0.01	0.15	0.11
Broendby	-0.02	-0.12	0.13
Dragoer	-0.05	0.94	-2.61
Gentofte	0.10	0.01	-0.46
Gladsaxe	0.06	-0.50	0.43
Glostrup	-0.01	-0.47	0.48
Herlev	-0.01	-0.01	0.08
Albertslund	0.00	0.16	0.00
Hvidovre	-0.01	-0.08	0.13
Hoeje-Taastrup	0.00	-0.11	0.21
Roedovre	0.02	-0.44	0.56
Ishoej	-0.02	0.09	-0.47
Taarnby	-0.03	0.00	-0.21
Vallensbaek	-0.01	0.08	-0.21
Rest Of Zealand	-0.20	-0.01	0.00

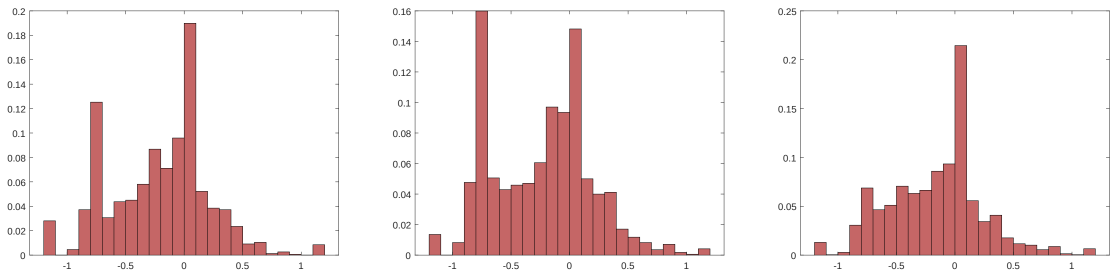
Note: Numbers are computed by subtracting baseline from counterfactual. Population share refers to the change in the share of all individuals who live in the region. E(inc) refers to the change in the average income of residents in the region. Std(inc) refers to the change in the standard deviation of income of residents in the region.

Figure 2.22: Simulated change in price per sqm in $t = 2$ in Counterfactual I (DKK)



Note: Numbers are computed by subtracting baseline from counterfactual.

Figure 2.23: Distribution of simulated change in commute time in $t = 2$ in Counterfactual II (hours)



(a) Low education (b) Medium education (c) High education

Note: Changes in commute times are computed by subtracting baseline from counterfactual.

allowance (a tax benefit for commuters), lower subsidies on ticket prices, increased gasoline prices or alternatively increasing congestion in a uniform manner across all regions.

As commute costs rise, households will intuitively want to be closer to their jobs. This may be achieved through relocating their jobs, relocating home location or opting out of working entirely. The overall effect on commuting in the counterfactual is visualized by the histograms in Figure 2.23 showing differences in commute time when comparing with the baseline simulation. To be clear, for each individual in the counterfactual simulation we have subtracted their commute time in the baseline simulation. Further, we condition on individuals who have i) chosen another home location than in baseline and ii) were in employment in simulated period 0. We condition on i) and ii) for the sake of exposition. In particular, ii) avoids a large mass at 0 as a consequence of no commute time in non-employment coupled with the high persistence of the non-employment state.

Figure 2.23 clearly shows more mass in the negative support stemming from individuals lowering commute time in response to the increased costs. The figures also show a mass point around -0.75 hours for low- and medium-educated individuals. This is caused by individuals who in baseline lived in Rest of Zealand and worked in Copenhagen, yet after the change in commute costs decided to opt for non-employment. The same mass point does not occur for the highly educated, see Figure 2.23c, as they have much less incentive to opt out of the labor market due to higher opportunity costs of not working.

We would expect the relocations of individuals to be particularly evident in the most peripheral regions and this is confirmed in Table 2.8. The first two columns show the percentage of individuals who relocated their job away from each municipality in baseline and counterfactual. Most strikingly, individuals working in Rest of Zealand change their work location to a relatively large degree. This can be concluded as it only holds 8 percentage points more workers than Copenhagen, cf. Table E1 in appendix, while the relative increase in job relocations is much higher than for Copenhagen. Like Rest of

Zeland, the municipality of Ishoej is associated with long commute times due to its location on the southern perimeter of the GCA. Being on the perimeter, Ishoej and Rest of Zealand are relatively close substitutes in terms of work locations and we therefore also see significant increase in the relocations of jobs away from Ishoej.

As a case study of the model predictions, Table 2.9 displays detailed moving statistics for Ishoej. The first column shows the initial (simulated period 0) residential locations of individuals working in Ishoej. Predominantly, workers of Ishoej lived in Rest of Zealand and many therefore had long commute times. In the counterfactual state of higher commute costs, those workers would be particularly discouraged from continuing to work. We see this pattern indirectly in column two, which displays the distribution of work locations for those who switched work location between periods 0 and 2. 24.5 percent of people employed in Ishoej in period 0 did not work in period 2. As the third column shows, 66.8 percent of these non-employed people lived in Rest of Zealand in period 0, hence underlining the discouraging effect of higher commute costs. Note also that Gentofte, which is located the furthest away from Ishoej, displayed the second-highest share of non-employed residents in period 2.

Column two also indicates which work regions are the closer substitutes to Ishoej. The municipalities of Taarnby, Roedovre and Hoeje-Taastrup attract the most workers from Ishoej, although Vallensbaek and Albertslund are closer to Ishoej than both Roedovre and Taarnby. However, these are relatively small labor markets so the probability of getting a new job prohibits workers from relocating there.

Returning to Table 2.8, column five shows the change in equilibrium prices in period 2 between counterfactual and baseline. We note that all regions except Rest of Zealand experience slightly increasing prices, while the prices in Rest of Zealand declines. This is a direct consequence of lower demand for residing in Rest of Zealand and a substitution towards the GCA where commute times are lower. Correspondingly, column three and four show the share of outmigrants from each home region in baseline and counterfactual. There is a net inflow to the CGA as the propensity to move away from Rest of Zealand is higher in the counterfactual while it is lower in Copenhagen. These two regions dominate the picture due to their sizes, see Table E1 in appendix.

Table 2.8: Simulated share of relocations of work (wl) and home (rl) and price change in $t = 2$ for Counterfactual II

	Baseline ($t = 2$) wl	Counterfactual ($t = 2$) wl	Baseline ($t = 2$) rl	Counterfactual ($t = 2$) rl	Δprice (DKK)
Copenhagen	3.66	3.98	1.15	1.00	59.93
Frederiksberg	6.51	6.28	1.98	1.99	86.24
Ballerup	7.21	7.84	3.44	3.08	73.96
Broendby	10.34	11.04	3.46	2.94	77.54
Dragoer	27.24	26.26	4.78	5.57	57.43
Gentofte	7.53	7.46	2.80	2.71	6.62
Gladsaxe	7.72	7.84	1.57	1.17	81.80
Glostrup	11.22	11.27	3.56	3.74	247.17
Herlev	11.94	11.98	2.72	2.94	248.75
Albertslund	12.99	13.30	2.37	1.75	51.89
Hvidovre	9.74	10.05	2.09	1.78	108.70
Hoeje-Taastrup	9.94	10.83	2.28	1.97	13.45
Roedovre	13.97	14.69	4.27	3.90	1.70
Ishoej	20.93	22.58	5.80	5.25	17.12
Taarnby	11.97	12.22	5.96	5.99	96.89
Vallensbaek	36.17	36.34	7.29	7.49	136.72
Rest of Zealand	0.87	1.41	0.30	0.45	-148.13
Non-employment	5.78	5.69	-	-	-

Table 2.9: Simulated distribution of locations in $t = 2$ for $t = 0$ workers in Ishoej in Counterfactual II

	Home region of workers in $t = 0$ (%)	New wl of job movers in $t = 2$ (%)	Home region of job movers when new $wl = \emptyset$ in $t = 2$ (%)
Copenhagen	16.8	7.3	0.6
Frederiksberg	3.8	0.7	2.6
Ballerup	1.4	0.5	0.6
Broendby	2.0	1.1	-
Dragoer	0.5	0.2	0.3
Gentofte	4.6	0.9	15.3
Gladsaxe	2.1	0.5	1.9
Glostrup	0.9	0.3	1.6
Herlev	0.9	0.5	1.3
Albertslund	1.0	2.1	1.3
Hvidovre	2.4	4.1	1.9
Hoeje-Taastrup	3.7	15.7	2.2
Roedovre	1.3	14.9	1.0
Ishoej	10.2	-	0.6
Taarnby	0.9	20.0	0.6
Vallensbaek	1.2	3.5	1.3
Rest of Zealand	46.4	3.0	66.8
Non-employment	-	24.5	-

8 Conclusion

In this chapter we developed and empirically estimated a structural dynamic equilibrium model of joint home and work location decisions for individuals and estimated it using Danish administrative data. We found that overall the empirical fit of the model is very good. We focused on the Greater Copenhagen Area (GCA) and analyzed the counterfactual effects of i) increasing the housing supply in Copenhagen and Frederiksberg by 5 percent and ii) increasing commute costs by 50 percent.

We found that the increase in housing supply resulted in relocations towards Copenhagen and Frederiksberg such that the degree of urbanization increased. The relocations did not completely offset the increase in the housing supply, so the average housing size also increased in those two regions. In total, the equilibrium prices dropped in all regions and especially in Copenhagen, where they fell by 2.8 percent. The sorting of individuals was also affected. Hence, Copenhagen and Frederiksberg were characterized by richer and more homoegeneous households on average after the policy change.

The increase in commute costs not only caused an anticipated relocation leading to a decrease in average commute times, but also to a significant labor supply effect. In particular, a significant share of residents of Zealand outside of the Copenhagen region who worked in Copenhagen ended up in non-employment in the counterfactual. This effect is more pronounced for low- and medium-educated workers. The downward movement of the equilibrium prices follows the decline in labor participations, which is in line with the higher incentive to live within the GCA where commute times were lower.

Overall, the model developed and estimated in the paper provides valuable insights into our understanding of the location and movement patterns among Danish households, which are driven by the cost of living and commuting, and are very heterogeneous in the population. The current implementation of the model is not free of strong simplifying assumptions, but even in their presence it proves to be a very valuable tool, capable of explaining important variation in the data, and enabling us to undertake interesting counterfactual experiments.

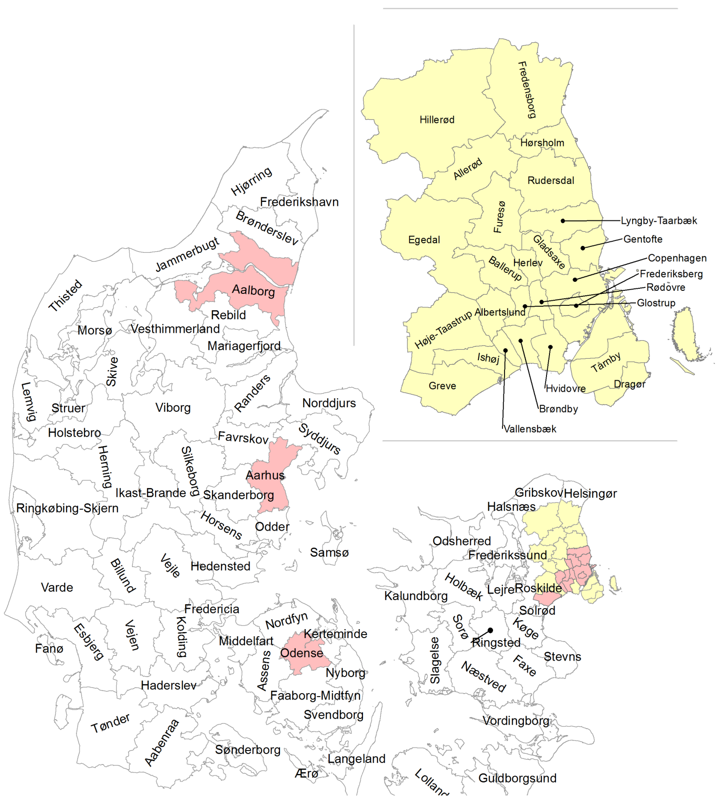
Among most significant limitations of the current implementation are the effective disregard of the time dimension of the data (especially in the dimension of developing amenities in different regions), and the static equilibrium house price calculations. The regions can be less aggregated, and a wider area of the country rather than the GCA can be used for estimation. Inclusion of the equilibrium wage settlement into the consideration is another obvious dimension for improvement. Even under the assumption of short term dynamics in the labor market similar to the housing market (so that the supply of jobs is constant) the wages can be treated similarly to house prices and be determined in the spatial equilibrium. All of these improvements, although requiring additional work and computational time, are straightforward to implement. Even though we do acknowledge all the limitations and relevant extensions mentioned above we leave the implementations for future research.

A Geographic Classifications

Table A1: Overview of geographical classifications in Denmark

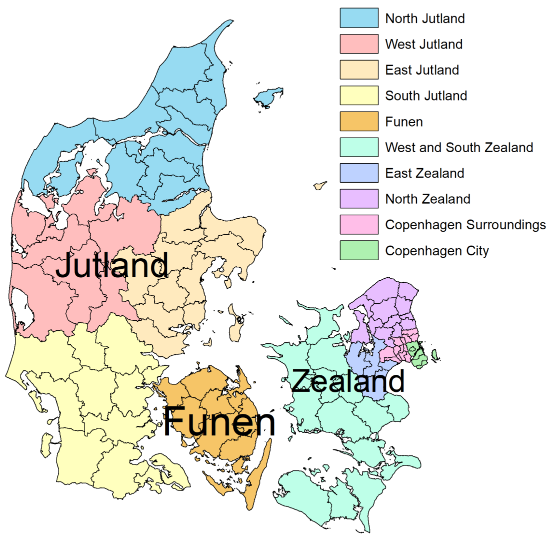
Danish	English	# units	Comment
Danmark	Denmark	1	
Regioner	Regions (states)	5	
Landsdele	Provinces	11	10 excl the island Bornholm
Amter	Counties	16	No longer exists
Valgkredse	Constituencies	92	
Kommuner	Municipalities	98	Reform in 2007: from 271 mun. to 98.
Trafikzoner	Traffic/LTM zones	907	Defined by DTU's traffic model
Sogne	Parish	2,201	

Figure A1: Municipalities of Denmark and urbanized areas



Note: Pink areas indicate an urban municipality. The yellow area corresponds to the main part of the greater Copenhagen area.

Figure A2: Provinces and main islands



B Amenities

Table B1: Summary statistics of job density by work region

Region	Low educ.	Medium educ.	High educ.
Copenhagen	1.0000	1.0000	1.0000
Frederiksberg	0.1117	0.1197	0.0966
Ballerup	0.1036	0.1517	0.0943
Broendby	0.0747	0.0976	0.0501
Dragoer	0.0074	0.0102	0.0064
Gentofte	0.0847	0.0962	0.1020
Gladsaxe	0.0915	0.1255	0.1042
Glostrup	0.0593	0.0886	0.0547
Herlev	0.0468	0.0721	0.0524
Albertslund	0.0639	0.0945	0.0463
Hvidovre	0.0783	0.1016	0.0605
Hoeje-Taastrup	0.0990	0.1351	0.0579
Roedovre	0.0474	0.0714	0.0310
Ishoej	0.0258	0.0332	0.0149
Taarnby	0.1011	0.1062	0.0353
Vallensbaek	0.0110	0.0147	0.0082
Rest Of Zealand	0.0508	0.0722	0.0397
Funen	0.0613	0.0893	0.0473
Jutland	0.0857	0.1211	0.0617

Note: Job density is defined as the number of jobs by education group and work region normalized by the value in Copenhagen. The numbers have been averaged over time.

C Wage Regressions

Table C1: Estimates from OLS of Log Real Wages for Low-Skilled Workers by Region

Work Region	*age*	*age*^{2}	$\mathbb{I}_{rw_{t-1}=\varnothing}$	*Constant*	R^2	N
Copenhagen	0.1587 (0.000)	-0.0017 (0.000)	-0.7765 (0.000)	9.0397 (0.000)	0.2332	410758
Frederiksberg	0.1554 (0.000)	-0.0016 (0.000)	-0.7561 (0.000)	9.0395 (0.000)	0.2327	43752
Ballerup	0.1313 (0.000)	-0.0014 (0.000)	-0.7811 (0.000)	9.8425 (0.000)	0.1976	47008
Broendby	0.1036 (0.000)	-0.0011 (0.000)	-0.7962 (0.000)	10.4932 (0.000)	0.1605	35442
Dragoer	0.1321 (0.000)	-0.0015 (0.000)	-0.6105 (0.000)	9.6267 (0.000)	0.1873	3412
Gentofte	0.1281 (0.000)	-0.0013 (0.000)	-0.7183 (0.000)	9.6782 (0.000)	0.1787	39708
Gladsaxe	0.1601 (0.000)	-0.0018 (0.000)	-0.8380 (0.000)	9.1874 (0.000)	0.2683	38132
Glostrup	0.1349 (0.000)	-0.0014 (0.000)	-0.6912 (0.000)	9.7162 (0.000)	0.2223	23925
Herlev	0.1306 (0.000)	-0.0014 (0.000)	-0.7055 (0.000)	9.7836 (0.000)	0.1958	19150
Albertslund	0.0957 (0.000)	-0.0010 (0.000)	-0.7419 (0.000)	10.5897 (0.000)	0.1527	28222
Hvidovre	0.1169 (0.000)	-0.0013 (0.000)	-0.6683 (0.000)	10.0812 (0.000)	0.1801	35960
Hoeje-Taastrup	0.1184 (0.000)	-0.0013 (0.000)	-0.7218 (0.000)	10.1158 (0.000)	0.1919	41393
Roedovre	0.1160 (0.000)	-0.0013 (0.000)	-0.6916 (0.000)	10.0640 (0.000)	0.1855	19878
Ishoej	0.1164 (0.000)	-0.0013 (0.000)	-0.7522 (0.000)	10.0493 (0.000)	0.1774	11720
Taarnby	0.1456 (0.000)	-0.0016 (0.000)	-0.7374 (0.000)	9.6359 (0.000)	0.2093	47816
Vallensbaek	0.1315 (0.000)	-0.0014 (0.000)	-0.7503 (0.000)	9.7778 (0.000)	0.2039	4907
Rest Of Zealand	0.1486 (0.000)	-0.0016 (0.000)	-0.6826 (0.000)	9.3810 (0.000)	0.2174	76475

Note: Standard errors in parentheses. $\mathbb{I}_{rw_{t-1}=\varnothing}$ means unemployed in $t-1$.

Table C2: Estimates from OLS of Log Real Wages for Medium-Skilled Workers by Region

Work Region	age	age^2	$\mathbb{I}_{rw_{t-1}=\varnothing}$	$Constant$	R^2	N
Copenhagen	0.1141 (0.000)	-0.0013 (0.000)	-0.7975 (0.000)	10.3419 (0.000)	0.1654	524053
Frederiksberg	0.1157 (0.000)	-0.0013 (0.000)	-0.8044 (0.000)	10.2659 (0.000)	0.1784	58739
Ballerup	0.1062 (0.000)	-0.0012 (0.000)	-0.7160 (0.000)	10.5949 (0.000)	0.1504	90715
Broendby	0.0961 (0.000)	-0.0011 (0.000)	-0.7564 (0.000)	10.8255 (0.000)	0.1517	53930
Dragoer	0.0925 (0.000)	-0.0011 (0.000)	-0.7435 (0.000)	10.7246 (0.000)	0.1376	5442
Gentofte	0.1021 (0.000)	-0.0012 (0.000)	-0.7419 (0.000)	10.6416 (0.000)	0.1561	59172
Gladsaxe	0.1247 (0.000)	-0.0014 (0.000)	-0.7785 (0.000)	10.2222 (0.000)	0.2258	69497
Glostrup	0.0985 (0.000)	-0.0011 (0.000)	-0.7235 (0.000)	10.7475 (0.000)	0.1556	45200
Herlev	0.0927 (0.000)	-0.0010 (0.000)	-0.6233 (0.000)	10.8365 (0.000)	0.1359	38145
Albertslund	0.0862 (0.000)	-0.0010 (0.000)	-0.7069 (0.000)	11.0337 (0.000)	0.1417	50102
Hvidovre	0.0894 (0.000)	-0.0010 (0.000)	-0.6742 (0.000)	10.8940 (0.000)	0.1324	56422
Hoeje-Taastrup	0.0905 (0.000)	-0.0010 (0.000)	-0.7153 (0.000)	10.8998 (0.000)	0.1334	71102
Roedovre	0.0911 (0.000)	-0.0010 (0.000)	-0.7055 (0.000)	10.8677 (0.000)	0.1451	36186
Ishoej	0.0979 (0.000)	-0.0011 (0.000)	-0.6555 (0.000)	10.7029 (0.000)	0.1548	17606
Taarnby	0.0900 (0.000)	-0.0010 (0.000)	-0.6377 (0.000)	10.9085 (0.000)	0.1303	57936
Vallensbaek	0.1047 (0.000)	-0.0012 (0.000)	-0.7684 (0.000)	10.5610 (0.000)	0.1578	8112
Rest Of Zealand	0.0974 (0.000)	-0.0011 (0.000)	-0.6999 (0.000)	10.7593 (0.000)	0.1682	102252

Note: Standard errors in parentheses. $\mathbb{I}_{rw_{t-1}=\varnothing}$ means unemployed in $t-1$.

Table C3: Estimates from OLS of Log Real Wages for High-Skilled Workers by Region

Work Region	age	age^2	$\mathbb{I}_{rw_{t-1}=\varnothing}$	$Constant$	R^2	N
Copenhagen	0.1687 (0.000)	-0.0018 (0.000)	-0.6917 (0.000)	9.1518 (0.000)	0.2132	657976
Frederiksberg	0.1634 (0.000)	-0.0017 (0.000)	-0.7167 (0.000)	9.1166 (0.000)	0.2159	71351
Ballerup	0.1444 (0.000)	-0.0015 (0.000)	-0.6027 (0.000)	9.8524 (0.000)	0.1779	62758
Broendby	0.1444 (0.000)	-0.0015 (0.000)	-0.6122 (0.000)	9.7670 (0.000)	0.1700	27438
Dragoer	0.1398 (0.000)	-0.0015 (0.000)	-0.7592 (0.000)	9.6609 (0.000)	0.1923	3559
Gentofte	0.1401 (0.000)	-0.0015 (0.000)	-0.7552 (0.000)	9.8928 (0.000)	0.1556	77232
Gladsaxe	0.1504 (0.000)	-0.0016 (0.000)	-0.6465 (0.000)	9.7113 (0.000)	0.1837	64861
Glostrup	0.1264 (0.000)	-0.0013 (0.000)	-0.6007 (0.000)	10.1571 (0.000)	0.1662	33987
Herlev	0.1094 (0.000)	-0.0011 (0.000)	-0.6071 (0.000)	10.4400 (0.000)	0.1397	31325
Albertslund	0.1337 (0.000)	-0.0014 (0.000)	-0.7433 (0.000)	10.0164 (0.000)	0.1529	21939
Hvidovre	0.1121 (0.000)	-0.0012 (0.000)	-0.6212 (0.000)	10.3636 (0.000)	0.1456	35407
Hoeje-Taastrup	0.1411 (0.000)	-0.0015 (0.000)	-0.6668 (0.000)	9.8275 (0.000)	0.1724	32240
Roedovre	0.1186 (0.000)	-0.0012 (0.000)	-0.6439 (0.000)	10.1954 (0.000)	0.1365	15335
Ishoej	0.1238 (0.000)	-0.0013 (0.000)	-0.5989 (0.000)	10.0317 (0.000)	0.1495	8132
Taarnby	0.1382 (0.000)	-0.0014 (0.000)	-0.7239 (0.000)	9.7441 (0.000)	0.1756	19408
Vallensbaek	0.1295 (0.000)	-0.0014 (0.000)	-0.6547 (0.000)	10.0063 (0.000)	0.1642	4764
Rest Of Zealand	0.1467 (0.000)	-0.0015 (0.000)	-0.5958 (0.000)	9.6913 (0.000)	0.2091	131409

Note: Standard errors in parentheses. $\mathbb{I}_{rw_{t-1}=\varnothing}$ means unemployed in $t-1$.

D Structural Estimates

Table D1: Regional Amenities

	Coeff. Estimates	Standard Error	Z-statistic
α_{rl} (1)	0.0153	0.00051	30.1
α_{rl} (2)	-0.9733	0.00145	-673.3
α_{rl} (3)	-1.2263	0.00178	-690.8
α_{rl} (4)	-0.6359	0.00268	-237.3
α_{rl} (5)	0.7848	0.00134	583.6
α_{rl} (6)	-0.2120	0.00085	-249.2
α_{rl} (7)	-1.0813	0.00196	-550.6
α_{rl} (8)	-0.8991	0.00178	-504.3
α_{rl} (9)	-1.5117	0.00217	-695.8
α_{rl} (10)	-0.8425	0.00128	-657.0
α_{rl} (11)	-1.5901	0.00196	-811.1
α_{rl} (12)	-0.7930	0.00138	-576.3
α_{rl} (13)	-1.7885	0.00252	-708.8
α_{rl} (14)	-0.7490	0.00136	-551.9
α_{rl} (15)	-1.4207	0.00249	-571.4
α_{rl} (16)	-1.1823	0.00159	-743.5

E Counterfactual

Table E1: Share of individuals in each region in baseline $t = 0$ (pct.)

	Baseline: wl	Baseline: rl
Copenhagen	20.1	19.5
Frederiksberg	2.3	4.1
Ballerup	2.9	1.7
Broendby	1.8	1.3
Dragoer	0.4	0.6
Gentofte	2.4	4.1
Gladsaxe	2.5	2.5
Glostrup	1.7	0.7
Herlev	1.4	0.9
Albertslund	1.6	0.9
Hvidovre	1.9	1.8
Hoeje-Taastrup	2.2	1.8
Roedovre	1.2	1.3
Ishoej	0.8	0.7
Taarnby	1.9	1.5
Vallensbaek	0.6	0.6
RestOfZealand	28.1	55.9
Non-employment	26.2	-

F Description of the Data Sources

This section provides details of how the sample we use is constructed from individual Danish registers.

F.1 Individual background characteristics

The population register BEF is posted on January 1st each year and lists all individuals who have their officially registered address in Denmark. Each individual in the register is represented by an anonymized version of their official social security number called PNR. PNR is used to merge BEF with other registers with individual-specific data. From BEF we know the age and gender of the individual, whether she has children, how old the children are, and if she lives with a partner (married or not). UDDAUPD is informative of the highest completed education of the individual on a very detailed level down to the field of study. By using a table from Statistics Denmark that translates finer codes into broader categories we can reduce the number of education categories to the three categories we use in the estimation of the model: low (no more than high school), medium (vocational or short-length further education) and high education (bachelor degree or more). The education register is updated every October. To make sure observations from BEF and UDDAUPD are as close in time as possible, we merge UDDAUPD in year t on to BEF from year $t + 1$ via PNR.

F.2 Addresses and home moves

Importantly, BEF also contains an anonymized version of the individual's home address and an unmasked code for the home municipality, parish and other administrative geographic regions. In 2007, a municipality reform took place in Denmark which reduced the number of municipalities from 271 to 98 municipalities. This caused a change in the home addresses in the register (as they are only unique within a municipality), but we have used a key file from Statistics Denmark that translates old addresses to their new version post 2007. We therefore use the definition of municipalties from 2007.

For each individual we also know when they moved into the address they are currently at. Since our model is formulated at an annual basis we define the individual to live at the address where she lived during most of the year. If she moves to a new address in e.g. May and stays there for the rest of the year, this end-of-period address will be her home region choice, but if she did not move until August, we would record her beginning-of-period address as her home choice.

F.3 Labor market information

Data on workplace and other workplace-related variables such as industry and occupation come from the Integrated Database for Labor Market Research (IDA). IDA consists of different panels: one for personal data on employees (IDAP), another one for employments (IDAN) and one for workplaces (IDAS). We mainly use IDAN which has a record for every combination of individual, employment and year. The information about an individual's employment in IDAN comes from the Central Tax Information Sheet Register (Centrale Oplysningsregister) until 2008 and from eIncome (also located at the tax authorities) for the remaining period.

An individual can have several employments during a year. The register is posted by the end of November each year and groups individuals' employments into either employed wage-earners, employer (A), self-employed (S) or co-working spouse (M). All groups are mutually exclusive. The group of wage-earners is then further divided into main occupation (H), sideline occupation or another November occupation or most important non-November occupation (the two latter categories only available from 2004 onwards). For each individual who has more than one wage-earner job in November, we use information from the main occupation (H) which is determined by the largest source of income. Individuals who do not work are either classified as unemployed or outside the labor force. We classify individuals as unemployed if they according to the register are coded as being on leave (including parental leave and sick leave), unemployed by the end of November, participating in unemployment activation (short-term jobs financially supported by the public sector) or on rehabilitation. This information comes from IDAP. We define people to be outside the labor force whenever an individual is not unemployed and is recorded as being outside the labor force, studying, retiree, early retiree, on transitional allowance or on social security benefits. IDAP also has a variable showing for how many days the person has been registered as (un)employed during the year and we use this as the individual's (un)employment rate.

Each workplace has an anonymized version of its address which is recorded in IDAN. This anonymous address can be linked to non-masked municipalities, parishes or traffic zones (LTM zones). In some cases employments in IDAN cannot be assigned to a registered workplace. Instead, Statistics Denmark assigns the home address as the workplace (a so-called fictious workplace). This is typically the case for employees who work from home or at several workplaces, e.g. cleaners or community nurses.

To model characteristics of the work regions, we compute a job density measure for each region. We define this measure as the number of jobs within three levels of education and normalize by the corresponding level in Copenhagen. The number of employees by region is available from IDAN and after merging with BEF we know the education group too.

Since IDAN and IDAP are posted in November, we merge IDAN and IDAP year t on to BEF year $t+1$ via PNR.

F.4 Income

Data on income is available from INDUPD for a given year but not split on different employments within the year[14]. We are able to distinguish between total income, wage income and transfer income before and after taxes though and use income measures before tax. For people whom we classified as working in November, we use their annual wage income divided by their employment rate in the year according to IDAP. For people who are unemployed in November we use their transfer income divided by their unemployment rate and for those outside the labor force we use their total income.

F.5 Commute time

Commute time data come from The Danish Traffic Model (LTM) which has been developed by researchers at The Technical University of Denmark (DTU). They have divided Denmark into 907 traffic zones (LTM zones) and modelled commute time between each pair of regions. They model commute time for different transport modes (car, public or walk/bike) and exploit information on the road network, speed limits, congestion, bus and train timetables including waiting times, and bike paths. The traffic model has been run for 2002 and 2010 using the road network and public transport schedules for each year. Since our model is formulated in terms of municipalties, not LTM zones, we compute a commute time measure by each transport mode between any pair of LTM zones within a municipality pair. For a given pair of LTM zones in a municipality pair, we use the commute time from the mode with the shortest commute time. We then weigh the commute time of each observed LTM pair in the municipality pair with its estimated number of trips by that mode from the traffic model and thereby get a trip-weigted average commute time between any pair of municipalities. The difference between the 2002 and 2010 simulations of the traffic model is due to changes in the road network or bus and train plans. Commute time by walking or biking is constant.

From LTM we also get data on travel distances between each zone. We do the same exercise for distance as we did for travel time to get an average distance measure between each pair of municipalities.

F.6 Property prices, home ownership and home characteristics

Information on property prices come from the sales transaction register EJSA. EJSA holds a record of each sale in Denmark including sales price, type of sale (e.g. single-family house or commercial), number of square meters sold, and type of post-sale ownership (e.g. private or business). We deflate all sales prices by the consumer price index with 2011 as the reference year. We use only private sales and disregard properties with commercial-only purpose. On top of that we clean the data on sales prices using the same criteria that

[14]There exists another register BFL which has wage income for each combination of employment, individual and month, but only since 2008.

Statistics Denmark uses for official statistics on property prices, i.e. property value must exceed the lot value[15], the property must not have been sold more than once on the same day and is sold on open market terms.

Data on home ownership come from the EJER register. It links every housing unit in Denmark with a PNR of the owner and define an owner as someone who owns more than zero percent of the property. In order to link EJER and EJSA, we exploit the unique housing unit identifier which is available in both registers. This enables us to merge EJSA with our household panel via PNR. Since EJSA is posted on January 1st and EJER on October 1st each year we merge EJSA year t with EJER year $t-1$ and then merge EJSA year t with the household panel of year t.

For home characteristics we use BOL which is based on BBR (The Central Register for Buildings and Dwellings). It is a register with a record of each property in Denmark and gives information on characteristics such as the number of rooms, bathrooms and most importantly square meter for living space. BOL in year t can be merged to the houseold panel in year t via the housing unit identifier.

[15]Property and land value measures come from the register EJVK.

Bibliography

Ahlfeldt, G. M. and D. P. McMillen (2015). The vertical city: The price of land and the height of buildings in chicago 1870-2010. SERC discussion papers, London School of Economics.

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

Albouy, D. and B. Stuart (2019). Urban Population and Amenities: The Neoclassical Model of Location. *Forthcoming International Economic Review*.

Alonso, W. et al. (1964). *Location and Land Use. toward a General Theory of Land Rent.* Publications of the Joint Center for Urban Studies of the Massachusetts Institute of Technology and Harvard University. Cambridge: Harvard University Press.

Attanasio, O. P., R. Bottazzi, H. W. Low, L. Nesheim, and M. Wakefield (2012). Modelling the demand for housing over the life cycle. *Review of Economic Dynamics 15*(1), 1–18.

Bayer, P., N. Keohane, and C. Timmins (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management 58*(1), 1–14.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica 84*(3), 893–942.

Bishop, K. C. (2012). A dynamic model of location choice and hedonic valuation. *Unpublished, Washington University in St. Louis*.

Borjas, G. J. (2000). The economic progress of immigrants. In *Issues in the Economics of Immigration*, pp. 15–49. National Bureau of Economic Research Conference Report. Chicago: University of Chicago Press.

Buchinsky, M., C. Gotlibovski, and O. Lifshitz (2014). Residential location, work location, and labor market outcomes of immigrants in israel. *Econometrica 82*(3), 995–1054.

Dahl, G. B. (2002). Mobility and the return to education: Testing a roy model with multiple markets. *Econometrica 70*(6), 2367–2420.

Favilukis, J., P. Mabille, and S. Van Nieuwerburgh (2019). Affordable housing and city welfare. Technical report, National Bureau of Economic Research.

Gallin, J. H. (2004). Net migration and state labor market dynamics. *Journal of Labor Economics 22*(1), 1–21.

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

Guglielminetti, E., R. Lalive, P. Ruh, and E. Wasmer (2017). Home sweet home? job search with commuting and unemployment insurance. Working paper.

Halket, J., L. Nesheim, and F. Oswald (2015). The housing stock, housing prices, and user costs: The roles of location, structure and unobserved quality. Technical report, Cemmap working paper, Centre for Microdata Methods and Practice.

Hicks, J. (1932). The theory of wages. *The Economic Journal 43*(171), 460–472.

Holt, F. M. (1997). *Family Migration Decisions: A Dynamic Analysis*. Ph. D. thesis, University of Virginia.

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

Kuminoff, N. V., K. V. Smith, and C. Timmins (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature 51*(4), 1007–62.

Mangum, K. (2015). Cities and labor market dynamics. *Andrew Young School of Policy Studies Research Paper Series* (15-05).

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

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics 3*(4), 303–328.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record* (673).

McMillen, D. P. (2006). Testing for monocentricity. *A Companion to Urban Economics. Oxford: Blackwell*, 128–140.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review 57*(2), 197–210.

Muth, R. F. (1969). *Cities and Housing: the Spatial Pattern of Urban Residential Land Use*. Studies in business & society / Univ. of Chicago. Graduate School of Business. Chicago University Press.

Nathanson, C. G. (2019). Trickle-down housing economics. Working paper.

Ogawa, H. and M. Fujita (1980). Equilibrium land use patterns in a nonmonocentric city. *Journal of regional science 20*(4), 455–475.

Oswald, F. (2019). The effect of homeownership on the option value of regional migration. *forthcoming Quantitative Economics*.

Ransom, T. (2016). The effect of business cycle fluctuations on migration decisions. *Available at SSRN 2741117*.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy 90*(6), 1257–1278.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of political economy 82*(1), 34–55.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, 999–1033.

Rust, J. (1988). Maximum likelihood estimation of discrete control processes. *SIAM Journal on Control and Optimization 26*(5), 1006–1024.

Schmutz, B. and M. Sidibé (2015). Job search and migration in a system of cities. *Economic Research Initiatives at Duke (ERID) Working Paper* (181).

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of political Economy 70*(5, Part 2), 80–93.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of political economy 64*(5), 416–424.

Topel, R. H. (1986). Local labor markets. *Journal of Political economy 94*(3, Part 2), S111–S143.

Tsivanidis, N. (2019). Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio. Working paper.

Tunali, I. (2000). Rationality of migration. *International Economic Review 41*(4), 893–920.

Chapter 3

A Structural Model of Couples' Joint Home and Work Decisions and the Intra-Household Allocation of Commuting

Maria Juul Hansen, University of Copenhagen

Abstract

When analyzing locations of job and residence as well as commuting, it is important to consider the fact that many households consist of two workers as these families face a co-location trade-off: they live in the same location but may choose to work in different places. Until now, the literature on location choice has mainly modeled households as single-person decision makers. Using high-quality Danish administrative data, I provide descriptive evidence that couples and singles do differ in terms of where they live, work and how far to commute and that the commuting patterns differ geographically. In particular, I investigate how intra-household differences in commute distance between men and women may affect the gender wage gap which also exhibits spatial variation. To do this I combine the literature on dynamic residential-work location choice and the collective model literature to build a collective dynamic model. Due to the computational complexity, I currently estimate a static version of the model for the Greater Copenhagen Area (GCA) and find that most households have an equal split of the bargaining weights. Hence, for couples in the GCA discrimination of women within the household cannot explain that women commute to less favorable jobs and as a result earn less on average. In a counterfactual where the number of jobs is increased in a region outside the most urbanized area, location decisions on all margins are affected and the intra-household difference in commute times and the gender wage gap are slightly reduced.

1 Introduction

When analyzing location of job and residence as well as commuting, it is important to consider that the dual-earner household has been representative for many couple households in the western part of the world (70 per cent of households in Denmark) for decades. When such households decide where to live they face a co-location problem where they choose a common home taking both spouses' work locations into account. This complicates the problem of minimizing commute time for each individual. Nevertheless, this issue has not been accounted for in the literature on home and work location choices. It has therefore not yet been possible to document how the intra-household decisions on allocation of commuting between spouses affect their earnings or career prospects in general. People do not only abstain from moving closer to better job opportunities due to moving costs and the trade-off between preferences for location-specific amenities and house prices, but also because their partner potentially will not agree to move. This calls for explicitly modeling preferences of each household member to better understand immobility and commuting.

In this chapter I consider the trade-off that occurs when the decision on residential location is made together with job location of the spouses: namely that the two spouses enter a bargaining process to decide where they want to live together while considering the implied commute time and earned salary for both of them. I set up a theoretical model where I allow the spouses to be forward-looking about aspects that affect their future wage prospects and preferences for certain locations. This is done to incorporate the concerns spouses might have about whether taking a lower-paid job or becoming unemployed today[1] to support the optimal choice for one's partner will harm them in the future where they might get divorced and have to live off of their own income. Ultimately, I use the model to simulate effects on location choices and the intra-household allocation of commuting from a counterfactual policy that increases the number of jobs in currently less attractive job regions. This is a simplified version of the Better Balance policy implemented by the Danish government during the last few years, where almost 8,000 public sector jobs have been or will be relocated away from the Copenhagen capital area to more rural locations. It is questionable whether couple households actually do respond to a policy like this and are willing to move to the rural locations as the government aims at, since many couples do need two jobs after all. The potential effects on intra-household allocation of commuting and earnings also have not been given any attention in the public debate. Moreover, I investiage whether the co-location problem and allocation of commuting within dual-earner households can explain the gender wage gap, which is particularly driven by a gap between wages of couple women and couple men. I.e. whehter the gender wage gap is related to women having lower bargaining power and therefore agree to live in regions where they are unable to reach their optimal jobs within an endurable commute time.

[1]In the current version, only the latter effect is accounted for.

To answer these questions while acknowleding the co-location trade-off, I break with the unitary and individual modelling tradition by using a collective model of couples' dynamic decisions on home and job locations. I incorporate the collective perspective in a dynamic structural model by letting each partner have a bargaining power that is affected by their outside options and determines how much weight the household attaches to the two individual current and future value functions. The two-person household chooses joint home location and separate work locations each period taking into account commute time and wages for each spouse, moving costs, amenities of the home and job regions and the effects that the current decisions have on future job prospects and bargaining power.

For estimation I use high-quality Danish administrative data where I observe the entire population of households and its members for the period 1994-2012. The richness of the data is crucial for identification: it allows me to link information on the background of each individual in the population with its decisions on home and job location over time and importantly also link all individuals who belong to the same household. However, introducing bargaining into a dynamic model is non-trivial. Estimating the full dynamic model for couples is currently infeasible for computational reasons and is left for future work. I therefore focus the estimation on a static version of the model and only estimate the dynamic model for singles. Together with the large state space that the dynamic model eventually allows for make full-solution methods unsuitable for estimation. Instead, I use the Conditional Choice Probability (CCP) methods originally developed by Hotz and Miller (1993).

In the literature on household location decisions the agent has typically been a household representative that ignores the dual-earner framework and rather assumes households can be characterized by a single utiltiy function in a unitary model or in a model of individual decision makers. To the best of my knowledge, Gemici (2011) is the only paper that estimates a structural dynamic model of two-person households' location decisions, but she focuses on the home location choice only. Buchinsky et al. (2014) and Buchinsky et al. (2017) estimate a dynamic residential-work location model but do not take the collective perspective into account. These are all very important contributions to the literature, but by combining the (dynamic) residential-work location model with the collective model we improve our understanding of the behaviour of couple households and how they will respond to policies that aim at affecting mobility, location choices and commuting.

I first find descriptive evidence that couples commute further than singles and men more than women. The tendency to commute, the wage return to commuting, the division of commuting within the households and the gender wage gap differ geographically with lower commute time differences and gender wage gaps in the most urban areas. Using a subsample of the regions in Denmark, I estimate the structural static model. I find that doubling the job density in a region outside Copenhagen has quite large effects on location choices. Both home and work locations for both spouses are affected. The difference between the husband's and wife's commute time is lowered and the average gender wage gap slightly reduced as women's wages grow by 2.2 percent on average and men's 1.9 percent. Since the subsample focuses on the Copenhagen capital area it is not representative for the entire country though. The results should therefore be interpreted in this light. Including more rural locations where the intra-household commute time differences and gender wage gap are more substantial may change the results. This chapter is therefore considerd a first step towards estimating a full dynamic model for Denmark as a whole.

The rest of the chapter is organized as follows: Section 2 gives an overview of the existing literature on household decisions that take the two-person structure of the family into account. Section 3 shows descriptive statistics used to assess the determinants of location choices, commuting and wages. This is used to motivate the development of the structural location choice model in Section 4. Section 5 goes over the details of the estimation method and Section 6 presents the results. Section 7 concludes.

2 Related Literature

This section provides an overview of the literature on household location decisions. First, unitary and non-unitary models are briefly explained. Next, I go over static models from the literature and end with dynamic models.

2.1 Unitary and non-unitary models

Household decision studies have originally used unitary models (Samuelson (1956); Becker (1962, 1981)) where the analysis treats the household as a single individual with one utility function and it cannot identify individual utility functions. It thus cannot predict how an individual will react in response to changes in a certain policy, only how the household as a whole will. The model has been criticized for not capturing how the actual decision process in a multiple-member household takes place. Some of the first attempts to deal with this were Manser and Brown (1980) and McElroy and Horney (1981) who introduced the first non-unitary models (for the basic theory of these models, see Chiappori (1988, 1992); Browning and Chiappori (1998)). Unlike unitary models, non-unitary models deal directly with separate utility functions for each spouse and these models fall into two broad categories: cooperative non-unitary models and non-cooperative non-unitary models.

The seminal papers by Manser and Brown (1980) and McElroy and Horney (1981) both belong to the cooperative framework where the household behaves as if it is maximizing a weighted sum of the spouses' utility functions. The weights can depend on both prices, wages and distribution factors (variables that do not enter preferences or the budget constraint but affect the bargaining power, e.g. the income ratio of the two spouses). The bargaining process over outcomes between the spouses lead to Pareto efficient outcomes. This is in contrast to non-cooperative models, introduced by Lundberg and Pollak (1993), where Pareto inefficient outcomes may occur. I.e. the household could have chosen differently and made at least one spouse better off without making the other spouse worse off. In this framework the household decision process is modelled as a game where each spouse maximizes his or her own utility while taking the decision of the partner as given.

There is no clear answer to whether cooperative or non-cooperative models are better at explaining household behaviour. Udry (1996) rejects Pareto efficiency while Bobonis (2009) does not. The empirical tests for the unitary against the non-unitary models are carried out in studies analysing mainly consumption by testing whether income pooling holds (Lundberg et al. (1997); Attanasio and Lechene (2002), among others), i.e. household decisions do not depend on whom of the members receives the income. They tend to reject the unitary models.

Generally, it has been recognized that it is important to account for the fact that many decisions within the household result from multiple agents reaching an agreement. Ignoring this and instead regarding the household as if it were a single individual will likely lead to models with biased estimates since the model is trying to rationalize a decision process characterized by two individuals' utilities and not just one[2].

2.2 Static models

The literature on household location decisions that takes into account this two-person structure of the households dates back to Mincer (1978). He introduced the concept of a tied family member - the one who compromises on his or her individually optimal moving decision and rather moves or stays because the partner's gain from doing so outweighs his or her disadvantage such that the family as a whole gains from that decision. He refers to this as negative personal externalities which may or may not be internalized by the household unit. He also points out that with the, at the time, growing labor force participation of women, both the wife and the husband might become tied. The reason is that they both might gain from living (and in this case also working) somewhere else. The process underlying the location decisions of couples can therefore differ from that of singles.

Costa and Kahn (2000) takes up Mincer's tied mover and stayer concept. They use a reduced-form model to analyse whether power couples (couples with two college-educated spouses) are increasingly likely to locate in big metropolitan areas, as observed in U.S. data, because of the co-location problem they face. Namely, that due to both partners' specialized skills and their active labor market participation, they find the thicker labor market in the cities relatively more attractive than part-power couples (couples with only one college-educated spouse), singles and low-power couples (couples with no college-educated spouse) who only look for one or zero job matches for specialized skills. This explanation is compared to the explanations that power couples are over-represented in

[2]See Picard et al. (2013) for a discussion of non-unitary models in urban economics.

these urban areas because they can share the cost of the high rents, returns to education in large cities has risen compared to small cities or because urban amenities are normal goods. They conclude that the co-location problem indeed is the most likely explanation. Freedman and Kern (1997) reaches a similar conclusion, namely that when both the wife and the husband have a professional career, they are more likely to reside and work in big cities and that wives' differing earnings potentials across the U.S. affect the chosen home location of the couple and thereby also the husband's work location. Other studies focus more on the commute times of each person in the household, e.g. Sermons and Koppelman (2001), and generally find that household decisions are more sensitive to increases in the wife's commute time. Since the effects are more pronounced for families with children they conjecture that this is due to the wife having more household responsibilities.

Modelling the actual bargaining process in location choice models is nevertheless still in its infancy. Chiappori et al. (2014) is an exception and one of the first that do an attempt to account for the effect individual-specific characteristics of members in the households have on decisions, both through affecting the bargaining power and via the individual's preferences for certain alternatives. They estimate a static multinomial Logit model for residential areas in France conditional on each spouse's workplace and find that taking bargaining power into account is important for getting unbiased estimates of the value of time. Additionally, they find evidence that residential location choices are Pareto optimal. This speaks in favor of using collective models as concluded in Section 2.1.

2.3 Dynamic models

In addition to the above considerations of the structural differences in the decision-making process between single- and multiple-person households, a discussion of dynamic versus static models within this regime has taken place in recent years. Such dynamic aspects are important when the researcher wants to evaluate policies with intertemporal dimensions.

Intertemporal unitary models can be, and have traditionally been, used to study how households allocate scarce resources to different purposes (e.g. income to different goods, time to labor, household production and leisure) and the intertemporal allocations of these (Scholz et al. (2006); Krueger and Perri (2006)). Just as the static unitary model, the dynamic unitary model is not suitable for studying allocation of goods across spouses. Moreover, it assumes bargaining power is constant across households and time.

When it comes to intertemporal collective models there are generally two types: those with limited (LIC) and those with full commitment (FIC), cf. Chiappori and Mazzocco (2017) for a thorough review of these models. The latter are those where the households can commit to future allocations. I.e. households formulate a plan of optimal decisions at the beginning of marriage and stick to this plan no matter the shocks they might experience during the course of life. The LIC models are more complex and have higher data requirements but relax this assumption by requiring households to make efficient decisions subject to the participation constraint of both spouses; i.e. both spouses must be at least as good off in the marriage as if they take their best outside option, typically the value of a divorcee. It also allows households to renegotiate the plan of allocations in cases where the participation constraints are no longer fulfilled due to changes in the outside options. Formally, the FIC is the LIC without the participation constraints. By extending the static framework of the collective model in Chiappori (1988, 1992), Mazzocco (2007) provides empirical evidence that the LIC is favoured over FIC in a study of consumption and savings. Similar conclusions are reached in Aura (2005) and Lise and Yamada (2018).

Even though the dynamic collective framework has become more popular, it has not been used very much in the location choice literature. Lundberg and Pollak (2003) point out from a theoretical viewpoint that choice of residence for couples is indeed a collective decision where bargaining power matters. Also, they argue that these decisions need not be Pareto efficient since one spouse might veto a move due to the expectation that this will deteriorate his or her bargaining position. This could be due to lower earnings potential in the area they consider relocating to, which would lower the outside option. Households therefore might find themselves in a situation where they could Pareto optimize but do not do so as they cannot commit to staying in the marriage and share the household income according to the current sharing rule.

Gemici (2011) is so far the only paper that estimates a dynamic collective model of couples' job location decisions. She employs a symmetric Nash bargaining framework where couples choose consumption, employment location, employment choice and divorce each period taking into account their outside options. She concludes women are more likely to be the tied spouse, hence could obtain better labor market outcomes had they made their decision individually. However, by assuming transferable utility the household is able to compensate the tied spouse such that all decisions are efficient. One paper that explicitly allows for a choice of both home and work locations for couple households is Buchinsky et al. (2017). Their goal is to investigate the source of differences in labor market outcomes between genders using data on the location decisions of engineer immigrants from the former Soviet Union to Isreal. To do this they estimate dynamic models for each gender where the individual chooses only work location conditional on the partner's work location and residence and another model where both margins are in the choice set. They conclude that the constrained model is better at explaining women's behaviour and the model with both choice of home and work locations suitable for explaining men's behaviour. However, they do not model the joint decisions of the wife and husband, but rather treats everyone as individual decision makers.

3 Data and Descriptive Results

Using data on the entire population of Denmark for the period 1994-2012, this section covers descriptive statsitics on the Danish population. I show that both couples and singles are more reluctant to move when they are attached to the labor market and that couples commute further compared to singles of the same gender. According to the descriptives, this can be explained by their tendency to live in more rural locations than singles. In general, the allocation of commute times within the household is also related to their home location. Couples living in rural parts of the country thus have a higher difference in commute time on average. The spouse who commutes the furthest earns a higher wage than the spouse who works near the home. These stylized facts indicate 1) a simultaneity in the choice of home and job locations and wages earned and 2) that singles and couples differ in their choices on commuting distance. Below I go into details with these descriptive findings starting with a short description of the data sources.

3.1 Data sources

The data come from Statistics Denmark's administrative panel registers which can all be linked by either a personal identification number, household identification number, address or property identification number. By merging the registers, I get information on everyone living in Denmark on characteristics such as education, family situation, employer-employee relationship, income, home ownership and house characteristics as well as corresponding information on the spouse if not living as a single. As the focus of the chapter is on the joint home and work location decisions, the data has been restricted to adults between age 25 to 64 as younger people are more likely to still be studying and older people start to retire. Appendix B gives a more detailed overview of the different data sources and the sample selection.

3.2 Overall summary statistics and marital sorting

Table 3.1 shows summary statistics for the pooled population data consisting of almost 53 million observations of individuals[3]. Most importantly, the male is only 2.4 years older than the wife on average, couples are much less likely to be living in big cities[4], have fewer children and are less mobile. Table 3.2 shows the combination of educational degrees within a couple[5]. There is evidence of some degree of assortative mating in terms of education, though it not perfect. Among individuals with a long-length education there is quite strong sorting to a partner with the same length of education: 47.2 percent of these women have a partner in the same education group, while 31.2 percent of the men

[3]Not all variables exist for all individuals why the total is lower for some variables

[4]Big cities (municipalities) in Denmark are Copenhagen, Frederiksberg, Broendby, Gentofte, Gladsaxe, Glostrup, Herlev, Albertslund, Lyngby-Taarbaek, Roedovre, Vallensbaek, Greve and Aarhus municipality according to Eurostat's definition of densely populated areas that make up an urbanized area, cf. Figure A1 for a map of Denmark and its municipalities, Table A1 for an overview of geographical definitions in a Danish context, Figure A2 for a map of the main islands and provinces of Denmark. This way of dividing Denmark into geographical regions will be used in the coming sections. I will use "region" interchangeably for the different types of geographical definitions.

[5]Short-, medium and long-length refer to higher education after high school. Short-length higher education corresponds to 1.5-2.5 years of study after high school, medium-length higher education to 3-4 years and long-length higher education to 5-6 years.

do so. 35.0 percent of these men have a partner with medium-length higher education, while this only holds for 18.9 percent of the women.

Table 3.1: Summary statistics of population data

	Singles			Couples		
	Mean	S.d.	N	Mean	S.d.	N
Mean age of household	43.352	11.35	14,549,716	45.307	10.77	38,265,702
Male - female age	.	.	0	2.414	4.44	38,178,382
Education						
Unskilled	0.336	0.47	14,549,716	0.252	0.43	38,268,353
Vocational (VET)	0.342	0.47	14,549,716	0.410	0.49	38,268,353
High school	0.081	0.27	14,549,716	0.053	0.22	38,268,353
Short-length	0.038	0.19	14,549,716	0.044	0.20	38,268,353
Medium-length	0.144	0.35	14,549,716	0.169	0.37	38,268,353
Long-length	0.059	0.24	14,549,716	0.072	0.26	38,268,353
Family real income (10,000 DKK)	29.695	44.85	13,736,314	70.503	89.34	36,089,373
# children	0.278	0.66	14,538,400	1.029	1.07	38,141,572
I[home big city]	0.397	0.49	14,549,716	0.253	0.43	38,268,353
I[move home]	0.176	0.38	14,475,545	0.086	0.28	38,190,790
Years in home	7.583	9.95	14,549,716	11.707	12.21	38,268,353
I[household exists next year]	0.806	0.40	14,549,716	0.937	0.24	38,268,353

Note: I[home big city] = 1 if the home location is among the urbanized areas defined in Figure A1. I[move home] = 1 if the household moves address, i.e. within-region moves are included. I[household exists nxt year] = 0 if a couple household divorces, a single household becomes a couple household or the individual or one of the spouses dies.

Table 3.2: Combination of educational degree within couple

Wife	**Husband** Unskilled	Vocational	High school	Short-length	Medium-length	Long-length	Total
	%	%	%	%	%	%	%
Unskilled	13.9	12.8	1.0	0.8	1.3	0.5	30.3
Vocational	7.6	18.2	1.5	1.6	2.9	1.1	32.9
High school	1.6	3.4	1.8	0.5	1.3	0.8	9.6
Short-length	0.5	1.5	0.3	0.4	0.6	0.4	3.6
Medium-length	2.2	5.8	1.5	1.1	5.0	2.8	18.3
Long-length	0.3	0.8	0.5	0.2	1.0	2.5	5.3
Total	26.0	42.5	6.6	4.6	12.2	8.0	100.0

3.3 Gender wage gap

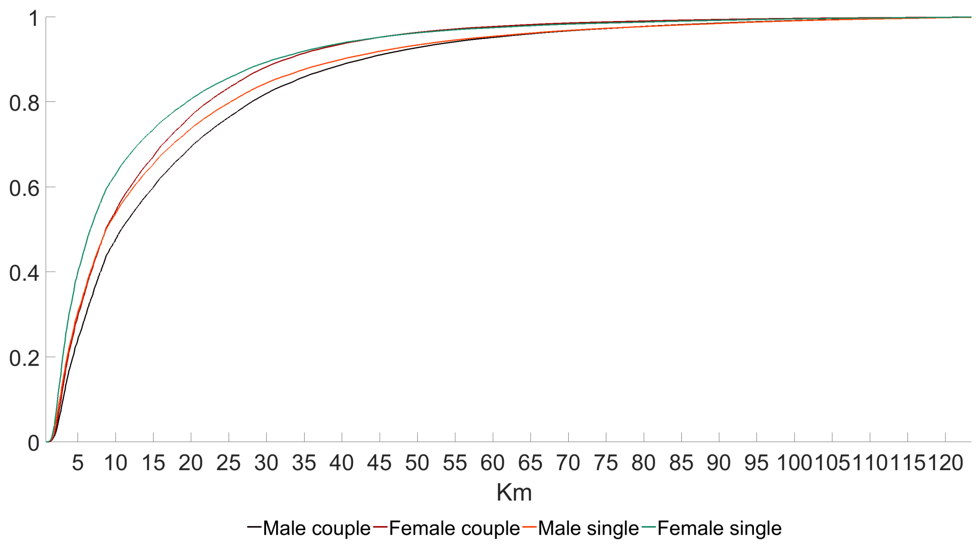
Having established that men and female in a couple tend to be quite similar in terms of age and education, it may be surprising that the men in couples earn 35.5 percent more than women in couples given controls for very detailed levels of education and age, cf.

column three of Table 3.3 which shows a regression of log annual wage income. Controlling for observable background characteristics essentially does not affect the raw gender wage gap from column 1. Looking at single individuals in column 4, the gender wage gap[6] is much smaller, namely 13.9 percent after controls for education and age. This difference is highly economically significant, so the research question of whether within-household decisions on home and job locations can explain this gap is also important in economic terms. I will return to the remaining specifications later[7].

3.4 Commuting and location choices

As a first look at commuting patterns, Figure 3.1 plots the cumulative probability distribution of commuting distance by gender and marital status. The figure reveals that men generally commute further than women, no matter the partnership status. Moreover, couples commute further than singles for both men and women. The figure as a whole is consistent with the hypothesis that couple households face a different trade-off in terms of choosing commute distance than singles do.

Figure 3.1: CDFs of commute distance (km) by marital status



Note: Commute distances come from the LTM model. Only including employed individuals.

[6]In this paper I refer to gender wage gap in annual wage income, not hourly wages.

[7]It should be noted that I do not account for neither gender differences in working hours, occupation nor industry which can affect wages. There is no clear answer to what the relevant measure of gender wage gap is and hence which controls to include. One should not control for factors that are considered a *result* of the gender wage gap. If for example women select into occupations that are lower-paid because they expect to earn a lower salary than similar males in the highest-paying occupations, and thus do not get sufficient compensation for the potentially heavier tasks carried out there, it is not obvious that occupation should be controlled for. By not controlling for occupation, I let gender differences in occupational choice translate into pay differences. The sensitivity of the results to inclusion of further controls will be explored in future work.

Table 3.3: OLS of log real annual wage income by marital status

	Model 1 Couples	Model 2 Singles	Model 3 Couples	Model 4 Singles	Model 5 Couples	Model 6 Singles	Model 7 Couples	Model 8 Singles	Model 9 Couples	Model 10 Singles
I[male]	0.344****	0.087****	0.355****	0.139****	0.354****	0.139****	0.354****	0.143****	0.235****	0.058****
	(0.0008)	(0.0016)	(0.0009)	(0.0019)	(0.0009)	(0.0019)	(0.0009)	(0.0019)	(0.0025)	(0.0033)
I[home big city]					0.034****	0.017****				
					(0.0008)	(0.0015)				
Constant	12.292****	12.180****	8.472****	9.335****	8.440****	9.318****	8.440****	9.352****	8.495****	9.389****
	(0.0005)	(0.0011)	(0.0103)	(0.0211)	(0.0103)	(0.0211)	(0.0103)	(0.0210)	(0.0103)	(0.0210)
Fixed effect controls										
Education	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Home region	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Home region × I[male]	No	No	No	No	No	No	No	No	Yes	Yes
N	29,499,306	9,680,302	29,499,306	9,680,302	29,499,306	9,680,302	29,499,302	9,680,299	29,499,302	9,680,299

Note: Standard errors clustered at the individual level in parentheses. sym* $p < 0.15$, ** $p < 0.10$, *** $p < 0.05$, **** $p < 0.01$ I[home big city] = 1 if the home location is among the urbanized areas defined in Figure A1 Education control is education at the finest level (length and field of study).

To get a more systematic insight into the reasons why couples have longer commute distances, Table 3.4 presents an OLS regression of commute distance on a number of observable demographics and individual fixed effects. The coefficients are therefore identified from time variation only. Model 1 shows that on average couples commute just slightly more than singles and controlling for differences in age, occupation and number of children do not change the coefficient on the couple dummy very much according to Model 2. Model 3 corresponds to Model 1 except fixed effects for home municipality have been added. This lowers the couple dummy coefficient. Model 4 adds back the controls for individual characteristics to the specification in Model 3, and in Model 5 the home municipality fixed effects have been replaced by a dummy for living in an urban area using the definition shown in Figure A1. The only coefficient that really stands out as economically significant is this dummy for living in an urban area. This suggests that the reason I see the distribution of commute distance for couples statistically dominating the ones for singles is that couples are less likely to live in the urban areas which are associated with shorter commute distances. Looking at Figure 3.2a there are indeed higher shares of couple households outside of the very Copenhagen center. This is also true for the remaining bigger cities in Denmark: Aarhus and Aalborg in Jutland and Odense on Funen.

To get a more detailed perspective on the effect on commute distance from living in various places in Denmark, Figure 3.3 plots the estimates of the home municipality fixed effects from Model 4 of Table 3.4. Copenhagen municipality is the reference group and is thus set to zero. Particularly on Zealand, there is a strong correlation between distance from the Copenhagen center and the average commute distance: the closer to Copenhagen, the shorter the commute distance. The degree of urbanization is therefore a very strong predictor of commute distance.

One thing is how the difference in commute distance between men, women, singles and couples look in the population and across the country, another is how the allocation of commute distance is *within* the household. Figure 3.4a shows the average of the maximum work distance within the household among the couples by their home location. The picture is very similar to Figure 3.3, i.e. higher maximum work distances the further from the most urbanized areas. Interestingly though, Figure 3.4b displays the average difference in work distance within the couple and again the picture is very similar to Figure 3.3. I.e. in locations where the average commute distance is high (among couples as well as singles), it is also very likely that one partner commutes considerably shorter than the other. I therefore conclude that couples who live in Copenhagen and pay the much higher prices for living there, cf. Figure 3.2b, both seem to exploit the access to the higher-paying jobs, cf. Figure 3.2d, that are also available within a short commute. A trade-off between prices and commute distance therefore appears relevant to account for when predicting where households decide to live and work.

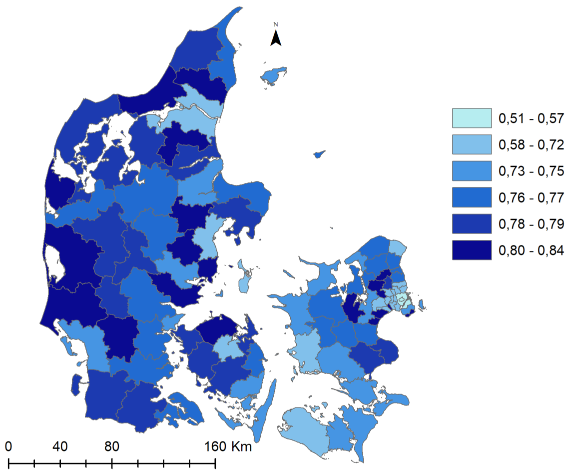
The households who live further away from Copenhagen are much more likely to divide

Table 3.4: OLS regression of commute distance (km) with individual fixed effects

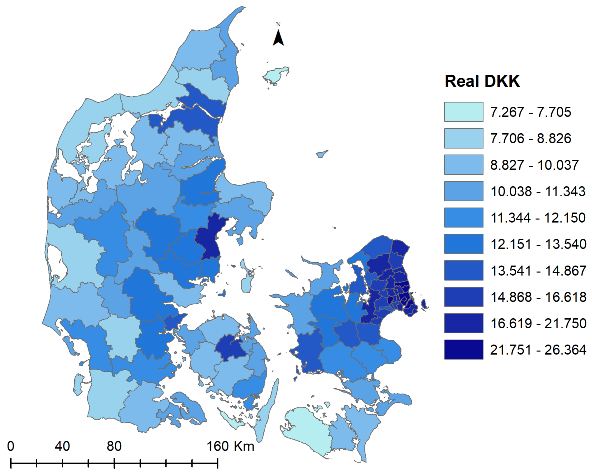
	Model 1	Model 2	Model 3	Model 4	Model 5
I[couple]	0.7***	1.0***	0.1***	0.6***	0.7***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Age		0.4***		0.3***	0.3***
		(0.0)		(0.0)	(0.0)
Age^2		-0.0***		0.0	0.0
		(0.0)		(0.0)	(0.0)
Occupation (ref. HS White)					
HS Blue		-2.3***		-2.4***	-2.3***
		(0.1)		(0.1)	(0.1)
LS Blue		-1.8***		-1.8***	-1.7***
		(0.0)		(0.0)	(0.0)
LS White		-1.9***		-1.9***	-1.8***
		(0.1)		(0.1)	(0.1)
Number of children (ref. 0)					
1		-0.6***		-0.9***	-0.9***
		(0.0)		(0.0)	(0.0)
2		-1.0***		-1.7***	-1.7***
		(0.0)		(0.0)	(0.0)
3 or more		-1.4***		-2.4***	-2.2***
		(0.1)		(0.1)	(0.1)
I[home big city]					-14.2***
					(0.1)
Constant	19.5***	4.4***	0.0	-9.2***	13.8***
	(0.0)	(0.3)	(0.2)	(0.3)	(0.3)
Home region	No	No	Yes	Yes	No
N	33,106,789	29,614,269	33,106,789	29,614,269	29,614,269

Note: Standard errors clustered at the individuals level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I[home big city] = 1 if the home location is among the urbanized areas defined in Figure A1. Occupations: HS Blue = high-skilled blue collar, LS Blue = low-skilled blue collar, HS White = high-skilled white collar, LS White = low-skilled white collar. The grouping of ISCO codes into the four groups follows the definitions from Eurofund: www.eurofound.europa.eu/surveys/ewcs/2005/classification.

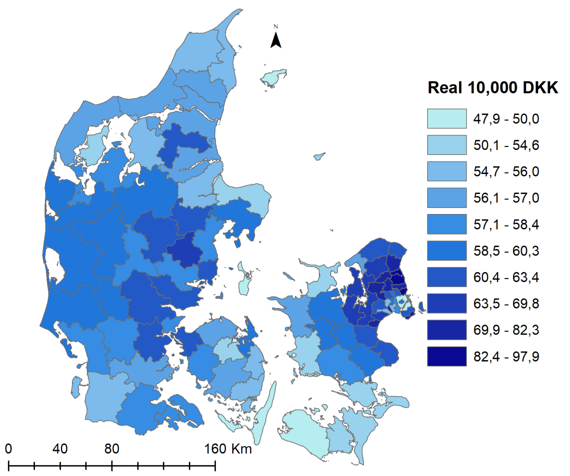
Figure 3.2: Share couple households, average prices per m^2, income and wages by region 1994-2012



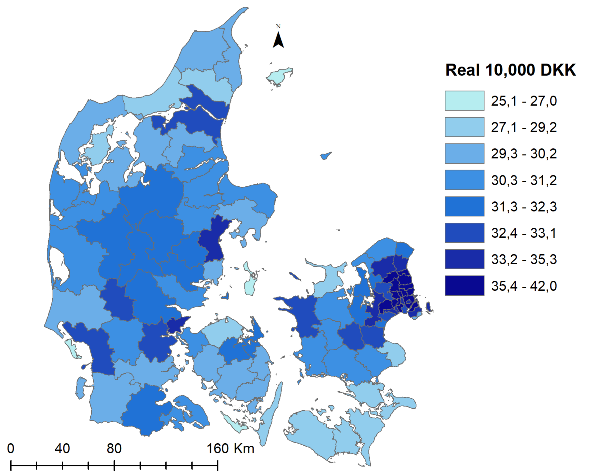
(a) Couple share by home region



(b) Sales prices per m^2 by home region



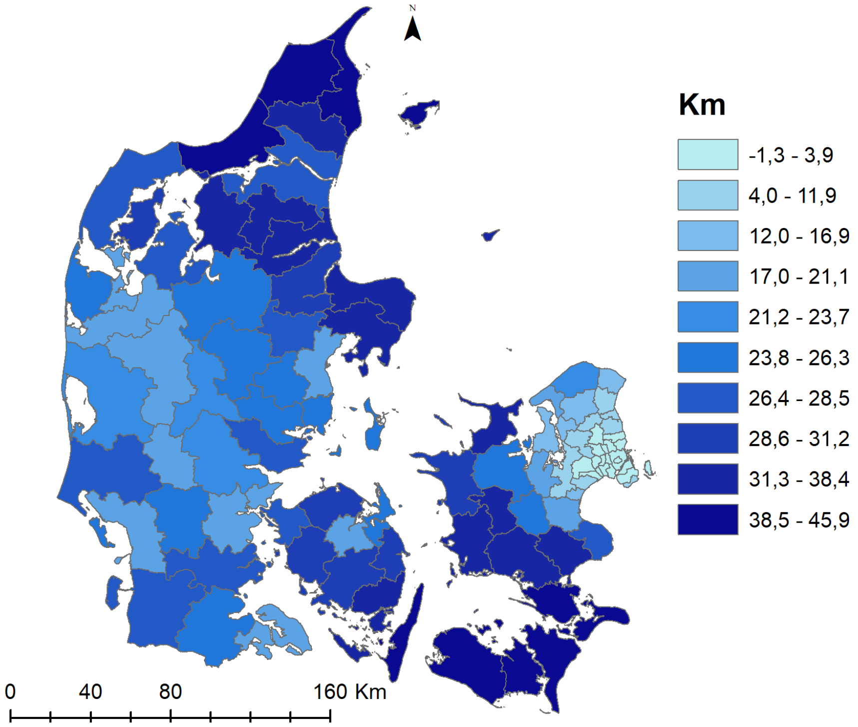
(c) Annual total household income by home region



(d) Annual individual wages by work region

Note: Intervals defined by Jenks natural breaks. Prices, income and wages measured in real 2011 DKK deflated by the consumer price index.

Figure 3.3: Home municipality fixed effects from OLS of commute distance

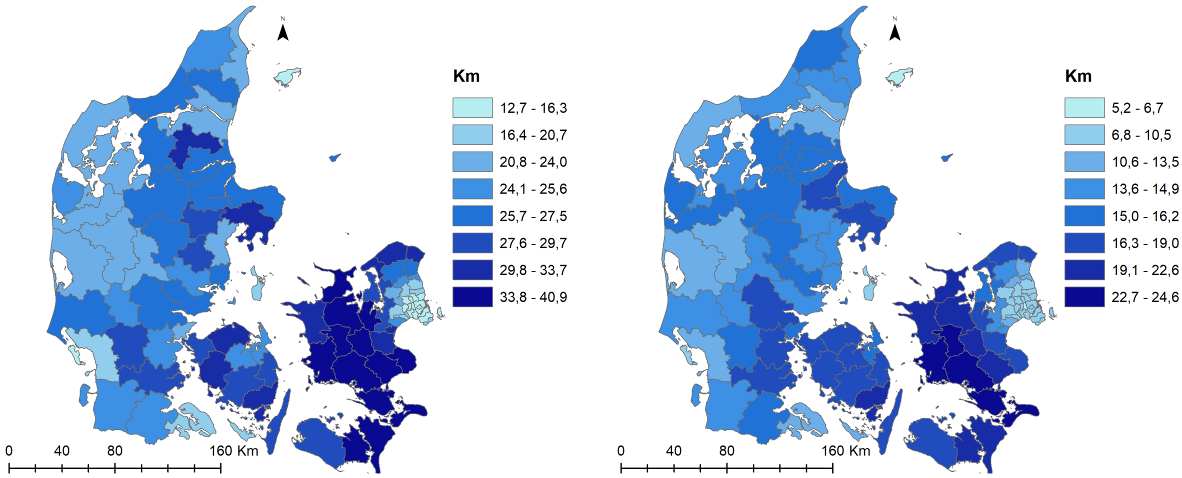


Note: Colors correspond to estimates of home municipality fixed effects in Model 4 of Table 3.4. Copenhagen municipality is the reference group. Km intervals defined by Jenks natural breaks.

the commute between the spouses such that one commutes short and the other long distances. For households living in the province West and South Zealand, the probability of working in the high-wage provinces Copenhagen City or Copenhagen Surroundings is indeed increasing in the commute distance for commute distances below 90 km. Thereafter it decreases slightly, cf. Figure 3.5. I.e. if a worker chooses to commute long, there is a tendency to commute to areas where the higher wages can potentially compensate for the longer commute. If long commutes are generally associated with higher earnings and men tend to be the ones who commute further within the couple which I established in Figure 3.1, this may be an important explanation for the reason women in couples earn much less on average all else equal.

To explore the latter question, Figure 3.6 plots the difference in commute distance between two spouses against the difference in annual wage income while taking account of differences in educational level between the individuals. The numbers are made such that a positive relationship corresponds to a positive wage return to being the one who commutes more and this return is not due to differences in educational attainment. Clearly, there is a positive relationship as soon as the difference in work distance exceeds a few kilometers. A person who commutes 10 km more to work earns almost 40,000 DKK more per year on average than his or her spouse. This is an economically significant difference and again supports the hypothesis that if the woman's better-paying jobs are generally (also)

Figure 3.4: Average maximum and difference in commute distance (km) within couple household by home region

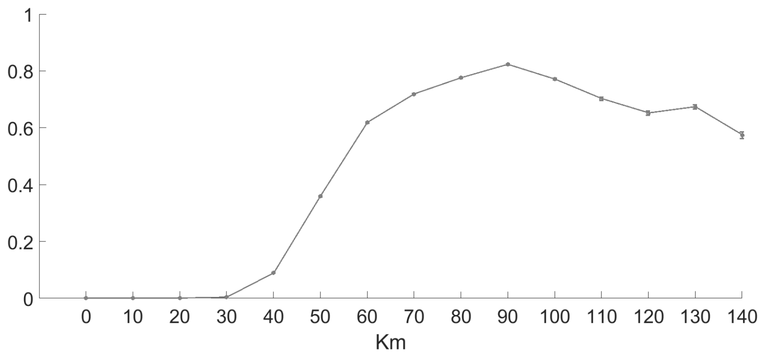


(a) Max

(b) Male - Female km

Note: Km intervals defined by Jenks natural breaks. .

Figure 3.5: Share commuting to Copenhagen City or Surroundings from West and South Zealand by commute distance

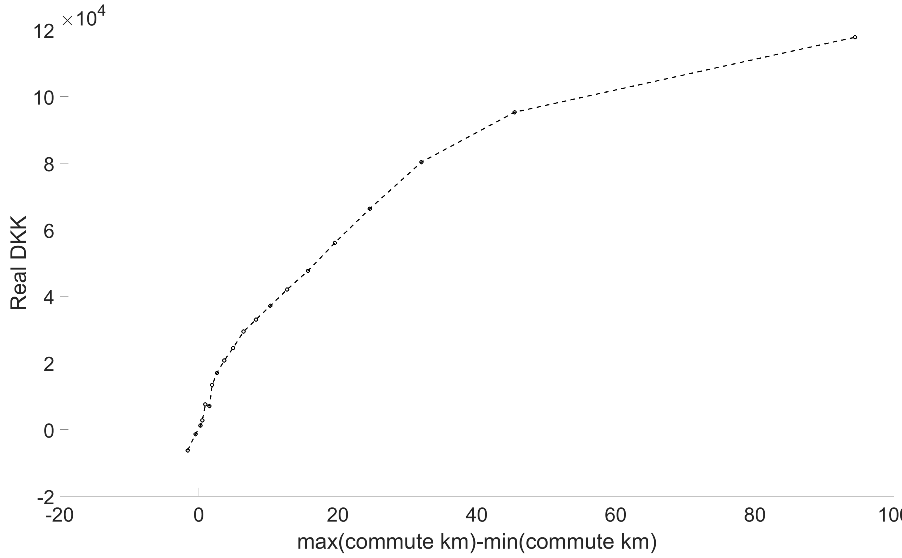


Note: The definition of provinces Copenhagen City, Copenhagen Surroundings and West and South Zealand are available in Figure A2.

located in the urban areas and if her career tends to be deprioritized in the household, this will cause women living in rural areas to earn less than their husbands on average.

Now that I have documented that the gender wage gap is much more pronounced among couple individuals, that the difference in work distance is higher the further away from the urban centres and that spouses who commute longer also earn better, it is informative to see if the gender wage gap is correlated with the difference in commute distance across the country too, i.e. not only within the household. Figure 3.7 therefore shows the estimates of the interaction between the dummy for male and home municipality fixed effects from the regression of log wages in Model 9 of Table 3.3 and thus uses variation in the difference between male and female wages within a home municipality (after controlling for age and education) to identify the spatial variation in the gender

Figure 3.6: Income difference by commute difference within couples



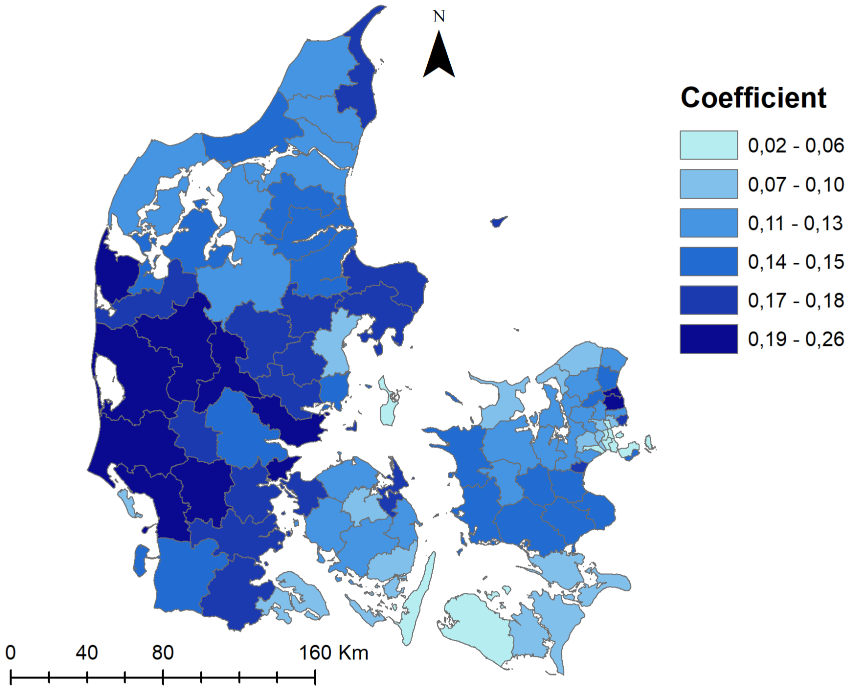
Note: Controls: difference in educational level between spouses. The binned scatter plot first regresses the y- and x-axis variables on the set of control variables and generates the residuals from those regressions. Then the residualized x-variables are grouped into equal-sized bins, whereafter the mean of the x-variable and y-variable residuals are computed within each bin. Then a scatterplot of these data points are made. A positive relationship means the spouse who commutes the longest also earns the most.

wage gap. As the figure shows there is a pattern similar to Figure 3.3 for the main part of Zealand, i.e. locations where the difference in work distance on average is high between two spouses have a higher average gender wage gap. This is consistent with the hypothesis that the allocation of commuting within the household can affect the gender wage gap if the woman is chosen to do the commute that is associated with a lower wage.

3.5 Spatial sorting

Generally, the high-price regions are also places where the average household income is high according to Figure 3.2c. Noticeable exceptions are the main cities in Denmark - Copenhagen and Odense in particular but also Aarhus and Aalborg - where the average household income is relatively low despite high prices per square meter. These were also locations with the lower couple shares in the country and since singles are expected to have a lower household income than couples this is not surprising. Nevertheless, this does not necessarily mean that households pay higher total prices for their home compared to other parts of the country since it is possible to adjust the number of square meters in response to the higher prices per square meter. The bigger cities have a much higher supply of apartments and the size of the homes are indeed also lower on average in the big cities, cf. Chapter 2. There does therefore seem to be not only a trade-off between lcoation and commuting distance, but also between location and house size. I will not

Figure 3.7: I[home region]·I[male] from OLS of log real wages



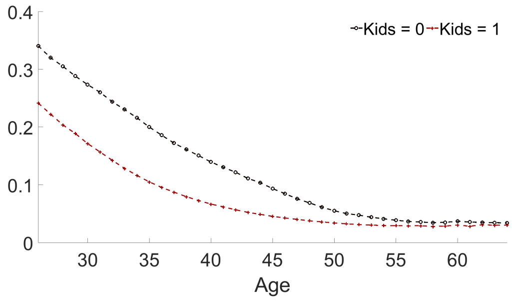
Note: Colors correspond to estimates of interactions between home municipality fixed effects and a dummy for being male in Model 9 of Table 3.3. Copenhagen municipality is the reference group. Intervals defined by Jenks natural breaks.

investigate the latter in the model of this chapter and rather consider house size as an exogenous amenity of the region, but the question was taken up in Chapter 2 where we estimated a model that incorporates endogenous square meter demand, home and work location choices of individuals in a dynamic setting.

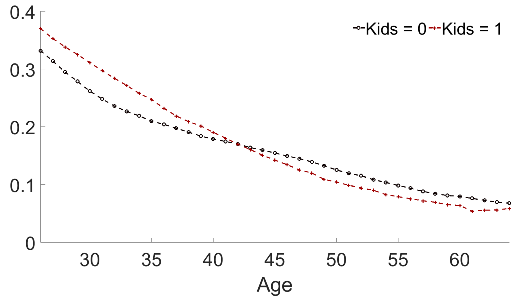
3.6 Home moving patterns

Lastly, I cover evidence on the decision to move home for singles and couples, respectively. As Figure 3.8 shows the probability of moving home address is monotonically declining in age. Conditioning on whether the individual has children or not makes a difference for the figure for couples, cf. Figure 3.8a. Those with children living at home are less likely to move until the age of 50 where most households no longer have children at home. This may indicate higher moving costs as soon as children are involved in the move, e.g. because the parents have to move the child to another school or daycare center. This can be incorporated into the structural model. The picture is a bit different for the singles according to Figure 3.8b: until age 43, the singles with children are slightly more mobile. After age 43 the ones with no children start to move a bit more often. However, one should keep in mind that singles with children are a special subgroup of the population and some of the moves are due to a couple with children splitting up and thereby forcing one of the spouses to move.

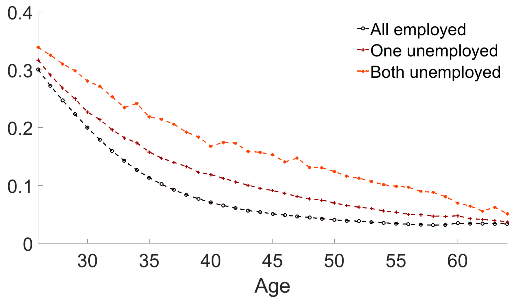
Figure 3.8: Moving probability by age



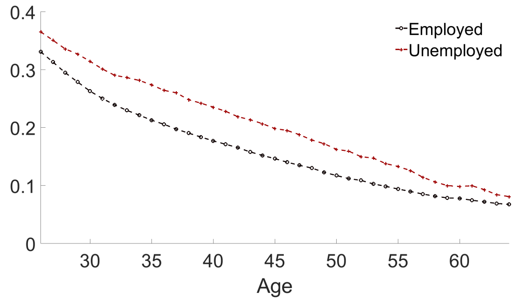
(a) Couples



(b) Singles



(c) Couples



(d) Singles

Note: Kids = 1 means having at least one child. Intra-regional moves included.

Looking at the lower panel of the figure, Figure 3.8c plots the probability of moving home address conditional on the number of working spouses in the household. This probability is much lower for couples where both have a job, especially compared to the case where both are unemployed. When that is the case, almost 30 percent of couples move home at age 30 and almost 20 percent at age 40. There are multiple reasons for moving when non-employed, one being that the household can no longer afford staying in the current home or they may move in order to live closer to more dense labor markets where chances for finding jobs can be higher. The structural model can incorporate both effects by allowing marginal utility of money to depend on the income (e.g. unemployed individuals with low income may have higher marginal utility of money and therefore look for cheaper housing) and by modelling the home and work location choice as a simultaneous decision, i.e. when choosing home the household also chooses how close to be to work locations that offer a better wage.

A reduced-form approach cannot separate all the above components that drive location choices. It is therefore essential to employ a structural model. Based on the above descriptive evidence, the next section outlines the model.

4 Model

This section develops a structural dynamic model of home and job location decisions for couples, where couples refer to households with two adults, no matter if they just live together as a couple or are married. I use "spouse" to refer to one's partner even if the couple is not legally married. First, I introduce some general notation of the model. Thereafter, I show how the decisions are made by the singles and couples according to the model[8].

4.1 Setting up the model

Each household consists of a wife and a husband. For clarity I let i denote the wife and j the husband. Each person $k \in \{i, j\}$ in household h gets flow utility $u^{d,M}_{k,t}$ in each period $t = \{t_0, ..., T\}$ if married and $u^{d,S}_{k,t}$ if single:

$$\begin{aligned} u^{d,M}_{k,t} &= u(d_{h,t}, \tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}) \text{ if married,} \\ u^{d,S}_{k,t} &= u(d_{k,t}, \tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}) \text{ if single,} \end{aligned} \tag{3.1}$$

where $d_{h,t}$ is the index of the decision made by household h at time t and holds the decisions on joint home location ($rh_{h,t} \in \mathfrak{D}^{rh} \equiv \{1, 2, ..., \bar{rh}\}$), wife's work location ($rw_{i,t} \in \mathfrak{D}^{rw} \equiv \{\varnothing, 1, 2, ..., \bar{rw}\}$), husband's work location ($rw_{j,t} \in \mathfrak{D}^{rw}$), where $rw_{k,t} = \varnothing$ denotes unemployment, and divorce $D_{h,t} \in \mathfrak{D}^D \equiv \{0, 1\}$. If the household chooses $D_{h,t} = 1$ it divorces in $t + 1$ ("time to build"). $d_{k,t}$ is the decision for singles and is given by home location $rh_{k,t} \in \mathfrak{D}^{rh}$ and $rw_{k,t} \in \mathfrak{D}^{rw}$. For now I do not take a stand on what the home and work locations are empirically, but they can be thought of as municipalities, parishes, provinces or commute zones for instance. Let the entire choice set for couples be $\mathfrak{D}^M \equiv \mathfrak{D}^D \times \mathfrak{D}^{rh} \times \mathfrak{D}^{rw} \times \mathfrak{D}^{rw}$ and for singles $\mathfrak{D}^S \equiv \mathfrak{D}^{rh} \times \mathfrak{D}^{rw}$.

$\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}} = (\boldsymbol{x}^{\boldsymbol{1}}_{\boldsymbol{k,t}}, \boldsymbol{x}^{\boldsymbol{2}}_{\boldsymbol{k,t}})$ is a subset of the state variables of individual k. It consists of observed state variables $\boldsymbol{x}^{\boldsymbol{1}}_{\boldsymbol{k,t}}$ and other observed state variables $\boldsymbol{x}^{\boldsymbol{2}}_{\boldsymbol{k,t}}$. $\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}} = (\boldsymbol{x}^{\boldsymbol{1}}_{\boldsymbol{i,t}} \setminus \{t_i\}, \boldsymbol{x}^{\boldsymbol{1}}_{\boldsymbol{j,t}}, \boldsymbol{x}_{\boldsymbol{h,t}})$ is a subset of the state variables of the couple household. It consists of observed state variables for either spouse ($\boldsymbol{x}^{\boldsymbol{1}}_{\boldsymbol{i,t}}$ and $\boldsymbol{x}^{\boldsymbol{1}}_{\boldsymbol{j,t}}$, except the wife's age t_i (explanation follows in section 4.4) and household-specific observed state variables ($\boldsymbol{x}_{\boldsymbol{h,t}}$)[9].

The specification in (3.1) allows for altruistic preferences or gains from marriage, i.e. that person k gets utility from choices that occur to the spouse when they are married. The exact content of $\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}$ and $\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}$ will be spelled out in section 4.4, but t indexes the age of the household which is given by the age of the husband when living in a couple and age of the individual when being single. $t = T + t_0 - 1$ is the maximum age of an

[8]In the empirical implementation the focus is on estimating a static version of the model. The restrictions are elaborated in Section 5.3

[9]In an earlier version of the chapter, unobserved moving and commute cost types were allowed for. Such models can be estimated using the methods in Arcidiacono and Miller (2011). However, it currently causes the estimation to be infeasible, so the extension with unobservable types is left for future research.

individual. To get the age of the wife in a couple I simply carry the age difference between the spouses as a constant state variable.

In addition to the deterministic flow utilities in (3.1), the household and individuals get an alternative-specific taste shock each period. For notational purpose, I introduce index a for agents, where agents can be either the couple household h or indvidual $k \in \{i, j\}$. The shocks are unique for agent a and independently and identically distributed over t, a and d according to the distribution f parametrized by $\boldsymbol{\theta_f}$:

$$\epsilon_{a,t}^{d} \sim f(\boldsymbol{\theta_f}),\ a \in \{h, k\} \tag{3.2}$$

where d refers to the choices described just above. This captures everything unobserved and not accounted for by the model that affects the location decsions. It can be interpreted as new information about the locations that is revealed to the agents each period. I assume $\boldsymbol{\epsilon_{a,t}}$ is multivariate IID Extreme Value Type I distributed and that $(\boldsymbol{\tilde{x}_{a,t}}, \boldsymbol{\epsilon_{a,t}})$ obeys a conditionally independent controlled Markov process with probability density

$$\begin{aligned} P(&\boldsymbol{\tilde{x}_{a,t+1}}, \boldsymbol{\epsilon_{a,t+1}} | d_{a,t}, \boldsymbol{\tilde{x}_{a,t}}, \boldsymbol{\epsilon_{a,t}}, \boldsymbol{\theta_f}, \boldsymbol{\theta_g})) \\ &= f(\boldsymbol{\epsilon_{a,t+1}} | \boldsymbol{\tilde{x}_{a,t+1}}, \boldsymbol{\theta_f}) g(\boldsymbol{\tilde{x}_{a,t+1}} | d_{h,t}, \boldsymbol{\tilde{x}_{a,t}}, \boldsymbol{\theta_g}), \end{aligned} \tag{3.3}$$

where $g(\cdot)$ is the p.d.f of $\boldsymbol{\tilde{x}_{a,t}}$ parametrized by $\boldsymbol{\theta_g}$.

4.2 1st stage: The single's planning problem

Let the single individual's state be $\boldsymbol{z_{k,t}} = (\boldsymbol{\tilde{x}_{k,t}}, \boldsymbol{\epsilon_{k,t}})$, where $\boldsymbol{\epsilon_{k,t}} = (\epsilon_{k,t}^{1}, \epsilon_{k,t}^{2}, ..., \epsilon_{k,t}^{\bar{d}})$ and $\bar{d}$ is the choice with highest choice index in the choice set $\mathfrak{D}^{S}$. A person entering period t as single optimizes with respect to $(rh_{k,t}, rw_{k,t})$ in each period until T under the assumption that he does not expect to find a new partner. I.e. transitions into marriage are considered completely random events and the model will not be used to shed light on whether singles tend to choose residential location with the thickness of the marriage market in mind. The Bellman equation for singles is

$$V_{k,t}^{S}(\boldsymbol{z_{k,t}}) = \max_{d_{k,t} \in \mathfrak{D}^{S}} \{u_{k,t}^{d,S} + \epsilon_{k,t}^{d} + \beta E[V_{k,t+1}^{S}(\boldsymbol{z_{k,t+1}}) | \boldsymbol{\tilde{x}_{k,t}}, d_{k,t}]\},\ k \in \{i, j\}. \tag{3.4}$$

β is the discount factor and $E[V_{k,t+1}^{S}(\boldsymbol{z_{k,t+1}}) | \boldsymbol{\tilde{x}_{k,t}}, d_{k,t}]$ is the conditionally expected value for a single who is in state $\boldsymbol{z_{k,t+1}}$ with the expectation taken over $\boldsymbol{z_{k,t+1}}$.

The Bellman equation can be rewritten to make the recursive structure explicit: define the alternative-specific value function for the individual as

$$v_{k,t+1}^{d,S}(\boldsymbol{\tilde{x}_{k,t+1}}) = u_{k,t+1}^{d,S} + \beta E[V_{k,t+2}^{S}(\boldsymbol{z_{k,t+2}} | \boldsymbol{\tilde{x}_{k,t+1}}, d_{k,t+1})]. \tag{3.5}$$

Exploiting the distributional assumption on the taste shocks and (3.3), the ex ante expected future value function defined as $\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}})$ is defined in the standard way:

$$\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}) = \sigma_\epsilon \log \left(\sum_{d_{k,t+1} \in \mathfrak{D}^S} \exp[v^{d,S}_{k,t+1}(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}})/\sigma_\epsilon] \right), \tag{3.6}$$

where $\sigma_\epsilon \in \boldsymbol{\theta_f}$ is the scale parameter of ϵ^d.

Using (3.6) and the assumption that $\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}$ follows the distribution g, (3.5) can be rewritten into

$$v^{d,S}_{k,t+1}(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}) = u^{d,S}_{k,t+1} + \beta \int_{\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}} \phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}) \cdot g(d\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}} | \tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}) \tag{3.7}$$

and (3.4) into

$$V^S_{k,t}(\boldsymbol{z}_{\boldsymbol{k,t}}) = \max_{d_{k,t} \in \mathfrak{D}^S} \{ v^{d,S}_{k,t}(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}) + \epsilon^d_{k,t})] \}, \tag{3.8}$$

which makes the recursive structure with respect to v explicit via (3.7). After starting at period T and solving the ex ante expected future value function (3.6) via backwards recursion[10] for all combinations of $\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}$ to period t, the individual's current age, let $\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}})$ be the solution to this problem. $\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}})$ is considered the expected outside option value at $t+1$ for the individual in a couple and is an unobserved counterfactual outcome. It can be found independetly from the second stage presented in the next section.

The alternative-specific shocks $\boldsymbol{\epsilon}_{\boldsymbol{k,t}}$ are observed for the agents but unobserved for the econometrician. This means the econometrician uses the conditional choice probability (CCP) when assessing how likely it is that $d_{k,t}$ is the optimal choice according to the model. It is given by

$$CCP(d_{k,t}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}) = \frac{\exp[v^{d,S}_{k,t}(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}})/\sigma_\epsilon]}{\sum_{m \in \mathfrak{D}^S} \exp[v^{m,S}_{k,t}(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}})/\sigma_\epsilon]} \tag{3.9}$$

by exploiting the distributional assumption of the taste shocks.

4.3 2nd stage: The couple's planning problem

The couple household maximizes the weighted sum of both spouses' utilities from a given choice $d \in \mathfrak{D}^M$ for each period t. t indexes the husband's age.

[10] Solving the model is not necessary for estimation, cf. Section 5. However, the full solution is still necessary for some types of counterfactuals, and this way of writing up the value function is useful for that purpose.

Bargaining weight

The weight on the utilitiy of each spouse in the household's utility function is denoted $W_{i,t}$ for wife i. I assume it is given by the following function of $(\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t}}), \phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t}}))$:

$$W_{i,t} = \frac{1}{1 + \exp(-(\Upsilon_0 + \Upsilon_1(\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t}}) - \phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t}}))))}, \tag{3.10}$$

where $\{\Upsilon_0, \Upsilon_1\}$ are parameters. This function ensures the weight is in the interval $(0, 1)$ and is increasing in $(\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t}}) - \phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t}}))$ for $\Upsilon_1 > 0$. Υ_0 measures the degree of discrimination towards the wife in a sitaution where $\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t}}) = \phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t}})$, i.e. where the wife's outside option equals her husband's. In that case, if $\Upsilon_0 = 0$ it implies $W_{i,t} = 0.5$ and if $\Upsilon_0 < 0$, the wife is being discriminated such that her bargaining weight is below 0.5. $W_{i,t}$ is thus increasing in Υ_0. The weights are endogenous since the outside options change in response to the state variables which are a function of the history of the decisions made in the household.

Divorce

I assume couples can decide to divorce and that the household behaves efficienctly in this respect, i.e. always makes the divorce decision that is optimal for the household as a whole when it has taken the bargaining weights into account. This means, the individuals as such do not themselves decide whether to stay in the marriage, but rather consider the benefits accruing to the households. This is done to make the optimization problem feasible[11].

Bellman equation

Considering $\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+1}})$ and $\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+1}})$ as given from the first stage, the household therefore solves for the value function given state $\boldsymbol{z_{h,t}} = (\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}, \boldsymbol{\epsilon_{h,t}})$. The Bellman equation is given by

$$\begin{aligned} V^M_{h,t}(\boldsymbol{z_{h,t}}) = \max_{d_{h,t} \in \mathfrak{D}^M} \{ & W_{i,t} u^{d,M}_{i,t} + (1 - W_{i,t}) u^{d,M}_{j,t} \\ & + \beta[(1 - D_{h,t}) E[V^M_{h,t+1}(\boldsymbol{z_{h,t+1}} | \tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}, d_{h,t})] \\ & + D_{h,t}[(0.5 \cdot E_{\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+1}}}[\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+1}}] + 0.5 \cdot E_{\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+1}}}[\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+1}})] - \Delta]] \\ & + \epsilon^d_{h,t} \} \end{aligned} \tag{3.11}$$

[11] Essentially, it is infeasible to solve the participation constraint as it requires computing the value or expected value of being in the marriage for an individual spouse. This cannot be done without having to simulate the Extreme Value taste shocks for each spouse and integrate them out. The closed-form solution for the expectation of the maximum extreme value shock cannot be applied as the individuals do not act as individually optimizing decisions makers as long as the marriage lasts.

where $E_{\tilde{\boldsymbol{x}}_{i,t+1}}[.]$ and $E_{\tilde{\boldsymbol{x}}_{j,t+1}}[.]$ denote the expectation over $\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+1}}$ and $\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+1}}$, respectively. As for the individuals, β is the household's discount factor and $E[V^M_{h,t+1}(.)]$ is the conditional expectation of next period's value function when staying married, where the expectation is with respect to the future taste shocks $\boldsymbol{\epsilon_{h,t+1}}$ and other state variables $\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}}$. Δ is a parameter representing the utility cost of divorce[12]. As seen from the second to last line in (3.11), I force the household to put equal weight on the wife's and husband's outside options when evaluating the divorce option. This is done to avoid that a spouse with a very high weight and high value of marriage can veto not divorcing just because the other spouse has a very low bargaining weight inside the marriage. In the specification above, a low bargaining weight does not affect the value of divorce directly.

The alternative-specific value function for the couple household is

$$\begin{aligned} v^{d,M}_{h,t+1}(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}}) = u^{d,M}_{h,t+1} & \\ & + \beta[(1 - D_{h,t+1})E[V^M_{h,t+2}(\boldsymbol{z_{ht+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}}, d_{h,t+1})] \\ & + D_{h,t+1}[(0.5 \cdot E_{\tilde{\boldsymbol{x}}_{i,t+2}}[\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+2}}] + 0.5 \cdot E_{\tilde{\boldsymbol{x}}_{j,t+2}}[\phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+2}})] - \Delta]] \end{aligned} \tag{3.12}$$

with $u^{d,M}_{h,t+1} = W_{i,t+1}u^{d,M}_{i,t+1} + (1 - W_{i,t+1})u^{d,M}_{j,t+1}$.

Along the same lines as for the singles, the ex ante expected future value function defined as $\phi^M(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}})$ can be written as

$$\phi^M(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}}) = \sigma_\epsilon \log\left(\sum_{d_{h,t+1}\in\mathfrak{D}^M} \exp[v^{d,M}_{h,t+1}(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}})/\sigma_\epsilon]\right). \tag{3.13}$$

Using this and the assumption that $\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}}$ follows the distribution g, (3.12) can be rewritten into

$$\begin{aligned} v^{d,M}_{h,t+1}(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}}) = u^{d,M}_{h,t+1} & \\ & + \beta[(1 - D_{h,t+1})\int_{\tilde{\boldsymbol{x}}_{\boldsymbol{ht+2}}} \phi^M(\tilde{\boldsymbol{x}}_{\boldsymbol{ht+2}}) \cdot g(d\tilde{\boldsymbol{x}}_{\boldsymbol{ht+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{h,t+1}}, d_{h,t+1}) \\ & + D_{h,t+1}(0.5 \cdot \int_{\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+2}}} \phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+2}})g(d\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{i,t+1}}, d_{h,t+1}) \\ & + 0.5 \cdot \int_{\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+2}}} \phi^S(\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+2}})g(d\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{j,t+1}}, d_{h,t+1}) - \Delta)] \end{aligned} \tag{3.14}$$

and (3.11) into

$$V_{h,t}(\boldsymbol{z_{h,t}}) = \max_{d_{h,t}\in\mathfrak{D}^M}\{v^{d,M}_{h,t}(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}) + \epsilon^d_{h,t})\}, \tag{3.15}$$

which makes the recursive structure with respect to v explicit via (3.14). After starting at period T and solving (3.13) backwards for all combinations of $\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}$ to period t, the

[12]Any gains from marriage that are not explicitly modelled are also soaked up by the divorce costs.

husband's current age, let $d^*_{h,t}$ be the solution to this problem. Hence, $d^*_{h,t}$ is the optimal decision given state $\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}$, which determines the bargaining power $W_{i,t}$ that the household then knows. This optimal solution takes into account how it will affect the future states of the household and how the optimal decisions may change as a consequence.

Like for the singles, the CCPs are given by

$$CCP(d_{h,t}|\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}) = \frac{\exp[v^{d,M}_{h,t}(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}})/\sigma_\epsilon]}{\sum_{m\in\mathfrak{D}^M}\exp[v^m_{h,t}(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}})/\sigma_\epsilon]}. \tag{3.16}$$

4.4 Utility specification

This section elucidates the exact content of the state variables and how the utility function looks for the individuals. To give a brief overview, the individual gets utility from income, local amenities in the home and work region, and disutility from work, housing costs, home and job moving costs and commute costs.

From a consumption-perspective and under the assumption that living together as a couple means sharing income to some extent at least, an individual would be expected to get utility from the spouse's income because it would imply a higher household income. Given that households also consume public goods (e.g. the quality of their home), higher household income would bring more ressources also for the public goods from which the individual gets utility. In this model, however, the way each person gets utility from income is not detailed in terms of how the income is spent. The above is just one example why household income and not just own income might matter for the individual.

Formally, the utility function for wife i is given by

$$\begin{aligned} u^{d_{h,t},M}_{i,t} =& \kappa(inc^{rw}_{i,t})\cdot(inc^{rw}_{i,t} + \chi\cdot inc^{rw}_{j,t} - hcost^{rh}_{h,t}) + taste^{rh}_{h,t} + taste^{rw}_{i,t} \\ &- pswcost^{rh,rh_{t-1}}_{i,t} - pswcost^{rw,rw_{t-1}}_{i,t} - comcost^{rh,rw}_{i,t} - cwork_{i,t}, \end{aligned} \tag{3.17}$$

and likewise for husband j. The utility for a single individual i is

$$\begin{aligned} u^{d_{i,t},S}_{i,t} =& \kappa(inc^{rw}_{i,t})\cdot(inc^{rw}_{i,t} - hcost^{rh}_{i,t}) + taste^{rh}_{i,t} + taste^{rw}_{i,t} \\ &- pswcost^{rh,rh_{t-1}}_{i,t} - pswcost^{rw,rw_{t-1}}_{i,t} - comcost^{rh,rw}_{i,t} - cwork_{i,t}, \end{aligned} \tag{3.18}$$

and similarly for a single male j. κ is marignal utility of money. $inc^{rw}_{i,t}$ is total earnings of the wife when she works in region $rw_{i,t}$ (including unemployment $rw_{i,t} = \varnothing$), and $inc^{rw}_{j,t}(\boldsymbol{x_{j,t}})$ $j's$ income from working in $rw_{j,t}$. $hcost^{rh}_{h,t}$ and $hcost^{rh}_{i,t}$ are the costs of living in region rh for household h and individual i, respectively. $taste^{rh}_{h,t}$ and $taste^{rh}_{i,t}$ control for amenities of residential location rh and $taste^{rw}_{i,t}$ for amenities in the work region. $pswcost^{rh,rh_{t-1}}_{i,t}$ are the psychological moving costs for residential moves, while $pswcost^{rw,rw_{t-1}}_{i,t}$ are psychological job moving costs including search costs. $comcost^{rh,rw_i}_{i,t}$ is commuting costs between residence $rh_{h,t}$ and work $rw_{i,t}$. Lastly, $cwork_{i,t}$ indexes disutility of working. The separate components of (3.17) and (3.18) will be elaborated below, but χ measures how much utility the individual gets from the spouses's income relative to its own income.

Specification of utility components

In the following I describe the separate components of (3.17) and (3.18) starting with an elaboration of the state variables: $\boldsymbol{x^1_{i,t}} = (t_i, rh_{it-1}, rw_{it-1}, educ_i)$ are state variables for individual i and include age, previous home location, previous work location, level of education $educ_i \in \{0, 1, 2\}$ corresponding to short, medium or long education[13], respectively. $\boldsymbol{x^1_{j,t}}$ is defined symmetrically. Other observed state variables for the individuals are $x^2_{i,t} = \mathbb{I}(kids_{i,t} > 0)$ and $x^2_{j,t} = \mathbb{I}(kids_{j,t} > 0)$ which indicate whether the individual has children living at home. When the houshold is a couple, it is the household's number of children that are relevant, hence $\boldsymbol{x_{h,t}} = (\mathbb{I}(kids_{h,t} > 0), t_{dif_h})$ which are a dummy for children in the household and the age difference between the husband and the wife. By not including t_i in the state variables for the household ($\boldsymbol{\tilde{x}_{h,t}} = (\boldsymbol{x^1_{i,t}} \setminus \{t_i\}, \boldsymbol{x^1_{j,t}}, \boldsymbol{x_{h,t}})$), but rather age of the husband and t_{dif_h}, the dimension of the state space is considerably reduced because age difference is constant over time within the marriage.

Moving on to the components of the utility function, I specify these for individual $k \in \{i, j\}$. $\kappa_{inc}(inc^{rw}_{k,t})$ is the marginal utility of money and $inc_{k,t}$ the predicted income. Letting κ depend on income allows high-income people to have a lower marginal utility of money which may affect their willingness to pay for certain goods, including the costs associated with moving. This is done to implicitly account for budget constraints in the model, which are not imposed explicitly. An individual with a high marginal utility of money will be less inclined to pay high house prices, all else equal, just like a person who is close to not satisfying his borrowing constraint is. I let

$$\kappa(inc^{rw}_{k,t}) = \kappa_0 + \kappa_y \cdot inc^{rw}_{k,t}.$$

When $\kappa_0 > 0$ and $\kappa_y < 0$, the highest marginal utility of money is attained by individuals with an income of zero.

The wage offer itself is specified by work location and education group for non-unemployment regions such that we get a set of coefficients for each combination of work location and education. Unemployment benefits are specified separately for each

[13]This could be replaced by occupation. However education is exogenous in this model where only people who are in the labor force, excluding students, are considered. Occupation, on the other hand, is a direct outcome of the work location choice and is therfore rather considered an (implicit) outcome of the model.

education and age group:

$$\ln(inc^{rw}_{k,t}) = \begin{cases} \delta_0^{rw,educ_k} + \delta_a^{rw,educ_k} age_{k,t} + \delta_{a^2}^{rw,educ_k} age^2_{k,t} + \delta_\varnothing \mathbb{I}(rw_{kt-1} = \varnothing) & \text{if } rw \neq \varnothing \\ b^{educ_{k,t}} & \text{if } rw = \varnothing. \end{cases} \tag{3.19}$$

$\delta_0^{rw,educ}$ is the constant, while $age_{k,t}$ is the age of the individual and equals $t + t_{dif_h}$ for the wife and t for the husband. I allow for positive, but decreasing returns to age when $\delta_a > 0$ and $\delta_{a^2} < 0$. Including a dummy for whether one was unemployed in the previous period allows the wage offers to differ between unemployed and employed job applicants[14]. The wage process does not distinguish between genders. This essentially means I model the individuals' decisions as if there were no wage discrimination on the labor market. If the model predicts a gender wage gap after all, it cannot attribute this to different wage offers conditonal on background characeristics, but rather to different choices of work location for each gender.

The remaining components of (3.17) and (3.18) are specified below, where $a \in \{h, k\}$ is used for certain subscripts to emphasize that the variable is specific to the household for married individuals. $a = h$ is used when considering a married individual and $a = k$ when a single:

$$taste^{rh}_{a,t} = \tau^{rh}_{rest} \cdot rest^{rh} + \tau_{nature} \cdot nature^{rh} + \tau_{thefts} \cdot thefts^{rh} \tag{3.20}$$

$$taste^{rw}_{k,t} = \psi_{dens} \cdot jobdens^{rw,educ_{k,t}} \tag{3.21}$$

$$hcost^{rh}_{a,t} = uc \cdot P^{rh} \cdot sqm^{rh} \tag{3.22}$$

$$pswcost^{rh_t,rh_{t-1}}_{k,t} = \mathbb{I}(rh_{kt-1} \neq rh_{a,t})(\gamma_0 + \gamma_a age_{k,t} + \gamma_{kids}\mathbb{I}(kids_{a,t} > 0)) \tag{3.23}$$

$$pswcost^{rw,rw_{t-1}}_{k,t} = \mathbb{I}(rw_{kt-1} \neq rw_{k,t})(o_0 + o_a age_{k,t} + o_\varnothing \mathbb{I}(rw_{kt-1} = \varnothing)) \tag{3.24}$$

$$cwork_{k,t} = \mathbb{I}(rw_{k,t} \neq \varnothing) \cdot \alpha_{work}. \tag{3.25}$$

$$comcost^{rh,rw}_{k,t} = \begin{cases} (\eta_0 + \eta_{kids}\mathbb{I}(kids_{a,t} > 0))time(rh_{a,t}, rw_{k,t}) \text{ if } a = k, \\ (\eta_0 + (\eta_{kids} + \eta_{male}\mathbb{I}(Male_{k,t})\mathbb{I}(kids_{a,t} > 0))time(rh_{a,t}, rw_{k,t}) \text{ if } a = h. \end{cases} \tag{3.26}$$

(3.20) controls for differences in number of restaurants and cafés ($rest$), nature capital index ($nature$) and number of victims of property crime ($thefts$). The function is specific

[14]Ideally, I would want to include a random wage component and the individual's wage income from $t-1$. The role of the former would allow individuals with the same state variables to not expect receiving the same wage in a given location. The role of the latter would be to capture that the quality of the previous jobs (quality of the acquired human capital), here measured as the realized wage, affects the individual's future earnings. This would allow the model to predict that individuals would hesitate to move to a certain area that can only provide relatively low-quality jobs for the person because that means it gets harder to get a high-paid job in the future even if applying for jobs in a generally high-paid area. However, it is infeasible to include other (continuous) state variable for now.

to the household and the amenities are considered constant over time. (3.21) controls for the differences in job density ($jobdens$) across the regions for a given education group. An individual is assumed to only care about the job density for the education group it belongs to, and the effect of job density on the taste for a work region is restricted to be constant across these groups.

Moving on to (3.22), this function controls for the costs of living in the region and is location-specific. Costs of living in a region are specified as a given share, uc, of the observed total average house prices in the region, where sqm^{rh} is the exogenous average square meters of housing in region rh and P^{rh} the exogenous square meter price. Everyone is considered renters in the model. Hence, capital gains from differences in house prices over time do not drive location decisions. Housing costs can be changed by moving to another region, but moving involves moving costs as specified in (3.23). They depend on age and whether there are children in the household. One can imagine that having children, especially children in school age, increases the moving costs since moving might mean changing school and friends for the children.

In reality, people do not change jobs every period. This might imply they keep a job which is not optimal. The reason for this can be job moving costs; it is costly to search for a new job for and change job region. To account for this, I include psychological job switching costs as described in (3.24) which depend on age and whether the individual was unemployed last period. The latter control is included to acknowledge that there is a notable difference between unemployed and employed individuals when it comes to searching for a job. Unemployed individuals have a higher incentive to find a "new" job, all else equal, as unemployment benefits are lower than earnings. (3.25) captures disutility of work through a constant.

For commute costs the first line of (3.26) specifies the costs for singles and the second line does so for couples. For both household types there is a base level of disutility η_0 and an effect of having children, η_{kids}, since being a parent might change your value of time at home. For couples, the effect of children is allowed to differ by whether the individual is the husband ($\mathbb{I}(Male_{k,t}) = 1$). All these variables are multiplied with the commuting time $time(rh_{a,t}, rw_{k,t})$ between home and workplace. I therefore allow for commuting costs both if the person does not work in the same region as where he lives and if he does, but it is zero per definition if he is unemployed.

4.5 Transition matrices

Household- and individual-specific state variables

In order for the household to optimize its locations it must have expectations about how the state variables evolve over time. Age of each spouse increases by one each year, i.e. mortality risk is disregarded. Naturally, the age difference is constant within the marriage. Previous work and home location are completely determined by previous period's choices.

Educational level is constant since I focus on individuals who are a part of the labor force and thus not studying anymore.

Though random transitions of children are not implemented in the empirical application but left for future research, the idea is that the number of children in the family can change over time. Since I do not model fertility decisions I would let the arrival of children be random shocks. It is only in rare events that more than one childbirth occurs within the same year so I would consider having an extra child as a 0/1 outcome. The number of children in the previous period affects the number of children in the current period, and I would assume only couple households may expect to have more children in the future. If singles have more children, it is an unexpected event. Furthermore, because children move out of home at some point, it is possible to go from having a positive number of children in the household to having less. I would let the age of the wife $t + t_{dif_h}$ affect the number of children as well to account for fertility and to predict when the last child is likely to move out. The distribution would be given by

$$kids_{h,t+1} \sim s(kids_{h,t}, t + t_{dif_h}, \boldsymbol{\theta_s}), \tag{3.27}$$

where $\boldsymbol{\theta_s}$ is a vector of parameters.

Location-specific variables

Observed amenities of the regions are average number of thefts and restaurants and cafés, nature capital index, property prices and housing size. They are all considered constant over time. The model therefore cannot say anything about whether households move in expectation of an area gentrifying in the future. This would not be a simple addition to the model since it would require carrying the amenities of *all* regions as state variables. That would seriously complicate the estimation of the model.

4.6 Solution method

The purpose of the model is to investigate whether the gender wage gap is affected by the intra-household decisions on locations and if policy interventions can help reducing the gap. To do so, one can carry out counterfactual policy experiments that asses how households change decisions after a new policy has been implemented. Simulating counterfactuals from a dynamic model requires solving it[15], while for estimation this is not necessary when applying the CCP algorithm from Hotz and Miller (1993).

With an assumption about the future expected value upon death at T (typically 0), the model should be solved by backwards induction starting at T and ending at t_0. For each possible state, the optimal solution is found after also computing the bargaining weights for couples according to (3.10). The same procedure is carried out for period $T-1$ with

[15] Section 6.5 explains circumstances under which the full solution is not needed

the solution for T in mind. Since future taste shocks and remaining states are unobserved to the household it computes the *expected* value function for T. It then considers how the decision at $T-1$ will affect the evolution of states and for couples also the bargaining power, and hence the expected value in the next period, and makes its decision. This procedure continues until t_0. In addition, I as the econometrician, must integrate over current taste shocks. I will then get the CCP of each decision being taken and can use this to assess the consequences of a counterfactual policy by simulating decisions from those CCPs.

5 Estimation Method

There exist different methods for estimating dynamic discrete choice problems like the one presented in this chapter. The traditional approach is the Nested Fixed Point (NFXP) algorithm developed by Rust (1987). That one falls under the category of full solution methods meaning that for every guess of the parameter vector in the maximization algorithm, the entire dynamic programming problem must be solved by either backwards induction (finite horizon problems) or value function iterations until a fixed point is reached (infinite horizon problems). The drawback of this estimation routine is that for compliated programming problems there is a high computational burden associated with solving the model several times until convergence is achieved.

Starting with Hotz and Miller (1993) another branch of estimation methods were introduced under the name Conditional Choice Probability (CCP) methods. The idea is to recognize that the expected future value terms can be expressed as a function of CCPs. Just like the policy functions in (3.9) and (3.16) are mappings from alternative-specific value functions (and thus expected future value functions) to CCPs, Hotz and Miller (1993) prove that the inverse mapping also exists. While this method avoids the full solution of the program for every trial value of the parameter vector, it does on the other hand imply other restrictions. Instead of exploiting the structural relationship between CCPs and alternative-specific value functions directly, one rather exploits that while the alternative-specific value functions cannot be observed by the econometrician, the CCPs can in principle. However, this comes with very high data requirements since one must be able to compute the CCP for every possible combination of states and choices in the model.

The methods above work well when there are no unobserved states as in the current empirical implementation. When that holds and the assumptions of additive separability of the errors and conditional independence hold, Rust (1987) noted that the parameters of the transition matrices and the preference parameters of the flow utilities, respectively, can be estimated separately starting with the estimation of transition matrix parameters. This can be seen below where $\theta \equiv [\theta_1, \theta_2]$ is the true parameter vector of the model and θ_2 governs the transition processes of the observed state variables (assuming all elements of $\tilde{\boldsymbol{x}}_{\boldsymbol{a,t}}$ are observed):

$$\hat{\theta} = \arg\max_{\theta} \sum_{a=1}^{A} \sum_{t=1}^{T} \ln[CCP(d_{a,t})|\tilde{\boldsymbol{x}}_{\boldsymbol{a,t}}, \theta] + \ln[g(\tilde{x}_{a,t+1}|\tilde{\boldsymbol{x}}_{\boldsymbol{a,t}}, d_{a,t}, \theta_2)].$$

Here, the log likelihood function is additively seperable into two components; one part that is concerned with the choice and the other concerned with the transitions of the state variables. θ_2 can be consistently estimated using data on the state variables and their transitions only. Thereafter $\hat{\theta}_2$ is known and θ_1 can be estimated by use of the choice data. The method is inefficient because information from choices is not exploited in the estimation of θ_2, but it is computationally lighter.

The coming sections will go through the CCP algorithm. After the introduction of the algorithm, I will show how to exploit that the model exhibits finite dependence to simplify the computation of the value functions. In the end of the section I will outline the restrictions I impose in the empirial implementation, the most important one that I focus the estimation on a static version of the model and leave the estimation of the full dynamic model for future research.

5.1 The CCP algorithm

Let $l_{d_{a,t}}$ denote agent a's likelihood contribution from his choice $d_{a,t}$. Given an estimate of the CCPs, $C\hat{C}P_0$, for example from a flexible Logit, the estimate $\hat{\theta}_1$ is found as

$$\hat{\theta}_1 = \arg\max_{\theta_1} \left\{ \sum_{a=1}^{A} \sum_{t=1}^{T} \ln[l_{d_{a,t}}(\tilde{x}_{a,t}, \theta_1, \hat{\theta}_2, C\hat{C}P_0)] \right\} \tag{3.28}$$

$$l_{d_{a,t}}(\tilde{x}_{a,t}, \theta_1, \hat{\theta}_2, C\hat{C}P_0) = \frac{\exp[v_{a,t}^{d_{a,t},q}(\tilde{\boldsymbol{x}}_{\boldsymbol{a,t}}, C\hat{C}P_0, \theta_1, \hat{\theta}_2)/\sigma_\epsilon]}{\sum_{m \in \mathfrak{D}^q} \exp[v_{a,t}^{m,q}(\tilde{\boldsymbol{x}}_{\boldsymbol{a,t}}, C\hat{C}P_0, \theta_1, \hat{\theta}_2)/\sigma_\epsilon]}, \; q \in \{S, M\}. \tag{3.29}$$

The right-hand side of (3.29) is not identical to the right-hand side of (3.9) and (3.16) but both are correct. The difference is that in (3.29), the future value terms entereing $v_{a,t}^{d,S}$ and $v_{a,t}^{d,M}$ have been rewritten in order to exploit i) the existence of a terminal choice that couples can take, namely divorce and ii) finite dependence of the state variables. When either of these two conditions hold for the problem at hand, Hotz and Miller (1993) and Hotz et al. (1994) show how to reduce the computational complexity a great deal. Arcidiacono and Miller (2011) extends the method to models with unobserved heterogeneity. In the next two paragraphs, the adjustment of the value functions for households and singles will be outlined. Papers that have used these methods include Bishop (2012) and Joensen and Mattana (2017), among others.

Estimating preference parameters

To estimate the preference parameters θ_1, the econometrician must know the alternative-specific value functions to get the likelihood contributions, cf. (3.29). The idea is that the future value function can be expressed as a function of flow utilities and CCPs. That is why the alternative-specific value functions have been written as a function of these in (3.29). The coming section explains in detail.

First, the parameters concerning singles are estimated according to (3.28) using data on singles only. Having obtained these parameters, I can compute the expected future value functions using (3.6). Next, using data on couples only, the dual-earner household's problem can be estimated. The possibility to first estimate parameters concerning the singles' problem and thereafter estimate the parameters of the dual-earner household's problem is because of the assumption that singles do not consider the option of remarrying or at least do not let that option affect their location decisions.

Estimation equation for singles For singles there is no terminal action that brings the singles into an absorbing state where the decision problem ends. However, the singles' problem exhibits finite dependence after two periods (like the couple household's problem does): when two choice sequences with different initial decisions lead to the same distribution of states after ρ periods, the problem exhibits ρ-period finite dependence. In this particular model $\rho = 2$ since conditional on $\tilde{x}_{k,t+1}$ only the choice at $t+1$, and not previous choices, affect the distribution of states at $t+2$. To be concrete, those state variables, that are directly affected by the choice in the last period are the previous home and work locations since these are both endogenous outcomes.

Observing this helps simplifying the computation of the alternative-specific value functions from (3.7) when imposing some (not necessarily optimal) choice $d_{k,t+1}^{imp1}$. Doing that means it is possible to rewrite the alternative-specific value function with respect to this choice, cf. Hotz et al. (1994):

$$
\begin{aligned}
v_{k,t}^{d,S} &= u_{k,t}^{S} \\
&+ \beta \int_{\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}} (v_{k,t+1}^{d^{imp1},S} - \ln[C\hat{C}P_0(d_{k,t+1}^{imp1}|\tilde{x}_{k,t+1})])g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t})
\end{aligned}
\tag{3.30}
$$

The fact that $E[V_{k,t+1}^{M}(\boldsymbol{z}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t})]$ can be replaced with $v_{k,t+1}^{d^{imp1}} - \ln[C\hat{C}P_0(d_{k,t+1}^{imp1}|\tilde{x}_{k,t+1})]$ has an intuitive explanation: the expected future value of being in state $\tilde{x}_{k,t+1}$ can be expressed as the sum of the alternative-specific value of taking an arbitrary decision d^{imp1} ($v_{k,t+1}^{d^{imp1}}(\tilde{x}_{k,t+1})$) and a non-negative adjustment term $\ln[C\hat{C}P_0(d_{k,t+1}^{imp1}|\tilde{x}_{k,t+1})]$)[16]. The latter adjusts for the possibility that d^{imp1} may be a disoptimal choice. If the probability of choosing d^{imp1} increases, the adjustment terms goes towards zero.

[16]In principle one should also add a constant representing the mean of the type I Extreme Value distribution, but since only differences in value functions are identified, this would cancel in estimation and is thus not relevant after all.

The second line of (3.30) can be furter expanded by imposing the choice $d_{k,t+2}^{imp2}$ for the second period. Again, this choice does not have to be optimal but can be any, potentially convenient, choice:

$$\begin{aligned}
&\Longleftrightarrow \\
v_{k,t}^{d,S} \quad &= \quad u_{k,t}^{d,S} \\
&+ \quad \beta \int_{\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}} (u_{k,t+1}^{d^{imp1},S} + \beta \int_{\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}} (v_{k,t+2}^{d^{imp2},S} - \ln[C\hat{C}P_0(d_{k,t+2}^{imp2}|\tilde{x}_{k,t+2})]) \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&- \quad \beta \int_{\tilde{x}_{k,t+1}} \ln[C\hat{C}P_0(d_{k,t+1}^{imp1}|\tilde{x}_{k,t+1})]) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \qquad (3.31)
\end{aligned}$$

Again, the second line of (3.31) can be expanded by imposing the choice $d_{k,t+3}^{imp3}$ for the third period:

$$\begin{aligned}
&\Longleftrightarrow \\
v_{k,t}^{d,S} \quad &= \quad u_{k,t}^{d,S} \\
&+ \quad \beta \int_{\tilde{x}_{k,t+1}} (u_{k,t+1}^{d^{imp1},S}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&+ \quad \beta^2 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} u_{k,t+2}^{d^{imp2},S} \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&+ \quad \beta^3 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \int_{\tilde{x}_{k,t+3}} v_{k,t+3}^{d^{imp3},S} \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+3}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}, d_{k,t+2}^{imp2}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&- \quad \beta^3 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \int_{\tilde{x}_{k,t+3}} \ln[C\hat{C}P_0(d_{k,t+3}^{imp3}|\tilde{x}_{k,t+3})] \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+3}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}, d_{k,t+2}^{imp2}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&- \quad \beta^2 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \ln[C\hat{C}P_0(d_{k,t+2}^{imp2}|\tilde{x}_{k,t+2})] \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&- \quad \beta \int_{\tilde{x}_{k,t+1}} \ln[C\hat{C}P_0(d_{k,t+1}^{imp1}|\tilde{x}_{k,t+1})] g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \qquad (3.32)
\end{aligned}$$

As will become clear below it is not necessary to expand further on (3.32).

Only differences in value functions matter (the location parameter is not identified in a Logit model). Hence, $v_{k,t}^{d,S}$ can be differenced with respect to another value function $v_{k,t}^{d',S}$ where the future component in $v_{k,t}^{d',S}$ has been expanded as above:

$$\begin{aligned}
v_{k,t}^{d,S} - v_{k,t}^{d',S} \quad &= \quad u_{k,t}^{d,S} - u_{k,t}^{d',S} \\
&+ \quad \beta \int_{\tilde{x}_{k,t+1}} u_{k,t+1}^{d^{imp1},S} g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t})
\end{aligned}$$

$$
\begin{aligned}
&- \quad \beta \int_{\tilde{x}_{k,t+1}} u_{k,t+1}^{d',S} g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d'_{k,t}) \\
&+ \quad \beta^2 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} u_{k,t+2}^{d^{imp2},S} \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&- \quad \beta^2 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} u_{k,t+2}^{d^{imp2},S} \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d'_{k,t}) \\
&- \quad \beta^2 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \ln[C\hat{C}P_0(d_{k,t+2}^{imp2}|\tilde{x}_{k,t+2})] \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&+ \quad \beta^2 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \ln[C\hat{C}P_0(d_{k,t+2}^{imp2}|\tilde{x}_{k,t+2})] \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d'_{k,t}) \\
&- \quad \beta \int_{\tilde{x}_{k,t+1}} \ln[C\hat{C}P_0(d_{k,t+1}^{imp1}|\tilde{x}_{k,t+1})] g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&+ \quad \beta \int_{\tilde{x}_{k,t+1}} \ln[C\hat{C}P_0(d_{k,t+1}^{imp1}|\tilde{x}_{k,t+1})] g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d'_{k,t})
\end{aligned} \tag{3.33}
$$

This part dropped out:

$$
\begin{aligned}
&+ \quad \beta^3 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \int_{\tilde{x}_{k,t+3}} v_{k,t+3}^{d^{imp3},S} \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+3}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}, d_{k,t+2}^{imp2}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&- \quad \beta^3 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \int_{\tilde{x}_{k,t+3}} v_{k,t+3}^{d^{imp3},S} \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+3}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}, d_{k,t+2}^{imp2}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d'_{k,t}) \\
&- \quad \beta^3 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \int_{\tilde{x}_{k,t+3}} \ln[C\hat{C}P_0(d_{k,t+3}^{imp3}|\tilde{x}_{k,t+3})] \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+3}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}, d_{k,t+2}^{imp2}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d_{k,t}) \\
&+ \quad \beta^3 \int_{\tilde{x}_{k,t+1}} \int_{\tilde{x}_{k,t+2}} \int_{\tilde{x}_{k,t+3}} \ln[C\hat{C}P_0(d_{k,t+3}^{imp3}|\tilde{x}_{k,t+3})] \\
&\qquad g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+3}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}, d_{k,t+2}^{imp2}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, d'_{k,t}) g(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}|\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1})
\end{aligned}
$$

To provide some intuition, $v_{k,t+3}^{d^{imp3},S}$ is the value of choosing d^{imp3} in period $t+3$ when the agent chose d^{imp2} in period $t+2$ and d^{imp1} in $t+1$. The reason dependence can be broken after two periods is because I conditioned on the same choice sequence $(d_{t+1}^{imp1}, d_{t+2}^{imp2})$ after period t and then the distribution of states at $t+3$ is independent of the period t choices d_t and d'_t: whenever the conditioning set is the same at $t+2$ $((\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}}, d_{k,t+1}^{imp1}))$ *and* at $t+3$ $(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+2}}, d_{k,t+2}^{imp2})$, the expected state at $t+3$ $(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+3}})$ is the same both when the period t decision is d_t and when it is d'_t. This is because memory only extends to the current $(t+2)$ and prior $(t+1)$ locations. Hence, the expected value at $t+3$ of choosing d_{t+3}^{imp3} is the same both when taking decision d_t and d'_t in period t under the imposition of similar choices for the coming periods.

Thus, by using (3.33) to replace $v_{k,t}^{d,S}(\tilde{\boldsymbol{x}}_{\boldsymbol{k,t}}, C\hat{C}P_0, \theta_1, \hat{\theta}_2)$ in (3.29) there is no need to solve the model via backwards induction since there are no expected future value terms anymore, only flow utilities and CCPs. The latter can be estimated by a flexibile reduced-form Logit or a frequency estimator, for example.

Estimation equation for couples The married household's problem also exhibits finite dependence. However, there is an additional feature of the married household's problem that is even more convenient. Namely that divorce is a terminal choice that leads to an absorbing state where the household no longer exists. By imposing a choice that involves divorce, e.g. $d_{h,t+1}^{imp1} = (rh_{h,t+1} = 1, rw_{i,t+1} = rw_{j,t+1} = \varnothing, D_{h,t+1} = 1)$ for period $t+1$, i.e. that the household decides it wants to divorce (effective from period $t+2$), lives in region 1 and lets both spouses be unemployed, the future value component of the alternative-specific value function can be written with respect to this choice:

$$
\begin{aligned}
v_{h,t}^{d,M} &= u_{h,t}^{d,M} \\
&+ \beta \cdot (1 - D_{h,t}) \int\limits_{\tilde{x}_{h,t+1}} [v_{h,t+1}^{d^{imp1},M} - \ln[C\hat{C}P_0(d_{h,t+1}^{imp1}|\tilde{x}_{h,t+1})]]g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&+ \beta \cdot D_{h,t}[(0.5 \cdot \int\limits_{\tilde{x}_{i,t+1}} [\phi(\tilde{x}_{i,t+1}]g(\tilde{x}_{i,t+1}|\tilde{x}_{h,t}, d_{h,t}) + 0.5 \cdot \int\limits_{\tilde{x}_{j,t+1}} [\phi(\tilde{x}_{j,t+1})]g(\tilde{x}_{j,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- \Delta]
\end{aligned} \tag{3.34}
$$

When imposing divorce in $t+1$ this approach simplfies the estimation a great deal since

$$
\begin{aligned}
v_{h,t+1}^{d^{imp1},M}(\tilde{x}_{h,t+1}) = u_{h,t+1}^{d^{imp1},M} \\
&+ 0.5 \cdot \beta \int\limits_{\tilde{x}_{i,t+2}} \left[\phi_{i,t+2}^S\right] g(x_{i,t+2}|\tilde{x}_{h,t+1}, d_{h,t+1}^{imp1}) \\
&+ 0.5 \cdot \beta \int\limits_{\tilde{x}_{j,t+2}} \left[\phi_{j,t+2}^S\right] g(x_{j,t+2}|\tilde{x}_{h,t+1}, d_{h,t+1}^{imp1}) \\
&- \beta\Delta.
\end{aligned} \tag{3.35}
$$

Divorce is an absorbing state for the household, hence no future decisions are made for the household as an agent from period $t+2$ onwards. This means there are no future value components left when computing $v_{h,t+1}^{d^{imp1},M}(\tilde{x}_{h,t+1})$ as $\phi_{i,t+2}$ and $\phi_{j,t+2}$ are pre-computed after estimating the singles' problem in the first step. Moreover, the transition densities g are estimated for singles as well. One just has to adjust the conditioning set from including $d_{i,t}$ and $d_{j,t}$, respectively, to include $d_{h,t}$. Since it is assumed that whether or not it is the household or the individual who has made a given choice for the individual (choices concerning the individuals are a subset of $d_{h,t}$), the transition densities are the same. So in order to compute $\int_{\tilde{x}_{i,t+2}}[\phi^S_{i,t+2}]g(x_{i,t+2}|\tilde{x}_{h,t+1}, d^{imp1}_{h,t+1})$ one has to do the integration using the transition density estimates from the singles' problem and apply the part of $d_{h,t}$ that concerns choices for individual i $((rw_{i,t}, rh_{h,t}))$ in the conditioning set. The same goes for the j counterpart. (3.34) can then be written as

$$
\begin{aligned}
v^{d,M}_{h,t} =& u^{d,M}_{h,t} \\
&+ (1-D_{h,t})\beta \cdot \int\limits_{\tilde{x}_{h,t+1}} \left[u^{d^{imp1}}_{h,t+1}\right] g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&+ (1-D_{h,t})\beta^2 \cdot 0.5 \int\limits_{\tilde{x}_{h,t+1}} \int\limits_{\tilde{x}_{i,t+2}} \left[\phi^S_{i,t+2}\right] g(\tilde{x}_{i,t+2}|\tilde{x}_{h,t+1}, d^{imp1}_{h,t+1}) g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&+ (1-D_{h,t})\beta^2 \cdot 0.5 \int\limits_{\tilde{x}_{h,t+1}} \int\limits_{\tilde{x}_{j,t+2}} \left[\phi^S_{j,t+2}\right] g(\tilde{x}_{j,t+2}|\tilde{x}_{h,t+1}, d^{imp1}_{h,t+1}) g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- (1-D_{h,t})\beta^2 \Delta \\
&- (1-D_{h,t})\beta \int\limits_{\tilde{x}_{h,t+1}} \left[\ln[C\hat{C}P_0(d^{imp1}_{h,t+1}|\tilde{x}_{h,t+1})]\right] g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&+ D_{h,t}\beta[(0.5 \cdot \int\limits_{\tilde{x}_{i,t+1}} [\phi(\tilde{x}_{i,t+1}] g(\tilde{x}_{i,t+1}|\tilde{x}_{h,t}, d_{h,t}) + 0.5 \cdot \int\limits_{\tilde{x}_{j,t+1}} [\phi(\tilde{x}_{j,t+1})] g(\tilde{x}_{j,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- \Delta]
\end{aligned}
\tag{3.36}
$$

When differencing $v^{d,M}_{h,t}$ and $v^{d',M}_{h,t}$ for some arbitrary $d' \in \mathfrak{D}^M$ that has been developed as in (3.36), we get

$$
\begin{aligned}
v^{d,M}_{h,t} - v^{d',M}_{h,t} =& u^{d,M}_{h,t} - u^{d',M}_{h,t} \\
&+ (1-D_{h,t})\beta \int\limits_{\tilde{x}_{h,t+1}} u^{d^{imp1},M}_{h,t+1} g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- (1-D'_{h,t})\beta \int\limits_{\tilde{x}_{h,t+1}} u^{d^{imp1},M}_{h,t+1} g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d'_{h,t}) \\
&+ (1-D_{h,t})\beta^2 \cdot 0.5 \int\limits_{\tilde{x}_{h,t+1}} \int\limits_{\tilde{x}_{i,t+2}} \phi^S_{i,t+2} g(\tilde{x}_{i,t+2}|\tilde{x}_{h,t+1}, d^{imp1}_{h,t+1}) g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- (1-D'_{h,t})\beta^2 \cdot 0.5 \int\limits_{\tilde{x}_{h,t+1}} \int\limits_{\tilde{x}_{i,t+2}} \phi^S_{i,t+2} g(\tilde{x}_{i,t+2}|\tilde{x}_{h,t+1}, d^{imp1}_{h,t+1}) g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d'_{h,t}) \\
&+ (1-D_{h,t})\beta^2 \int\limits_{\tilde{x}_{h,t+1}} \int\limits_{\tilde{x}_{j,t+2}} \phi^S_{j,t+2} g(\tilde{x}_{j,t+2}|\tilde{x}_{h,t+1}, d^{imp1}_{h,t+1}) g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- (1-D'_{h,t})\beta^2 \int\limits_{\tilde{x}_{h,t+1}} \int\limits_{\tilde{x}_{j,t+2}} \phi^S_{j,t+2} g(\tilde{x}_{j,t+2}|\tilde{x}_{h,t+1}, d^{imp1}_{h,t+1}) g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d'_{h,t}) \\
&- (1-D_{h,t})\beta^2 \Delta
\end{aligned}
$$

$$
\begin{aligned}
&+ (1 - D'_{h,t})\beta^2\Delta \\
&- (1 - D_{h,t})\beta \int\limits_{\tilde{x}_{h,t+1}} \ln[C\hat{C}P_0(d^{imp1}_{h,t+1}|\tilde{x}_{h,t+1})]g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&+ (1 - D'_{h,t})\beta \int\limits_{\tilde{x}_{h,t+1}} \ln[C\hat{C}P_0(d^{imp1}_{h,t+1}|\tilde{x}_{h,t+1})]g(\tilde{x}_{h,t+1}|\tilde{x}_{h,t}, d'_{h,t}) \\
&+ D_{h,t}\beta \cdot 0.5 \int\limits_{\tilde{x}_{i,t+1}} \phi^S_{i,t+1} g(\tilde{x}_{i,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- D'_{h,t}\beta \cdot 0.5 \int\limits_{\tilde{x}_{i,t+1}} \phi^S_{i,t+1} g(\tilde{x}_{i,t+1}|\tilde{x}_{h,t}, d'_{h,t}) \\
&+ D_{h,t}\beta \cdot 0.5 \int\limits_{\tilde{x}_{j,t+1}} \phi^S_{j,t+1} g(\tilde{x}_{j,t+1}|\tilde{x}_{h,t}, d_{h,t}) \\
&- D'_{h,t}\beta \cdot 0.5 \int\limits_{\tilde{x}_{j,t+1}} \phi^S_{j,t+1} g(\tilde{x}_{j,t+1}|\tilde{x}_{h,t}, d'_{h,t}) \\
&- D_{h,t}\beta\Delta \\
&+ D'_{h,t}\beta\Delta
\end{aligned} \tag{3.37}
$$

Since $C\hat{C}P_0(d^{imp1}_{h,t+1}|\tilde{x}_{h,t+1})$ is considered data, it is straightforward to plug them into (3.37) and then insert that expression in the likelihood contribution function (3.29) in place of $v^{d,M}_{h,t}(\tilde{\boldsymbol{x}}_{\boldsymbol{h,t}}, C\hat{C}P_0, \theta_1, \hat{\theta}_2)$. The likelihood function can then be computed and maximized with respect to the preference parameters for the cohabiting households.

5.2 Identification

In this section I list a set of informal arguments why the model parameters are identified. Generally, I use either spatial variation only, individual variation only or both of them togehter.

Starting with κ_0, the constant of marginal utility of money, it is identified from the correlation between spatial variation in house prices or wages and the distribution of individuals across home and work regions (higher prices and lower wages are unattractive, all else equal). The effect of income on marginal utility of money, κ_y, is identified from the sorting of households with higher income to regions of higher prices. User costs, uc, are separately identified from the variation in house prices only.

For a given value of marignal utility of money and user costs, the returns to scale in the couple household, χ, is identified from the variation in prices and *total* income in the household weighted by χ. So if households tend to locate in regions that are too expensive from the individual's perspective with its individual income, χ will be more positive.

For the taste of regional amenities ($\tau_{rest}, \tau_{nature}, \tau_{thefts}$) they are identified from spatial variation in house prices and values of these amenities. Admittedly, they also soak up any unobservable regional-specific effects that make the region more or less attractive. In order to obtain stronger identification, time variation in the amenities would be helpful, since that allows me to exploit within-region variation. I will return to this in future work. For the job taste parameters, the identification comes from the transitions of workers across work regions with different job density.

The constants of psychological switching cost for residential and job moves, γ_0 and o_0, are identified by cross-sectional variation in the share moving home and job, respectively. The effects of age, γ_a and o_a, are determined by the evolution of moving probabilities over the life cycle. γ_{kids} can be identified from the difference in moving propensity for households with and without children and $o_{\varnothing}$ from the gap between job transition probabilities for unemployed and employed people. It should be noted though that I do not account for unobserved fixed preferences for living or working in a certain region which may cause spurious state dependence. This can lead to an upward bias in the estimates of moving costs. To deal with this I would have to model individuals' initial conditions more carefully and not just start at age 25 ignoring that initial locations are not random at this point. The unemployment share identifies the disutiliy of work, α_{work}, while the average travel time identifies the commute cost parameter η_0. The difference in commute distance between individuals with and without children identifies η_{kids}, and males' commute time relative to females' identifies η_{male}.

Differences in the value of locations for men compared to women and whether the household tends to locate close to the male's or female's optimal place identifies Υ_0 of the bargaining weight. If households generally locate where the female would prefer, she is assumed to have a higher bargaining weight, all else equal, which increases Υ_0. The variation in the intra-household difference in outside options identifies Υ_1. I.e. for a given value of Υ_0, if the difference between wife's and husband's outside option increases and the household as a response chooses to locate closer to the female's optimal location, this informs that Υ_1 should be increased in the positive direction. The divorce costs Δ are identified from the share of couples who divorce for reasons not explained by differences in values of being single and in a couple.

5.3 Restrictions for empirical implementation

As I alluded to in the beginning of the section, I am currently not able to estimate the full dynamic model for couples because the state-choice space grows by a power of seven in the number of locations $(rh, rw_m, rw_m, rhp_m, rhp_f, rwp_m, rwp_f)$. The main specification for the structural estimation therefore sets $\beta = 0$, meaning future values do not matter for the agents' decisions. This also means I do not include the decision to divorce for the couples, since that is a decision which only affects future values. This also helps to reduce the state-choice space. Consequently, the evolution of future states is irrelevant for the estimation. The same goes for the initial CCPs that would be estimated by a flexible reduced form specification. The estimation equations for singles is therefore:

$$v_{k,t}^{d,S} \quad = \quad u_{k,t}^{d,S}, \tag{3.38}$$

and for couples

$$v_{h,t}^{d,M} = W_{i,t}u_{i,t}^{d,M} + (1 - W_{i,t})u_{j,t}^{d,M}. \tag{3.39}$$

When the future values are not relevant for the decision problem of couples it also means the evolution of the bargaining weight as a response to the decisions is not taken into account by the household. Each household member rather enters period t with their individual-specific states and the couple computes the bargaining weight according to (3.10). This corresponds to the specification for bargaining weights in a dynamic model with full commitment, i.e. where the evolution of the spouses' states cannot affect future bargaining weights.

The computational burden associated with the estimation for couples does not carry over to singles to the same extent. I do therefore show estimates from the the dynamic version of the single model, but restrict the evolution of future states to be non-random. Define $\hat{\boldsymbol{x}}_{\boldsymbol{k,t+1}} \equiv (\tilde{\boldsymbol{x}}_{\boldsymbol{k,t+1}} \setminus \{t+1, rhp_{k,t+1}, rwp_{k,t+1}\})$. The assumption then is

$$g(\hat{\boldsymbol{x}}_{\boldsymbol{k,t+1}}|\hat{\boldsymbol{x}}_{\boldsymbol{k,t}}, \boldsymbol{\theta_g}) = \begin{cases} 1 & \text{if } \hat{\boldsymbol{x}}_{\boldsymbol{k,t+1}} = \hat{\boldsymbol{x}}_{\boldsymbol{k,t}} \\ 0 & \text{if } \hat{\boldsymbol{x}}_{\boldsymbol{k,t+1}} \neq \hat{\boldsymbol{x}}_{\boldsymbol{k,t}} \end{cases} \tag{3.40}$$

such that the states in $t+1$ are expected to stay constant, except for age t which increases by one as in the original set-up and location-specific state variables which are determined directly by $d_{k,t}$.

6 Results

This section presents and interprets the results from the estimation of the static version of the model outlined in Section 4. I restrict attention to the Greater Copenhagen Area (GCA) which is a selected subset of the population data presented in Section 3. I find that bargaining weights are around 0.51 for the women in all regions and the effect of the outside option on the weight is negligible. It is therefore not discrimination of the woman within the household that causes them to choose unfavorable commutes and earn less according to the model of this chapter. Increasing the job density in Albertslund outside of Copenhagen causes more households to live there and the neighbouring municipalities. On the job side, there is a large effect on the probability of working in Ballerup, close by Albertslund, because the wages in this region are very favourable. The female response is stronger than the males'. Hence, Albertslund becoming more attractive as a work region induces households to live there, let at least one spouse work there and then exploit the now shorter commute to the high-paying Ballerup work region. The intra-household difference in commute and the gender wage gap are slightly reduced as a consequence. However, the initial differences in commute and pay between genders is very low in the GCA compared to the rest of the country. This calls for estimating the model on a larger and more representative sample of the Danish economy to gain more policy-relevant insight into the discussion of the effects of relocating jobs from urban to rural regions.

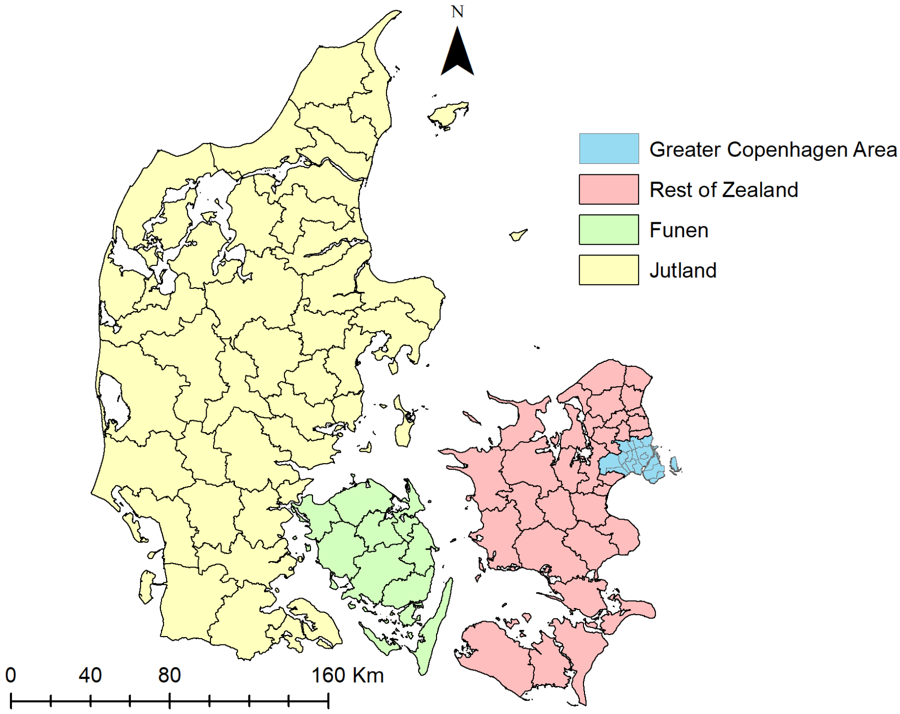
Below I present the results in detail. First, I define the estimation dataset and the dimension of the state space. Next, I present parameter estimates from the single model including model fits and do the same for couples thereafter. During that presentation, I translate parameters into monetary terms using a standard single person of age 40 with an annual income of 300,000 DKK. For couples I use a couple of the same age, where each spouse earns 300,000 DKK and the bargaining weights are 0.5 on each of them. This allows me to compute marginal utility of money used to do the transformation from utils to monetary terms. I then show how residential and work location choices, differences in commute time within the couple household and wage incomes are affected from a counterfactual increase of the job density in Albertslund. In the end I show preliminary results from a dynamic version of the model for singles.

6.1 Estimation sample and state space

The data used for estimation is a selected subsample of the original data described in section 3. The main change is the definition of regions. Instead of estimating the model for the entire Denmark, I choose a subsample of households who either live or work (or both) in the Greater Copenhagen Area (GCA) which is defined in Figure 3.9. However, I do still allow for households choosing to live or work in the remaining outside options Rest of Zealand, Funen and Jutland which are defined on the map too, but I do not include someone who e.g. lives and works in Rest of Zealand. For couples living outside the GCA, at least one of the spouses must work within the GCA to be included in the sample or the household must live there if both spouses work outside the area. Amenities for the outside options such as house prices, wages and nature are defined as the average across the municipalities in that region. The reason I restrict the number of regions is because the state and choice space grows exponentially in that dimension. The value function must be computed for every combination of observed states and all choices in order to compute the logsum which is used in the calculation of household $h's$ choice probability. I currently need to restrict the dimensions of the state and choice space for that operation to be computationally feasible. In that sense, the computational burden of NFXP has not been completely solved just by the use of CCP methods.

For the rest of the state space, I define children to be a dummy for having a child and schooling to take three values: zero if low education (no more than high school), 1 if medium education (vocational or short-length further education) and 2 for high education (bachelor degree or more). The model is estimated for individuals where the male is

Figure 3.9: Definition of regions in estimation



between 25 and 64 years old[17]. Since the $t-1$ home location for couples only potentially differ in the year where they move in together and the focus of this chapter is not specifically on the time around the start of cohabitation, I define the male and female previous home location to equal the male's previous home. This significantly reduces the state space. The age difference t_{dif} can take one value of 0 implying that I do not distinguish between the male's and female's ages. If her age is different from her husband's age in the dataset, I restrict it to be the same as his. This is not expected to make a significant difference in this sample, since on average the age difference is no more than 1.6 years, cf. Table 3.6 which shows summary statistics of the estimation sample of couples. Table 3.5 shows the corresponding figures for singles. The distribution of home and work locations in t and $t-1$ are presented in Table D1 for singles and Table D2 for couples.

The dimension of the stace space for singles is 91,200 and for couples 5,472,000. The dimension of the choice space is 380 different alternatives for singles and 7,600 for couples (in the dynamic model it would be 15,200 to allow for the decision to divorce) implying a 34,656,000 dimensional state-choice space for singles and more than $41 \cdot 10^9$ dimensional state-choice space for couples. To avoid computing the value functions for states which are not observed in the data and to reduce the time it takes to evaluate the likelihood function, I use a 1 percent random sample from the dataset of singles and couples, respectively. I then find the unique states that are observed in either sample and evaluate the value

[17]Even if estimating the dynamic model, the terminal value is irrelevant for $t > 64$ as finite dependence ensures only value functions up to two periods ahead are included in the expected value function.

function for each observed state-all choices combination. For the static model, I do not have to evaluate the value function for all the potential future states that might be reached. For the dynamic model, I would have to compute the value function for all possible states two periods ahead for singles and one period ahead for couples which is currently infeasible. Using the 1 percent random sample, I observe 13,004 states for singles and 25,804 states for couples. Hence, the value function must be computed for the state-choice space of dimension 4,941,520 for singles and 196,110,400 for couples.

The disadvantage of focusing on the GCA is that it is a region where the difference among the spouses is relatively negligible, e.g. in terms of differences in education according to Table 3.7. More importantly, however, the difference in commute time wihthin the household is the lowest throughout the country and there is not much variation across the municipalities within the region, cf. Section 3. The same goes for the gender wage gap. I do still include the Rest of Zealand region, where the intra-household differences in commute time and the gender wage gap shows more variation, but I group all the municipalities in that region into one. That region is therefore not very accurately described and the municipalities in that region do in reality differ in terms of house prices, wages, household income and commute distance. The estimates must therefore be interpreted in light of this and rather perceived a first step towards estimating a home-work location choice model for couples with endogenous bargaining weights. In order to better trust the estimates and predictions, I would have to disaggregate at least the Rest of Zealand regions into finer subregions and include dynamics as moving decisions are inherently dynamic. This is left for future work.

Table 3.5: Summary statistics of estimation data for singles

	Mean	S.d.	N
Age	40.069	10.91	3,703,096
I[kids]	0.163	0.37	3,703,096
Education			
Low	0.343	0.47	3,703,096
Medium	0.311	0.46	3,703,096
High	0.346	0.48	3,703,096
Commute time t (hours)	0.577	0.61	3,507,359
Commute time t-1 (hours)	0.578	0.57	3,264,517
I[move home]	0.076	0.26	3,703,096
Home move distance (hours)	0.433	0.52	3,703,096
I[move job]	0.142	0.35	3,703,096
Job move distance (hours)	0.427	0.57	3,191,874

Note: Commute time and job move distances only available for individuals in employment. Low education is no more than high school, medium education is vocational or short-length further education and high is bachelor degree or more. I[move home] = 1 and I[move job] = 1 if moving address, i.e. intra-regional moves included.

Table 3.6: Summary statistics of estimation data for couples

	Mean	S.d.	N
Age male	42.752	9.25	3,379,748
Age male - Age female	1.546	3.16	3,379,748
I[kids]	0.637	0.48	3,379,748
Commute time male t (hours)	0.767	0.68	3,318,801
Commute time female t (hours)	0.704	0.56	3,296,957
Commute time male t-1 (hours)	0.750	0.63	3,231,537
Commute time female t-1 (hours)	0.692	0.52	3,122,637
I[move home male]	0.045	0.21	3,379,748
I[move home female]	0.044	0.21	3,379,748
Home move distance male (hours)	0.599	0.51	3,379,748
Home move distance female (hours)	0.599	0.51	3,379,748
I[move job male]	0.135	0.34	3,379,748
I[move job female]	0.104	0.30	3,379,748
Job move distance male (hours)	0.479	0.58	3,202,540
Job move distance female (hours)	0.497	0.53	3,088,889

Note: Commute time and job move distances only available for individuals in employment. I[move home] = 1 and I[move job] = 1 if moving address, i.e. intra-regional moves included.

Table 3.7: Combination of educational degree within couple in estimation sample

	Husband			
Wife	Low %	Medium %	High %	Total %
Low	8.2	9.7	5.5	23.4
Medium	8.1	18.1	7.6	33.9
High	6.7	9.4	26.6	42.7
Total	23.0	37.2	39.7	100.0

Note: Low education is no more than high school, medium education is vocational or short-length further education and high is bachelor degree or more.

6.2 First-stage estimates and fixed parameters

Before estimating the structural models, I fix the parameters in Table 3.8. Below I explain each set of parameters.

Preference parameters

The scale of alternative-specific taste shocks is not identified and therefore fixed. All other parameters should be interpreted relative to the scale.

The constant of the marginal utility of money (κ_0) should in principle be identified, cf. Section 5.2. However, I experienced problems of poor identification and suspect the relatively lower variation in prices and wages across the regions in the estimation

Table 3.8: Fixed parameters

Parameter	Value
σ_ϵ	1.0
κ_0	0.176
κ_y	-0.00124
uc	0.1035

Note: κ_0 and κ_y come from Chapter 2 and are rescaled to the money scale of 10,000 in this chapter. uc is from Chapter 2.

sample to be the reason. I therefore use the estimates from Chapter 2[18] where there was also variation in housing costs *within* regions through the individual-specific demand for square meters. In this chapter, I consider square meters exogenous and the choice of a home region means choosing the average number of square meters for that region for any household. The value for user costs (uc) also comes from Chapter 2. Since I denote all monetary terms in 10,000s and Chapter 2 used 100,000s, κ_0 must be divided by 10 and κ_y by 100 compared to the estimates from the previous chapter.

The issues with keeping these parameters fixed is that they might affect the estimates of the remaining structural parameters. If the true parameters of marginal utility of money and user costs in the current model are not the same as the ones in the dynamic model from Chapter 2, the rest of the parameters will change if they were actually estimated.

Wage equations

The parameters of the wage equation in (3.19) are the same as in Chapter 2. I.e., they have been estimated in a first stage using the population data of all individuals. For the non-work region, log of non-work income was estimated on a constant for each age and schooling level. Non-work income was defined as transfer income if the individual was unemployed and total income if the individual was retiree, both weighted by the share of the year one had been in the non-work state. For the remaining work regions log of annual earned income was estimated on a constant, second-order polynomial in age and a dummy for being non-employed in the previous period for each combination of work region and schooling level such that coefficients are specific to these groups. The estimates for the employment regions are presented in Table E1, Table E2 and Table E3 for each of the skill groups, respectively[19].

Keeping parameters for wages fixed during the estimation of the preference parameters means wages do not respond to individuals' choices of where to live and work. Also, the current wage estimates do not deal with Roy sorting (Roy, 1951). I therefore disregard that an individual chooses to work in some location, which may seem odd in terms of

[18] Estimates as of August 2019. They have changed in the latest version.

[19] Estimates from the non-employment region are available upon request.

Table 3.9: Utility Cost of Moving Home Region

Parameter	Estimate	Standard Error	t-statistic
γ_0	2.1995	0.0730	30.15
γ_a	0.0815	0.0020	40.20
γ_{kids}	0.7713	0.0666	11.58

Note: Estimates of Equation 3.23.

observables, but does so be because (s)he got an unusually good job offer for that location. The implementation of wages in the current set-up implies that each individual is assumed to get the average predicted wage that people of its (observable) type get, and the randomness is reserved to the Logit taste shock. In other words, I do not deal with self-selection. I could consider adding an individual-specific random wage shock, but adding regional-specific wage shocks come at a high computational cost as I would have to integrate over the distribution of wage shocks for all regions.

6.3 Preference parameters for singles

Residential moving costs

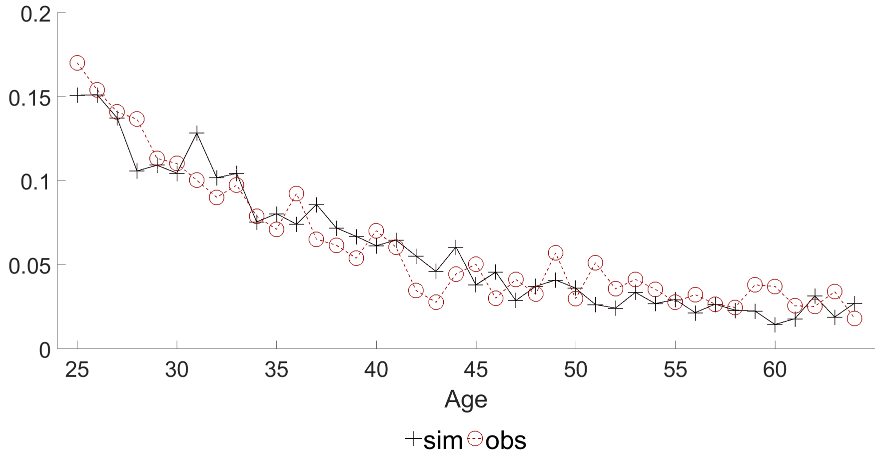
The structural parameters which index residential moving costs for singles in (3.23) are displayed in Table 3.9. The coefficients have the expected signs such that moving costs are higher, the older the individual and for those who have children. The utility of the standard single person without children who moves to another home region is thus $2.2 + 0.08 \cdot 40 = 5.4$ units lower compared to a similar person who does not move, all else equal. Using the estimates of κ_0 and κ_y, the moving costs for such a person is $5.4/(0.176 - 0.00124 \cdot 30) = 389,049$ DKK. Figure 3.10 shows the model fit of the share of individuals moving by age and this is very well-captured by the moving cost parameters[20].

Residential amenities

While the utility cost of moving will make the model fit moving behaviour, I include regional amenities for the home location to better capture exactly where households would like to locate - not just whether they move or stay. Table 3.10 shows the estimates associated with parameters in (3.20). They all have the expected signs, i.e. better access to nature is associated with higher utility. The same holds for more restaurants, while more thefts lower the utility. A region like Copenhagen which has 0.0352 thefts per inhabitants is therefore associated with 1.06 lower utility than a region like Hoeje-Taastrup,

[20]To compare the model predictions with the observed data, I use the estimated parameters to compute the CCPs of all choices by all observed states in the data. I then use the observed state data points and simulate a decision for each individual one period ahead by drawing from the model CCPs. Supplementary model fits are available in Appendix F.1

Figure 3.10: Model fit: Share of singles moving home region



Note: Choice data is simulated for 1 period ahead using the states in the observed data.

where there are 0.0073 thefts per inhabitant. On the other hand, Copenhagen has $1,257$ restuarants compared to 210 in Hoeje-Taastrup. This implies Copenhagen is 2.90 utils more attractive in terms of thefts[21]. For nature, Copenhagen has a value of 25 while Hoeje-Taastrup has 18, implying a utility difference of 0.2506 in favor of Copenhagen. In total, Copenhagen therefore performs better than Hoeje-Taastrup in terms of amenities. Using the marginal utility of money, the standard single would be willing to pay 16,571 DKK for 100 additional restaurants. These estimates should not be taken at face value as there is a risk they soak up region-specific fixed effects which will be included when time varying amenities are allowed for. Figure 3.11a shows the model fit of the probability of living in Copenhagen. Generally, it is well-captured except for the youngest cohorts. This is because the model expects the younger individuals to live in cheaper regions, because young people have a relatively low predicted income according to the wage regression estimates. The reason many young people can afford to live in Copenhagen is because they share apartments with roommates or have had to move there to study. This, and house size demand in general, has not been accounted for since everyone in a region will receive the average housing size and hence pay the average housing price for their home.

Turning to the probability of moving to Copenhagen, the model has a harder time predicting that behaviour. The very good fit of moving in general is to some extent driven by the fact that moving costs prevent people from moving, but when they move it is harder to predict exactly where they move. Figure 3.11b shows that the model does a pretty good job capturing the behaviour of singles in their late 30s, but up until then it cannot explain why people move to Copenhagen. Adding more amenities that describe the different regions and exploiting time-variation in them can help capture why certain regions are preferred over others. Interacting them by individual characteristics like age

[21] This number is computed as $(1.257 - 0.0210) \cdot 2.3348$ since restaurants are measured in 1,000s.

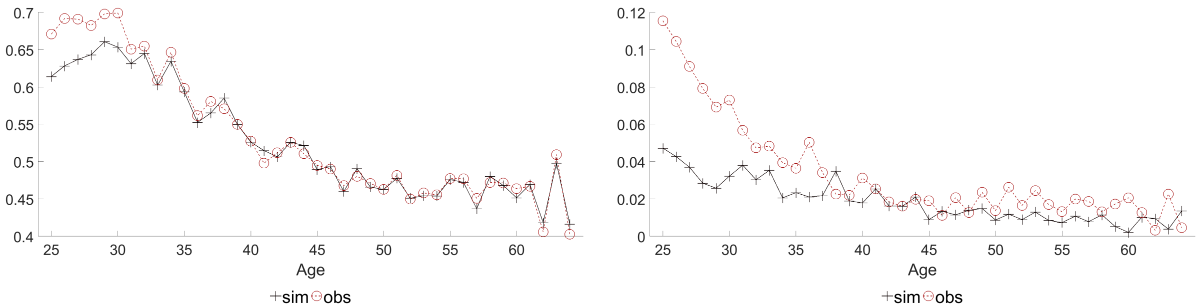
Table 3.10: Regional Amenities of Home Region

Parameter	Estimate	Standard Error	t-statistic
τ_{nature}	0.0358	0.0026	13.93
τ_{rest}	2.3348	0.1111	21.02
τ_{thefts}	-38.1157	5.3438	-7.13

Note: Estimates of Equation 3.20. Nature is nature capital index. Restaurants in 1,000s and thefts is number of thefts per inhabitants in the region. See Table C1 for summary statistics of amenities by region.

to pick up that some amenities are more important to some subgroups of the population might also be helpful going forward.

Figure 3.11: Model fit: Share of singles living in or moving home to Copenhagen



(a) Home in Copenhagen (b) Move to Copenhagen

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Job moving costs and disutility of work

The other part of the choice problem is the work location choice. Table 3.11 displays the parameters associated with job moving costs in (3.24). Again, the coefficients have the expected signs so older individuals are less likely to move job, all else equal, while those who were unemployed are more likely to shift into employment and thereby change job region. Figure 3.12a shows that the model predicts job moving behaviour well in general, but slightly underpredcits for the younger individuals. The job moving costs for a standard employed single person is 334,290 DKK, while an unemployed person with the same age and income is willing to pay 73,087 DKK to get a job. For someone below age 75, i.e. everyone in the sample, the job moving costs are negative if the person was unemployed in the previous period. The disutility of work from (3.25) is shown in Table 3.12. This is a utility gain when someone is unemployed. Nevertheless, an unemployed person would still be willing to pay money in order to get a job because the compensation in terms of wage earnings being higher than unemployment benefits is sufficiently high, even if he avoids paying commute costs as long as he is unemployed. This is not surprising since the

Table 3.11: Utility Cost of Moving Work Region

Parameter	Estimate	Standard Error	t-statistic
o_0	2.4575	0.0502	48.92
o_a	0.0427	0.0013	33.25
$o_\varnothing$	-5.6544	0.0570	-99.17

Note: Estimates of Equation 3.24.

Table 3.12: Utility Cost of Working

Parameter	Estimate	Standard Error	t-statistic
α_{work}	3.4750	0.0408	85.16

Note: Estimates of Equation 3.25.

estimates of commute costs from (3.26) and presented in Table 3.13 are quite low. For every additional hour of commuting, utility falls by just 0.3768 utils for someone without children.

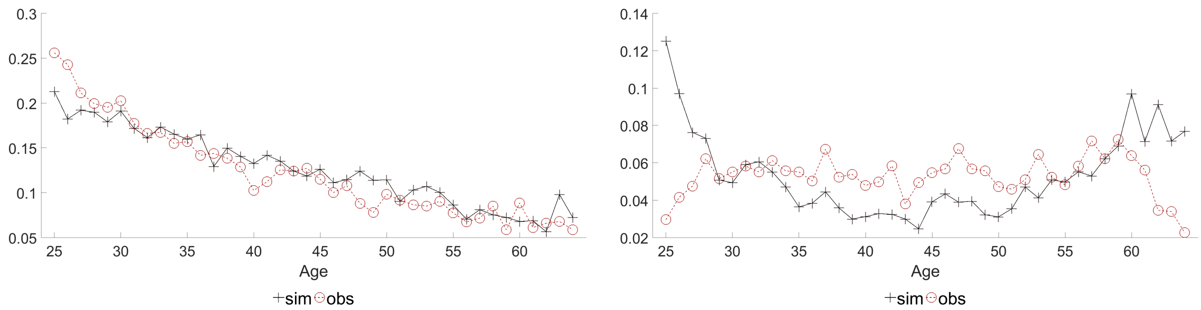
Figure 3.12b depicts the probability of being unemployed over the life cycle and Figure 3.13 the employment rate by home region. The model predicts a U-shaped unemployment probability which is not found in the data. The high predicted unemployment probability for young people is because they have a relatively low predicted wage. The incentive to take a job is therefore not very high for this group according to the model. The opposite holds for the mid-life individuals: their predicted wage income is high, so the model would expect them to work. These issues could be reduced by modelling wages more flexibly and especially allowing for unobservables and an AR process, i.e. include wage at $t-1$ as a state variable, but on the other hand this involves a huge increase in the dimension of the already large state space. Another option would be to model job search as we did in Chapter 2 since that would allow for involuntary unemployment. In the current model unemployment is always voluntary.

Table 3.13: Utility Cost of Commuting

Parameter	Estimate	Standard Error	t-statistic
η_0	0.3768	0.0125	30.12
η_{kids}	0.1610	0.0419	3.84

Note: Estimates of Equation 3.26. Commute time is measured in hours.

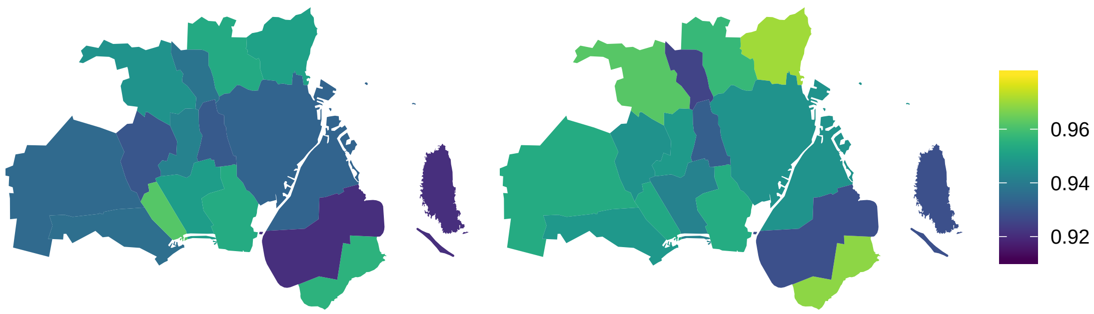
Figure 3.12: Model fit: Share of singles moving work region or being unemployed



(a) Move work region (b) Unemployment

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure 3.13: Model fit: share of singles in employment by home region



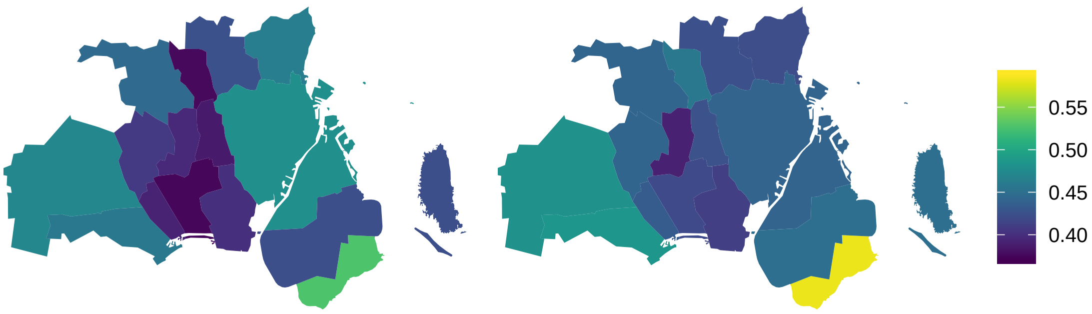
(a) (obs.) (b) (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

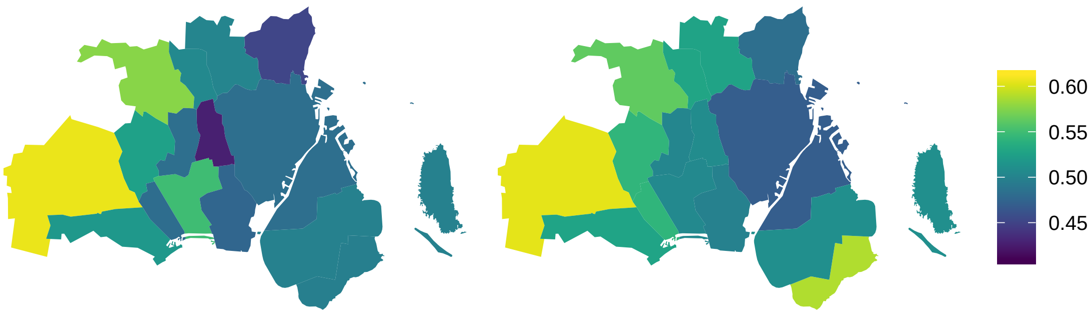
Commute costs

The ability to predict the spatial variation in commute times is shown in Figure 3.14. The top panel shows the commute time by home region. There is a slight underprediction for those who live in Copenhagen. I.e. the model cannot explain why these individuals would want to commute longer now that they pay a high price for housing there and there are many jobs available within a short distance. For Dragoer the model overpredicts the commute time, but in general the error margins are quite low. For work regions, the model predicts commute time to be more consistently increasing in distance from the Copenhagen center. In the data, however, the picture is not as clear-cut, especially for the regions on the border of the GCA. This can be explained by the fact that it is more likely that someone in the data who works there lives in the region Rest of Zealand. That region, however, is characterized by having relatively long commutes in the model, because it is a mix of commute times from all the municipalities in that region. Hence, the model does not distinguish between living just at the border between GCA and Rest of Zealand or far from that border. In reality though, people do live just outside the border of GCA and work in e.g. Albertslund, and the associated commute time is not very high. Making the Rest of Zealand region more disaggregate could improve on this in future research.

Figure 3.14: Model fit: commute time (hours) by home and work region



(a) Home (obs.) (b) Home (sim.)



(c) Work (obs.) (d) Work (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Job taste

While the commute cost parameters do help predicting exactly which work region to choose conditional on the home location, they cannot be used to distinguish between two regions located equally close to the home. Apart from the differences in income across regions, the job taste function in (3.21) serves this purpose despite not modelleing job search behaviour explicitly as in Chapter 2. As Table 3.14 shows, a one unit increase in job density increases utility by 2.2 units. So while Copenhagen does offer higher wages than most other parts of the country, the predicted wage is lower than the one in e.g. Ballerup, but still many more people work in Copenhagen. In that sense Copenhagen with a job density normalized to 1.0 outperforms Ballerup with a job density of 0.1036 for low-skilled, 0.1517 for medium-skilled and 0.0943 for high-skilled jobs. That attracts workers to Copenhagen.

Distinguishing by home region Figure 3.15 shows that the model overpredicts the share who lives on the western border between Rest of Zealand and GCA and commute to

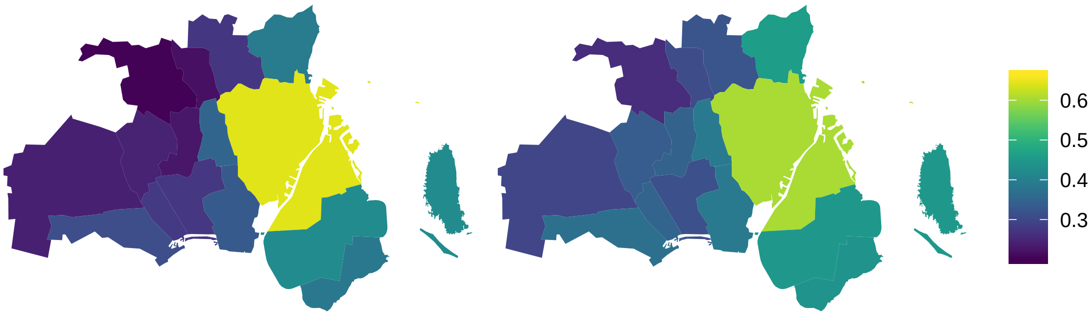
Table 3.14: Regional Amenities of Work Region

Parameter	Estimate	Standard Error	t-statistic
ψ_{dens}	2.2296	0.0180	123.73

Note: Estimates of Equation 3.21. Job density is defined as number of jobs by education group for the decision maker and normalized by the level of the education group in Copenhagen. Summary statistics of job density by region and education is available in Table C2.

Copenhagen, because it cannot explain why a relatively large share would rather commute to Rest of Zealand where commute time is much higher and wage income lower according to the model inputs.

Figure 3.15: Model fit: share of singles working in Copenhagen by home region



(a) (obs.) (b) (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

With the structural parameters from single model at hand, the next section goes over the parameters from the couple model.

6.4 Preference parameters for couples

Having estimated the model for singles and shown that the model provides a reasonable fit, I use these estimates to compute the outside option (the expected value) for each spouse in the couple dataset. For each household I then calculate the difference in outside options between the two spouses, use this as data[22] and estimate the static model for couples.

Residential moving costs

The parameters associated with moving residence are displayed in Table 3.15. They all have the expected signs such that older households and those with children are more

[22] As of now I do not account for the fact that the outside options are subject to variability as the coeffcients in the single model are estimated, not observed.

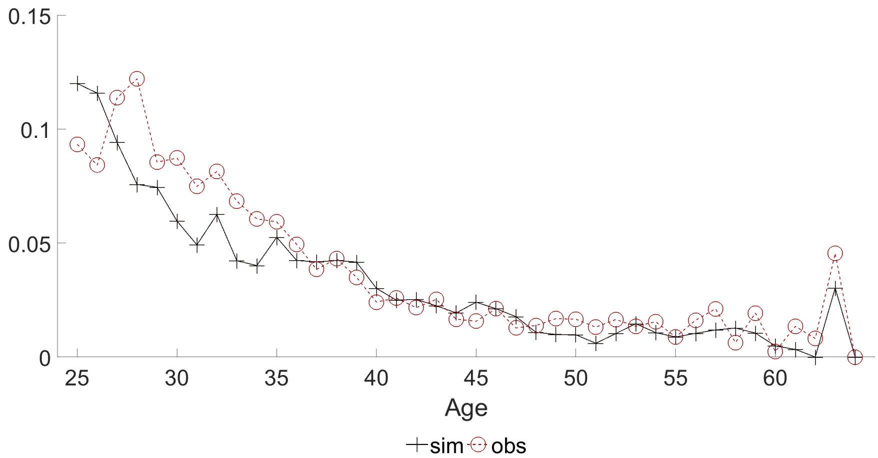
Table 3.15: Utility Cost of Moving Home Region

Parameter	Estimate	Standard Error	t-statistic
γ_0	3.3320	0.1499	22.23
γ_a	0.0740	0.0037	19.96
γ_{kids}	0.2712	0.0622	4.36

Note: Estimates of Equation 3.23.

reluctant to move. Compared to the singles, γ_0 is higher while the coefficient on age, γ_a, is slightly lower. On the other hand, the effect of children, γ_{kids}, is almost three times as large for singles. For a standard couple with no children the moving costs are 6.3 utils compared to 5.4 for singles. This corresponds to 453,890 DKK[23]. This reflects that couples are generally less likely to move over the life cycle. As Figure 3.16 shows, the probability of moving by age is captured well for couples above age 35. For the younger households the model underpredicts the moving probability a bit.

Figure 3.16: Model fit: Share of couples moving home region



Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Regional amenities

With a fairly satisfying fit of the moving probabilities, the regional amenities are supposed to help predicting *where* couples would like to live and move to (commuting costs have a similar role). The parameters are presented in Table 3.16 and they all have the expected signs: more nature, more restaurants and less property crime is attractive. In monetary terms, a standard couple is willing to pay 13,523 DKK for 100 extra restaurants. Com-

[23] Computed as $6.3/(0.5\cdot(0.176-0.00124\cdot 30)+0.5\cdot(0.176-0.00124\cdot 30))\cdot 10,000$ as money is measured in 10,000s

Table 3.16: Regional Amenities of Home Region

Parameter	Estimate	Standard Error	t-statistic
τ_{nature}	0.0847	0.0038	22.54
τ_{rest}	1.8771	0.1498	12.53
τ_{thefts}	-103.7621	7.1319	-14.55

Note: Estimates of Equation 3.20. Nature is nature capital index. Restaurants in 1,000s and thefts is number of thefts per inhabitants in the region. See Table C1 for summary statistics of amenities by region.

pared to the singles, couples get 2.4 more utils from a unit increase in the nature capital index, but only 0.8 of the utility gain from another restaurant in the neighborhood. On the contrary, they get 2.7 more disutility from an increase in the number of thefts. This may be related to the fact that couples are more likely to have children than singles. Having children means the household might use nature more, use restaurants less and be more nervous about bringing up a child in a shady neighborhood. To explore this further, one could interact the parameters with a dummy for having children.

One way to review this in the current model set-up is by looking at the probability of living in Copenhagen. As Figure 3.17a shows, this probability is declining in age and much faster than for the singles presented in the previous section. Looking at Figure 3.17b, the couples with children are indeed less likely to live in Copenhagen, at least until they reach age 50 where most households no longer have young children. Since Copenhagen is characterized by more thefts, more restaurants and less nature than several other regions, this is at least consistent with the idea that couples' preferences for amenities change when they have children.

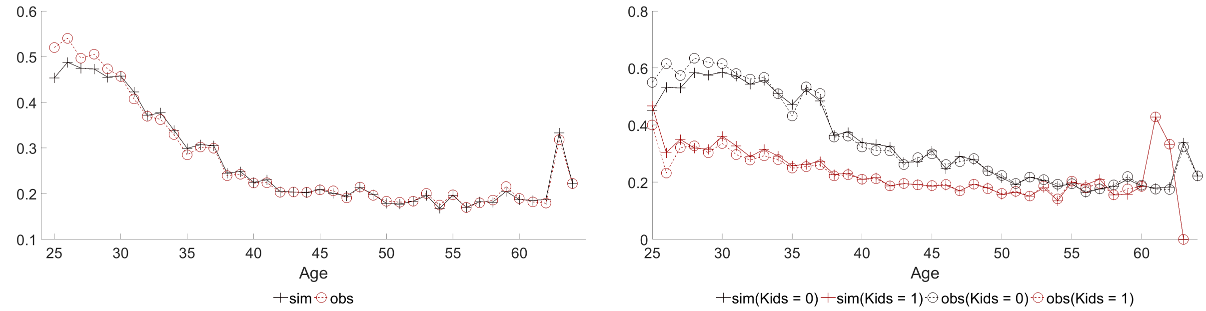
Overall, the fit for living in Copenhagen over the life cycle is very satisfying. For comparison, Figure 3.18 plots the probability of living in Rest of Zealand and shows that the life cycle trend is reversed: as the couple ages, the probability of living in Zealand increases and the gap between families with and without children is less pronounced than for Copenhagen. Disregarding the life cycle perspective, Figure 3.19 also illustrates that the model predicts the spatial allocation of households well[24].

Job moving, unemployment and income sharing

Now that the parameters indexing the decision to move and where to live have been presented, Table 3.17 shows the estimates for the job moving costs in (3.23). As for the singles, the parameters have the expected signs, but are of much larger magnitudes. Thus, a standard couple person who was not unemployed in $t-1$ gets a disutility of 8.8 if (s)he decides to move work region, whereas the corresponding single person only gets 4.2. This corresponds to job moving costs of 643,006 DKK, i.e. much higher than the singles. This

[24]See Appendix F.2 for supplementary model fits.

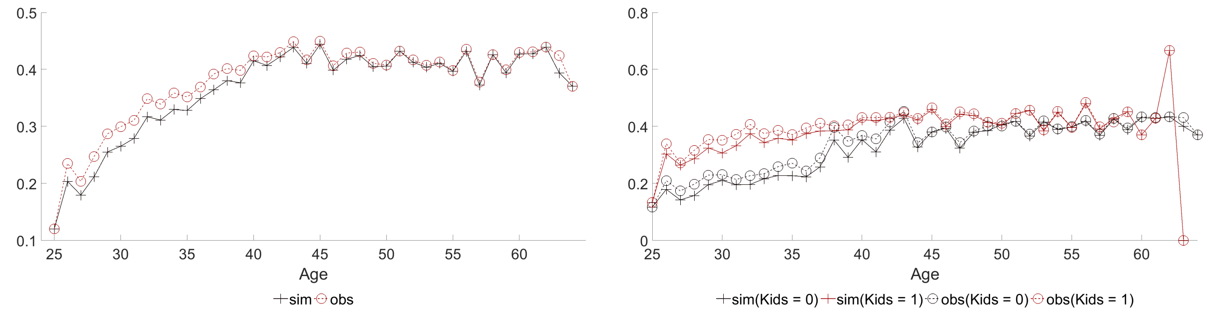
Figure 3.17: Model fit: Share of couples living in Copenhagen



(a) Aggregate (b) By children

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

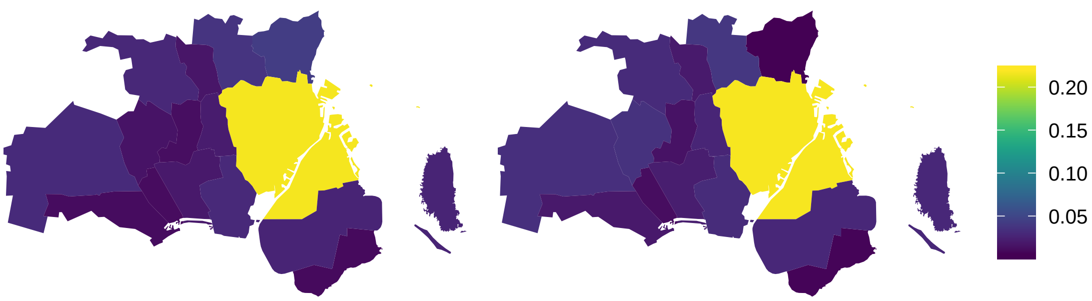
Figure 3.18: Model fit: Share of couples living in Rest of Zealand



(a) Aggregate (b) By children

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure 3.19: Model fit: share of couples living in each home region



(a) (obs.) (b) (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

points to the fact that couples are more constrained in their mobility decisions because they cannot as easily move their residence if they decide to move work region as their partner might not benefit from that. Figure 3.20a illustrates that the work-related moving

Table 3.17: Utility Cost of Moving Work Region

Parameter	Estimate	Standard Error	t-statistic
o_0	5.5717	0.1017	54.79
o_a	0.0823	0.0024	33.76
$o_{\varnothing}$	-9.6011	0.1219	-78.75

Note: Estimates of Equation 3.24.

Table 3.18: Utility Cost of Working

Parameter	Estimate	Standard Error	t-statistic
α_{work}	3.5572	0.0863	41.22

Note: Estimates of Equation 3.25.

probability is well-captured both for males and females.

Whereas the job moving costs were negative for all singles who were unemployed in $t-1$, this is not the case for the couples: individuals above age 49 have positive utility costs associated with going from unemployment to employment. This means couples are more reluctant to take a job when they are unemployed compared to singles despite the fact that disutility from work is of the same size as singles, cf. Table 3.18. This can be explained by income sharing in the household. As Table 3.19 shows, they get positive utility from the spouse's income. The coefficent on their own income is normalized to one, so an estimate of 0.3621 means a couple individual considers 36.21 percent of the spouse's income as if it was his or her own. Since earning a higher income in employment compared to unemployment is the main driver for taking a job, that incentive is reduced for someone who has a working spouse. On average though, the unemployment probabilities are low for both genders, and lower than for singles, but the fit is quite good, cf. Figure 3.20b. This can be explained by couples being higher educated than singles, cf. Section 6.1, since the wages are higher for high-skilled types. This improves the incentives for taking a job, all else equal.

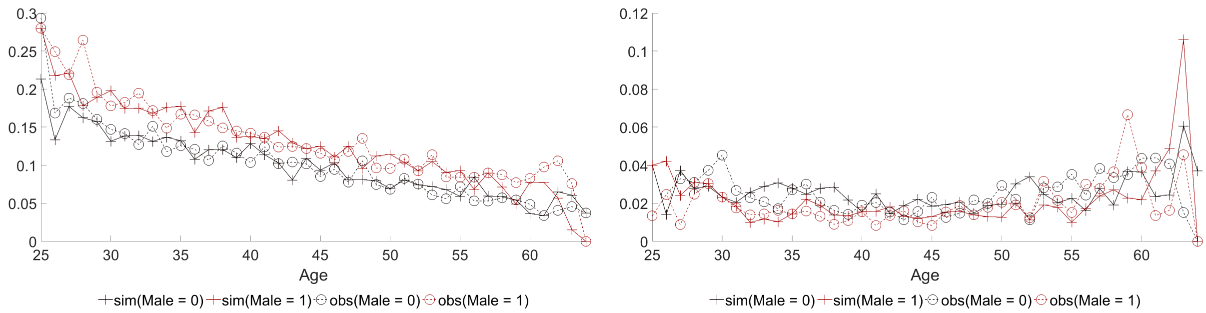
Commute costs

Next, the parameters indexing commute time are shown in Table 3.20. η_0 is almost double the value for singles, but the effect of having children has the opposite sign: for women,

Table 3.19: Income Sharing

Parameter	Estimate	Standard Error	t-statistic
χ	0.3621	0.0421	8.59

Figure 3.20: Model fit: Share of couple individuals moving work region and in unemployment



(a) Work move

(b) Unemployment

Note: Choice data is simulated for 1 period ahead using the states in the observed data. $Male = 1$ is men and $Male = 0$ is women.

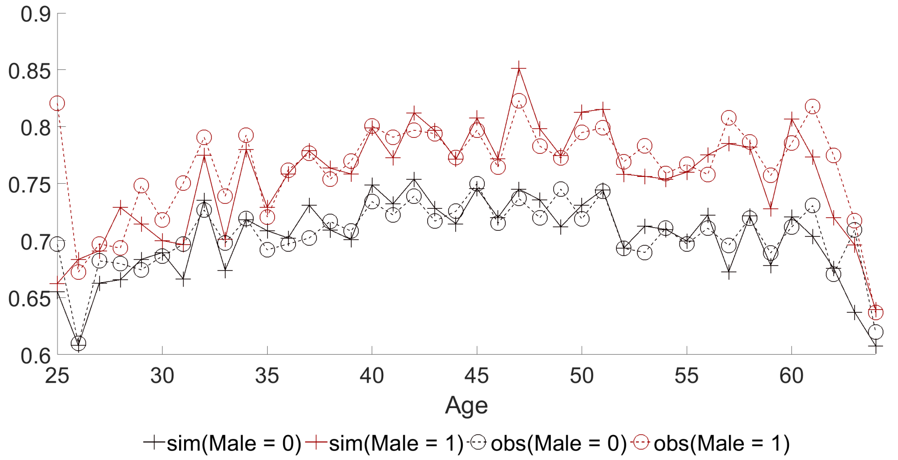
Table 3.20: Utility Cost of Commuting

Parameter	Estimate	Standard Error	t-statistic
η_0	0.6789	0.0321	21.15
η_{kids}	-0.0052	0.0523	-0.10
$\eta_{kids} \cdot \mathbb{I}[male]$	-0.3131	0.0610	-5.13

Note: Estimates of Equation 3.26. Commute time is measured in hours.

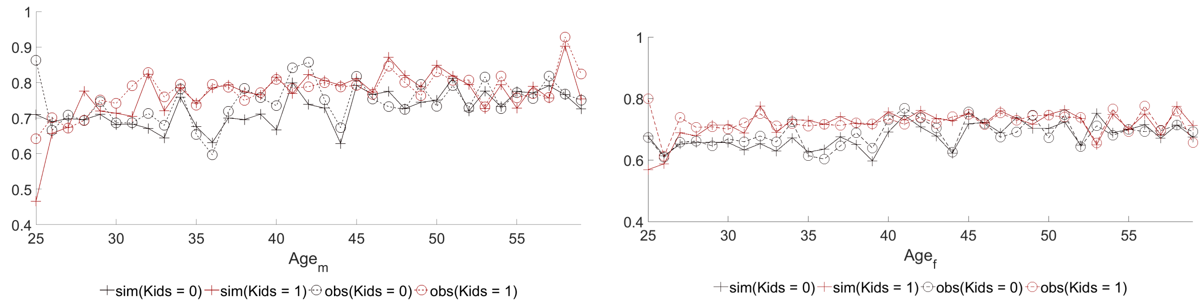
there is a negative but insignificant effect of children on her commute time, while for males the effect is significantly negative. This means men's disutility of commuting is lower when they have children. This can be explained by the following: men with children who are in a couple per definition have a wife with whom they can share the responsibilities for e.g. picking up the child from school. This is not the case for singles who are therefore more constrained in their work location choice conditional on their home if they want to be within a certain distance from the child's school. Men are therefore expected to commute longer when they have children, all else equal, according to the model. Figure 3.21 shows that the model nicely fits that commute time exhibits a slight inverse U-shape over the life cycle for both men and women and that women commute a bit less than men. Nevertheless, as Figure 3.22a and Figure 3.22b illustrate, both men and women commute a bit longer when they have children. These plots, however, do not condition on anything else than age and children, so the fact that even women commute further when they have children should be understood as a combination of the fact that they also live elsewhere than couples with no children. The structural parameters indicate that *all else equal* she would not change her commute time at the arrival of children, but the man would. A standard couple individual with no children would require a compensation of $48,912$ DKK if (s)he should increase the commute time by one hour.

Figure 3.21: Model fit: Average commute time (hours) for couples



Note: Choice data is simulated for 1 period ahead using the states in the observed data. $Male = 1$ is men and $Male = 0$ is women.

Figure 3.22: Model fit: Average commute time (hours) for couples by children



(a) Male by children

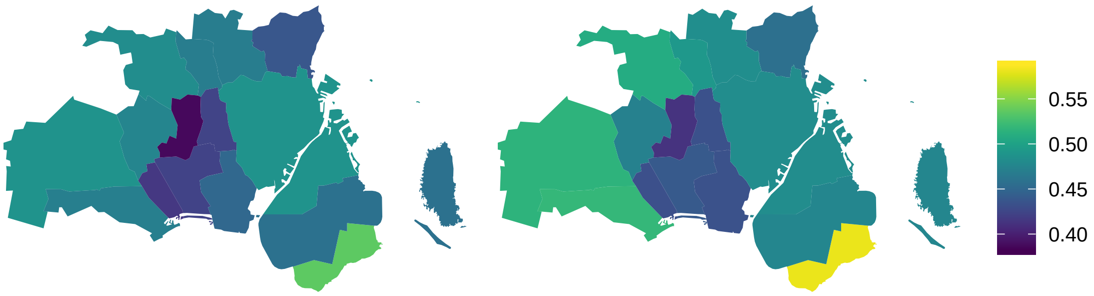
(b) Female by children

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

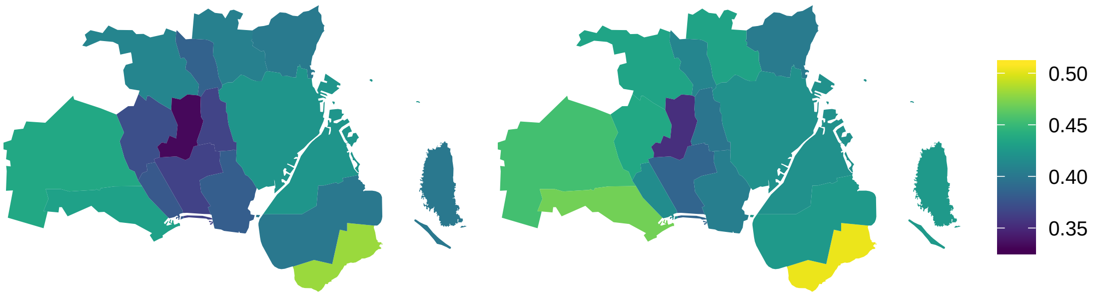
Considering the distribution of average commute time across home regions, Figure 3.23 shows that the fit of male and female commute times across space is good. For both genders, the commute time is predicted to be the longest for households who live on the border of the GCA and especially for Dragoer. The shortest commute time is in Glostrup which is also true in the data.

The difference in commute time within the household is less well predicted, cf. Figure 3.24, but in general the male commute time is no more than 5 minutes higher than the female's both in the data and the simulation. This is a consequence of the sample selection. It is therefore hard to tell if the model would be able to predict the commute time differences had I used a dataset where the regions were more disaggregate and hence would have more variation in commute time differences.

Figure 3.23: Model fit: commute time (hours) by home region



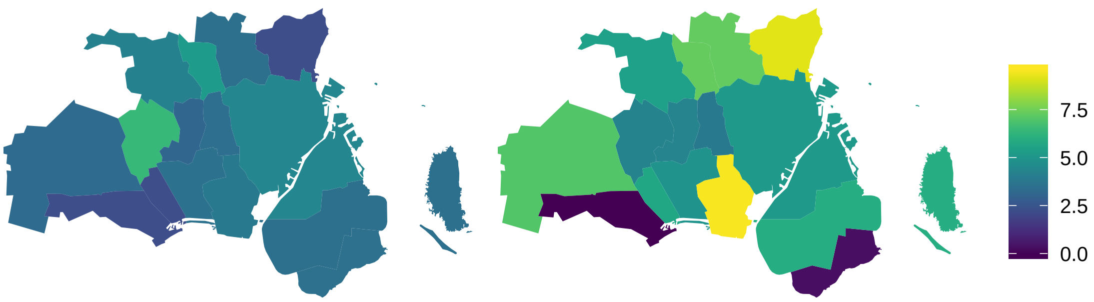
(a) Males (obs.) (b) Males (sim.)



(c) Females (obs.) (d) Females (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure 3.24: Model fit: Difference in commute time (male-female, minutes) by home region



(a) (obs.) (b) (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Table 3.21: Regional Amenities of Work Region

Parameter	Estimate	Standard Error	t-statistic
ψ_{dens}	3.3676	0.0312	108.07

Note: Estimates of Equation 3.21. Job density is defined as number of jobs by education group for the decision maker and normalized by the level of the education group in Copenhagen. Summary statistics of job density by region and education is available in Table C2.

Table 3.22: Bargaining Weight Parameters

Parameter	Estimate	Standard Error	t-statistic
Υ_0	0.0576	0.0049	11.69
Υ_1	-0.0008	0.0005	-1.41

Note: Estimates of subsection 3.10.

Job taste

Compared to the singles, couple households place more value on working in regions with a higher job density, cf. Table 3.21. The share of couples working in Copenhagen is quite similar for the two genders, cf. Figure 3.25. However, there is a more clear gradient by children for the females. Looking instead at the lower panel of Figure 3.25, the model also fits well that highly educated individuals are more likely to work in Copenhagen whose supply of high-skilled jobs is much higher than anywhere else. The model also captures that the distinction between schooling levels in the likelihood of working in Copenhagen is more pronounced for males than females. The same can be said about the fit of the spatial distribution of work places for each gender, cf. Figure 3.26. The job moving costs have a significant role in capturing this pattern.

Bargaining weight

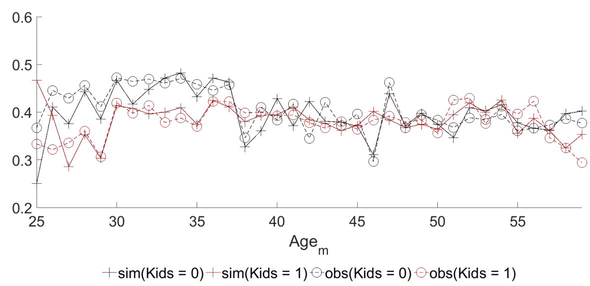
The last set of parameters are the ones indexing the bargaining weights. They are presented in Table 3.22. Only Υ_0 turned out to be significantly different from zero. It is positive implying that a couple where both spouses have the same value for their outside options have a weight of $1/(1 + exp(-0.0576)) = 0.514$. I.e. the households places a 0.514 weight on the woman's and 0.486 on the male's value of a given choice on home and work locations for both spouses. The effect of the difference in outside options, Υ_1 is negligible, though negative. The bargaining weights are therefore also very similar across home regions, but the difference in outside options between the two spouses is also very close to zero for most home regions according to Table 3.23. The wives have a higher outside option than their husbands on average in 8 out of 19 regions.

Table 3.23: Summary statistics of difference in outside options (female minus male) by home region

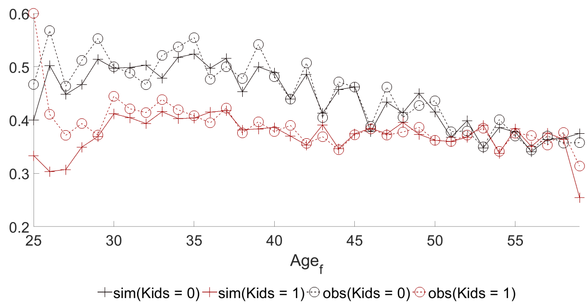
Home Region	Mean	S.e.	Min.	Max.	N
Copenhagen	0.0634	2.7298	-27.51	25.80	7327
Frederiksberg	-0.1588	3.0389	-26.07	21.38	1564
Ballerup	0.3735	3.5383	-16.61	23.77	978
Broendby	0.0618	2.7285	-8.56	25.64	644
Dragoer	0.0246	2.1951	-6.73	7.13	345
Gentofte	-0.0201	3.0453	-22.89	23.07	1452
Gladsaxe	-0.0405	2.8518	-25.53	14.46	1271
Glostrup	-0.0223	3.0159	-8.87	24.31	439
Herlev	-0.0973	2.7425	-7.53	22.93	571
Albertslund	0.0202	2.8918	-20.37	10.43	533
Hvidovre	-0.1381	2.5202	-23.01	8.47	1010
Hoeje-Taastrup	-0.0729	3.3818	-23.50	28.33	1034
Roedovre	0.0165	3.1681	-24.82	20.23	707
Ishoej	0.0600	2.5190	-7.91	12.18	390
Taarnby	0.0078	2.8553	-29.73	8.58	902
Vallensbaek	0.0879	2.4124	-6.62	9.36	393
Rest Of Zealand	-0.0032	3.0366	-28.12	25.38	13058
Funen	-0.0360	2.4454	-7.59	7.36	278
Jutland	-0.3137	3.1776	-24.79	18.27	962

Note: Outside options for each spouse computed as the expected value.

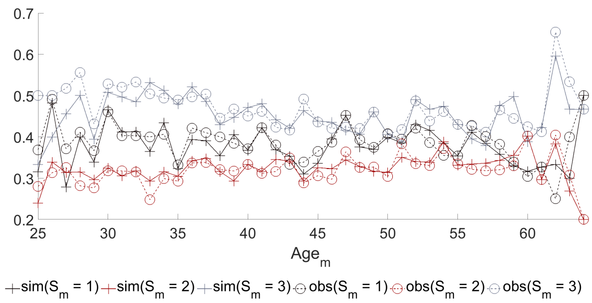
Figure 3.25: Model fit: Share of couples working in Copenhagen



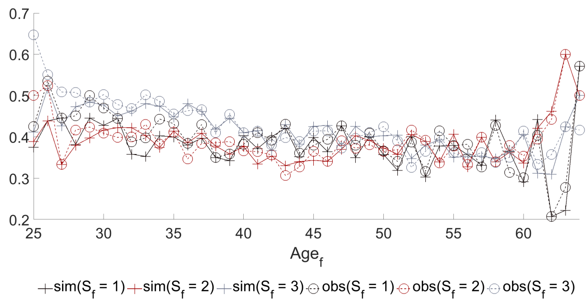
(a) Male by children



(b) Female by children



(c) Male by schooling



(d) Female by schooling

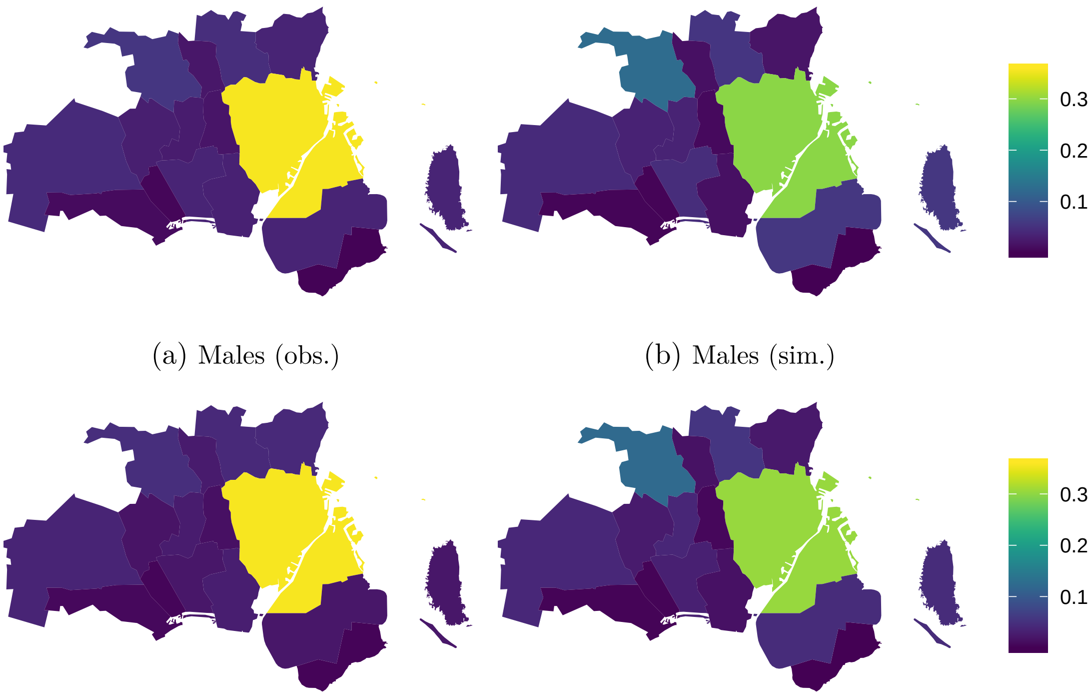
Note: Choice data is simulated for 1 period ahead using the states in the observed data. $S = 1$ is low education, $S = 2$ is medium education and $S = 3$ is high education. $Kids = 0$ means no children in the household, $Kids = 1$ means at least one child in the household.

Gender wage gap

These results allow me to give an answer to whether it is discrimination of the wife within the household that can explain the gender wage gap documented in Section 3. The answer is no under the current model set-up, though that set-up and thus the answer should be considered preliminary. If anything, the male is the one being discriminated since he has a lower bargaining weight on average. In the estimation of the current model, wages were kept fixed across genders conditional on the state variables. I.e. even if there might be gender differences in wages not explained by state variables, including discrimination on the labor market, I disregard those here and attribute any gender wage gap to gender differences in work location choices conditional on background characteristics. The model does indeed predict that men and women earn very similar wages on average across home regions, cf. Figure 3.27, because there is no real difference in the choice of work location between genders in the sample and they are very similar in terms of observables. The higher bargaining weight for women is therefore a reflection of them getting the *right* to choose the slightly shorter commute than their husbands.

The model predits that females get a fair weight, or even a bit more, in the household decision process, so the observed gaps from Section 3 cannot be attributed to intra-

Figure 3.26: Model fit: share of couples working in each region



(a) Males (obs.)

(b) Males (sim.)

(c) Females (obs.)

(d) Females (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

household discrimination of the wife. Explanations of the gap is rather that employers do discriminate women or that the model is not flexible enough. The latter is a valid critique since I do not model choice of occupation or hours worked. Even for the same age and degree of education, there are considerable differences in wages across occupations, and men and women may select differently on that margin. To explore this, the model would have to be extended with an endogenous choice of occupation as Buchinsky et al. (2014) did in an individual decision model. The same goes for hours of work, but both are left for future work. Other important points to make are both that the sample of households used in the current estimation is not necessarily representative of the entire country and many of the couples are more similar in terms of background characteristics and hence outside options than can be said about the couples in the population data. Having more variation in differences in commute times and outside options may therefore change the results. Moreover, the lack of dynamics in the model is a serious limitation since it ignores that couples anticipate not moving again next period. The immobility is because moving is costly and involves an irreversible investment. In a static world the wife might agree to move with the husband to another region even though her outside option may deteriorate if she gets less attached to the labor market as a consequence. The household does not take this future deteroration into account.

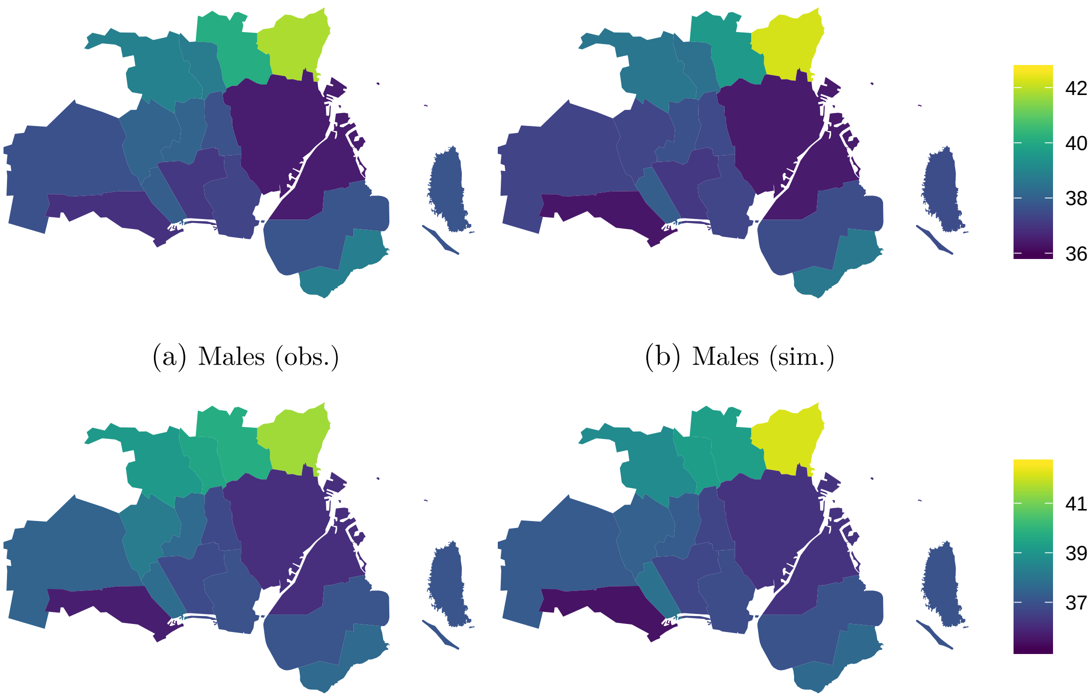
The risk of divorce is also not taken into account in the estimation above. This means couples assume they can share income in all relevant periods, namely the current one. They do not care about the risk that in the future they might divorce and therefore have to live off of their own incomes and that the income prospects are affected by how their outside options have evolved during the marriage. The conclusion about bargaining weights and the effect of the co-location trade-off on the wage gap should therefore be interpreted with caution until these effects have been modelled.

6.5 Counterfactual policy experiment

The main advantage of estimating structural models compared to reduced form models is the possibility to run counterfactual experiments, where the economic environment is changed exogenously, but individuals' preference parameters are kept fixed at the estimated values. This gives insights into the effects of such a change in the environment on individuals' behaviour. For dynamic models this usually requires solving the structural model once given the new economic environment and keeping other parameters fixed.

One situation where counterfactual experiments can be run on dynamic models without having to do the full backwards induction is for short-run and unexpected interventions which do not have general equilibrium effects. Imagine a policy that is being implemented at time t and came as a complete surprise. Moreover, the policy is only in place for n periods. After those n periods, the economic environment returns to the $t-1$ version and everything is as before the intervention except individuals may have ended up in other states in period $t+n$ than under the original policy regime. The essential point here

Figure 3.27: Model fit: Predicted income (10,000 DKK) of couples by home region



(a) Males (obs.)

(b) Males (sim.)

(c) Females (obs.)

(d) Females (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

is that *given* the states at $t+n$, the structural decision rule has not changed (and the state values at $t+n$ are spanned by the values in $t-1$). Acknowledging that means the researcher only has to solve the structural model n periods ahead. From period $t+n$ onwards, the initial CCPs are again valid for describing how individuals will act in future periods. For the dynamic model in Section 4 the number of state points I would have to solve the model for is still huge even if I assumed the policy was only in place for, say, two periods. In the full dynamic set-up with 19 home regions, 20 work regions, $T = 40$, 9 number of age differences, 3 education levels and a dummy for having a child there are 935,712,000 state points and the choice space is of dimension 15,200. The state space for couples in the post policy experiment world would be the original state space lowered by a factor 40/2, i.e. the dimension would be 46,785,600 and 711,141,120,000 when multiplying by the choice space. This would be the number of loops I would have to go through in order to solve for the expected values for period $t+1$ and $t+2$. When the individuals wake up in $t+3$, they are back in the old policy regime and therefore start acting according to the old policy rule. Running through this number of states, however, is still extremely time-consuming and currently infeasible.

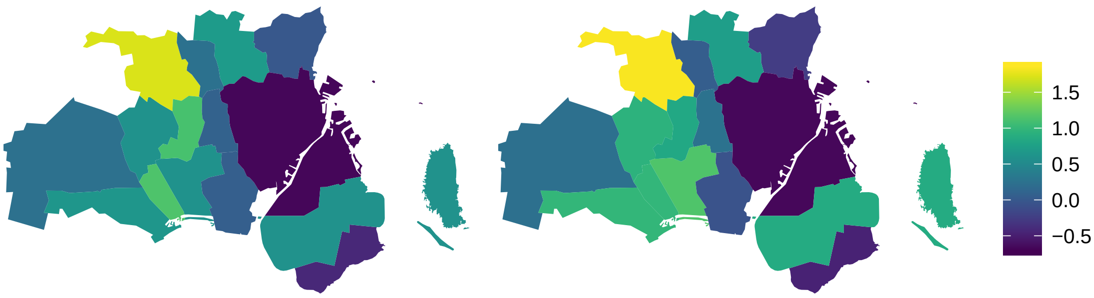
This is all ignored in the static model. I therefore run a counterfactual where I double the job density in Albertslund for all education groups and consider the effects on residential and work location choices for each spouse and the difference in commute times within the households. To summarize the attributes of Albertslund, it is characterized by fewer thefts, fewer restaurants, lower property prices, and shorter commute time than the rest of the regions on average (which is dominated by Copenhagen municipality), cf. Table G7 which shows standardized measures of the amenities. As a work region, Albertslund originally has a lower job density of all types of jobs and higher predicted wages. To interpret the effects of the policy experiment, I compare to the predictions from the original model, i.e. where the job density was at its original value. First off, Figure 3.28 shows how the distribution of chosen work regions for males and females, respectively, change in response to the policy[25]. For both genders, the elasticity of the probability of working in Copenhagen with respect to changing the job density in Albertslund is negative. When the job density in Albertslund changes by 1 percent, 0.6742 percent fewer men choose to work in Copenhagen. For females this number is 0.6658. For both men and women the probability of choosing unemployment increases as the elasticity is 1.1432 and 1.5417, respectively. For working in Albertslund itself, the probability increases by 0.5946 for men and 0.9330 for women. The largest positive effect is seen on the propensity to work in Ballerup though with an elasticity of 1.7099 for men and 1.8317 for women. This may seem surprising, because the amenities of Ballerup itself were not changed. This is where the simultaneous modelling of home and work locations and commuting time show its strength.

Looking at Figure 3.29a, the probabilities of choosing each home location namely

[25] Tables with the numbers are available in Appendix G.2

also changes. With an elasticity of 0.8860 more couples are residing in Albertslund, while Copenhagen has a negative elasticity of 0.0273. The higher number of couples who choose to live in Albertslund also means more households reduce the commute time to Ballerup where the wages are the highest in the GCA for all skill types. The reason fewer couples worked there before the policy change is that house prices were too high, and commuting thereto from other regions would take too long. The fact that Albertslund itself now becomes attractive as a work region means it is beneficial for the couples to reside there at low house prices. They can let at least one spouse work there, with the option of letting the other spouse commute to Ballerup and earn a high wage. Especially the females show a stronger tendency to work in Ballerup. The largest positive effect on home locations is on the probability of living in Ishoej (elasticity of 0.9384) which shares borders with Albertslund. Ishoej has much fewer thefts than Albertslund, approximately the same number of restaurants, worse nature and a bit higher prices. The average travel time from Ishoej is lower than in Albertslund though. Couples therefore seem to exploit that the travel time between Ishoej and Albertslund is short so they can live in Ishoej and benefit from the lower crime levels there at about the same prices.

Figure 3.28: Elasticity of distribution of work locations



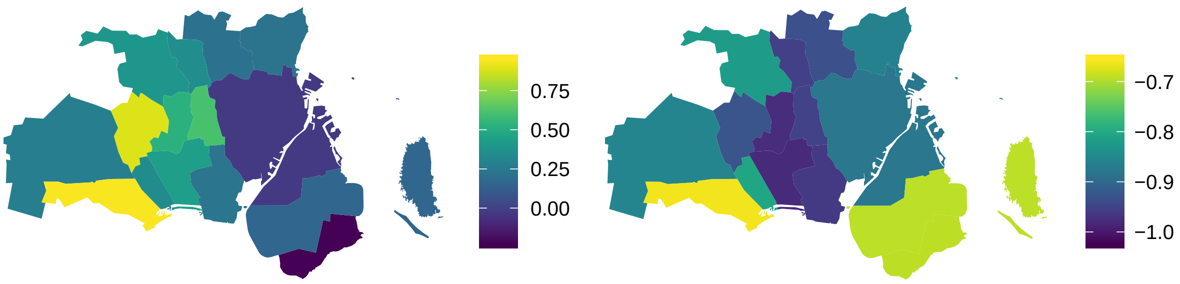
(a) Males (b) Females

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Due to the shifts in work locations, the intra-household difference in commute time is also affected. Figure 3.29 shows the elasticity of commute time difference by home region. Generally, the effects are small, but this is also a result of the generally low travel times within the GCA. The largest negative effect in minutes is seen in Glostrup, where the average difference is reduced by 8.8 minutes corresponding to an elasticity of -0.98. Glostrup also shares a border with Albertslund and also experienced an increase in the probability of couples residing there.

For the wage growth by each spouse Figure 3.30 shows that for all home regions, couples experience an increase in their wage income and in 13 out of the 16 regions in GCA, the female wage growth is higher than the males'. The lowest wage growth is for males when they reside in Dragoer, while the most significant boost in wage income

Figure 3.29: Elasticity of home location and intra-household commute time difference (male-female)

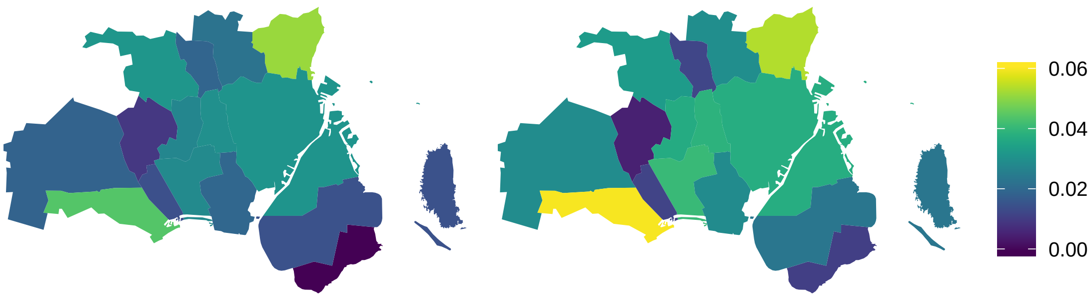


(a) Home (b) Male-Female commute time

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

goes to wives who live in Ishoej. They have an average wage growth of 6.0 percent (corresponding to an elasticity of 0.06) because they are now motivated to commute to Albertslund where the wages are higher than in Ishoej. The same effect appears to men, but by 4.4 (corresponding to an elasticity of 0.04) percent.

Figure 3.30: Income growth by home location (%)



(a) Male (b) Female

Note: Choice data is simulated for 1 period ahead using the states in the observed data. To get the elasticity, the numbers should be divided by 100.

The increase in wages across the home regions primarily stems from the fact that spouses are more likely to work in Albertslund compared to before. This is evident because among those who live in Albertslund, the wage growth is much smaller of 0.01 percent for males and 0.00 percent for females. The slight growth in wages for males in Albertslund that is observed after all is because some of the residents of Albertslund now choose to work there instead of commuting to one of the surrounding regions where wages can be lower. In conclusion, the women experience an average wage growth of 2.2 percent and the men 1.9 percent (corresponding to an elasticity of 0.022 and 0.019, respectively). This means, the policy would reduce the gap between men and female wages marginally according to the predictions of the model.

6.6 Dynamic structural model for singles

Estimating the dynamic version of the model outlined in Section 4 for couples was currently infeasible due to the need for evaluating the value function for all observed states and the entire choice set. Estimating the dynamic model for singles is not as computationally demanding though because the state space is smaller. In this section I therefore present the estimates of such a model where I have estimated the initial CCPs with a reduced-form Logit and use these estimates to approximate future value functions two periods ahead. I still fix the parameters in Table 3.8 and set the discount factor $\beta = 0.95$.

Precomputing future states

As explained in section 5, I need to integrate over period $t+1$ and $t+2$ future states for singles to estimate the model. The potentially random future states in this model are children. For now, however, I do not consider it a random state, but rather assume households have static expectations such that they expect the number of children at $t+1$ to equal the number of children at t. In that sense children arriving or moving out of the home come as a surprise and the household does not choose locations with the expectation of having more of fewer children in the future in mind. I do this because it simplifies the estimation as it allows me to precompute all future states of the household before the estimation takes place. Even though the number of children is implemented as a dummy for having children such that the evolution of this dummy can only take three values (-1 if going from having children to not, 0 if staying in the same child state in $t+1$ as in t and +1 if going from having no children in t to having children in $t+1$), this still increases the number of potential future states by a factor three compared to assuming no change in the child state.

Initial conditional choice probabilities

In order to estimate the model using the method outlined in section 5, I recover an estimate of the CCPs, particularly the CCPs of the imposed choices in $t+1$ and $t+2$ for singles. I choose the imposed choice for $t+1$ to be $rh_{t+1} = Copenhagen, rw_{t+1} = \varnothing$ and the same for $t+2 : rh_{t+2} = Copenhagen, rw_{t+2} = \varnothing$. To estimate these CCPs, I estimate a conditional Logit for all alternatives of the form

$$CCP(d_{it} = l|w_{ilt}) = \frac{\exp(w_{ilt}\Pi)}{\sum_{m=1}^{L}\exp(w_{imt}\Pi)}, \tag{3.41}$$

where l is the alternative considered by individual i, w_{il} is a $1 \times KW$ vector of regressors that must vary over l and potentially over i and Π is a $KW \times 1$ vector of coefficients.

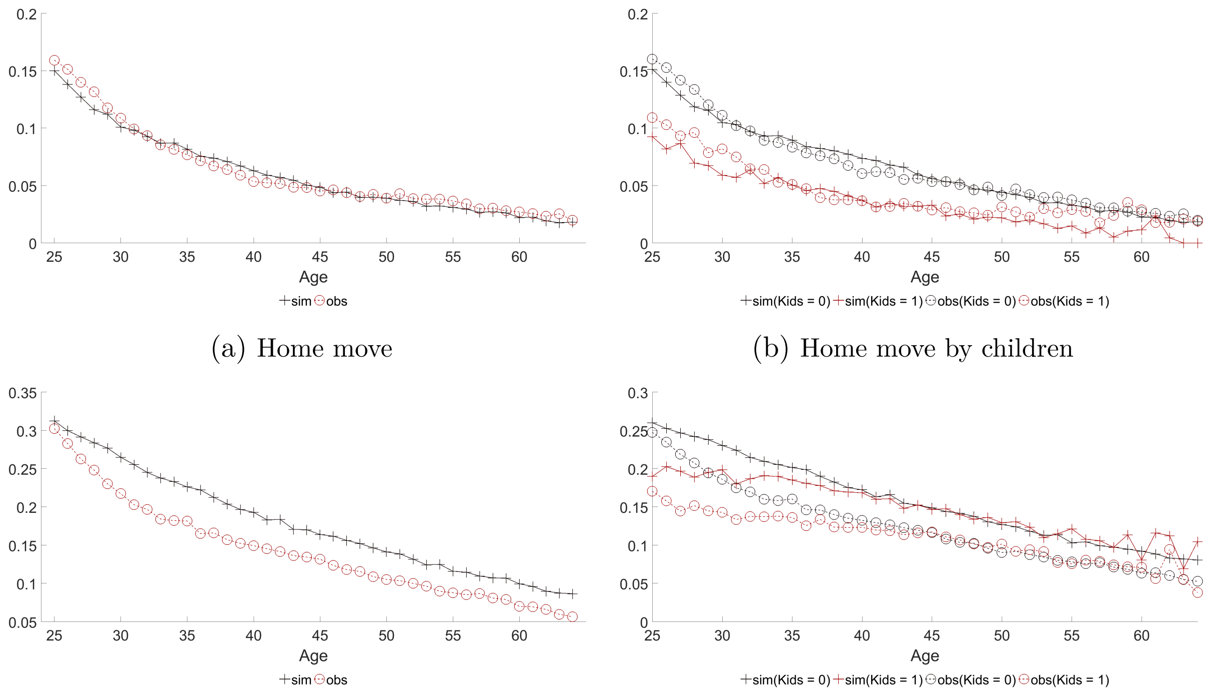
Alternatively, I could have used a frequency estimator and compute the number of observations in each $(state, choice)$ cell which amounts to 34,656,000 cells. The problem with that method is that even though I work with population data, only 1.08 percent of the cells are observed in the data, because so many transitions are very unlikely. The Logit model is a way to use our observed cells and smooth over the rest of the $(states, choice)$ space to get predictions for non-observed cells too.

The focus in this part of the estimation is only prediction, not causalal interpretation, though I aim at including predictors with high predictive power to get as precise predictions as possible. To estimate the model, I use the same 1 percent random subsample from the population of singles as in the estimation of the static models, where w_{ilt} contains the following variables: nature-capital index, number of restaurants (measured in 10s), square meter prices in real 2011 10,000 DKK, number of square meters in 100s, number of victims of property crime in 100s, travel time in quarters between home and work location associated with the choice, interaction between travel time in quarters and dummy for having a child, a dummy indicating if the chosen home region considered is the same as the previous home region and this dummy interacted with age, a dummy indicating if the chosen work region considered is the previous work region and this dummy interacted with age, a dummy indicating if the work region considered is Copenhagen and lastly a dummy indicating if the work region considered is unemployment. The coefficient estimates for singles can be seen in Table G1 and most have the expected signs.

As mentiond above, prediction is the most essential part of this exercise. Hence, I simulate decisions from the estimated CCPs using the state variables in the estimation dataset. I.e. I keep the states of the individuals fixed, draw a random uniform number between 0 and 1 for each individual and simulate a decision using the Logit CCPs as the decision rule. This enables me to check the model fit and I do so for selected moments in the figures below.

The fit in terms of moving home probability over the life cycle is very good, also when conditioning on whether the indivdual has children or not, cf. Figure 3.31a and Figure 3.31b. Looking at the probability of moving job region over the lifecycle, Figure 3.31c shows that the reduced form model overpredicts this probablity a bit. The same is seen when conditioning on schooling. Since the imposed choice involves living in Copenhagen, it is informative to take a look at the predicted probability of living in exactly Copenhagen. This is what Figure 3.32 does. The fit is very good over the life cycle and also when conditioning on having children or level of education. Lastly, I look at the model fit of the probability of being unemployed, cf. Figure 3.33. Again, the fit is good over the life cycle, though the model slightly underpredicts for the younger cohorts and overpredicts for the oldest cohorts. Generally though, it looks good also when separating by whether or not the individual has children or the level of schooling.

Figure 3.31: Model fit: Share of singles moving home and job region



(a) Home move

(b) Home move by children

(c) Job move

(d) Job move by children

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

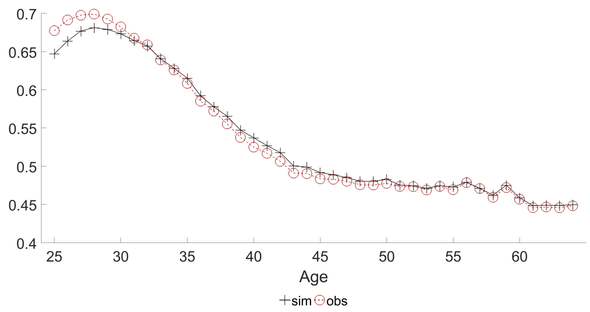
Preference parameters

With the initial CCPs estimated, I estimate the structural dynamic model. The estimates are presented in Table 3.24. They all have the expected signs and are surprisingly similar to the estimates from the static model in Section 6.3. The model fit is very similar too[26]. This indicates that either the expected future values are actually close to zero such that allowing for a positive discount factor does not change the results or the model is misspecified. The former seems unlikely since there obviously is a life cycle profile in moving behaviour and location choices. The latter cannot be ruled out though.

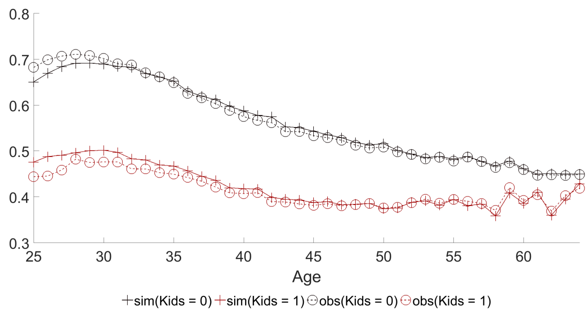
In order to estimate the dynamic model using the CCP method, I need consistent estimates of the probabilities of the imposed choices. It can be that even though the fit with respect to moving home and job, living in Copenhagen and being unemployed looks reasonable, the Logit CCPs are not good approximations of the future value components after all. Admittedly, the predictive power of regressors in Table G1 which do not index a home or job moving cost or whether the choice is associated with unemployment is quite low. The initial CCPs may therefore mainly fit transitions of home and job rather than the decision on where to locate conditional on transitioning. To model this more accurately, I need to capture the differences across the regions better. This can be achieved by including

[26]Not presented, but available upon requst.

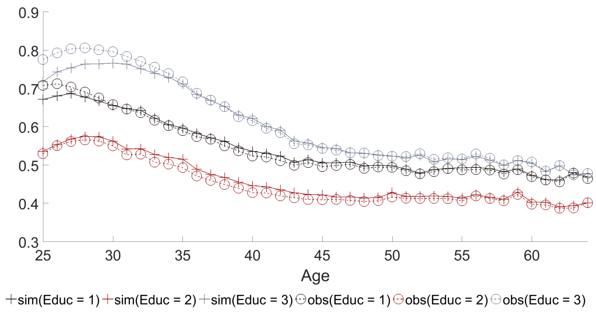
Figure 3.32: Model fit: Share of singles living in Copenhagen



(a) Aggregate



(b) By children



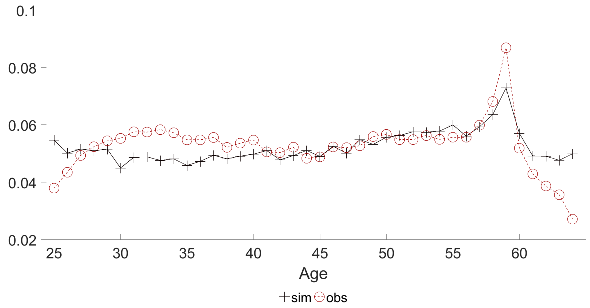
(c) By schooling

Note: Choice data is simulated for 1 period ahead using the states in the observed data. $S = 1$ is low education, $S = 2$ is medium education and $S = 3$ is high education.

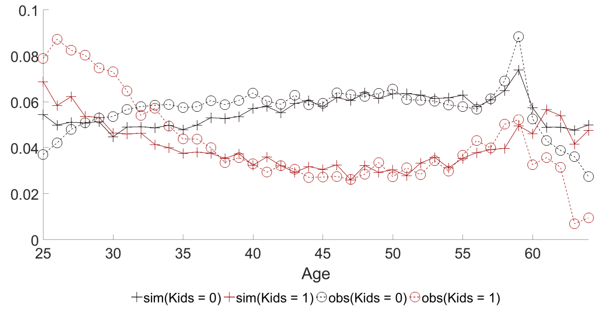
more amenities or by estimating regional-specific coefficients to certain variables. The latter quickly gets infeasible when I work with many regions. If I do not capture the differences in values across regions, the dynamic effects become negligible because there is no difference in the expected value when choosing one region instead of another conditional on moving according to the prediction, so the model would essentially assume there were no dynamic effects.

Another issue is the fact that regions within the GCA do not exhibit much variation in amenities compared to variation observed across the entire country. Allowing for a more disaggregate division of Rest of Zealand and consider them actual regions instead of just outside options may in itself be helpful, even with the current set of regressors. At least there is a chance that the reduced-form model can better capture the value of regional-specific amenities like nature, less crime and restaurants. The most compelling task for future research is therefore to implement a finer specification of the regions and to handle the evaluation of the value function for very large state spaces such that the dynamic model can be estimated for couples too.

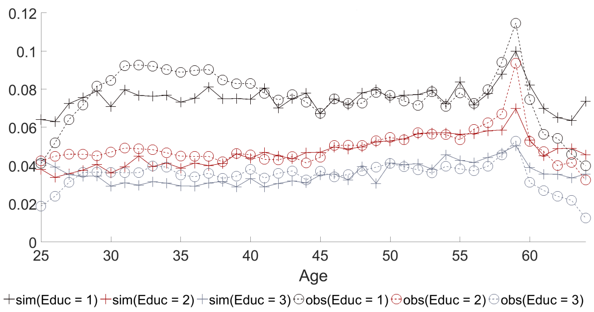
Figure 3.33: Model fit: Share of singles in unemployment



(a) Aggregate



(b) By children



(c) By schooling

Note: Choice data is simulated for 1 period ahead using the states in the observed data. $S = 1$ is low education, $S = 2$ is medium education and $S = 3$ is high education.

Table 3.24: Preference parameters from dynamic model for singles

Parameter	Estimate	Standard Error	t-statistic
γ_0	2.1804	0.0750	29.07
γ_a	0.0822	0.0021	39.09
γ_{kids}	0.7742	0.0668	11.59
τ_{nature}	0.0332	0.0026	12.83
τ_{rest}	2.0792	0.1114	18.67
τ_{thefts}	-25.7881	5.3625	-4.81
o_0	2.4921	0.0519	47.99
o_a	0.0418	0.0013	31.16
o_{unemp}	-5.6696	0.0574	-98.72
α_{work}	3.5282	0.0411	85.90
η_{ttime}	0.3730	0.0126	29.69
η_{kids}	0.1640	0.0419	3.91
$\psi_{jobdens}$	2.2372	0.0181	123.35

Note: Estimates of Equation 3.20. Nature is nature capital index. Restaurants in 1,000s and thefts is number of thefts per inhabitants in the region. Job density is defined as number of jobs by education group and normalized by the level of the education group in Copenhagen. Commute time is in hours.

7 Conclusion

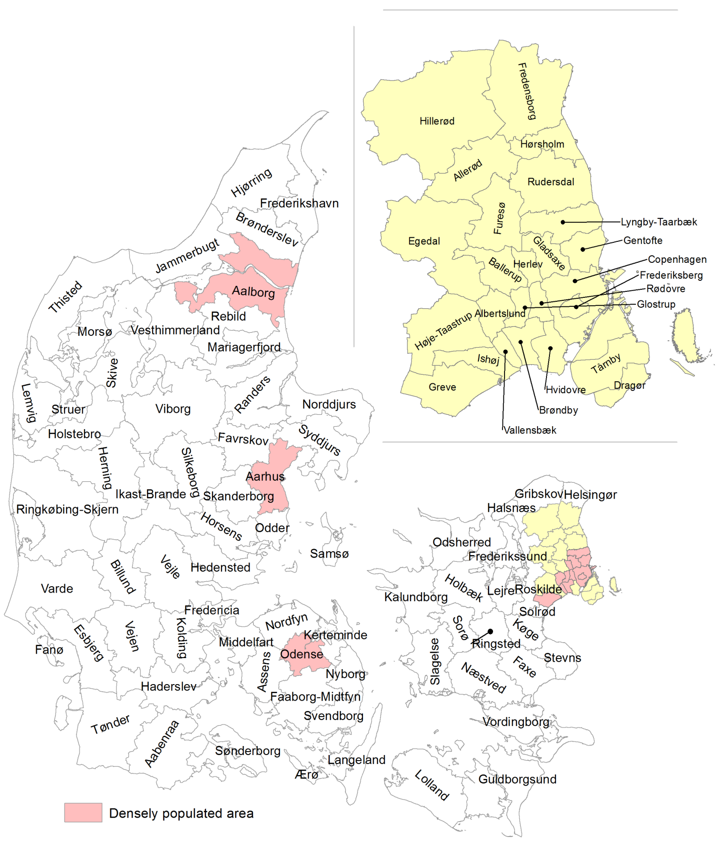
In this chapter I set up a theoretical dynamic structural model of multiple-member households' decisons on home and work locations for each spouse. I looked specifically at the intra-household allocation of commuting and raised the question whether an unfavorable bargaining power of the woman could explain the gender wage gap that was documented for individuals in couples in particular. I estimated a static version of the model since the computational burden associated with the dynamic model was too heavy and therefore left for future research. The estimation was restricted to a subsample of Denmark concentrated around the Greater Copenhagen Area and I found that since spouses are very similar in this region, they have an almost equal bargaining power. The gender wage gap observed in the data can therefore not be explained by women being discriminated within the household. In a counterfactual I increased the job density in Albertslund which is a relatively unattractive work region of the Greater Copenhagen Area. The simulation showed that both home and work location choices were altered and intra-household commute time differences slightly reduced. The predicted gender wage gap before the policy change was negligible but still marginally reduced in the counterfactual set-up. The results should be considered preliminary though since I did not account for the dynamic incentives of moving in the estimation or the occupational choice that may differ structurally across genders. Extending the estimation to allow for dynamics and use a more representative sample of Denmark as whole, i.e. not only focus on the most urbanized area, is an important task for future research and may affect the conclusions.

A Appendix: Geographic Classifications

Table A1: Overview of geographical classifications in Denmark

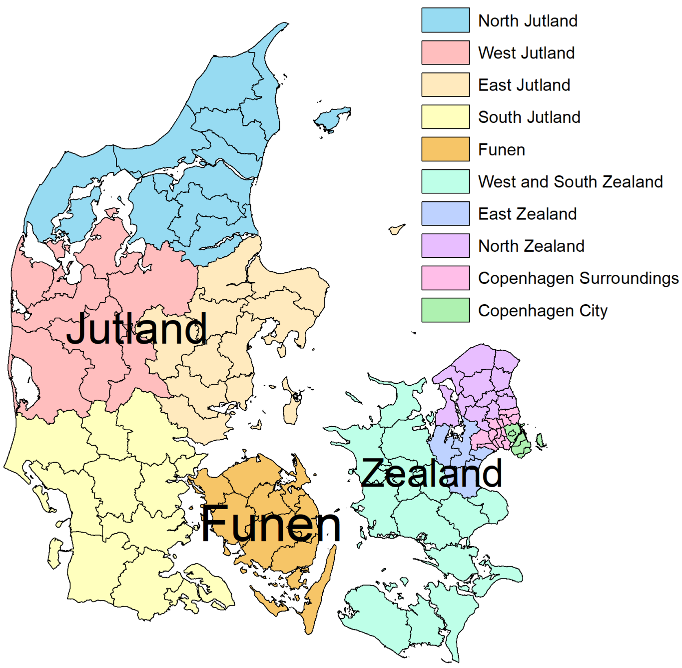
Danish	English	# units	Comment
Danmark	Denmark	1	
Regioner	Regions (states)	5	
Landsdele	Provinces	11	10 excl the island Bornholm
Amter	Counties	16	No longer exists
Valgkredse	Constituencies	92	
Kommuner	Municipalities	98	Reform in 2007: from 271 mun. to 98.
Trafikzoner	Traffic/commute zones	907	Defined by DTU's traffic model
Sogne	Parish	2,201	

Figure A1: Municipalities and urban areas



Note: Pink areas indicate an urban municipality. The yellow area corresponds to the official main part of the greater Copenhagen area (not the same definition used in the estimation).

Figure A2: Provinces and main islands



B Appendix: Data Sources

This section describes the datasets and the variables that I use in the empirical analysis. Statistics Denmark has provided access to administrative registers containing information of the entire population of Denmark for 1994-2012[27]. The Technical University of Denmark (DTU) has computed commute times between home and workplaces. Below I go over all these datasets.

Individual and household identifiers

The population register BEF holds data on a masked version of the unique and time-constant social security numbers of all individuals with residence in Denmark according to the National Registers of Persons (Folkeregisteret). This is a key variable since it enables me to link information from other registers. Useful for this analysis, the BEF register also contains a household identifier where a household can be either a single person or a couple with or without children. Children living at home belong to their parents' household as long as they live at the same address as one of the parents, are below 25 years old, have not been married or lived in a civil partnership, do not have children themselves and do not make up one part of a cohabiting couple. The household ID is stable over time as long as the person either stays single or the two people in the couple stay as a couple. If a couple splits up, both of them will get a new household ID as long as they do not still live together. The same holds if one of the partners die. In addition to household IDs, BEF thus also tells whether the person lives in a couple or not. A couple can be either a married couple, a civil partnership or cohabiting people with or without children.

Addresses and moves

From BEF I also get the home address of each person in the year. The addresses consist of identifiers for street name, house number, floor and door side and are unique within a municipality, so combined with information on municipality I get the exact address of the individual in a masked version.

The timing of moves can be identified from the date of official change in address. The law requires people to let the municipality know about their change of address no later than 5 days after the move. There is a fine for not complying with these rules and since not registering one's new address means mail is not delivered at the new home, very few people probably do not change their official address very fast. In the register, it is the individual's address as of January 1st that is recorded.

[27] Can be extended to 2017.

Personal characteristics

In addition to residence information, other personal characteristics from BEF that I use include age, and mother's and father's social security numbers to be able to match children and their parents. From the income registers INDH and INDK I get information on wage income, taxes and total income and the education register UDDUPD holds information of the educational degree the individual has obtained and details such as field of study.

Labor market information

The population register can be merged on to the Integrated Database for Labour Market Reasearch (IDA). This is a panel of all employments regarding persons living in Denmark since the end of the year since 1980. The database allows me to link individuals and firms and get data on the start year of the employment and number of days employed by the employer who is also occupied with an employer ID. These are based on the start and end dates of the employment that come from the Central Tax Information Sheet Register (Centrale Oplysningsregister) until 2008 and after then from eIncome which is also located at the tax authorities.

Individuals can have several jobs during a year. The register is posed in November of the year[28] and uses the register-based labor force statistics (Registerbaseret Arbejdsstyrkestatistik) to group individuals into either employed wage-earners, employer (A), self-employed (S) or co-working spouse (M). These four groups are mutually exclusive and the difference between A and S is that being a type A means having employees. The category of employed wage-earners can be further divided into main occupation (H), sideline occupation, another November occupation, and most important non-November job. The two latter, however, are only defined from 2004 onwards. The type variable is important when determining the main job to which the individual spends most of the time commuting to.

It may be, however, that one's H or A job in November is not the job that the individual has had for the major part of the year. In that case the most important non-November job should be considered the job of the year. Since this type is not defined until 2004, I looked at data from 2004-2012 to check how restrictive it would be to define the job in the year as the H or A job no matter what other job categories might be present. I found that for 90% of the population the most important non-November job had been the main employment for less than half a year and vice versa for H and A type jobs. I therefore decided to use the November employment to define the job of the year. In general, I aim at defining home and job location such that the probability that I model the commute that took place during most of the year is high. In this regard there is a trade-off since I also want consistency in the data over time which is why I do not exploit the non-November employments from the point in time where it was defined.

[28]From 2008 there is another register, BFL, where the frequency is monthly.

In order to model the commute I need information about the location of workplaces. Fortunately, the database allows me to not only link employees and employers but also employees and workplaces (and workplaces and employers). The workplace has an address code attached to it from which I can get province, municipality, parish and traffic zone. In those cases where an employment cannot be assigned to a registered workplace Statistics Denmark will assign a so-called fictitious workplace and the address will be the residence of the individual. This is often the case for people who conduct their work from or near their home or at several different workplaces. The latter concerns, in particular, workplaces for cleaners, insurance and for people working in the social- and healthcare system as for instance a community nurse[29]. Of course this gives rise to problems when calculating the travel time or commute distance for these workers and is something one must have in mind. It can be regarded a measurement error in the workplace variable. The workplace variable is used to get travel times between residential and work location. The travel time estimates will thus tend to be downward biased for people with fictitious workplaces.

Commute times and distances

The data on travel times come from The Danish Traffic Model (LTM) developed by researchers at the DTU. In LTM, Denmark as a country has been divided into 907 zones[30]. The number of trips by use of different transport modes between pairs of zones are estimated in the model. The definition of zones are based on the parish borders which can be linked to the addresses from BEF.

There are different definitions of travel time, namely both by public transportation, car and walk or bike. Travel time by car is given by the sum of free time (minutes in car with free flow, i.e. where the speed equals the allowed speed), congestion time (minutes with congestion, i.e. where the speed is less than the allowed speed), ferry time (minutes sailing by ferry), ferry wait time and pre-departure arrival time that take wait time into account. Travel time by public transport is given by summing waiting time, walk time in connection with shifts between different buses, trains and other public transport modes, walk time to and from first and last stop, respectively, and lastly travel time by other vehicles. In addition, travel time by walk and bike is calculated by LTM for pairs of zones where walk and biking trips actually exist according to the model.

It is important to note that the LTM has been run for 2002 and 2010 only. Walk and bike time is invariant, but travel time by car and public transportation may change over the years. This is due to for instance new stations/stops being established or existing ones closed, just like construction of a new roads influences travel time.

[29] See www.dst.dk/da/TilSalg/Forskningsservice/Dokumentation/hoejkvalitetsvariable/ida-arbejdssteder/lbnr for a more elaborate explanation.

[30] There are 4 different zone levels in LTM. This corresponds to level 2, which is rather detailed, and still ensures that the data complies with Statistics Denmark's rules about discretion. See www.landstrafikmodellen.dk for more information (in Danish).

To compute travel times between municipalities instead of between zones from LTM I follow the recommendations from researchers at DTU and calculate a weighted average of travel times using the zone pairs within the municipalities in question and weighing travel time in a given zone pair in the municipality by the number of trips made according to LTM. This ensures that when someone is observed to live in municipality A and work in municipality B, I assign the travel time that the average trip takes when accounting for the fact that it is most likely (unconditionally) that this person starts and ends her trip in the zones characterized by most trips.

To get the travel time for a given person and year in the dataset, I calculate both travel time by public transport, car, walk and bike for 2002 and 2010. I want one measure of travel time only for each person in the year and use the minimum of all 4 travel times as the representative travel time. This is done both for the 2002 and 2010 versions. The line of thought behind this rule is that I do not observe how people commute. I could, in principle, observe if a person owns a car and thus decide if it is likely that she commutes using the car. However, since cars can be shared within a household and most households own only 1 car, if any, I would have to guess which of the household members used the car. Also, since there can be a huge difference in travel time when using car instead of public transport it would not make much sense to use the mean of the two travel times when calculating travel time by transport. Of course, by taking the minimum of the two travel times, I do underestimate travel time for some people. However, if there is a very huge difference in travel time, it makes sense to assume the fastest mode is used. On the other hand, if the two travel times are not too different, the mistake is not too serious. Another argument for using the minimum of the travel times is that it represents the fastest possible way to get from home to job. I therefore implicitly assume that this is an amenity that people attach to the locations and base their decisions on that rather than also considering whether to buy a car, go by train or bus or choose to walk or bike.

LTM also provides a measure of the distance between zones. There is a choice between using work distance or travel time to measure the burden associated with the commute. The arguments for using travel time is that two jobs located in the same distance from some home location may give rise to very different travel times dependent on congestion, speed limits and public transport availability. On the other hand, travel times are only available for 2002 and 2010. When I use travel time I use the 2002 measures.

Home ownership and home characteristics

From the register of home characteristics, BOL, I observe characteristics of all dwellings in Denmark such as number of rooms, whehter the property has toilet and kitchen, living area, basement area, ground area, year of construction, number of buildings on the lot, type of home (e.g. townhouse, apartment, summer cottage), whether the dwelling is listed and its address. Additionally, EJSA tells what the transaction price of the house was at the time of a sale as well as the public evaluation of the home. One does not, of course, have to own a home. This information comes from the register EJER that contains information about who owns a given building or part of a building.

Local amenities

Data on local amenities include data on property crime (thefts, robberies and blackmailing), number of restaurants and quality-adjusted nature[31].

The crime data come from the register KROF which holds a record of the number of victims of criminal offense by the type of offense and the address where the crime took place. The crime level is computed as number of victims per parish code that have been reported to the police in a year and are available from 2005-2017. I have aggregated the numbers by municipalities. Since crime levels are not available for each year in the population data, I compute an average level of crime over the years and use this measure to describe the crime level in the municipality.

To capture some of the positive attributes that characterize a region, I calculate the number of restaurants in each municipality. This number is obtained by scraping the number of restaurants by municipality using Open Street Map[32]. Another amenity of interest is proximity to beautiful nature or green space. Researchers from The Danish Centre for Environment and Energy at Aarhus University[33] have used all available data on biodiversity including a bio score which gives a municipality a ranking on different categories of biodiversity. A high bio score means the type of nature has a high quality. The types of nature considered include proximity to the coast, sceneries with high nature density or prevalence of rare species of plan or forests. The nature capital index is computed as a weighted mean of the various bio scores in the municipality weighted by how much nature of the type there is. Heaths and open grazing land has a high bio score, while fields have a low bio score. Municipalities where a large share of the area is fields may therefore have a higher nature capital index than a municipality where there are a lot of open grazing land, but the open grazing land constitutes a low share of the total municipality area.

From IDA I access information on the number of workers of each of three schooling levels in a given municipality. I therefore calculate the average number of low-, medium- and high-skilled workers in a region from 1998-2010 and normalize the the number of workers within a skill type by the numbers for Copenhagen such that Copenhagen works as a reference grups. This is done to capture the fact that the Copenhagen area is described by many more jobs compared to the other regions, especially for high-skilled workers[34].

[31]See Table C1 for summary statistics of these amenities by region.
[32]The scraping took place on 25.02.2019.
[33]See www.biodiversitet.nu/naturkapital for access to data and information about methodology.
[34]See Table C2 for summary statistics of job density across regions and education.

C Appendix: Amenities

Table C1: Average value of amenities by home region (rh) or work region (rw)

Region	inc^{rh}	inc^{rw}	$Price/sqm$	Sqm	$Nature$	$Thefts$	$Restaurants$
Copenhagen	3.043	4.043	27.9	90.9	25	17936	1257
	(1.886)	(2.268)	.	(93.3)	.	.	.
Frederiksberg	3.452	3.680	29.4	108.9	17	1458	150
	(2.218)	(2.053)	.	(181.8)	.	.	.
Ballerup	3.163	4.442	24.4	104.4	23	391	21
	(1.848)	(2.070)	.	(39.8)	.	.	.
Broendby	2.898	4.068	22.4	106.0	20	361	22
	(1.599)	(1.947)	.	(61.3)	.	.	.
Dragoer	3.673	3.270	26.8	133.3	41	276	18
	(2.352)	(1.846)	.	(115.6)	.	.	.
Gentofte	4.205	4.142	36.1	130.3	31	620	88
	(2.956)	(2.432)	.	(129.1)	.	.	.
Gladsaxe	3.320	4.210	27.9	106.9	29	452	26
	(1.972)	(2.102)	.	(158.8)	.	.	.
Glostrup	3.142	4.171	23.3	99.5	20	301	20
	(1.749)	(2.013)	.	(38.5)	.	.	.
Herlev	3.109	3.989	24.7	102.3	25	313	15
	(1.765)	(1.947)	.	(66.2)	.	.	.
Albertslund	2.999	4.007	21	104.2	31	254	5
	(1.653)	(1.897)	.	(30.4)	.	.	.
Hvidovre	3.071	3.818	24.6	101.1	18	363	21
	(1.655)	(1.810)	.	(65.6)	.	.	.
Hoeje-Taastrup	3.139	4.045	20.5	110.8	14	403	40
	(1.730)	(1.909)	.	(72.5)	.	.	.
Roedovre	3.035	3.682	25.6	106.6	22	200	16
	(1.654)	(1.770)	.	(242.0)	.	.	.
Ishoej	2.958	3.600	19.8	111.8	15	100	10
	(1.606)	(1.657)	.	(50.5)	.	.	.
Taarnby	3.139	3.967	24.9	103.2	41	802	46
	(1.667)	(1.883)	.	(41.4)	.	.	.
Vallensbaek	3.506	3.732	21.8	118.5	28	227	6
	(1.955)	(1.955)	.	(53.2)	.	.	.
Rest Of Zealand	4.637	4.037	21.6	132.6	28	363	38
	(2.316)	(2.091)	.	(56.1)	.	.	.
Funen	4.257	3.981	15.1	137.0	17	387	43
	(2.326)	(2.447)	.	(67.0)	.	.	.
Jutland	4.230	4.026	15.4	139.8	26	418	59
	(2.390)	(2.357)	.	(62.0)	.	.	.

Note: Standard errors in parentheses. Prices measured in 1,000 DKK and income in 100,000s, nature is nature-capital index, thefts is total number of thefts by region, restaurants is total number of restaurants by region. Income comes from prediction of wages using estimates in Appendix E.

Table C2: Summary statistics of job density by work region

Region	Low educ.	Medium educ.	High educ.
Copenhagen	1.0000	1.0000	1.0000
Frederiksberg	0.1117	0.1197	0.0966
Ballerup	0.1036	0.1517	0.0943
Broendby	0.0747	0.0976	0.0501
Dragoer	0.0074	0.0102	0.0064
Gentofte	0.0847	0.0962	0.1020
Gladsaxe	0.0915	0.1255	0.1042
Glostrup	0.0593	0.0886	0.0547
Herlev	0.0468	0.0721	0.0524
Albertslund	0.0639	0.0945	0.0463
Hvidovre	0.0783	0.1016	0.0605
Hoeje-Taastrup	0.0990	0.1351	0.0579
Roedovre	0.0474	0.0714	0.0310
Ishoej	0.0258	0.0332	0.0149
Taarnby	0.1011	0.1062	0.0353
Vallensbaek	0.0110	0.0147	0.0082
Rest Of Zealand	0.0508	0.0722	0.0397
Funen	0.0613	0.0893	0.0473
Jutland	0.0857	0.1211	0.0617

Note: Job density is defined as the number of jobs by education group and work region normalized by the value in Copenhagen. The numbers have been averaged over time.

D Appendix: Estimation Sample

Table D1: Distribution of locations in estimation data for singles

	Home t	Home $t-1$	Work t	Work $t-1$
Copenhagen	0.484	0.470	0.446	0.403
Frederiksberg	0.086	0.083	0.053	0.049
Ballerup	0.023	0.023	0.041	0.037
Broendby	0.017	0.017	0.027	0.024
Dragoer	0.005	0.005	0.003	0.003
Gentofte	0.040	0.039	0.042	0.037
Gladsaxe	0.035	0.035	0.041	0.037
Glostrup	0.013	0.012	0.024	0.022
Herlev	0.014	0.014	0.020	0.018
Albertslund	0.016	0.016	0.023	0.021
Hvidovre	0.027	0.027	0.031	0.028
Hoeje-Taastrup	0.025	0.024	0.032	0.029
Roedovre	0.021	0.021	0.018	0.016
Ishoej	0.011	0.011	0.009	0.008
Taarnby	0.021	0.021	0.033	0.031
Vallensbaek	0.006	0.006	0.004	0.003
Rest of Zealand	0.144	0.152	0.089	0.097
Funen	0.004	0.006	0.002	0.004
Jutland	0.011	0.019	0.008	0.016
Unemployment			0.053	0.118
Total	1.0	1.0	1.0	1.0
Observations	3,703,096	3,703,096	3,703,096	3,703,096

Table D2: Distribution of locations in estimation data for couples

	Home t	Home$_m$ $t-1$	Home$_f$ $t-1$	Work$_m$ t	Work$_m$ $t-1$	Work$_f$ t	Work$_f$ $t-1$
Copenhagen	0.217	0.227	0.227	0.357	0.341	0.357	0.334
Frederiksberg	0.048	0.049	0.049	0.036	0.035	0.044	0.041
Ballerup	0.030	0.030	0.030	0.057	0.053	0.047	0.043
Broendby	0.019	0.019	0.019	0.035	0.033	0.023	0.021
Dragoer	0.010	0.010	0.010	0.003	0.002	0.003	0.003
Gentofte	0.042	0.041	0.041	0.035	0.033	0.042	0.039
Gladsaxe	0.039	0.038	0.038	0.047	0.044	0.041	0.038
Glostrup	0.013	0.013	0.013	0.027	0.026	0.027	0.026
Herlev	0.017	0.017	0.017	0.022	0.021	0.024	0.022
Albertslund	0.016	0.016	0.016	0.032	0.030	0.019	0.018
Hvidovre	0.031	0.031	0.031	0.032	0.030	0.031	0.029
Hoeje-Taastrup	0.030	0.030	0.030	0.042	0.039	0.035	0.033
Roedovre	0.021	0.021	0.021	0.021	0.020	0.015	0.014
Ishoej	0.012	0.012	0.012	0.010	0.009	0.008	0.007
Taarnby	0.028	0.027	0.027	0.036	0.035	0.023	0.022
Vallensbaek	0.011	0.010	0.010	0.005	0.004	0.004	0.004
Rest of Zealand	0.382	0.372	0.372	0.165	0.172	0.205	0.198
Funen	0.008	0.008	0.008	0.004	0.005	0.006	0.007
Jutland	0.027	0.029	0.029	0.016	0.023	0.021	0.025
Unemployment				0.018	0.044	0.024	0.076
Total	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Observations	3,379,748	3,379,748	3,379,748	3,379,748	3,379,748	3,379,748	3,379,748

E Appendix: Wage Regressions

Table E1: Estimates from OLS of Log Real Wages for Low-Skilled Workers by Region

Work Region	age	age^2	$\mathbb{I}_{rw_{t-1}=\varnothing}$	$Constant$	R^2	N
Copenhagen	0.1587	-0.0017	-0.7765	9.0397	0.2332	410758
	(0.000)	(0.000)	(0.000)	(0.000)		
Frederiksberg	0.1554	-0.0016	-0.7561	9.0395	0.2327	43752
	(0.000)	(0.000)	(0.000)	(0.000)		
Ballerup	0.1313	-0.0014	-0.7811	9.8425	0.1976	47008
	(0.000)	(0.000)	(0.000)	(0.000)		
Broendby	0.1036	-0.0011	-0.7962	10.4932	0.1605	35442
	(0.000)	(0.000)	(0.000)	(0.000)		
Dragoer	0.1321	-0.0015	-0.6105	9.6267	0.1873	3412
	(0.000)	(0.000)	(0.000)	(0.000)		
Gentofte	0.1281	-0.0013	-0.7183	9.6782	0.1787	39708
	(0.000)	(0.000)	(0.000)	(0.000)		
Gladsaxe	0.1601	-0.0018	-0.8380	9.1874	0.2683	38132
	(0.000)	(0.000)	(0.000)	(0.000)		
Glostrup	0.1349	-0.0014	-0.6912	9.7162	0.2223	23925
	(0.000)	(0.000)	(0.000)	(0.000)		
Herlev	0.1306	-0.0014	-0.7055	9.7836	0.1958	19150
	(0.000)	(0.000)	(0.000)	(0.000)		
Albertslund	0.0957	-0.0010	-0.7419	10.5897	0.1527	28222
	(0.000)	(0.000)	(0.000)	(0.000)		
Hvidovre	0.1169	-0.0013	-0.6683	10.0812	0.1801	35960
	(0.000)	(0.000)	(0.000)	(0.000)		
Hoeje-Taastrup	0.1184	-0.0013	-0.7218	10.1158	0.1919	41393
	(0.000)	(0.000)	(0.000)	(0.000)		
Roedovre	0.1160	-0.0013	-0.6916	10.0640	0.1855	19878
	(0.000)	(0.000)	(0.000)	(0.000)		
Ishoej	0.1164	-0.0013	-0.7522	10.0493	0.1774	11720
	(0.000)	(0.000)	(0.000)	(0.000)		
Taarnby	0.1456	-0.0016	-0.7374	9.6359	0.2093	47816
	(0.000)	(0.000)	(0.000)	(0.000)		
Vallensbaek	0.1315	-0.0014	-0.7503	9.7778	0.2039	4907
	(0.000)	(0.000)	(0.000)	(0.000)		
Rest Of Zealand	0.1486	-0.0016	-0.6826	9.3810	0.2174	76475
	(0.000)	(0.000)	(0.000)	(0.000)		
Funen	0.1834	-0.0018	-0.8162	8.4091	0.2774	1594
	(0.000)	(0.000)	(0.000)	(0.000)		
Jutland	0.1645	-0.0017	-0.6684	8.9870	0.2352	6311
	(0.000)	(0.000)	(0.000)	(0.000)		

Note: Standard errors in parentheses. $\mathbb{I}_{rw_{t-1}=\varnothing}$ means unemployed in $t-1$.

Table E2: Estimates from OLS of Log Real Wages for Medium-Skilled Workers by Region

Work Region	age	age^2	$\mathbb{I}_{rw_{t-1}=\varnothing}$	$Constant$	R^2	N
Copenhagen	0.1141 (0.000)	-0.0013 (0.000)	-0.7975 (0.000)	10.3419 (0.000)	0.1654	524053
Frederiksberg	0.1157 (0.000)	-0.0013 (0.000)	-0.8044 (0.000)	10.2659 (0.000)	0.1784	58739
Ballerup	0.1062 (0.000)	-0.0012 (0.000)	-0.7160 (0.000)	10.5949 (0.000)	0.1504	90715
Broendby	0.0961 (0.000)	-0.0011 (0.000)	-0.7564 (0.000)	10.8255 (0.000)	0.1517	53930
Dragoer	0.0925 (0.000)	-0.0011 (0.000)	-0.7435 (0.000)	10.7246 (0.000)	0.1376	5442
Gentofte	0.1021 (0.000)	-0.0012 (0.000)	-0.7419 (0.000)	10.6416 (0.000)	0.1561	59172
Gladsaxe	0.1247 (0.000)	-0.0014 (0.000)	-0.7785 (0.000)	10.2222 (0.000)	0.2258	69497
Glostrup	0.0985 (0.000)	-0.0011 (0.000)	-0.7235 (0.000)	10.7475 (0.000)	0.1556	45200
Herlev	0.0927 (0.000)	-0.0010 (0.000)	-0.6233 (0.000)	10.8365 (0.000)	0.1359	38145
Albertslund	0.0862 (0.000)	-0.0010 (0.000)	-0.7069 (0.000)	11.0337 (0.000)	0.1417	50102
Hvidovre	0.0894 (0.000)	-0.0010 (0.000)	-0.6742 (0.000)	10.8940 (0.000)	0.1324	56422
Hoeje-Taastrup	0.0905 (0.000)	-0.0010 (0.000)	-0.7153 (0.000)	10.8998 (0.000)	0.1334	71102
Roedovre	0.0911 (0.000)	-0.0010 (0.000)	-0.7055 (0.000)	10.8677 (0.000)	0.1451	36186
Ishoej	0.0979 (0.000)	-0.0011 (0.000)	-0.6555 (0.000)	10.7029 (0.000)	0.1548	17606
Taarnby	0.0900 (0.000)	-0.0010 (0.000)	-0.6377 (0.000)	10.9085 (0.000)	0.1303	57936
Vallensbaek	0.1047 (0.000)	-0.0012 (0.000)	-0.7684 (0.000)	10.5610 (0.000)	0.1578	8112
Rest Of Zealand	0.0974 (0.000)	-0.0011 (0.000)	-0.6999 (0.000)	10.7593 (0.000)	0.1682	102252
Funen	0.1866 (0.000)	-0.0021 (0.000)	-0.7140 (0.000)	8.9670 (0.000)	0.3002	1529
Jutland	0.1340 (0.000)	-0.0015 (0.000)	-0.7068 (0.000)	10.0440 (0.000)	0.1983	7080

Note: Standard errors in parentheses. $\mathbb{I}_{rw_{t-1}=\varnothing}$ means unemployed in $t-1$.

Table E3: Estimates from OLS of Log Real Wages for High-Skilled Workers by Region

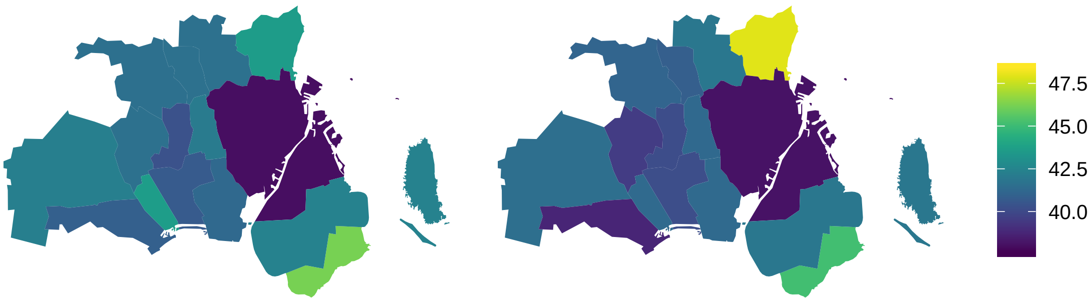
Work Region	age	age^2	$\mathbb{I}_{rw_{t-1}=\varnothing}$	$Constant$	R^2	N
Copenhagen	0.1687	-0.0018	-0.6917	9.1518	0.2132	657976
	(0.000)	(0.000)	(0.000)	(0.000)		
Frederiksberg	0.1634	-0.0017	-0.7167	9.1166	0.2159	71351
	(0.000)	(0.000)	(0.000)	(0.000)		
Ballerup	0.1444	-0.0015	-0.6027	9.8524	0.1779	62758
	(0.000)	(0.000)	(0.000)	(0.000)		
Broendby	0.1444	-0.0015	-0.6122	9.7670	0.1700	27438
	(0.000)	(0.000)	(0.000)	(0.000)		
Dragoer	0.1398	-0.0015	-0.7592	9.6609	0.1923	3559
	(0.000)	(0.000)	(0.000)	(0.000)		
Gentofte	0.1401	-0.0015	-0.7552	9.8928	0.1556	77232
	(0.000)	(0.000)	(0.000)	(0.000)		
Gladsaxe	0.1504	-0.0016	-0.6465	9.7113	0.1837	64861
	(0.000)	(0.000)	(0.000)	(0.000)		
Glostrup	0.1264	-0.0013	-0.6007	10.1571	0.1662	33987
	(0.000)	(0.000)	(0.000)	(0.000)		
Herlev	0.1094	-0.0011	-0.6071	10.4400	0.1397	31325
	(0.000)	(0.000)	(0.000)	(0.000)		
Albertslund	0.1337	-0.0014	-0.7433	10.0164	0.1529	21939
	(0.000)	(0.000)	(0.000)	(0.000)		
Hvidovre	0.1121	-0.0012	-0.6212	10.3636	0.1456	35407
	(0.000)	(0.000)	(0.000)	(0.000)		
Hoeje-Taastrup	0.1411	-0.0015	-0.6668	9.8275	0.1724	32240
	(0.000)	(0.000)	(0.000)	(0.000)		
Roedovre	0.1186	-0.0012	-0.6439	10.1954	0.1365	15335
	(0.000)	(0.000)	(0.000)	(0.000)		
Ishoej	0.1238	-0.0013	-0.5989	10.0317	0.1495	8132
	(0.000)	(0.000)	(0.000)	(0.000)		
Taarnby	0.1382	-0.0014	-0.7239	9.7441	0.1756	19408
	(0.000)	(0.000)	(0.000)	(0.000)		
Vallensbaek	0.1295	-0.0014	-0.6547	10.0063	0.1642	4764
	(0.000)	(0.000)	(0.000)	(0.000)		
Rest Of Zealand	0.1467	-0.0015	-0.5958	9.6913	0.2091	131409
	(0.000)	(0.000)	(0.000)	(0.000)		
Funen	0.1759	-0.0018	-0.5868	8.8772	0.2565	3214
	(0.000)	(0.000)	(0.000)	(0.000)		
Jutland	0.1964	-0.0021	-0.5550	8.5656	0.2339	9330
	(0.000)	(0.000)	(0.000)	(0.000)		

Note: Standard errors in parentheses. $\mathbb{I}_{rw_{t-1}=\varnothing}$ means unemployed in $t-1$.

F Appendix: Structural Model Fits

F.1 Structural model fit for singles

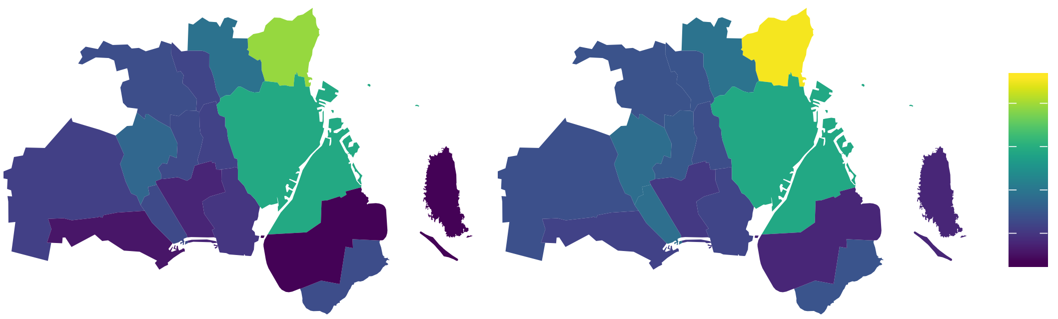
Figure F1: Structural model fit: Average age by home region



(a) Age (obs.) (b) Age (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure F2: Structural model fit: Share of highly educated singles by home and work region

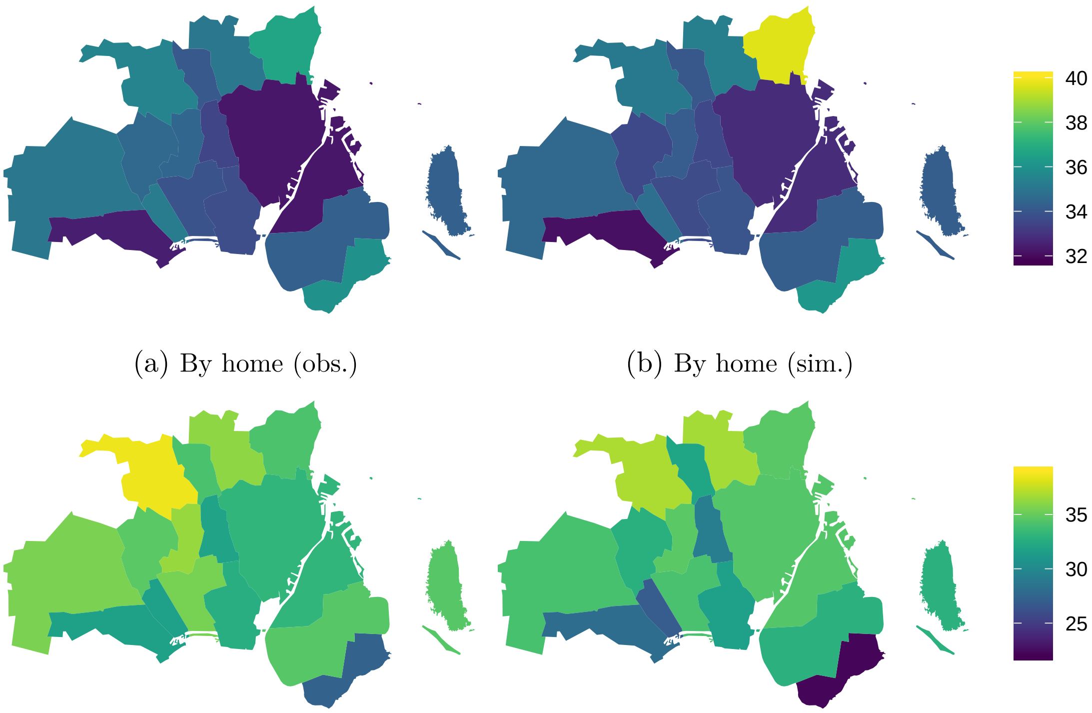


(a) Home (obs.) (b) Home (sim.)

(c) Work (obs.) (d) Work (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure F3: Structural model fit: Predicted income (10,000 DKK) by home and work region



(a) By home (obs.)

(b) By home (sim.)

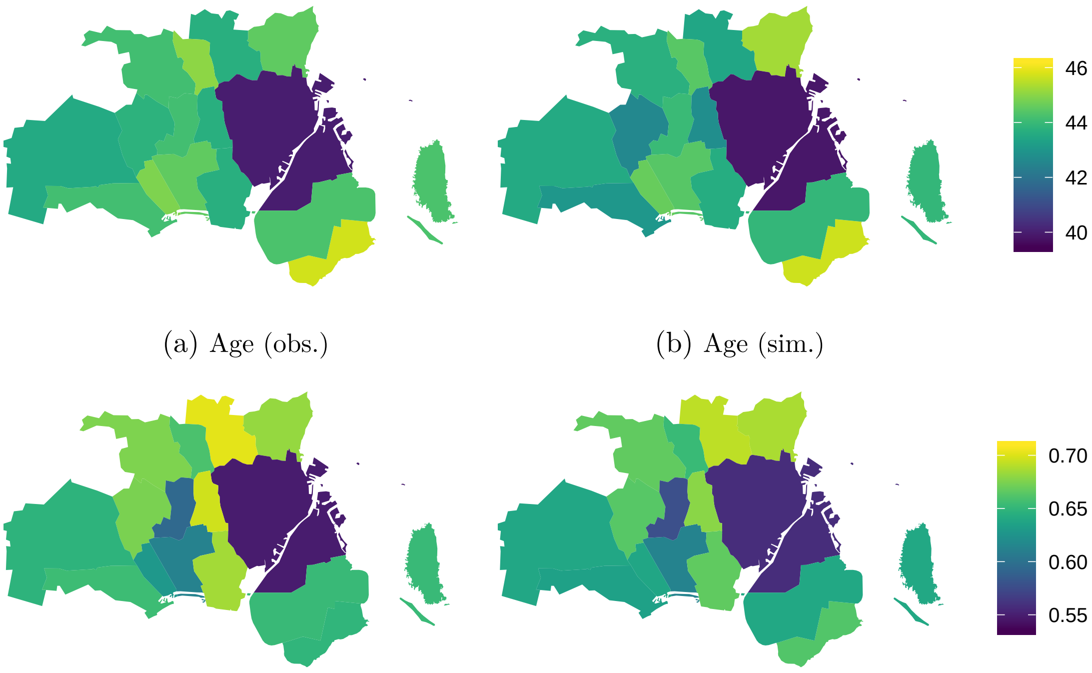
(c) By work (obs.)

(d) By work (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

F.2 Structural model fit for couples

Figure F4: Structural model fit: Average age and children by home region

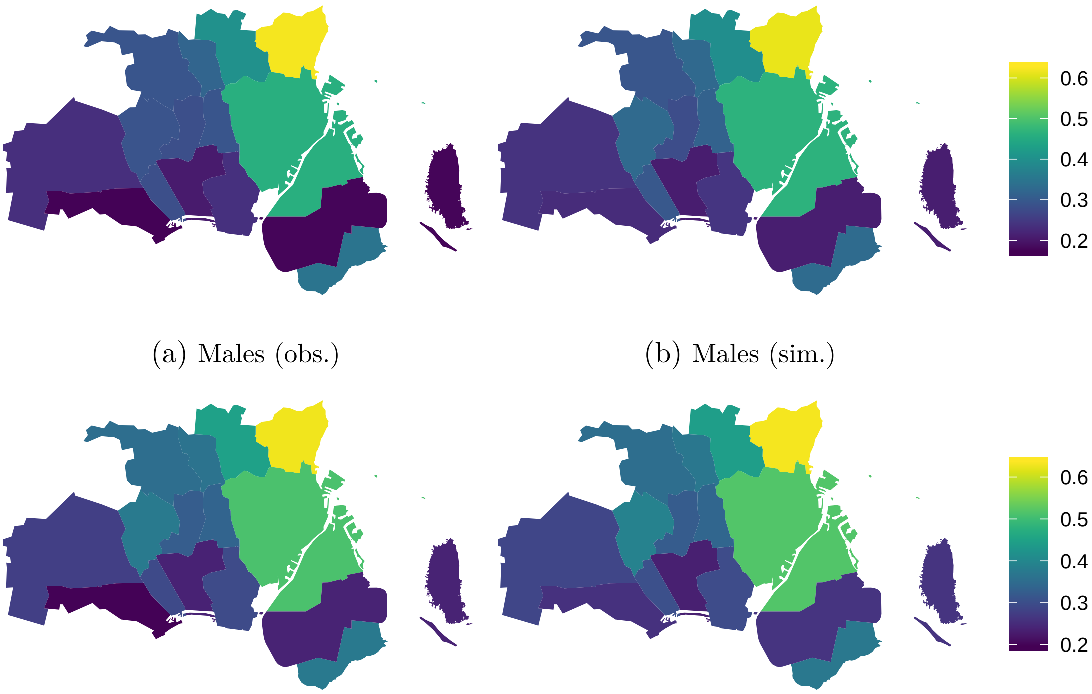


(a) Age (obs.) (b) Age (sim.)

(c) Children (obs.) (d) Children (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure F5: Structural model fit: Share of highly educated couples by home region

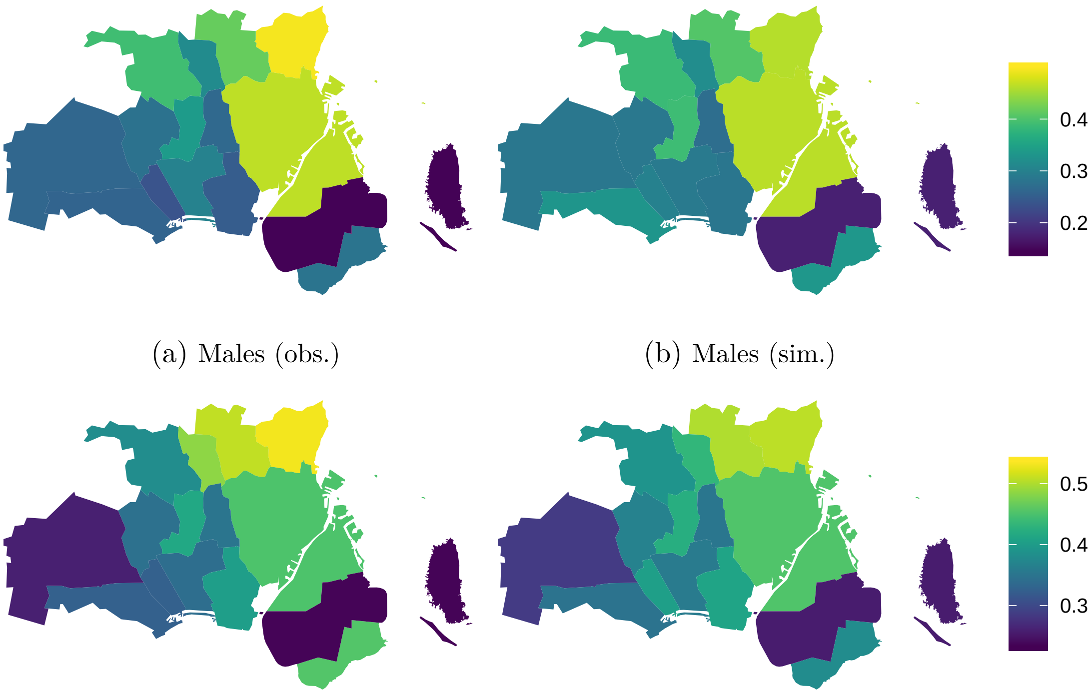


(a) Males (obs.) (b) Males (sim.)

(c) Females (obs.) (d) Females (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure F6: Structural model fit: Share of highly educated couples by work region



(a) Males (obs.)

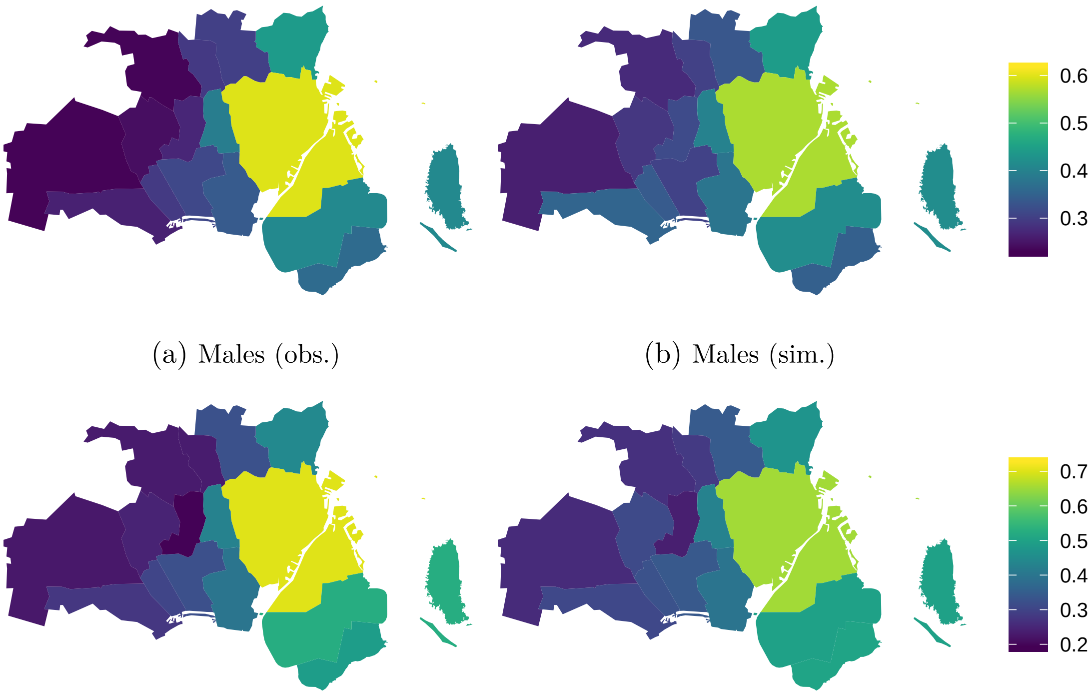
(b) Males (sim.)

(c) Females (obs.)

(d) Females (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

Figure F7: Structural model fit: Share of couples working in Copenhagen by home region



(a) Males (obs.)

(b) Males (sim.)

(c) Females (obs.)

(d) Females (sim.)

Note: Choice data is simulated for 1 period ahead using the states in the observed data.

G Appendix: Supplementary Results

G.1 Estimation of dynamic model for singles

Table G1: Parameter estimates from conditional Logit for singles

Parameter	Estimate	Standard Error	t-statistic	5% p value
# restaurants (10s)	0.0978	0.0138	7.0723	0.0049
# bars (10s)	-0.2295	0.0361	-6.3628	0.0060
# thefts pr km2 (10s)	0.0470	0.0090	5.2149	0.0087
Total house price (100,000s)	-0.0137	0.0080	-1.7219	0.0568
Travel time (quarters)	-0.0274	0.0037	-7.4813	0.0044
Avg. income (work region)	0.1816	0.0035	51.7058	0.0001
Travel time (quarters) × I[Kids]	-0.0493	0.0117	-4.2025	0.0131
I[stay home]	3.0819	0.0922	33.4131	0.0002
I[stay home]× Age	0.0700	0.0021	32.7905	0.0002
I[stay home]× I[Kids]	0.9695	0.0711	13.6270	0.0013
I[stay home]× Schooling	-0.0545	0.0247	-2.2053	0.0396
I[stay job]	3.5789	0.0229	156.5657	0.0000
I[stay job]× Age	0.0382	0.0013	29.0839	0.0003
I[unemployment]	2.4012	0.1287	18.6519	0.0007
I[unemployment]× Age	0.0153	0.0023	6.6924	0.0054
I[unemployment] × I[Kids]	-0.7832	0.0786	-9.9659	0.0025
I[unemployment] × Schooling	-0.1683	0.0301	-5.5862	0.0076
Log lik.	-1.58479			
N	37155			

Note: Initial CCPs used to approximate future expected value function in dynamic model.

G.2 Counterfactual

Table G2: Difference in distribution of male work regions

Region	Counterfactual	Original	Difference	Elasticity
Copenhagen	0.0968	0.2973	-0.2005	-0.6742
Frederiksberg	0.0035	0.0054	-0.0019	-0.3552
Ballerup	0.3391	0.1251	0.2140	1.7099
Broendby	0.0721	0.0452	0.0268	0.5937
Dragoer	0.0004	0.0006	-0.0003	-0.4286
Gentofte	0.0191	0.0193	-0.0003	-0.0153
Gladsaxe	0.0941	0.0558	0.0383	0.6870
Glostrup	0.0737	0.0351	0.0387	1.1036
Herlev	0.0178	0.0144	0.0035	0.2402
Albertslund	0.0538	0.0337	0.0201	0.5946
Hvidovre	0.0153	0.0146	0.0008	0.0527
Hoeje-Taastrup	0.0525	0.0426	0.0099	0.2330
Roedovre	0.0090	0.0084	0.0007	0.0813
Ishoej	0.0071	0.0043	0.0028	0.6395
Taarnby	0.0897	0.0567	0.0330	0.5812
Vallensbaek	0.0073	0.0034	0.0039	1.1379
Rest Of Zealand	0.0157	0.1206	-0.1048	-0.8695
Funen	0.0077	0.0532	-0.0455	-0.8557
Jutland	0.0014	0.0531	-0.0517	-0.9744
Unemployment	0.0239	0.0111	0.0127	1.1432

Table G3: Difference in distribution of female work regions

Region	Counterfactual	Original	Difference	Elasticity
Copenhagen	0.1016	0.3039	-0.2024	-0.6658
Frederiksberg	0.0026	0.0087	-0.0061	-0.7007
Ballerup	0.3473	0.1227	0.2247	1.8317
Broendby	0.0683	0.0319	0.0364	1.1437
Dragoer	0.0003	0.0005	-0.0002	-0.4706
Gentofte	0.0177	0.0237	-0.0061	-0.2550
Gladsaxe	0.0947	0.0553	0.0394	0.7117
Glostrup	0.0705	0.0385	0.0319	0.8284
Herlev	0.0173	0.0166	0.0007	0.0446
Albertslund	0.0503	0.0260	0.0243	0.9330
Hvidovre	0.0152	0.0160	-0.0008	-0.0517
Hoeje-Taastrup	0.0480	0.0393	0.0087	0.2216
Roedovre	0.0086	0.0068	0.0017	0.2554
Ishoej	0.0070	0.0035	0.0035	0.9833
Taarnby	0.0814	0.0439	0.0375	0.8528
Vallensbaek	0.0060	0.0030	0.0030	0.9806
Rest Of Zealand	0.0146	0.1530	-0.1384	-0.9043
Funen	0.0060	0.0450	-0.0390	-0.8669
Jutland	0.0011	0.0451	-0.0440	-0.9751
Unemployment	0.0414	0.0163	0.0251	1.5417

Table G4: Difference in distribution of home regions

Region	Counterfactual	Original	Difference	Elasticity
Copenhagen	0.2111	0.2171	-0.0059	-0.0273
Frederiksberg	0.0358	0.0339	0.0019	0.0575
Ballerup	0.0432	0.0310	0.0122	0.3924
Broendby	0.0319	0.0223	0.0096	0.4286
Dragoer	0.0056	0.0073	-0.0016	-0.2236
Gentofte	0.0065	0.0053	0.0012	0.2222
Gladsaxe	0.0476	0.0392	0.0084	0.2142
Glostrup	0.0230	0.0151	0.0079	0.5264
Herlev	0.0268	0.0198	0.0069	0.3497
Albertslund	0.0674	0.0357	0.0317	0.8860
Hvidovre	0.0389	0.0316	0.0073	0.2306
Hoeje-Taastrup	0.0432	0.0339	0.0093	0.2735
Roedovre	0.0441	0.0275	0.0166	0.6054
Ishoej	0.0372	0.0192	0.0180	0.9384
Taarnby	0.0329	0.0285	0.0044	0.1554
Vallensbaek	0.0162	0.0121	0.0041	0.3407
Rest Of Zealand	0.2822	0.3659	-0.0837	-0.2288
Funen	0.0040	0.0187	-0.0147	-0.7867
Jutland	0.0024	0.0359	-0.0335	-0.9334

Table G5: Difference in distribution of male-female commute time (minutes)

Region	Counterfactual	Original	Difference	Elasticity
Copenhagen	0.6475	5.2694	-4.6219	-0.8771
Frederiksberg	-0.0431	3.3279	-3.3711	-1.0130
Ballerup	1.3435	7.5529	-6.2095	-0.8221
Broendby	0.1482	7.5591	-7.4109	-0.9804
Dragoer	2.0979	6.8217	-4.7238	-0.6925
Gentofte	-1.1351	-8.3331	7.1980	-0.8638
Gladsaxe	0.5374	8.2066	-7.6692	-0.9345
Glostrup	0.2072	8.9786	-8.7714	-0.9769
Herlev	0.2652	5.7130	-5.4478	-0.9536
Albertslund	0.4967	6.9076	-6.4109	-0.9281
Hvidovre	0.2560	6.7106	-6.4546	-0.9619
Hoeje-Taastrup	1.0860	7.5299	-6.4439	-0.8558
Roedovre	0.4212	8.6712	-8.2500	-0.9514
Ishoej	1.2565	3.7267	-2.4702	-0.6628
Taarnby	1.2315	4.0073	-2.7758	-0.6927
Vallensbaek	1.6486	8.5048	-6.8562	-0.8062
Rest Of Zealand	1.7509	3.2569	-1.5061	-0.4624
Funen	9.3145	7.4810	1.8335	0.2451
Jutland	11.5388	22.4637	-10.9249	-0.4863

Table G6: Summary statistics of wage growth by home region

Region	Male	Female
Copenhagen	0.0309	0.0374
Frederiksberg	0.0244	0.0264
Ballerup	0.0313	0.0326
Broendby	0.0282	0.0406
Dragoer	-0.0007	0.0106
Gentofte	0.0512	0.0535
Gladsaxe	0.0225	0.0293
Glostrup	0.0274	0.0367
Herlev	0.0189	0.0119
Albertslund	0.0091	0.0047
Hvidovre	0.0195	0.0286
Hoeje-Taastrup	0.0186	0.0287
Roedovre	0.0298	0.0386
Ishoej	0.0442	0.0599
Taarnby	0.0147	0.0234
Vallensbaek	0.0140	0.0115
Rest Of Zealand	0.0278	0.0301
Funen	0.0907	0.0626
Jutland	-0.1235	-0.1444

Table G7: Standardized values of amenities by region

Region	**Nature**	**Thefts**	**Restaurants**	**Total price**	**Commute time**	**Job dens. (low)**	**Job dens. (medium)**	**Job dens. (high)**	**Wages (low)**	**Wages (medium)**	**Wages (high)**
Copenhagen	-0.1394	2.8612	3.7230	-0.2983	0.1092	3.7140	3.6928	3.7165	-0.0334	-0.0391	0.1838
Frederiksberg	-0.7148	0.3436	0.1296	0.5034	0.0169	-0.0580	-0.1087	-0.0705	-0.4187	-0.4080	-0.2760
Ballerup	-0.2832	-0.5941	-0.2891	-0.2495	-0.0284	-0.0925	0.0293	-0.0802	0.6154	0.5243	0.6051
Broendby	-0.4990	-0.2858	-0.2859	-0.4751	-0.0729	-0.2152	-0.2044	-0.2656	0.2586	0.2919	0.2052
Dragoer	1.0116	1.0115	-0.2988	1.3650	0.0984	-0.5012	-0.5817	-0.4487	-0.8522	-0.8515	-0.8636
Gentofte	0.2922	-0.5042	-0.0716	3.1411	0.0534	-0.1730	-0.2104	-0.0479	-0.1222	-0.0854	0.0629
Gladsaxe	0.1484	-0.7354	-0.2729	0.1741	0.0336	-0.1440	-0.0839	-0.0387	-0.0044	0.0729	0.4216
Glostrup	-0.4990	0.2152	-0.2923	-0.5064	-0.0905	-0.2808	-0.2431	-0.2462	0.2422	0.2585	0.3690
Herlev	-0.1394	-0.1445	-0.3086	-0.3324	0.0040	-0.3336	-0.3141	-0.2557	-0.0320	0.0130	0.0585
Albertslund	0.2922	-0.4785	-0.3410	-0.8405	-0.0258	-0.2610	-0.2176	-0.2812	0.1476	0.3181	0.1385
Hvidovre	-0.6429	-0.7225	-0.2891	-0.3441	-0.0281	-0.1998	-0.1870	-0.2219	-0.0334	-0.0526	-0.0768
Hoeje-Taastrup	-0.9306	-0.5684	-0.2274	-0.7435	0.0131	-0.1121	-0.0422	-0.2328	0.2602	0.2611	0.2615
Roedovre	-0.3552	-0.9537	-0.3053	-0.4779	-0.0246	-0.3311	-0.3174	-0.3456	-0.1551	-0.1187	-0.2676
Ishoej	-0.8587	-1.0437	-0.3248	-0.7925	-0.0561	-0.4231	-0.4822	-0.4130	-0.3485	-0.1133	-0.3408
Taarnby	3.2415	0.9088	-0.2079	-0.0399	0.0788	-0.1030	-0.1672	-0.3277	0.6417	0.1469	-0.2405
Vallensbaek	0.0764	0.6904	-0.3378	-0.0836	-0.0809	-0.4856	-0.5621	-0.4410	-0.1658	-0.2181	-0.2406

Note: Nature is nature capital index. Job density is the number of jobs by education group in the work region normalized by the level of the education group in Copenhagen and averaged over time. Wages are predicted using the estimates in Appendix E. Mean and standard deviation for the standardized measure are computed using all regions in the Greater Copenhagen Area.

Bibliography

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

Attanasio, O. and V. Lechene (2002). Tests of income pooling in household decisions. *Review of economic dynamics 5*(4), 720–748.

Aura, S. (2005). Does the balance of power within a family matter? The case of the retirement equity act. *Journal of Public Economics 89*(9), 1699–1717.

Becker, G. (1981). A treatise on the family. *Harvard University Press, Cambridge, MA 30.*

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy 70*(5, Part 2), 9–49.

Bishop, K. (2012). A dynamic model of location choice and hedonic valuation.

Bobonis, G. J. (2009). Is the allocation of resources within the household efficient? New evidence from a randomized experiment. *Journal of political Economy 117*(3), 453–503.

Browning, M. and P.-A. Chiappori (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, 1241–1278.

Buchinsky, M., C. Gotlibovski, and O. Lifshitz (2014). Residential location, work location, and labor market outcomes of immigrants in israel. *Econometrica 82*(3), 995–1054.

Buchinsky, M., C. Gotlibovski, and O. Lifshitz (2017, February). The role of women and men in choices of residential and work locations in israel.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica: Journal of the Econometric Society*, 63–90.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of political Economy 100*(3), 437–467.

Chiappori, P.-A., A. de Palma, N. Picard, and I. A. Inoa (2014). Couple residential location and spouses workplaces.

Chiappori, P.-A. and M. Mazzocco (2017). Static and intertemporal household decisions. *Journal of Economic Literature 55*(3), 985–1045.

Costa, D. L. and M. E. Kahn (2000). Power couples: changes in the locational choice of the college educated, 1940–1990. *The Quarterly Journal of Economics 115*(4), 1287–1315.

Freedman, O. and C. R. Kern (1997). A model of workplace and residence choice in two-worker households. *Regional Science and Urban Economics 27*(3), 241–260.

Gemici, A. (2011). Family migration and labor market outcomes.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies 61*(2), 265–289.

Joensen, J. S. and E. Mattana (2017). Student aid, academic achievement, and labor market behavior. Working paper.

Krueger, D. and F. Perri (2006). Does income inequality lead to consumption inequality? evidence and theory. *The Review of Economic Studies 73*(1), 163–193.

Lise, J. and K. Yamada (2018). Household sharing and commitment: Evidence from panel data on individual expenditures and time use. *The Review of Economic Studies 86*(5), 2184–2219.

Lundberg, S. and R. A. Pollak (1993). Separate spheres bargaining and the marriage market. *Journal of political Economy 101*(6), 988–1010.

Lundberg, S. and R. A. Pollak (2003). Efficiency in marriage. *Review of Economics of the Household 3*, 153–167.

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997). Do husbands and wives pool their resources? evidence from the united kingdom child benefit. *Journal of Human resources*, 463–480.

Manser, M. and M. Brown (1980). Marriage and household decision-making. *International economic review 1*, 31–44.

Mazzocco, M. (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. *The Review of Economic Studies 74*(3), 857–895.

McElroy, M. B. and M. J. Horney (1981). Nash-bargained household decisions. *International economic review 2*, 333–349.

Mincer, J. (1978). Family migration decisions. *Journal of political Economy 86*(5), 749–773.

Picard, N., A. De Palma, and S. Dantan (2013). Intra-household discrete choice models of mode choice and residential location. *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti*, 419–445.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers 3*(2), 135–146.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, 999–1033.

Samuelson, P. A. (1956). Social indifference curves. *The Quarterly Journal of Economics 70*(1), 1–22.

Scholz, J. K., A. Seshadri, and S. Khitatrakun (2006). Are americans saving optimally for retirement? *Journal of political economy 114*(4), 607–643.

Sermons, M. W. and F. S. Koppelman (2001). Representing the differences between female and male commute behavior in residential location choice models. *Journal of transport geography 9*(2), 101–110.

Udry, C. (1996). Gender, agricultural production, and the theory of the household. *Journal of political Economy 104*(5), 1010–1046.