UNIVERSITY OF COPENHAGEN
DEPARTMENT OF ECONOMICS

PhD Thesis

Esben Scriver Andersen

Essays on structural microeconometrics

Perturbed utility and equilibrium models

Supervisors: Mogens Fosgerau & Bertel Schjerning

Submitted on: February 1, 2022

This page has been left blank intentionally.

Contents

Dansk Resumé

Denne afhandling består af tre uafhængige kapitler. Hvert kapitel analyserer ligevægtsudfald på arbejds- og boligmarkedet ved at beskrive adfærden af deltagerne i markederne ved *the perturbed utility model*, da denne klasse af modeller tillader for mere generelle substitutionsmønstre end den mest anvendte klasse af modeller, *the additive random utility model*. Substitutionsmønstre er vigtige, da lønninger og boligpriser ikke kun afspejler henholdsvis produktiviteten af arbejdskraften og kvaliteten af boligerne, men også afspejler tilgængeligheden af tætte substitutter.

Det første kapitel *Multidimensional matching and labor market complementarity* er skrevet i samarbejde med Young Jun Lee. Vi foreslår et empirisk framework for matching i et friktionsløst arbejdsmarked, hvor arbejdere og virksomheder forhandler lønninger såvel som andre kontraktuelle forholds (som f.eks. at arbejde deltid eller fuldtid). Arbejdernes og virksomhedernes adfærd er beskrevet af en klasse af statiske perturbed utility models, hvor alternativerne kan være både komplementer eller substitutter. Vi viser, at der eksisterer en unik ligevægt, og hvordan denne kan findes. Som et proof-of-concept estimerer vi modellen på aggregeret data for det danske arbejdsmarked, og vi finder, at arbejdere med forskelligt uddannelsesniveau kan være komplementer in produktionen, hvilket ikke kan være tilfældet for the additive random utility model.

I det andet kapitel *A perturbed spatial equilibrium model* foreslås et empirisk framework for en spatial ligevævgtsmodel for bolig- og arbejdsmarkedet, hvor husholdninger omkostningsfyldt kan flytte og pendle mellem lokationer. Husholdningernes adfærd er beskrevet af en statisk perturbed utility model, der indebærer at ethvert fremadskuende motiv er ignoreret. Den foreslåede statiske ligevægtsmodel muliggør en analyse af de spatiale substitutionsmønstre for huspriser og lønninger, og på grund af modellens simple struktur er det muligt at løse modellen for et stort antal lokationer. Som et proof-of-concept er modellen estimeret på dansk data, hvor det er antaget at husholdninger kan vælge at bo og arbejde i 92 kommuner. Baseret på den estimerede model er der foretaget en

kontrafaktisk analyse af en 1 procent stigning i udbuddet af boliger i København. Som forventet viser analysen, at kvadratmeterpriserne og lønningerne er mest påvirkede i København og omegnskommunerne. På grund af det højere udbud af boliger og arbejdskraft falder de gennemsnitlige kvadratmeterpriser og lønninger med henholdsvis 10,3 og 0,5 procent i København.

At flytte bolig er forbundet med store finansielle og ikke-finansielle omkostninger. Derfor er boligvalget i sagens natur et dynamisk valg, da husholdningerne tager deres fremtidige velfærd med i betragtning, når de skal træffe beslutningen. Det tredje kapitel *A dynamic spatial equilibrium model for the housing and labor market* viser at perturbed utility kan anvendes til at beskrive dynamiske valg. På baggrund af en dynamisk perturbed utility model foreslås en spatial ligevægtsmodel svarende til kapitel 2. Da dynamiske modeller er beregningstunge at estimere, fokuserer den specificerede model på *hovedstadskommunerne*, der består af 23 kommuner. Den kontrafaktiske analyse viser, at de gennemsnitlige kvadratmeterpriser og lønninger i København falder med henholdsvis 10,8 og 0,2 procent. At de kontrafaktiske effekter er relativt ens i de to kapitler, afspejler til dels at antagelsen om statiske forventninger til boligpriser og lønninger er fasthold i kapitel 3. Dette fører til at priser og lønninger reagerer for kraftigt til at begynde med. At indbygge fremadskuende forventninger i modellen er efterladt til fremtidig forskning.

Summary

This thesis consists of three independent chapters. Each chapter analyzes the equilibrium outcomes of the labor and housing market by describing the behavior of the participants in the markets by the *perturbed utility model*. This class of models allows for a more general substitution patterns as opposed to the mostly used model class in the literature, the *additive random utility model*. Substitution patterns is important, as wages and housing prices not only reflect the productivity of labor and the quality of the properties, but also reflect the availability of close substitutes.

The first chapter *Multidimensional matching and labor market complementarity* is written in cooperation with Young Jun Lee. We propose an empirical framework for matching in a frictionless labor market, where workers and firms negotiate over wages and other contractual terms (e.g. whether to work part time or full time). The behavior of workers and firms is described by a class of static perturbed utility models in which alternatives may be complements or substitutes. We show that a unique equilibrium exists and how the equilibrium of the model is obtained. As a proof of concept, we estimate the model based on aggregated data for the Danish labor market and find that workers with different educational levels can be complements in production, which the additive random utility model rules out.

The second chapter *A perturbed spatial equilibrium model* proposes an empirical spatial equilibrium framework for the housing and labor markets, where households can only move and commute between locations at a cost. Household behavior is described by a static perturbed utility model, which implies that any forward-looking motive is ignored. The proposed static equilibrium model allow spatial substitution patterns of housing prices and wages to be analyzed, and due to the simplicity of the model it is possible to solve it for a large choice set of locations. As a proof of concept, the model is estimated on Danish data, where the households are assumed to choose their residential and work location among 92 municipalities. Based on the estimated model, a counterfactual analysis

of an increase of 1 percent in the supply of housing in the capital municipality, *Copenhagen*, is conducted. As expected this shows analysis that the square meter prices and wages are most affected in Copenhagen and the nearby municipalities. Due to the higher supply of housing and labor the average square meter prices and wages decrease by 10.3 and 0.5 percent in Copenhagen.

Moving house is associated with large financial and none-financial costs. Hence, the housing decision is inherently a dynamic decision, as the households take their future welfare into account when they making a decision. The third chapter *A dynamic spatial equilibrium model for the housing and labor market* shows that perturbed utility can be used to describe dynamic decision. Based on a dynamic perturbed utility model, the chapter then proposes a spatial equilibrium model similar to the model in chapter 2. As dynamic models are computationally costly to estimate, the proposed model focuses on the *capital municipalities* that constitutes of 23 municipalities. The counterfactual analysis shows that the average square meter prices and wages in Copenhagen decrease by 10.8 and 0.2 percent, respectively. The similar sizes of the counterfactual changes in the two chapter are partly due to the maintained assumption of static expectations with respect to square meter prices and wages in chapter 3. Initially this leads prices and wages to respond too strongly. Incorporating forward-looking expectations into the model is left for future research.

This page has been left blank intentionally.

Chapter 1

Multidimensional matching and labor market complementarity

Multidimensional Matching and Labor Market Complementarity*

Esben Scriver Andersen† Young Jun Lee‡

February 1, 2022

Abstract

We propose an empirical framework for multidimensional matching in a frictionless labor market, where the employment level can be identified, and workers and firms are allowed to negotiate over other contractual terms than wages and who matches with whom. The behavior of workers and firms is described by a class of perturbed utility models in which alternatives may be complements or substitutes. We show that a unique equilibrium exists and how the equilibrium of the model is obtained. As a proof of concept, we estimate the model based on aggregated data for the Danish labor market and find that workers with different educational levels can be complements in production.

KEYWORDS: multidimensional matching, complementarity, perturbed utility, similarity.

---

*This research is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 740369). The authors thank Mogens Fosgerau for comments and suggestions.

†Department of Economics, University of Copenhagen, Denmark, esben.scriver.andersen@econ.ku.dk

‡Department of Economics, University of Copenhagen, Denmark, yjl@econ.ku.dk

1 Introduction

In the field of macro and labor economics, a fundamental question is whether different workers are substitutes or complements in production. As this affects wages and the allocation of workers, the answer to this question has important empirical and policy implications. For instance, fierce debates in the Western world have centered on whether immigrants and native workers are substitutes (see, e.g., Borjas (2003); Card (2005); Ottaviano and Peri (2012)), as the policy recommendation - whether to restrict immigration policies or not - from the native workers' point of view depends on this. In contrast, the empirical literature on matching models primarily describes the behavior of workers and firms by the random utility (RU) model, which restricts inputs to be substitutes. Further, this branch of the literature has mostly studied firms' composition of workers, thereby ignoring how the number of matches is determined and the fact that similar workers can be employed in different occupations.

This paper proposes an empirical framework for two-sided matching in a frictionless labor market for representative workers and firms, where wages and allocation of workers are determined in equilibrium. The proposed framework addresses three issues partly ignored by this branch of the literature: (1) alternatives in the choice sets of workers or firms are allowed to be complements, (2) the number of matches can be determined in equilibrium, and (3) matching is multidimensional as workers and firms can negotiate over other contractual terms than wages and whom to match with.

Most of the existing empirical work describes the choices of the workers and firms by the RU model, which entails overly simplistic substitution patterns between the different alternatives that constitute the choice sets of the workers and firms. In particular, the RU model restricts alternatives to be substitutes. This is especially restrictive for the firm side since it is well-known that different inputs can be complements[1] in production. Instead, this paper describes the behavior of the workers and firms by the perturbed utility (PU) model. This class of behavioral models rationalizes behavior as maximizing a utility function that consists of a linear term and a concave perturbation function. Due to the concavity, the perturbation function can be interpreted as representing preferences for taste-for-variety of the workers. In the presence of maximal quotas - such that the representative firms cannot hire more workers than are present in the economy - we show that the perturbation function can be interpreted as the production function of the firm. The PU model can be seen as a more general class of models than the class of RU models. Consequently, Hofbauer and Sandholm (2002) showed that the choice

[1] We use the classical definition of complementarity that labels two alternatives as complements if the cross-price derivative of compensated demand is negative.

probabilities generated by any RU model could be derived from a PU model. Allen and Rehbeck (2019a), Fosgerau *et al.* (2020, 2021b), and Fosgerau and Nielsen (2021) all showed that some PU models can allow for complementarity.

In the existing literature, workers only choose which firm they want to work for, and firms only choose which workers to hire. However, in practice, most firms hold a variety of occupations making matching multidimensional as the worker and firm have to negotiate over other contractual terms than wages and who to match with. For instance, occupations can vary with respect to working hours, job-tasks, managerial responsibilities, etc. By allowing for multidimensionality, the use of matching models can be extended to analyze occupation composition and substitution patterns between different labor inputs along multiple dimensions.

An obvious empirical issue for one-to-one matching models of the labor market is that unmatched firms are not observed in the data. For instance, Dupuy *et al.* (2020) only consider matched college football coaches when studying the deadweight loss due to taxation. As a result, the aggregate demand for given types of football coaches is exogenous, which by construction implies that the employment level is unaffected by changes in the tax rate. Instead, this paper proposes a many-to-one matching model where several workers can match to the same firm and where the firm chooses how many workers to employ.

The proposed framework is closely related to Galichon *et al.* (2019) as matching is frictionless and wages can be seen as an imperfect transfer of utility between workers and firms. However, as alternatives are not restricted to be substitutes in our framework, the existence and uniqueness of an equilibrium can not be obtained from the gross substitutability condition of Kelso and Crawford (1982). Instead, we use the convexity of the indirect utility and profit functions to show the unique existence of an equilibrium. In equilibrium, workers and firms behave optimally, and there is market clearing within each type of worker. The equilibrium wages and allocation of workers reflect the distribution of workers, workers' preferences, and firms' production functions.

For our empirical application we specify the perturbation functions in terms of the similarity function as suggested by Fosgerau and Nielsen (2021). We then propose a maximum likelihood estimator closely related to Dupuy and Galichon (2022) who, unlike Galichon *et al.* (2019), allow estimation to be based on both observed choices and observed wages.

As a proof of concept, we estimate the model based on administrative data for the Danish labor market and find complementarity between workers of different education level. Based on our estimated model, we study of the welfare consequences of a counterfactual skill-bias technological change and a counterfactual decrease in the bottom tax rate. Our welfare analysis is inspired by Dupuy *et al.* (2020), who analyze the deadweight loss from matching distortion[2]. However, as our framework allows the total employment to be an equilibrium outcome, our welfare analysis includes changes along the extensive margin.

The counterfactual study of a skill-bias technological change analyzes the welfare effects of increased productivity of highly educated workers matched with firms in the manufacturing industry in occupations that demand a high level of skills. The analysis shows that wages for matches in the manufacturing industry tend to increase if they are complements to the match that experiences increased productivity. In contrast, wages tend to decrease for matches that are substitutes. This illustrates that skill-bias technological changes can have very different welfare effects for different labor inputs, and ignorance of the possibility of complementarities can mislead the effect of technological changes on wages. On the aggregate level, the welfare of skilled and unskilled workers decreases, whereas the welfare increases for medium and highly educated workers increases.

The welfare analysis of a counterfactual reduction in the bottom tax rate of a one percentage point implies that wages fall across matches, as firms do not have to pay as high wages for compensating the workers for matching with them. However, the average wages are predicted to increase across worker types as workers sort into more productive matches with higher wages. As a result, the reduced tax rate implies the the total employment increases by 1,300 workers, and total welfare increases by 1,550 million DKK. However, the government budget is worsened by 4,850 million DKK due to the lower tax revenue.

Relation to the literature. This review focuses on the empirical framework for one-to-one matching. This framework was initiated by the pioneering work of Becker (1973), who analyzes the equilibrium outcomes in a frictionless matching market, where heterogenous tastes over partners are taken as primitives. Choo and Siow (2006) introduced an empirical equilibrium framework with perfect transferable utility, where the choices of the participants were described by the RU model of McFadden (1973). This was followed by Galichon and Salanie (2021), who showed the existence and uniqueness of the matching equilibrium are guaranteed under more general forms of unobserved heterogeneity and proposed an empirical framework. Fox *et al.* (2018) study the nonparametric identification of distributions of unobserved match heterogeneity and agent heterogeneity. Galichon *et al.* (2019) unified the matching model of Becker (1973) and the collective model of Chiappori (1992) in order to allow for imperfect transfers of utility between agents when matched.

[2]Matching distortion is when, e.g., workers choose to work for a firm they like instead of choosing to work for a firm at which they are most productive.

This branch of the literature assumes a finite number of types of agents on each side of the market and a continuum of each type exists. Dupuy and Galichon (2014) extended this framework to account for possible continuous multivariate attributes of the participants. For an application to the labor market, see Bojilov and Galichon (2016) who use distributional assumptions to derive closed-form formulas for the equilibrium matching problem. Even though wages were included in this literature, estimation was only based on observed choices, which made this framework less appealing for analyzing the labor market. Dupuy and Galichon (2022) proposed a maximum likelihood estimator based on observed matches and observed wages.

Various extensions of the framework of Choo and Siow (2006) have been applied to study, for example, the marriage market, the formation of supply chains, gender differences in educational attainment, and discrimination in the US (see Chiappori *et al.* (2009); Fox (2018); Dupuy *et al.* (2020); Dupuy and Galichon (2022)). This list is necessarily very incomplete.

Organization of the paper. The remainder of this paper is organized as follows. Section 2 provides a class of two-sided matching framework for labor market. Section 3 shows the unique existence of equilibrium. Section 4 introduces the similarity model used for estimation, and Section 5 specifies the model and sets up the applied maximum likelihood estimator. This is followed by Section 6 that describes the data used for estimation and presents the estimation results. Finally, based on the estimated model a counterfactual study of the welfare effects of a skill-biased technology change and a reduced bottom tax rate is conducted in Section 7, and Section 8 concludes.

2 The model

We study a two-sided many-to-one matching of representative workers and firms in which alternatives may be substitutes or complements. Unlike the existing literature[3], matching between workers and firms can be multidimensional as both worker and firm, when matched, have to mutually agree on the characteristics of the match. We assume that the type of the match is defined over a finite set, for example, full-time or part-time jobs, or jobs requiring cognitive, manual, or social skills. Without loss of generality, we will assume that the type of the match, which workers and firms have to agree on, is one-dimensional. We will refer to this additional dimension as occupation. Hence, workers have to choose which firm to match with and which occupation to perform. Similar firms

[3] Azevedo and Leshno (2016) provide a theoretical framework for the model with transfers and other contractual terms.

have to choose which workers to employ and which occupation the worker should perform. The results from our model are not restricted to a labor market. For example, the model may be applied to matching of students to schools with majors in specific subjects.

This section is organized as follows. Subsection 2.1 defines the participants of the matching model. Subsections 2.2 and 2.3 set up the optimization problem for workers and firms, respectively. This is followed by Subsection 2.4 that defines the aggregate equilibrium.

2.1 Participants

The economy consists of firms and workers, and only the type of the participants can be observed. Let $x \in \mathcal{X}$ index the types of workers; let $y \in \mathcal{Y}$ index the types of firms. We further assume that the number of types, $|\mathcal{X}|$ and $|\mathcal{Y}|$, are finite. Further, let N_x denote the number of workers of type x, and N denote the total number of workers, $\sum_x N_x = N$. Finally, let $z \in \mathcal{Z}$ index the types of occupations and let there be $|\mathcal{Z}|$ different occupations that any worker can perform for any firm.

2.2 Worker's problem

The worker of type x can either choose a pair of firm and occupation types $(y, z) \in \mathcal{Y} \times \mathcal{Z}$ or choose to become unemployed. When working in the pair (y, z), the worker receives the wage w_{xyz}, and when unemployed the worker receives the exogenous unemployment benefits w_{x0}. Let $W_{x\cdot\cdot} := (w_{xyz})_{(y,z)\in\mathcal{Y}\times\mathcal{Z}}$ denote the vector of wages. The worker's base their choices on their optimal choice probabilities, $p^x\left(W_{x\cdot\cdot}\right) \in \triangle^{|\mathcal{Y}||\mathcal{Z}|+1}$ where $\triangle^{|\mathcal{Y}||\mathcal{Z}|+1} = \left\{p \in \mathbb{R}_+^{|\mathcal{Y}||\mathcal{Z}|+1} : \sum_{y,z} p_{yz} + p_0 = 1\right\}$, that maximize the worker's perturbed utility:

$$\max_{p\in\triangle^{|\mathcal{Y}||\mathcal{Z}|+1}} \left\{\sum_{y,z} p_{yz} u_{xyz}\left(w_{xyz}\right) + p_0 u_{x0}\left(w_{x0}\right) + G_x\left(p\right)\right\} \tag{1}$$

where u_{xyz} and u_{x0} are utility indexes that depend on the wage w_{xyz} and unemployment benefit w_{x0}. $G_x : \mathbb{R}_+^{|\mathcal{Y}||\mathcal{Z}|+1} \to \mathbb{R}$ is the perturbation function of the worker of type x. If G_x is strictly concave and satisfies that $\nabla_{p_{yz}} G_x\left(p\right) \to \infty$ as $p_{yz} \to 0$, then there exists a unique optimal $p^x\left(W_{x\cdot\cdot}\right)$ for Eq. (1). This perturbation function represents taste for variety as it penalizes small values of p_{yz}. Together with utilities, the perturbation function encodes whether the alternatives are substitutes or complements. However, as each worker only chooses one pair of (y, z), complementarity between two pairs is not a desirable scenario.

As noted by Allen and Rehbeck (2019a), the choice probabilities $p^x\left(W_{x\cdot\cdot}\right)$ can be the result of aggregating the discrete choices of workers of type x. Hence, the perturbed utility model can be seen as a representative agent model[4].

2.3 Firm's problem

It is assumed that the representative firm of each type has the same quota, N, which implies that the firm at maximum can hire all workers. It is assumed that firms of type y have preferences regarding how many workers of each type to employ and how to allocate them across the types of occupations, $(n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}$. As a result, workers of identical type are perfect substitutes. The preference for labor inputs for firms of type y is described by the production function $F_y\left((n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}\right)$, which is assumed to be strictly increasing, strictly concave, and satisfies the boundary condition, $\nabla_{n_{xyz}}F_y(n)\to\infty$ as $n_{xyz}\to 0$. For a given wage schedule $(w_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}$, the representative firm of type y chooses its bundle of workers $(n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}$ in order to maximize its profit given by the production minus wage costs:

$$\max_{(n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}}\left\{F_y\left((n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}\right)-\sum_{x,z}w_{xyz}n_{xyz}\right\},$$

subject to the constraint that $n_{xyz}\geq 0$. The first-order condition is $\frac{\partial F_x\left((n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}\right)}{\partial n_{xyz}}=w_{xyz}$ for all x and z.

This specification of a firm's problem allows different labor inputs in the production to be substitutes or complements[5], and allows total employment to be endogenously determined under the following weak assumptions.

Assumption 1 (Technology)**.** *Each representative firm y has the same quota for labor and its production depends on the unused quota:*

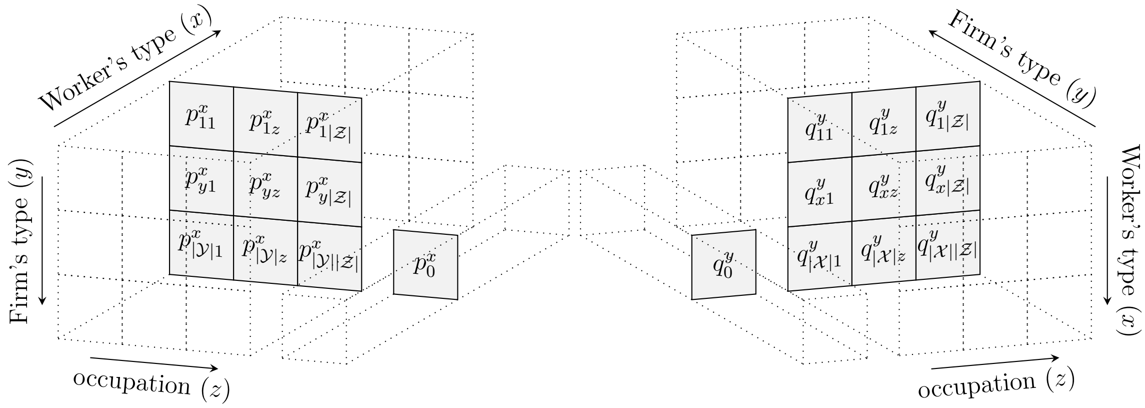
$$F_y\left((n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}\right)=\bar{F}_y\left((n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}},N-\sum_{x,z}n_{xyz}\right),$$

and F_y satisfies for any constant $C>0$ that

$$\bar{F}_y\left((n_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}},N-\sum_{x,z}n_{xyz}\right)=C\bar{F}_y\left(\left(\frac{n_{xyz}}{C}\right)_{(x,z)\in\mathcal{X}\times\mathcal{Z}},\frac{N-\sum_{x,z}n_{xyz}}{C}\right)+f_y(C).$$

[4]We can alternatively interpret $p^x(W_{x\cdot\cdot})$ as generated by a preference for randomization (Fudenberg *et al.*, 2015). A worker of type x chooses an optimal probability over $(\mathcal{Y}\times\mathcal{Z})\cup\{0\}$, and the worker uses this to choose among alternatives randomly.

[5]Our output function F_y is familiar from the one in Eeckhout and Kircher (2018), but we do not impose the additive separability across its types. As Eeckhout and Kircher (2018) mentioned, additive separability is restrictive because it does not allow for interactions between workers.



(a) Worker of type x **(b)** Firm of type y

Figure 1: Worker's and firm's choice probabilities

Our assumption of quotas is related to Eeckhout and Kircher (2018), who include scarce managerial resources in a firm's production function. Assumption 1 allow the firm's profit maximization problem to be rewritten as:

$$\max_{q\in\triangle^{|\mathcal{X}||\mathcal{Z}|+1}} N\left\{\bar{F}_y\left(q\right)-\sum_{x,z} w_{xyz}q_{xz}\right\}+f_y\left(N\right), \tag{2}$$

with $q=\left(\left(q_{xz}\right)_{(x,z)\in\mathcal{X}\times\mathcal{Z}}, q_0\right)$, where $q_{xz}=n_{xyz}/N$ and $q_0=\left(N-\sum_{x,z} n_{xyz}\right)/N$. For a given wage schedule, $W_{\cdot y\cdot}:=\left(w_{xyz}\right)_{(x,z)\in\mathcal{X}\times\mathcal{Z}}$, let $q^y\left(W_{\cdot y\cdot}\right)$ denote the solution to the profit maximization problem. As $q^y\left(W_{\cdot y\cdot}\right)$ belongs to the probability simplex, $\triangle^{|\mathcal{X}||\mathcal{Z}|+1}$, it can be interpreted as the optimal choice probabilities of the representative firm of type y. As N is a constant, it does not affect the optimal choice probabilities of the firm. Hence, the maximization problem, $\max_{q\in\triangle^{|\mathcal{X}||\mathcal{Z}|+1}}\left\{\bar{F}_y\left(q\right)-\sum_{x,z} w_{xyz}q_{xz}\right\}$, is a perturbed utility model.

As $\bar{F}_y$ is strictly increasing, strictly concave, and satisfies the boundary condition, $\nabla_{q_{xz}}F_y\left(q\right)\to\infty$ as $q_{xz}\to 0$, the optimal vector of choice probabilities,$q^y\left(W_{\cdot y\cdot}\right)$, is in the interior of $\triangle^{|\mathcal{X}||\mathcal{Z}|+1}$. The implications of our boundary conditions are similar to the implications of the Inada conditions, as both imply that when any input is absent, no output is produced.

2.4 Market clearing and aggregate equilibrium

This subsection defines the equilibrium of the model. Without loss of generality, we normalize the mass of workers to unity, $\sum_x r_x=1$, where $r_x>0$ denotes the mass of workers of type x. However, as the mass of firms equals to unity for all types $y\in\mathcal{Y}$ by assumption, it can be ignored. For a given wage schedule W, the optimal choices of the worker of type x and firm of type y are given by the choice probabilities $p^x(W_{x\cdot\cdot})$ and $q^y(W_{\cdot y\cdot})$, respectively. This is illustrated by Figure 1. In equilibrium, markets must clear such that the mass of workers of type x choosing a pair of industry and occupation, (y,z), should coincide with the mass of workers of type x in occupation z chosen by a firm of type y:

$$r_x p^x_{yz}(W_{x\cdot\cdot}) = q^y_{xz}(W_{\cdot y\cdot}), \quad \forall (x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}. \tag{3}$$

Let μ^*_{xyz} be the equilibrium mass of workers of type x assigned to the pair (y,z). Let $\mathcal{M}(r)$ be the set of distributions, $\mu = \left((\mu_{xyz})_{(y,z)\in\mathcal{Y}\times\mathcal{Z}}, \mu_{x0}\right)_{x\in\mathcal{X}}$, over $(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}) \cup (\mathcal{X}\times\{0\})$ with first margin $r = \left(r_1,\ldots,r_{|\mathcal{X}|}\right)$:

$$\mathcal{M}(r) = \left\{ \mu \in \triangle^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|+|\mathcal{X}|} : \sum_{y,z} \mu_{xyz} + \mu_{x0} = r_x, \forall x \in \mathcal{X} \right\} \tag{4}$$

This leads to the following definition:

Definition 1 (Aggregate equilibrium)**.** An aggregate equilibrium is characterized by a set of wages $W^* = \left(w^*_{xyz}\right)_{(x,y,z)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}}$ and mass of workers $\mu^* = \left(\left(\mu^*_{xyz}\right)_{(y,z)\in\mathcal{Y}\times\mathcal{Z}}, \mu^*_{x0}\right)_{x\in\mathcal{X}}$, satisfying:

(i) (Market clearing) $\mu^* \in \mathcal{M}(r)$.

(ii) (Optimality for worker) A representative worker of type x solves the following perturbed utility model for the wage schedule $W^*_{x\cdot\cdot}$:

$$\mu^*_{\cdot|x} \in \underset{p\in\triangle^{|\mathcal{Y}||\mathcal{Z}|+1}}{\arg\max} \left\{ \sum_{y,z} p_{yz} u_{xyz}\left(w^*_{xyz}\right) + p_0 u_{x0}(w_{x0}) + G_x(p) \right\}.$$

(iii) (Optimality of firm) A representative firm of type y solves the following perturbed utility model for the wage schedule $W^*_{\cdot y\cdot}$:

$$\left(\mu^*_{\cdot y\cdot}, 1 - \sum_{x,z} \mu^*_{xyz}\right) \in \underset{q\in\triangle^{|\mathcal{X}||\mathcal{Z}|+1}}{\arg\max} \left\{ \bar{F}_y(q) - \sum_{x,z} w^*_{xyz} q_{xz} \right\}.$$

If a $|\mathcal{X}|\times|\mathcal{Y}|\times|\mathcal{Z}|$ wage matrix W^* exists such that (μ^*, W^*) is an aggregate equilibrium for a given vector of unemployment benefits, $w_0 = \left(w_{10},\ldots,w_{|\mathcal{X}|0}\right)'$, we say that μ^* is stable and W^* supports μ^*. Hence, in equilibrium, no workers or firms want to deviate. This concept of equilibrium corresponds to the competitive equilibrium in Kelso and Crawford (1982) and Dupuy *et al.* (2020).

3 Aggregate equilibrium: existence and uniqueness

We prove the unique existence of the equilibrium under the following regularity conditions.

Assumption 2. *The perturbation function $G_x(p)$ (and $\bar{F}_y(q)$) satisfies the following conditions for all $x \in \mathcal{X}$ (and $y \in \mathcal{Y}$):*

(i) $G_x(p)$ *is a proper, closed strictly concave function with domain* $\triangle^{|\mathcal{Y}||\mathcal{Z}|+1}$,

(ii) For all $p \in \operatorname{Int}\left(\triangle^{|\mathcal{Y}||\mathcal{Z}|+1}\right)$, $G_x(p)$ *is twice differentiable with invertible* $\nabla^2 G_x(p)$, *and*

(iii) For any sequence $\left(p^\ell\right) \subset \operatorname{Int}\left(\triangle^{|\mathcal{Y}||\mathcal{Z}|+1}\right)$, *we have* $\nabla_{p_{yz}} G_x\left(p^\ell\right), \nabla_{p_0} G_x\left(p^\ell\right) \to \infty$ *if and only if* $p^\ell_{yz}, p^\ell_0 \to 0$, *respectively.*

Fosgerau *et al.* (2021a) show that $G_x(p)$ and $\bar{F}_y(q)$, and their conjugates are essentially smooth under Assumption 2. Then, it follows that $\nabla G_x(p)$ and $\nabla \bar{F}_y(q)$ are homeomorphic from $\mathbb{R}^{|\mathcal{Y}||\mathcal{Z}|+1}_{++}$ and $\mathbb{R}^{|\mathcal{X}||\mathcal{Z}|+1}_{++}$, respectively. With functions H_x and M_y defined by

$$H_x(p) = \nabla G_x(p) + \iota_{|\mathcal{Y}||\mathcal{Z}|+1}, \quad M_y(q) = \nabla \bar{F}_y(q) + \iota_{|\mathcal{X}||\mathcal{Z}|+1},$$

the optimal choice probabilities can be expressed in closed form

$$p^x(W_{x\cdot\cdot}) = H_x^{-1}\left(-U_x\left(W_{x\cdot\cdot}, w_{x0}\right)\right) e^{-G_x^*(U_x(W_{x\cdot\cdot}, w_{x0}))} = \frac{H_x^{-1}\left(-U_x\left(W_{x\cdot\cdot}, w_{x0}\right)\right)}{\iota'_{|\mathcal{Y}||\mathcal{Z}|+1} H_x^{-1}\left(-U_x\left(W_{x\cdot\cdot}, w_{x0}\right)\right)}, \tag{5}$$

$$q^y(W_{\cdot y\cdot}) = M_y^{-1}\left(W_{\cdot y\cdot}\right) e^{-\bar{F}_y^*(W_{\cdot y\cdot})} = \frac{M_y^{-1}\left(W_{\cdot y\cdot}\right)}{\iota'_{|\mathcal{X}||\mathcal{Z}|+1} M_y^{-1}\left(W_{\cdot y\cdot}\right)}, \tag{6}$$

where $G_x^*(\cdot)$ and $\bar{F}_y^*$ are the convex conjugates,

$$G_x^*\left(U_x\left(W_{x\cdot\cdot}, w_{x0}\right)\right) = \max_{p \in \triangle^{|\mathcal{Y}||\mathcal{Z}|+1}} \left\{ \sum_{y,z} p_{yz} u_{xyz}\left(w_{xyz}\right) + p_0 u_0\left(w_{x0}\right) + G_x(p) \right\},$$

$$\bar{F}_y^*\left(W_{\cdot y\cdot}\right) = \max_{q \in \triangle^{|\mathcal{X}||\mathcal{Z}|+1}} \left\{ \bar{F}_y(q) - \sum_{x,z} w_{xyz} q_{xz} \right\},$$

and $U_x\left(W_{x\cdot\cdot}, w_{x0}\right) := \left(\left(u_{xyz}\left(w_{xyz}\right)\right)_{(y,z)\in\mathcal{Y}\times\mathcal{Z}}, u_{x0}\left(w_{x0}\right)\right)$ is the vector of utilities for a worker of type x given a wage schedule $W_{x\cdot\cdot}$. Note that the optimal choice probabilities are given by the multinomial logit choice probabilities when the perturbation function is given by Shannon entropy, $H(p) = -\log p$.

The aggregate equilibrium is affected by the total indirect utility and profit of workers and firms defined as

$$G^*\left(U\left(W, w_0\right)\right) := \sum_x r_x G_x^*\left(U_x\left(W_{x\cdot\cdot}, w_{x0}\right)\right), \quad \bar{F}^*(W) := \sum_y \bar{F}_y^*\left(W_{\cdot y\cdot}\right),$$

where $U(W, w_{x0}) = (U_x(W_{x\cdot\cdot}, w_{x0}))_{x\in\mathcal{X}}$. Since the convex conjugates $G_x^*(U_x(W_{x\cdot\cdot}, w_{x0}))$ and $\bar{F}_y^*(W_{\cdot y\cdot})$ are convex in U_x and concave in $W_{\cdot y\cdot}$, then $G^*(U(W, w_{x0}))$ and $\bar{F}^*(W)$ are also convex in U and concave in W, respectively. From the Williams-Daly-Zachary theorem, it follows that the mass of workers of type x choosing a pair (y, z) and the mass of firms of type y choosing a pair (x, z) are given by the quantities

$$p_x p_{yz}^x(W_{x\cdot\cdot}) = \frac{\partial G^*(U(W, w_{x0}))}{\partial u_{xyz}} = r_x \frac{\partial G_x^*(U_{x\cdot\cdot}(W_{x\cdot\cdot}, w_{x0}))}{\partial u_{xyz}},$$
$$q_{xz}^y(W_{\cdot y\cdot}) = \frac{\partial \bar{F}^*(W)}{\partial w_{xyz}} = \frac{\partial \bar{F}_y^*(W_{\cdot y\cdot})}{\partial w_{xyz}},$$

respectively. This leads to the following equivalence between our equilibrium outcome and the one in Galichon *et al.* (2019).

Proposition 1. *Let Assumption 2 hold. Then there exists an aggregate equilibrium outcome (μ^*, W^*) given the utility index, perturbation functions, and unemployment benefit if and only if*

(i) (Market clearing) $\mu^* \in \mathcal{M}(r)$;

(ii) (Optimality) Given the vector of equilibrium wages W^*,

$$\mu_{xyz}^* = r_x \left.\frac{\partial G_x^*(U_{x\cdot\cdot}(W_{x\cdot\cdot}, w_{x0}))}{\partial u_{xyz}}\right|_{W_{x\cdot\cdot}=W_{x\cdot\cdot}^*} = \left.\frac{\partial \bar{F}_y^*(W_{\cdot y\cdot})}{\partial w_{xyz}}\right|_{W_{\cdot y\cdot}=W_{\cdot y\cdot}^*}$$

and

$$\mu_{x0}^* = r_x \left.\frac{\partial G_x^*(U_{x\cdot\cdot}(W_{x\cdot\cdot}, w_{x0}))}{\partial u_{x0}}\right|_{W_{x\cdot\cdot}=W_{x\cdot\cdot}^*}.$$

Furthermore, $\mu^* \in \operatorname{Int}\left(\triangle^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|+|\mathcal{X}|}\right)$.

We follow the results in Galichon *et al.* (2019) to study the aggregate equilibria in terms of a demand system. The matching of types (x, y, z) is treated as a good, with workers as producers and firms as consumers. Each worker of type x chooses to produce one of the goods of type (x, y, z), where $(y, z) \in \mathcal{Y}\times\mathcal{Z}$ or the outside option 0; similarly, each industry of type y chooses to consume one of the goods of type (x, y, z), where $(x, z) \in \mathcal{X} \times \mathcal{Z}$ or the outside option 0. Then, $r_x \partial G_x^*(U_x(W_{x\cdot\cdot}, w_{x0}))/\partial u_{xyz}$ and $\partial \bar{F}_y^*(W_{\cdot y\cdot})/\partial w_{xyz}$ can be interpreted as supply and demand of the good (x, y, z), respectively, if the price matrix is W. Let D_{xyz} define the excess demand function for the good (x, y, z)

$$\begin{aligned} D_{xyz}(W_{x\cdot\cdot}, W_{\cdot y\cdot}) &:= \partial \bar{F}^*(W)/\partial w_{xyz} - \partial G^*(U(W, w_{x0}))/\partial u_{xyz} \\ &= \partial \bar{F}_y^*(W_{\cdot y\cdot})/\partial w_{xyz} - r_x \partial G_x^*(U_x(W_{x\cdot\cdot}, w_{x0}))/\partial u_{xyz}. \end{aligned}$$

As a result, $D(W) = \{D_{xyz}(W_{x\cdot\cdot}, W_{\cdot y\cdot}) : x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}\}$ can be interpreted as the excess demand function and W as a matrix of market prices: if w_{xyz} increases and all the other entries of W remain constant, the utility $u_{xyz}(w_{xyz})$ of the worker in match (x, y, z) increases. Hence, the excess demand, D_{xyz}, for the match (x, y, z) increases.

We now state and prove a theorem that ensures the existence and uniqueness of an equilibrium using the characterization of aggregate equilibrium as a demand system. We show there exists a unique price matrix W^*, at which the value of excess demand is 0.

Definition 2. A function $D : \mathbb{R}^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|} \to \mathbb{R}^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|}$ is called monotone if it has the property that $(D(W^1) - D(W^2))'(W^1 - W^2) \geq 0$ for all W^1 and W^2, and strictly monotone if this inequality is strict when $W^1 \neq W^2$.

The monotonicity is a valuable tool in the solution mapping and is connected to the complementarity condition in convex analysis[6]. However, as the wages satisfying the complementarity condition do not guarantee $D(W)$ to be zero in $\mathbb{R}_+^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|}$, it is required to extend our domain to $\mathbb{R}^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|}$ and specify the worker's utility function for the extended domain.

Assumption 3. *$u_{xyz}(w_{xyz}) = \alpha_x w_{xyz} + \beta_{xyz}$ with nonnegative α_x for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.*

Lemma 1. *Under Assumption 2 and 3, $-D$ is monotone.*

Our results can be employed in the classical matching model with perfectly transferable utility, where the total welfare of the match is unaffected of the utility transfer. Although recent studies such as Galichon and Salanie (2021) do not constrain the model to the generalized extreme value class, alternatives are always substitutes in an additive random utility model. Our model may allow for complementarity.

By applying the perturbed utility model, we allow for more general substitution patterns, including complementarity. However, we need to limit $u_{xyz}(\cdot)$ for workers of type x to be linear in wages for matching models with imperfectly transferable utility. This linearity condition on deterministic utilities is more restricted than in Galichon *et al.* (2019)[7], requiring monotonicity in w. However, we can still apply the following unique existence result to the wage market with proportional taxation or agent-specific taxation in Dupuy *et al.* (2020).

[6]The monotonicity shares the same definition with the stable game in Hofbauer and Sandholm (2007) and we can think of the existence of equilibrium in Hofbauer and Sandholm (2007). In the stable game, the Nash equilibrium is characterized in terms of invasion.

[7]Under the gross substitutability in the class of an ARUM, both $\partial G^*(U(W))/\partial U$ and $-\partial \bar{F}^*(W)/\partial W$ are strictly diagonally isotone and in the type of an M-function, which is off-diagonally antitone and inverse isotone. This allows us to use theorems in Rheinboldt (1970) and Berry *et al.* (2013) for the unique existence of equilibrium.

Theorem 1. *Under Assumptions 2 and 3, there exists a unique matrix W^* such that $D(W^*) = 0$.*

Corollary 1. *Under Assumptions 2 and 3, there exists a unique equilibrium (μ^*, W^*).*

We established a new existence result for the aggregate equilibrium outcome that is not based on the gross substitutability condition of Kelso and Crawford (1982). We note from Fosgerau and Nielsen (2021) that

$$\nabla q^y (W_{\cdot y \cdot}) = \nabla M_y (q^y (W_{\cdot y \cdot}))^{-1} - q^y (W_{\cdot y \cdot}) \, q^y (W_{\cdot y \cdot})' .$$

For the logit model, $M_y (q) = \log q$ and the gross substitutability condition holds. However, we obtain the result for the model with the general class of perturbation functions G_x and $\bar{F}_y$, which leads to a model in which complementarity may arise. Fosgerau *et al.* (2021b) give a simple example of the model with taste for variety at the group level.

4 Complementarity and substitutability with similarity model

The perturbed utility model is a general class of behavioral models that nests the class of additive random utility models (ARUM). Hence, our matching model allows for more general forms of unobserved heterogeneity compared to Galichon and Salanie (2021).

By specifying the perturbation function in terms of a similarity function, defined below, Fosgerau and Nielsen (2021) introduced a class of perturbed utility models that allow alternatives to be both substitutes and complements.

Definition 3 (Similarity function)**.** Let $\mathcal{C} = \{1, \ldots, C\}$ index a finite set of characteristics among a menu $\mathcal{J} = \{1, \ldots, J\}$ and let $\Phi = (\phi_{cj})_{c \in \mathcal{C}, j \in \mathcal{J}}$ with nonnegative entries $\phi_{cj} \geq 0$ and columns that sum to 1, $\sum_{c \in \mathcal{C}} \phi_{cj} = 1$. A similarity function $S : \mathbb{R}^J_+ \to \mathbb{R}$ has the form

$$S(q) = \sum_{c \in \mathcal{C}} \lambda_c \sum_{j \in \mathcal{J}} \phi_{cj} q_j \log \left(\frac{\phi_{cj} q_j}{\sum_{j' \in c} \phi_{cj'} q_{j'}} \right) - \sum_{j \in \mathcal{J}} q_j \log q_j ,$$

where $\lambda = (\lambda_1, \ldots, \lambda_C)' \in \mathbb{R}^C$ is a vector of parameters associated with the characteristics in $\mathcal{C}$ satisfying $\sum_{c \in \mathcal{C}} \max\{\lambda_c, 0\} \phi_{cj} < 1$ for all $j \in \mathcal{J}$.

Fosgerau and Nielsen (2021) show that the similarity function is strictly concave when $\sum_{c \in \mathcal{C}} \max\{\lambda_c, 0\} \phi_{cj} < 1$ for all $j \in \mathcal{J}$ and satisfies Assumption 2.

λ_c can be interpreted as measuring the degree of substitutability within a set of alternatives, $\mathcal{J}_c = \{j \in \mathcal{J} : \phi_{cj} > 0\}$, compared to the logit model. If $\lambda_c > 0$, alternatives in $\mathcal{J}_c$ are more substitutable compared to the logit model. For $\lambda_c < 0$, alternatives in $\mathcal{J}_c$ are less substitutable and may be complementary. Below we give examples for different similarity functions that nest well-known ARUM models.

Example 1 (Multinomial logit model)**.** Given that $\lambda_c = 0$ for all $c \in \mathcal{C}$, the similarity function reduces to the Shannon entropy:

$$S(q) = -\sum_j q_j \log q_j.$$

Then, as shown by Fosgerau and Nielsen (2021), the optimal choice probabilities are given by the multinomial logit choice probabilities

Example 2 (Nested logit model)**.** Consider a simple two-level nested logit model that partitions a menu of alternatives into nests, $\{\mathcal{J}_1, \ldots, \mathcal{J}_C\}$. Then the similarity function is

$$\begin{aligned} S(q) &= -\sum_c \lambda_c \left(\sum_{j\in\mathcal{J}_c} q_j\right) \log\left(\sum_{j\in\mathcal{J}_c} q_j\right) - \sum_c \sum_{j\in\mathcal{J}_c} (1-\lambda_c)\, q_j \log q_j \\ &= -\sum_{j\in\mathcal{J}} q_j \log q_j + \sum_c \lambda_c \sum_j 1_{\{j\in\mathcal{J}_c\}} q_j \log\left(\frac{1_{\{j\in\mathcal{J}_c\}} q_j}{\sum_{j'} 1_{\{j'\in\mathcal{J}_c\}} q_{j'}}\right). \end{aligned}$$

with the nesting parameters $\{\lambda_c : c = 1, \ldots, C\}$ and $0 \le \lambda_c < 1$.

For $\lambda_c < 1$, the model extends to the generalized nested choice model by Allen and Rehbeck (2019b), where complementarity can occur between any two alternatives in any nest J_c with parameter $\lambda_c < 0$.

Fosgerau *et al.* (2021b) introduced the Inverse Product Differentiation Logit (IPDL) model, which allows for complementarity. As another example of similarity function, we consider the IPDL model for the following simple firm's problem: two characteristics divide the set of workers into four groups.

Example 3. We consider the representative firm's problem with a menu $\mathcal{J} = (\mathcal{X} \times \mathcal{Z}) \cup \{0\}$ where $\mathcal{X} = \mathcal{Z} = \{1, 2\}$. Let $\mathcal{C} = \left\{c_{\mathcal{X}|1}, c_{\mathcal{X}|2}, c_{\mathcal{Z}|1}, c_{\mathcal{Z}|2}\right\}$ where

$$c_{\mathcal{X}|z} = \{(x,z) : x \in \mathcal{X}\}, \quad c_{\mathcal{Z}|x} = \{(x,z) : z \in \mathcal{Z}\},$$

and $\phi_{cj} = 1_{\{j\in c\}}/2$ indicates that the alternative j belongs to nest c. Then the firm faces the maximization problem

$$\max_{q=(q_0,q_{11},q_{12},q_{21},q_{22})\in\triangle^5} \left\{\bar{F}(q) - \sum_{x,z=1}^{2} w_{xz} q_{xz}\right\},$$

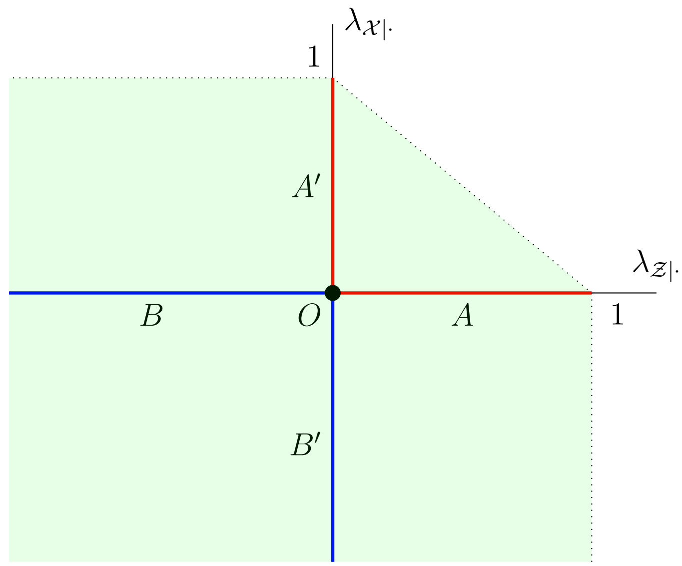


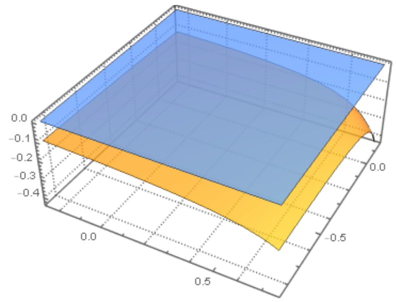
Figure 2: Nesting parameter region for similarity function

where $\triangle = \left\{ q \in \mathbb{R}^5_+ | q_0 + \sum_{x,z=1}^{2} q_{xz} = 1 \right\}$ is the probability simplex and

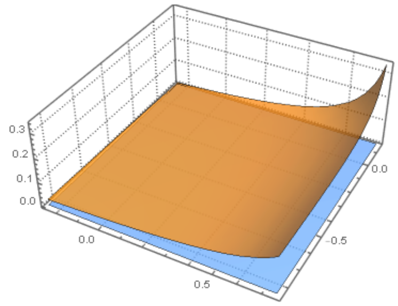
$$\begin{aligned}\bar{F}(q) = &- q_0 \log q_0 - \sum_{x,z=1}^{2} q_{xz} \log (q_{xz}) \\ &+ \sum_{x,z=1}^{2} q_{xz} \left[\lambda_{\mathcal{X}|z} \log \left(\frac{q_{xz}}{\sum_{x'=1}^{2} q_{x'z}} \right) + \lambda_{\mathcal{Z}|x} \log \left(\frac{q_{xz}}{\sum_{z'=1}^{2} q_{xz'}} \right) \right].\end{aligned}$$

This similarity model includes the multinomial logit and (generalized) nested choice models as special cases. The region for the nesting parameter of the similarity model is shown in Figure 2. When $\lambda_{\mathcal{X}|1}$ and $\lambda_{\mathcal{X}|2}$ are zero, the model reduces to the nested choice model with a hierarchical nesting structure, where the worker type is the first layer and the occupation type is the second layer. Conversely, when $\lambda_{\mathcal{Z}|1}$ and $\lambda_{\mathcal{Z}|2}$ are zero, the occupation type is the first layer and the worker type is the second layer. The lines A and A' indicate the regions of nesting parameters for the nested logit model, and $A \cup B$ and $A' \cup B'$ indicate the regions of nesting parameters for the generalized nested choice model. Hence, the similarity model has a more general structure as it allows the nesting structure to be non-hierarchical.

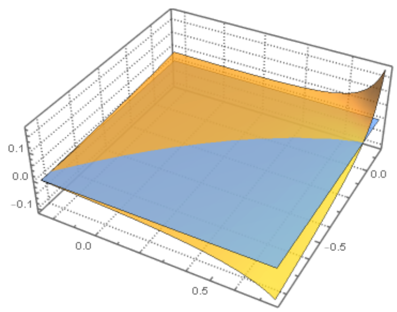
We now illustrate the substitution patterns for different values of the nesting parameters. Let $\lambda_{\mathcal{X}|1} = \lambda_{\mathcal{X}|2} = \lambda_{\mathcal{X}}$ and $\lambda_{\mathcal{Z}|1} = \lambda_{\mathcal{Z}|2} = \lambda_{\mathcal{Z}}$ and let the wages be identical across alternatives, $W = (w_{xz})_{x \in \mathcal{X}, z \in \mathcal{Z}} = (w, w, w, w)'$. Then the choice probabilities are given by $q_{xz}(W) = \frac{1}{4 + e^{w} 2^{(\lambda_{\mathcal{X}} + \lambda_{\mathcal{Z}})}}$ for all alternatives. However, the substitution between the different alternatives are not identical. This is illustrated by Figure 3 that, for different values of $(\lambda_{\mathcal{X}}, \lambda_{\mathcal{Z}})$, plots the derivatives of the choice probabilities, $\partial q_{xz}(W) / \partial w_{11}$, for



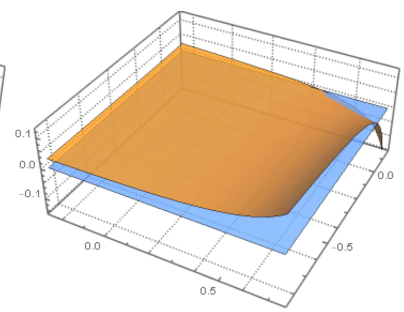
(a) $\partial q_{11}^{*}(W)/\partial w_{11}|_{W=(1,1,1,1)}$



(b) $\partial q_{12}^{*}(W)/\partial w_{11}|_{W=(1,1,1,1)}$



(c) $\partial q_{21}^{*}(W)/\partial w_{11}|_{W=(1,1,1,1)}$



(d) $\partial q_{22}^{*}(W)/\partial w_{11}|_{W=(1,1,1,1)}$

Figure 3: Demand substitutability and complementarity over $(\lambda_{\mathcal{X}}, \lambda_{\mathcal{Z}}) \in [-0.25, 0.75] \times [-0.85, 0.15]$: A blue area represents the values of zero.

each $(x, z) \in \mathcal{X} \times \mathcal{Z}$. Note $\partial q_{11}(W)/\partial w_{11}$ is negative for all pairs of $(\lambda_{\mathcal{X}}, \lambda_{\mathcal{Z}})$, which means that the probability of choosing alternative $(1,1)$ decreases as its own wage increases. Figure 3.(b) shows that alternatives $(1,1)$ and $(1,2)$ become closer substitutes as $\lambda_{\mathcal{Z}}$ increases. However, interestingly, Figure 3.(c) shows that alternatives $(1,1)$ and $(2,1)$ become complements as $\lambda_{\mathcal{X}}$ decreases. This illustrates that the substitution pattern depends on $\lambda_{\mathcal{Z}}$ as well as $\lambda_{\mathcal{X}}$ and the model can allow for complementarity.

5 Model specification and estimation

The choice specific payoff, U_x, and the perturbation functions, $\left(G_x, \bar{F}_y\right)$ represent the preferences of the workers and the technology of the firms. This section specifies workers' preferences and firms' technologies, and proposes a maximum likelihood estimator.

5.1 Model specification

The utility of the worker linearly increases in post-tax income

$$u_{xyz}(w_{xyz}) = \alpha\left(w_{xyz} - \tau\left(w_{xyz}\right)\right) + \beta_{xyz}, \tag{7}$$

$$u_{x0}(w_0) = \alpha (w_0 - \tau (w_0)),$$

and the parameters (α, β_{xyz}) represent the marginal utility of post-tax income and the utility of working in a given match, respectively. It should be noted that α has to be nonnegative in order to obey Assumption 3 and to guarantee existence and uniqueness of the equilibrium. The tax function, $\tau(\cdot)$, is a smooth and convex function which in turn implies that worker's utility is concave in pre-tax income[8].

The perturbation function is assumed to be identical across worker types and, as suggested by Fosgerau and Nielsen (2021), it is given by the similarity function,

$$G_x(p) = \sum_{y,z} p_{yz} h_{yz}(p_{y\cdot}, p_{\cdot z}) + p_0 h_0(p_0),$$

where the functions $\left\{ (h_{yz}(p_{y\cdot}, p_{\cdot z}))_{(y,z)\in\mathcal{Y}\times\mathcal{Z}}, h_0(p_0) \right\}$ describe workers' taste-for-variety over different types of firms and occupations

$$\begin{aligned} h_{yz}(p_{y\cdot}, p_{\cdot z}) &= -\log(p_{yz}) + \eta_{\mathcal{Y}} \log\left(\frac{p_{yz}}{\sum_{y'} p_{y'z}} \right) + \eta_{\mathcal{Z}} \log\left(\frac{p_{yz}}{\sum_{z'} p_{yz'}} \right), \\ h_0(p_0) &= -\log(p_0). \end{aligned}$$

The nesting parameters $(\eta_{\mathcal{Y}}, \eta_{\mathcal{Z}})$ determine the degree of substitution between similar types of firms and similar types of occupations, respectively. As we illustrated in Example 3, the perturbation function $G_x(p)$ satisfies Assumption 2 if $\max\{\eta_{\mathcal{Y}}, 0\} + \max\{\eta_{\mathcal{Z}}, 0\} < 1$. If $\eta_{\mathcal{Y}} = \eta_{\mathcal{Z}} = 0$, $h_{yz}(p_{y\cdot}, p_{\cdot z}) = -\log(p_{yz})$ and $h_0(p_0) = -\log(p_0)$, then the similarity function reduces to the Shannon entropy and the optimal choice probabilities are given by the multinomial logit choice probabilities.

As already noted the perturbation function for the firms represents the production function

$$\bar{F}_y(q) = \sum_{x,z} q_{xz} (\pi_{xyz} + m_{xz}(q_{x\cdot}, q_{\cdot z})) + q_0 (\pi_{y0} + m_0(q_0)),$$

where the functions $\left\{ (m_{xz}(q_{x\cdot}, q_{\cdot z}))_{(x,z)\in\mathcal{X}\times\mathcal{Z}}, m_0(q_0) \right\}$ are defined as

$$\begin{aligned} m_{xz}(q_{x\cdot}, q_{\cdot z}) &= -\log(q_{xz}) + \lambda_{\mathcal{X}} \log\left(\frac{q_{xz}}{\sum_{x'} q_{x'z}} \right) + \lambda_{\mathcal{Z}} \log\left(\frac{q_{xz}}{\sum_{z'} q_{xz'}} \right), \\ m_0(q_0) &= -\log(q_0). \end{aligned}$$

$(\pi_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}$ reflects the productivity of matches and the functions $m_{xz}(\cdot)$ and $m_0(\cdot)$ describe how the firms' production is affected by the input shares of different worker

[8] The applied tax function is described in Subsection 6.3.

and occupation types. Note that $\left((\pi_{xyz})_{(x,z)\in\mathcal{X}\times\mathcal{Z}}, \pi_{y0}\right)$ and $\left((m_{xz}(\cdot))_{(x,z)\in\mathcal{X}\times\mathcal{Z}}, m_0(\cdot)\right)$ represent the linear and nonlinear parts of production, respectively. For identification, π_{y0} is normalized to zero for all $y \in \mathcal{Y}$.

Higher positive values for $\lambda_{\mathcal{X}}$ imply stronger substitution between matches with similar worker types relative to the logit model. In contrast, smaller values lead to weaker substitution, and workers can become complements when $\lambda_{\mathcal{X}}$ is sufficiently negative. The interpretation of $\lambda_{\mathcal{Z}}$ is similar to $\lambda_{\mathcal{Z}}$, but with respect to matches with similar occupations. Similar to workers' perturbation function, the firms' production function $\bar{F}_y(q)$ satisfies Assumption 2 if $\max\{\lambda_{\mathcal{X}}, 0\} + \max\{\lambda_{\mathcal{Z}}, 0\} < 1$.

5.2 Maximum likelihood estimation

Given the observed number of matched workers and observed average wages, $\left\{\tilde{n}, \tilde{W}\right\}$, the vector of structural parameters, $\theta = \{\beta, \alpha, \eta, \pi, \lambda\}$, can be estimated by maximum likelihood estimation.

Equations (5) and (6) can not be used to compute the optimal choice probabilities for workers and firms as $H^{-1}(\cdot)$ and $M^{-1}(\cdot)$ cannot be expressed in closed form even though they exist. Following Fosgerau and Nielsen (2021), it is shown in Appendix A that the optimal choice probabilities can be expressed as fixed points

$$p^x = \frac{p^x \circ e^{U_x(W_{x\cdot\cdot})+H(p^x)}}{\left\langle p^x, e^{U_x(W_{x\cdot\cdot})+H(p^x)}\right\rangle}, \quad q^y = \frac{q^y \circ e^{(\pi_{\cdot y\cdot} - W_{\cdot y\cdot})+M(q^y)}}{\left\langle q^y, e^{(\pi_{\cdot y\cdot} - W_{\cdot y\cdot})+M(q^y)}\right\rangle}. \tag{8}$$

Fosgerau and Nielsen (2021) also provide a contraction mapping and prove that, for any initial guess of the choice probabilities, the contraction mapping converges to the optimal choice probabilities. The contraction mapping is given in Appendix.B. For later use define the logarithm of the denominator of the fixed points

$$ws_x = \log\left\langle p^x, e^{U_x(W_{x\cdot\cdot})+H(p^x)}\right\rangle, \quad fs_y = \log\left\langle q^y, e^{(\pi_{\cdot y\cdot} - W_{\cdot y\cdot})+M(q^y)}\right\rangle. \tag{9}$$

These equal respectively the worker and firm surplus up to a constant. By inserting the fixed point of the optimal choice probabilities, Eq. (8), into the market clearing condition given by Eq. (3) we can derive an expression for equilibrium wages

$$w_{xyz} = \frac{1}{1+\alpha}\Big\{ \log(q_y/p_x) + (\pi_{xyz} - fs_y) - (\beta_{xyz} - \alpha\tau(w_{xyz}) - ws_x) + (\log q^y_{xz} + m_{xz}(q^y)) - \left(\log p^x_{yz} + h_{yz}(p^x)\right) \Big\}. \tag{10}$$

Equation (10) is a fixed point for equilibrium wages since $(\tau(w_{xyz}), ws_x, fs_y, p^x, q^y)$ are also functions of equilibrium wages. This fixed point based on the similarity model generalizes the expression for equilibrium wages that Dupuy *et al.* (2020) derived for the logit model. Let $w_{xyz}(\theta)$ denote the derived equilibrium wage for a given vector of structural parameters, θ.

When estimating the structural parameters, we allow observations of matches and wages across several time periods. Hence, let superscript $t = 1, 2, ..., T$ denote the time period of any given variable. Similar to Dupuy *et al.* (2020), we assume that the equilibrium wage is observed with an independent identical normal distributed measurement error, $\varepsilon_{xyz}^t \sim N(0, \sigma^2)$,

$$\tilde{w}_{xyz}^t = w_{xyz}^t(\theta) + \varepsilon_{xyz}^t,$$

Let $\tilde{n}_{x0}^t$ denote the observed number of unmatched workers and let $\tilde{n}_{xyz}^t$ denote the observed number of workers in a given match. Due to the assumption of independence between the measurement errors the likelihood contribution of any matched worker type is given by the product of the likelihood of the observed number of matches and the likelihood of the observed wage

$$\ell_{xyz}^t(\theta) = Pr\left(\mu_{xyz}^t, \tilde{w}_{xyz}^t|\theta\right) = p_{yz}^{x,t}\left(W_{x\cdot\cdot}^t(\theta)|\theta\right)^{\tilde{n}_{xyz}^t} f\left(\tilde{w}_{xyz}^t|\theta\right),$$
$$\ell_{x0}^t(\theta) = Pr\left(\mu_{x0}^t|\theta\right) = p_0^{x,t}\left(W_{x\cdot\cdot}^t(\theta)|\theta\right)^{\tilde{n}_{x0}^t},$$

where $p_{yz}^{x,t}(W_{x\cdot\cdot}^t(\theta)|\theta)$ is the choice probability of the worker[9] and $f\left(\tilde{w}_{xyz}^t|\theta\right)$ is the density function of the observed wage. Note that, for any type of unmatched worker, the likelihood contribution is simply given by the likelihood of the observed number of unemployed, since it is assumed that the benefit payment, w_{x0}^t, is known with certainty. The corresponding likelihood function is given by the product of the individual likelihood contributions

$$L(\theta) = \left\{\prod_{t=1}^{T}\prod_{x}\ell_{0x}^t(\theta)\right\}\left\{\prod_{t=1}^{T}\prod_{x,y,z}\ell_{xyz}^t(\theta)\right\}.$$

As a result, we can express the log-likelihood function as the sum of the log-likelihood function of observed matches, $\log L_M(\theta)$, and the log-likelihood function of observed wages, $\log L_W(\theta)$

$$\log L(\theta) = \log L_M(\theta) + \log L_W(\theta), \tag{11}$$

where the likelihood function of the observed matches is

$$\log L_M(\theta) = \sum_{t=1}^{T}\sum_{x}\mu_{x0}^t \cdot \log p_0^{x,t}\left(W_{x\cdot\cdot}^t(\theta)|\theta\right) + \sum_{t=1}^{T}\sum_{x,y,z}\mu_{xyz}^t \cdot \log p_{yz}^{x,t}\left(W_{x\cdot\cdot}^t(\theta)|\theta\right).$$

[9]Since the fixed points of the optimal choice probabilities, Eq. (8), are not ensured to converge, we cannot use these to calculate the choice probabilities. However, Fosgerau and Nielsen (2021) provide a contraction mapping and prove that for any initial guess of the choice probabilities, the contraction mapping converges to the true choice probabilities. Hence, when evaluating the likelihood function we use this contraction mapping to find optimal choice probabilities of firms and workers.

Since the measurement errors are assumed to be normal distributed the log-likelihood function of the observed wages is

$$\log L_W(\theta) = -\sum_{t=1}^{T}\sum_{x,y,z}\left(\tilde{w}_{xyz}^t - w_{xyz}^t(\theta)\right)^2 \frac{1}{2\sigma^2(\theta)} - \frac{T|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|}{2}\log\left(\sigma^2(\theta)\right),$$

and as it can be shown that the variance of the measurement error that maximizes the likelihood function can be expressed in closed form

$$\sigma^2(\theta) = \frac{1}{T|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|}\sum_{t=1}^{T}\sum_{x,y,z}\left(\tilde{w}_{xyz}^t - w_{xyz}^t(\theta)\right)^2,$$

we can concentrate this parameter out of the log-likelihood function and thereby reduce the number of parameters we maximize the likelihood function with respect to.

Note that, for each guess of θ, the model has to be solved T times in order to find the equilibrium wages for each period. This is by far the most computational costly operation when evaluating the full likelihood function. Two methods are available when solving for equilibrium wages. The equilibrium wages can for an initial guess for W^t be found by successful approximation by iterating on Eq. (10), or the equilibrium wages can be found by the use of gradient-based numerical methods in order to set excess demand to zero. The latter method is applied during estimation since we find that this is the most efficient method.

6 Data and estimation results

This section describes the data source used for estimation and presents the estimation results. The section is organized as follows. Subsection 6.1 outlines our data sources and construction of the data set used for estimation and Subsection 6.2 presents some descriptive statistics for the Danish labor market. Subsection 6.3 presents the used tax function that mimics the key features of the Danish tax system. Finally, Subsection 6.4 presents the estimation results and Subsection 6.5 graphically illustrates the fit of the model.

6.1 Data construction

Our data is drawn from several administrative registers from Statistics Denmark that cover the entire population of individuals. Since all individuals are associated with a unique identifier, the individual information can be combined across registers.

Observed matches between firms and workers are based on the register-based labor force statistics (Registerbaseret Arbejdststyrke Statistik, RAS). RAS records the labor market status in week 48 each year. For employed individuals, this register also contains information on occupation and industry. By excluding individuals with the labor market status of employer, self-employed, or self-employed spouses, we can divide the population into employed wage-earners and non-employed individuals. For employed wage-earners, we have detailed information on industry affiliation and occupation. Inspired by Traiberman (2019) and Hummels *et al.* (2014), we define four industries and six occupations. We also define four types of workers based on the individual's highest attained education. In turn, we consider 96 different triples of worker types, industries, and occupations.

The definition of the four industries is based on the Nomenclature of Economic Activities (NACE) two-digit industry code. *1. Manufacturing* (including utilities and construction) contains industries with NACE code 1-45; *2. FIRE* (including R&D) contains industries with code 46-82; *3. Public services* contains industries with code 84-90; and *4. Other services* contains the residual group of industries.

For the definition of the six occupation groups (*1. Management; 2 High level skills; 3 Medium level skill; 4. Basic level skills; 5. Other wage earners; 6. Missing observations*), we use a definition of Statistics Denmark that is based on the International Standard Classification of Occupations (ISCO). We treat missing observations as its own group, since this is a relatively big group of employed individuals, and we expect that these missing observations differ substantially from the remaining groups of occupation.

Lastly, the definition of the four types of workers is based on the International Standard Classification of Education (ISCED) two-digit code. Information on education comes from the register of education attainment (Uddannelsesregistret, UDDA). *1. Unskilled workers* include individuals with high school completion or less (ISCED code 10-20 or missing); *2. Skilled workers* include individuals with vocational education (ISCED code 30); *3. Medium educated workers* include individuals with short- or medium-cycle higher education or a bachelor degree (ISCED code 35-60); *4. Highly educated workers* include individuals with a master or PhD degree (ISCED code 70-80).

Observed wages are based on the register of employment for wage-earners (Beskæftigelse for lønmodtagere, BFL) that contains individual information on wages, industry affiliation, and occupation across employment spells. By merging BFL with UDDA, we can calculate the total wage payments across the 96 triples of educational, industry, and occupational groups. We then calculate the average wages for the 96 triples by dividing the corresponding total wage payments from BFL by the number of matches from RAS. When calculating observed matches and wages, we only include individuals in the age interval 18-65, where 65 is the official pension age in Denmark throughout the period we consider.

6.2 Descriptive statistics of the Danish labor market

Table 1 shows the aggregated numbers of individuals of the four education groups from 2010 through 2018. From the table, we see that most of the labor force is comprised of unskilled and skilled labor. Further, we see that the general level of education has increased for the labor force. More specifically, the number of medium and highly educated individuals has increased. Especially the relative increase among highly educated individuals has been dramatical. From 2010 to 2018, the number of highly educated individuals has increased from 242,000 to 367,000, corresponding to a 52 pct. increase. The number of medium educated individuals has also increased considerably. This group has increased from 676,000 to 782,000 corresponding to an increase of 16 pct. In contrast, the numbers of unskilled and skilled individuals have slightly decreased over the period.

	2010	2011	2012	2013	2014	2015	2016	2017	2018
1. Unskilled	1,240	1,233	1,225	1,217	1,210	1,200	1,196	1,194	1,185
2. Skilled	1,108	1,098	1,086	1,073	1,061	1,052	1,043	1,032	1,021
3. Medium educ.	676	687	697	706	720	737	757	767	782
4. Highly educ.	242	253	264	276	291	308	327	347	367
Total	3,266	3,271	3,273	3,272	3,281	3,297	3,322	3,341	3,356

Table 1: Number of individuals across years and education groups, 1,000. The table only includes individuals between 18-65 years old; employer, self-employed, and self-employed spouses have been excluded.

Table 2 shows the number of individuals across industries or unemployment. The largest industry is the residual group *Other Services,* with 971,000 employed individuals in 2018 followed by *Public Services* and *Manufacturing.* The smallest industry is *FIRE,* with 249,000 employed individuals in 2018. Further, we note that *Non-employed* is the largest group as this group constituted 885,000 individuals in 2018. However, this group has decreased over the period as it constituted 962,000 individuals in 2010. The number of individuals employed in the public sector has been almost constant throughout the period, whereas the number of individuals employed in the three other industries has increased. With an increase of 90,000 employed workers *Other services* has experienced the largest increase in employment. However, the percentage increase has been of similar magnitude across these three industries.

	2010	2011	2012	2013	2014	2015	2016	2017	2018
0. Non-employed	962	972	987	979	964	948	932	913	885
1. Manufacturing	489	494	487	481	490	498	510	522	532
2. FIRE	223	223	223	225	227	232	238	242	249
3. Public Services	713	703	699	704	705	708	707	711	718
4. Other Services	880	879	878	884	896	912	935	952	971
Total	3,266	3,271	3,273	3,272	3,281	3,297	3,322	3,341	3,356

Table 2: Number of workers in the labor force across years and industries, 1,000. The table only includes individuals between 18-65 years old; employer, self-employed, and self-employed spouses have been excluded.

Table 3 shows the number of employed individuals across occupation groups of varying size. The largest group consists of occupations that require *Basic Level* of skills, which in 2018 constituted more than a million employed individuals, whereas *Management* only constituted 104,000 individuals. Further, we see that the number of individuals employed in jobs classified as demanding *High level of skills* has increased the most from 573,000 to 686,000 employed individuals, corresponding to an increase of 20 pct. However, the number of individuals with missing information shows a similar relative increase of 17 pct. The only occupation group where the number of employed individuals has decreased is occupations that demand *Medium Level* of skills. This occupation group has decreased from 301,000 to 274,000 employed individuals, corresponding to a decrease of 9 pct.

	2010	2011	2012	2013	2014	2015	2016	2017	2018
1. Management	100	101	98	97	102	101	103	103	104
2. High Level	573	586	592	602	618	637	653	669	686
3. Medium Level	301	283	279	275	265	267	270	274	274
4. Basic Level	980	974	957	951	972	983	997	1,007	1,015
5. Other	197	196	191	193	198	200	206	211	213
6. Missing	154	158	169	175	162	160	161	164	180
Total Employment	2,305	2,299	2,286	2,293	2,317	2,349	2,391	2,428	2,471

Table 3: Number of employed workers across occupations and year, 1,000. The table only includes individuals between 18-65 years old; employer, self-employed, and self-employed spouses have been excluded.

Table 4 shows the empirical choice probabilities for the four worker types for 2018 (excluding the outside option of non-employment). The tables show that unskilled and skilled individuals are mostly concentrated in occupations that only require basic level of skills and primarily work in *Manufacturing* and *Other Services*. Medium educated individuals are especially concentrated in *Public Services*, where they are mostly employed in jobs that require *High level* of skills. Finally, highly educated individuals are to a great extent employed in occupations that require a high level of skills and are mostly concentrated in *Public Services* and *Other Services*.

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	0.3	0.1	0.1	0.7
2. High Level	0.3	0.8	2.6	1.6
3. Medium Level	0.6	0.7	0.7	2.1
4. Basic Level	7.6	1.4	6.8	16.0
5. Other	2.4	0.4	1.3	5.7
6. Missing	1.6	0.6	0.7	4.0

(a) Unskilled workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	1.0	0.2	0.1	1.5
2. High Level	0.6	1.4	1.4	1.6
3. Medium Level	2.1	1.6	1.9	4.6
4. Basic Level	17.6	2.1	13.0	16.4
5. Other	2.3	0.2	1.1	3.3
6. Missing	1.8	0.7	0.4	2.8

(b) Skilled workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	1.2	0.6	1.2	1.5
2. High Level	2.6	4.3	30.5	6.9
3. Medium Level	2.8	2.4	2.2	4.6
4. Basic Level	2.8	1.3	2.8	7.3
5. Other	0.4	0.1	0.3	1.8
6. Missing	0.9	0.7	0.3	2.0

(c) Medium educated workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	1.2	1.5	0.9	3.0
2. High Level	6.1	12.1	23.7	19.6
3. Medium Level	1.2	2.2	0.8	3.4
4. Basic Level	1.1	1.0	1.4	3.6
5. Other	0.2	0.1	0.1	1.2
6. Missing	0.6	0.6	0.2	1.3

(d) Highly educated workers

Table 4: Empirically choice probabilities in 2018 for each of the four worker types, pct. None of the tables sum to 100 pct. since non-employed workers are excluded.

Table 5 shows the average wages for the four educational groups across the 24 pairs of industries and occupations for 2018. The tables show that the average wages, in general, are highest for *Management* and *Missing observations*. This strongly indicates that information on occupations are not missing at random and serves as an additional justification for these observations to be treated as their own group of occupations. Further, we see that in general, the wages deceases as skill level decreases, as the wages tend to be higher for individuals in jobs that require *High level of skills* than jobs that require *Medium* or *Basic level of skills*. The tables also show that the average wages, in general, are highest for individuals employed in *Manufacturing* and *FIRE*. Further, the average wages tend to be lowest in the *Public Services*. Finally, the tables also show that average wages tend to increase with education level.

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	699	1,067	602	680
2. High Level	521	509	274	418
3. Medium Level	444	455	318	438
4. Basic Level	322	282	243	257
5. Other	334	248	253	258
6. Missing	444	490	292	375

(a) Unskilled workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	623	870	626	657
2. High Level	528	566	361	519
3. Medium Level	468	489	333	463
4. Basic Level	358	379	341	344
5. Other	368	340	304	340
6. Missing	1,028	638	507	676

(b) Skilled workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	846	1,077	563	830
2. High Level	596	549	409	488
3. Medium Level	502	483	435	456
4. Basic Level	406	370	306	396
5. Other	369	301	277	283
6. Missing	874	1,099	699	779

(c) Medium educated workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	1,194	1,485	767	1,012
2. High Level	684	659	583	546
3. Medium Level	594	711	460	574
4. Basic Level	480	577	346	718
5. Other	393	423	317	315
6. Missing	925	2,259	2,325	1,247

(d) Highly educated workers

Table 5: Average wages in 2018 for each of the four worker types, DKK 1,000 wages are measured in 2018 numbers.

For simplicity, the benefit payments are set to DKK $134,000$, corresponding to the maximum yearly payments of the unemployment insurance.

6.3 Tax function

Recall from Section 5.1 that the utility of the worker is increasing in post-tax income. Hence, in order to estimate the model we need to specify the tax function, $\tau(\cdot)$, from Eq. 7. As we use Danish register data for estimation the specified tax function mimics the key features of the Danish tax system which is a progressive and highly complicated tax system. Since we only include four different types of workers several aspects of the tax system have to be ignored. Hence, our tax function ignores, for example, pension savings, whether the worker is married, or whether the worker is a house owner.

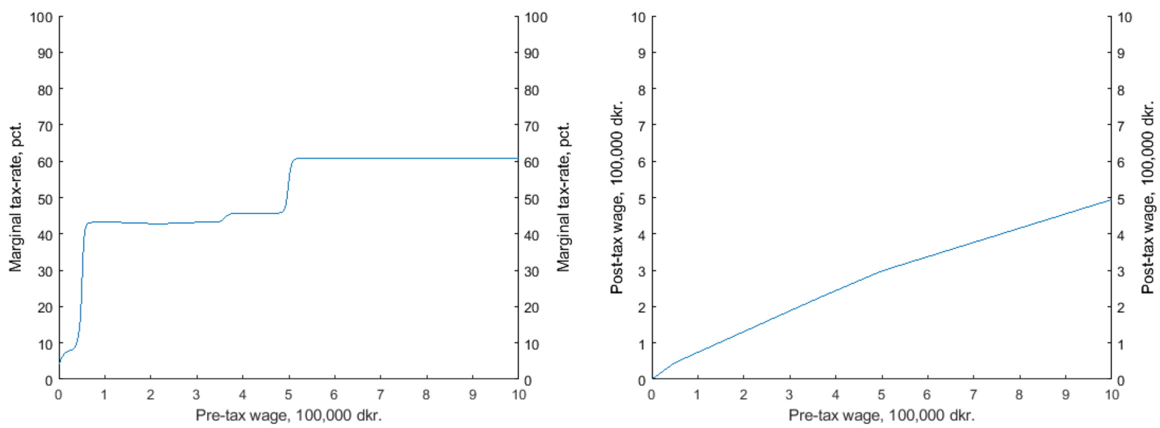
Figure 4 illustrates for 2018 the marginal tax rate of wage income and the corresponding post-tax wage income based on the specified tax function. Note that we have used the LogSum function to smooth out any kink of the tax-function.[10] We do so in order to apply gradient-based numerical methods when maximizing the log-likelihood function

[10] The LogSum function is specified as

$$T = \epsilon \log \left(e^{ry} + e^{m}\right) + m,$$
$$m = \left(B - max(ry, B)\right) / \epsilon,$$

where r is the tax rate paid when income, y, is above the tax-bracket, B. Finally, ϵ is a smoothing parameter. In turn T is the paid taxes after smoothing.

with respect to the structural parameters of the model. The figures show that a wage earner pays the labor market contribution of 8 percentage points for income less than the personal allowance of DKK 46,000. Above this threshold, the marginal tax rate increases to approximately 43 percentage points as the wage earner starts to pay the bottom tax rate, the municipal income tax and health contribution. For wage income that surpass DKK 361,000. the wage earner exempts her employment deduction and the marginal tax rate increases to approximately 46 percentage points. Finally, for income that exceed DKK 498,900. the wage earner starts to pay the top tax rate of 15 percentage points. As a result the marginal tax rate increases to approximately 60 percentage points.



(a) Marginal tax rate, 2018. **(b)** Post-tax wage income, 2018.

Figure 4: Post-tax wage and marginal rate of taxation for employed workers, 2018

6.4 Estimation results

Based on the Danish administrative register data we estimate the structural parameters of the described model in Subsection 5.1 by maximum likelihood. The used data set contains almost 30 million observations for the time period 2010 to 2018. Table 6 shows the estimated key parameters for two different specifications of the model. For the *logit model*, all the nesting parameters are restricted to zero, $\eta_{\mathcal{Y}} = \eta_{\mathcal{Z}} = \lambda_{\mathcal{X}} = \lambda_{\mathcal{Z}} = 0$, which implies that alternatives are substitutes. In contrast, for the *similarity model*, only the nesting parameters describing the preferences of the workers are restricted to zero, $\eta_{\mathcal{Y}} = \eta_{\mathcal{Z}} = 0$. The nesting parameters describe the technology of the firms and are estimated freely. Hence, this model nests the *logit model*. Based on a likelihood ratio test statistic of size $2,018$, we find that the inclusion of the two additional parameters in the *similarity model* increases the log-likelihood significantly. Since the estimated nesting parameter, $\hat{\lambda}_{\mathcal{Z}} = 0.8963$, is positive, matches with different occupations are closer substitutes than for the logit model. In other words, $\lambda_{\mathcal{Z}} > 0$ implies that similar worker types are close substitutes as matches with different occupations share the same worker type. In contrast, the negative estimate of the nesting parameter, $\hat{\lambda}_{\mathcal{X}} = -0.1944$, implies that matches of workers with different education level are less substitutes and can even be complements. In other words, matches with similar occupations are complements.

	Logit model	Similarity model
$\hat{\alpha}$	2.1185	2.1906
	(0.0001)	(0.0001)
$\hat{\lambda}_{\mathcal{Z}}$	-	0.8963
	-	(0.0001)
$\hat{\lambda}_{\mathcal{Y}}$	-	-0.1944
	-	(0.0001)
σ	0.1956	0.2746
$\log L$	-68,435,700	-68,434,691
$\log L_M$	-68,437,542	-68,436,240
$\log L_W$	1,842	1,549
Obs.	29,678,958	29,678,958

Table 6: Maximum likelihood estimates based on observed matches and wages for the time period 2010 through 2018. wages are measured in DKK 1,000,000 in 2018 numbers. Standard error, calculated from the scores of the log-likelihood, are in parentheses.

The estimated marginal utility of consumption, $\hat{\alpha} = 2.1606$, is relatively low compared to the estimated utilities of the matches, $\hat{\beta}_{xyz}$, given in Table 7. To see this, note that income is measured in million DKK, and recall from Table 5 in Subsection 6.2 that the average pre-tax wages for most matches are less than a million DKK and the pre-tax benefit paiments are exogenous set to DKK $134,000$. Hence, for most matches, the worker derives more utility from the matches than the induced additional utility of consumption. Due to the relatively low estimates of the marginal utility of consumption any wage change only leads to modest changes in the supply of labor.

Note that the estimated utility parameters of the matches from Table 7 are highly correlated with the empirical choice probabilities in Table 4. For instance unskilled workers are typically employed in *Other services* and in occupations that require *Basic level of skills* and Table 7 shows that for for unskilled workers the estimated $\hat{\beta}_{xyz}$ takes its largest value for this match. Similarly, skilled workers are typically employed in manufacturing or other services in occupations that require *Basic level of skills* and the two largest estimates of $\hat{\beta}_{xyz}$ for this type of worker is for these matches. For medium and highly educated workers, the estimated $\hat{\beta}_{xyz}$ takes its largest value for matches in the public sector where the occupations, in which these types of workers are typically employed, require *high level skills*. Note that almost all the estimates of $\hat{\beta}_{xyz}$ are negative. This is a result of non-employed workers being the biggest group even though they receive the lowest income, see Table 2.

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	-5.57	-6.74	-6.82	-4.66
2. High Level	-5.45	-4.52	-2.51	-3.73
3. Medium level	-4.58	-4.37	-4.16	-3.21
4. Basic Level	-1.48	-3.18	-2.04	-0.71
5. Other	-2.67	-4.93	-3.47	-1.77
6. Missing	-3.61	-4.86	-4.27	-2.21

(a) Unskilled workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	-3.55	-5.38	-5.81	-3.10
2. High Level	-4.16	-3.31	-2.53	-3.09
3. Medium level	-3.06	-3.13	-2.25	-3.18
4. Basic Level	0.03	-2.60	-0.27	-0.02
5. Other	-2.07	-4.76	-2.85	-1.68
6. Missing	-3.36	-4.07	-4.64	-2.57

(b) Skilled workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	-3.52	-4.45	-2.86	-3.35
2. High Level	-2.55	-2.05	0.66	-1.58
3. Medium level	-2.09	-2.34	-1.66	-1.52
4. Basic Level	-2.52	-3.23	-2.42	-1.03
5. Other	-4.58	-5.26	-4.16	-3.10
6. Missing	-3.79	-4.21	-4.72	-2.97

(c) Medium educated workers

	1. Manufac.	2. FIRE	3. Public	4. Other
1. Management	-3.36	-3.26	-3.24	-2.27
2. High Level	-1.39	-0.70	0.40	-0.23
3. Medium level	-3.03	-2.39	-3.40	-1.93
4. Basic Level	-3.30	-3.25	-2.88	-2.03
5. Other	-5.14	-4.82	-4.76	-3.19
6. Missing	-3.88	-5.02	-5.64	-3.23

(d) Highly educated workers

Table 7: Estimated utility of the matches, $\hat{\beta}_{xyz}$, of the preferred model for each type of worker.

Similarly, the productivity parameters of the matches, $\hat{\pi}_{xyz}$, are relatively big compared to the costs of the matches given by the pre-tax wages, w_{xyz}. Hence, the choices of the firms are to a larger extent driven by differences in productivity than differences in pre-tax wages. For instance, from Table 8 we see that the estimates of $\hat{\pi}_{xyz}$ on average are relatively large for *Other services*, which captures that this is the industry where the most workers are employed, see Table 2. Further, note that all estimates of $\hat{\pi}_{xyz}$ are negative. This is because the maximum quotas are very large relative to the number of matched workers in the different industries, $N \gg L_y$.

	1. Uskilled	2. Skilled	3. Medium	4. Highly
1. Management	-2.53	-2.01	-2.61	-3.36
2. High Level	-2.92	-2.31	-2.87	-3.60
3. Medium level	-2.76	-1.60	-3.16	-4.13
4. Basic Level	-2.78	-2.13	-3.22	-4.52
5. Other	-2.79	-2.40	-3.53	-4.78
6. Missing	-2.46	-1.61	-3.18	-4.27

(a) Manufacturing

	1. Uskilled	2. Skilled	3. Medium	4. Highly
1. Management	-3.45	-3.01	-2.79	-2.62
2. High Level	-3.90	-3.19	-3.07	-3.06
3. Medium level	-3.84	-2.97	-3.34	-3.49
4. Basic Level	-4.07	-3.27	-3.33	-3.75
5. Other	-3.81	-3.50	-3.69	-4.14
6. Missing	-3.69	-3.19	-3.03	-2.26

(b) FIRE

	1. Uskilled	2. Skilled	3. Medium	4. Highly
1. Management	-2.82	-2.58	-1.81	-2.88
2. High Level	-3.05	-3.07	-1.77	-2.82
3. Medium level	-2.84	-2.62	-2.29	-3.18
4. Basic Level	-2.25	-2.38	-1.90	-3.34
5. Other	-2.54	-2.65	-2.50	-4.03
6. Missing	-2.42	-1.85	-1.16	-1.60

(c) Public Services

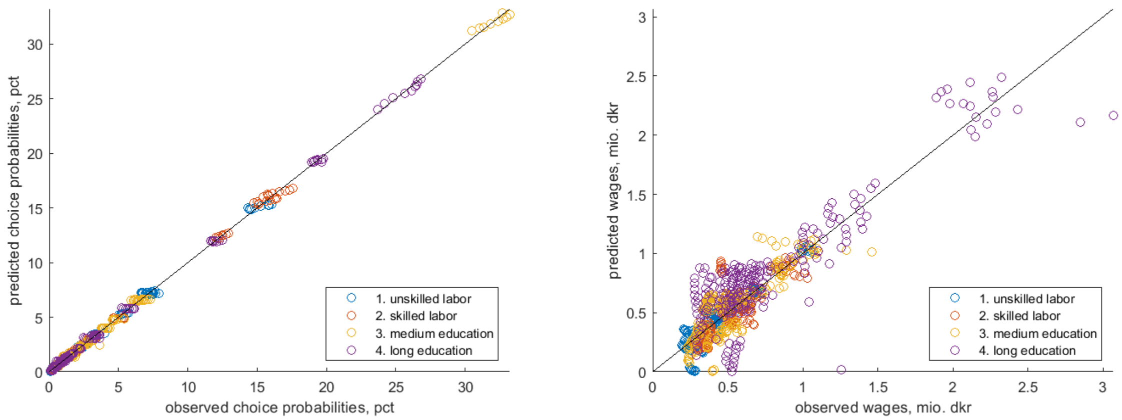
	1. Uskilled	2. Skilled	3. Medium	4. Highly
1. Management	-1.39	-1.63	-1.77	-2.18
2. High Level	-1.69	-1.83	-1.93	-2.21
3. Medium level	-1.65	-0.02	-2.34	-2.76
4. Basic Level	-1.73	-1.86	-2.61	-3.00
5. Other	-1.78	-2.09	-2.35	-3.11
6. Missing	-1.86	-1.60	-2.16	-2.63

(d) Other Services

Table 8: Estimated productivity of the matches, $\hat{\pi}_{xyz}$, of the preferred model for each type of industry.

6.5 Model fit

This sub-section illustrates the model fit of the *similarity model*. Figure 5a compares the observed choice probabilities of the workers to the model predictions. Since the choice probabilities are relatively constant across years and due to the very flexible specification of the payoff functions, it is not surprising that the model does a relatively good job in predicting the observed choice probabilities. The model does a worse job when predicting the observed wages, see figure 5b. The reason why the estimated model does a better job in explaining the choice probabilities compared to wages may be that we assume that the measurement errors are on the aggregate level. Recall that the full log-likelihood function is given by the sum of the log-likelihood function of the wages and the log-likelihood function of matches. Hence, there exists a trade-off between improving predictions of matches and improving the predictions of the wages. Since, we only have a single observation for the wage of any match, we weight this log-likelihood contribution of wages with just one. In contrast, we weight the log-likelihood contribution of a given match with the number of employed workers in that match. Hence, it is not surprising that the model does a worse job when predicting the observed wages than when it predicts observed choice probabilities.



(a) Choice probabilities of workers, pct.

(b) wages, DKK 1,000

Figure 5: Model fit of the preferred model.

7 Counterfactual studies

In this section we conduct two counterfactual studies based on the *similarity model*. Firstly, in order to illustrate the implied substitution patterns of the model we study a counterfactual increase in the productivity of highly educated workers in occupations that require high level skills in the manufacturing industry. Secondly, we study a counterfactual reduction of the bottom tax rate of 1 percentage point. For both studies we analyze the changes in equilibrium employment and wages across the different types of workers and firms. For both studies we analyze the changes in equilibrium employment and wages across the different types of workers and firms. Since we assume that utility is linear in consumption, we can measure utility in monetary units. Hence, we can compare changes in the welfare of the different types of workers and firms by the total welfare of the worker, WW_x, and total welfare of the firm, WF_y, defined as

$$WW_x = \tfrac{1}{\alpha} ws_x N_x, \tag{12}$$
$$WF_y = fs_y N, \tag{13}$$

where the the worker and firm surplus, (ws_x, fs_y), is given by Eq. (9). In order to study the total welfare changes we include a government sector that collects taxes and pays benefits to unemployed workers. For each type of worker, we measure their contribution to the government's budget, GB_x,

$$GB_x = \left(p_0^x \left(\tau \left(w_0 \right) - w_0 \right) + \sum_{y,z} p_{yz}^x \cdot \tau \left(w_{xyz} \right) \right) N_x. \tag{14}$$

The total welfare is given by the sum of the welfare of the workers and firms and the government budget

$$W = \sum_x \left(WW_x + GB_x \right) + \sum_y WF_y. \tag{15}$$

The change in welfare is measured as the difference between the welfare in the counterfactual scenario and the model consistent welfare in the baseline scenario. This definition of welfare is related to Dupuy *et al.* (2020), who study the deadweight loss due to matching distiortion.[11] However, as our empirical model includes unmatched workers who receive an exogenous benefit payment, we can identify the level of the wages and not just the wage differences. Further, the inclusion of unmatched firms and workers implies that the total employment level in each industry and for each type of work is determined in equilibrium. Hence, we can include changes along the extensive margin in our analysis. Changes along the extensive margin reflect changes in the number of workers and firms that match.

7.1 Increased productivity

In this counterfactual study we analyze the welfare consequences of an increase in the linear productivity parameter π_{421} of DKK $100,000$, ($\Delta\pi_{421} = 0.1$). More specifically, we consider an increase in the productivity for highly educated workers, ($x = 4$), employed in the manufacturing sector, ($y = 1$), in occupations that require high level of skills, ($z = 2$), and we will refer to this particular match as *match P*. The purpose of this counterfactual study is twofold. Firstly, this counterfactual study illustrates substitution and complementarity between the different workers and occupations of the preferred model. Secondly, it illustrates how skill-bias technology changes can have different effects on employment, wages and welfare for different types of workers and firms. The degree of substitutability and complementarity between different matches determines the welfare effects of the skill-bias technology change. For comparison we have repeated the welfare analysis of this counterfactual study for the *logit model* in section E in the appendix.

[11] As Dupuy *et al.* (2020) restrict $\alpha = 1$ and do not consider unmatched workers or firms, their welfare measure can be expressed as

$$\sum_{x,y,z} p_{yz}^x \left(\beta_{xyz} - \pi_{xyz} \right) N_x,$$

where wages and taxation are ignored since the total welfare is independent of how the workers, firms and goverment split the wages.

For fixed wages Table 9a shows the excess demand for labor across matches due to the increase in productivity. When keeping the wages fixed changes in productivity only affect the payoffs of the firms. Hence, excess demand merely reflects changes in firms' demand. Note that excess demand only differs from zero for firms in the manufacturing industry, $(y = 1)$, since only this industry experiences the increase in productivity. The increase in productivity for *match P* implies that excess demand becomes positive for this match. For the remaining matches in the manufacturing industry the sign of excess demand reflects whether the match is substitute or complement to *match P*. From the table we see as expected, that excess demand is largest for *match P* and that excess demand is negative for all the remaining matches where highly educated workers are employed, $(x = 4)$. This is partly due to the positive estimate of $\lambda_{\mathcal{Z}} > 0$, which implies that matches with similar worker types are substitutes. For the remaining worker types, we see that the excess demand is positive for matches that require high skill level, $(z = 2)$. This is due to the negative estimate of $\lambda_{\mathcal{X}} < 0$, which implies that similar occupations are complements in firms' production. The analysis of the sign of excess demand is complicated by the fact that the gradient of the firms' choice probabilities is a non-linear function of the choice probabilities. Hence, for most of the remaining matches excess demand becomes negative. This is mainly due to the fact that the absolute size of $\lambda_{\mathcal{Z}}$ is larger than the absolute size of $\lambda_{\mathcal{X}}$, which implies that matches with different occupations and different worker types tend to be substitutes. As illustrated in table 18a in appendix E, the excess demand of the estimated *logit model* is negative for all matches in the manufacturing industry except for the match that experiences the shock, which is an implication of the restrictive IIA property of the logit model.

Tables 9c and 9b show the changes in wages and employment across matches. The adjustment to equilibrium is a complicated process, since the model includes multiple types of workers, firms and occupations. In a simplified description of the equilibrium adjustment the increased productivity of *match P* increases the demand for this match by 1,200 workers and increases the wage DKK 66,500. Since the remaining matches of highly educated workers in the manufacturing industry are substitutes to *match P*, the demand for these matches are decreased which in turn leads to a decrease in wages by DKK 15-17,000. Due to the large wage increase of *match P*, firms in the remaining industries have to increase the wages for highly educated workers to compensate them. However, as wages for highly educated workers increase in these industries, the demand for highly educated works is reduced. Hence, the employment decreases by 0-200 and increases the wages by DKK 2-3,000. for these matches. Lastly, matches that require high skill level performed by the remaining types of workers are complements. Hence, the increased demand for these matches drives up the employment by 0-100 and wages by roughly DKK 3,500. The effects on wages and employment for the remaining matches are small.

	y=1	y=2	y=3	y=4
$z=1, x=1$	-0.21	0.00	0.00	0.00
$z=2, x=1$	0.36	0.00	0.00	0.00
$z=3, x=1$	-0.20	0.00	0.00	0.00
$z=4, x=1$	0.06	0.00	0.00	0.00
$z=5, x=1$	0.03	0.00	0.00	0.00
$z=6, x=1$	-0.23	0.00	0.00	0.00
$z=1, x=2$	-0.58	0.00	0.00	0.00
$z=2, x=2$	0.59	0.00	0.00	0.00
$z=3, x=2$	-0.57	0.00	0.00	0.00
$z=4, x=2$	0.29	0.00	0.00	0.00
$z=5, x=2$	0.05	0.00	0.00	0.00
$z=6, x=2$	-0.18	0.00	0.00	0.00
$z=1, x=3$	-0.69	0.00	0.00	0.00
$z=2, x=3$	1.94	0.00	0.00	0.00
$z=3, x=3$	-0.80	0.00	0.00	0.00
$z=4, x=3$	-0.17	0.00	0.00	0.00
$z=5, x=3$	-0.02	0.00	0.00	0.00
$z=6, x=3$	-0.12	0.00	0.00	0.00
$z=1, x=4$	-1.14	0.00	0.00	0.00
$z=2, x=4$	5.49	0.00	0.00	0.00
$z=3, x=4$	-0.85	0.00	0.00	0.00
$z=4, x=4$	-0.58	0.00	0.00	0.00
$z=5, x=4$	-0.09	0.00	0.00	0.00
$z=6, x=4$	-0.36	0.00	0.00	0.00

(a) Changes to excess demand, 1,000

	y=1	y=2	y=3	y=4
$z=1, x=1$	-0.00	-0.00	-0.00	-0.00
$z=2, x=1$	0.01	-0.00	-0.00	-0.00
$z=3, x=1$	-0.00	-0.00	0.00	0.00
$z=4, x=1$	-0.02	0.00	0.00	0.01
$z=5, x=1$	-0.01	0.00	0.00	0.00
$z=6, x=1$	-0.01	0.00	0.00	0.00
$z=1, x=2$	-0.01	-0.00	-0.00	-0.00
$z=2, x=2$	0.02	-0.00	-0.00	-0.00
$z=3, x=2$	-0.01	0.00	0.00	0.00
$z=4, x=2$	-0.04	0.00	0.01	0.01
$z=5, x=2$	-0.01	0.00	0.00	0.00
$z=6, x=2$	-0.00	0.00	0.00	0.00
$z=1, x=3$	-0.01	-0.00	-0.00	-0.00
$z=2, x=3$	0.06	-0.00	-0.01	-0.01
$z=3, x=3$	-0.02	0.00	0.00	0.00
$z=4, x=3$	-0.01	0.00	0.00	0.00
$z=5, x=3$	-0.00	0.00	0.00	0.00
$z=6, x=3$	-0.00	0.00	0.00	0.00
$z=1, x=4$	-0.09	-0.01	-0.01	-0.03
$z=2, x=4$	1.17	-0.11	-0.17	-0.17
$z=3, x=4$	-0.09	-0.02	-0.01	-0.02
$z=4, x=4$	-0.06	-0.01	-0.01	-0.02
$z=5, x=4$	-0.01	-0.00	-0.00	-0.01
$z=6, x=4$	-0.04	-0.00	-0.00	-0.01

(b) Changes to employment, 1,000

	y=1	y=2	y=3	y=4
$z=1, x=1$	-0.71	-0.13	-0.05	-0.06
$z=2, x=1$	3.59	-0.13	-0.03	-0.14
$z=3, x=1$	-0.40	-0.01	0.02	0.01
$z=4, x=1$	-0.13	0.02	0.03	0.03
$z=5, x=1$	-0.13	0.03	0.04	0.03
$z=6, x=1$	-0.25	0.00	0.01	0.02
$z=1, x=2$	-0.69	-0.15	-0.08	-0.08
$z=2, x=2$	3.55	-0.16	-0.05	-0.16
$z=3, x=2$	-0.45	-0.03	-0.01	-0.01
$z=4, x=2$	-0.15	0.00	0.00	0.01
$z=5, x=2$	-0.15	0.01	0.01	0.01
$z=6, x=2$	-0.29	-0.02	-0.02	0.00
$z=1, x=3$	-1.00	-0.08	-0.00	-0.00
$z=2, x=3$	3.53	-0.08	0.01	-0.09
$z=3, x=3$	-0.67	0.04	0.05	0.06
$z=4, x=3$	-0.46	0.08	0.09	0.09
$z=5, x=3$	-0.46	0.08	0.08	0.09
$z=6, x=3$	-0.56	0.06	0.06	0.08
$z=1, x=4$	-17.28	2.73	2.52	2.81
$z=2, x=4$	66.46	2.70	2.35	2.71
$z=3, x=4$	-16.18	2.81	2.54	2.77
$z=4, x=4$	-16.18	2.83	2.58	2.88
$z=5, x=4$	-15.48	2.66	2.42	2.84
$z=6, x=4$	-16.73	2.86	2.61	2.89

(c) Changes to wages, DKK 1,000

Table 9: Changes to excess demand, equilibrium employment and equilibrium wages due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level skill occupations.

Table 10 summarizes the welfare effects of the four different types of workers. As shown in the previous tables the increase in productivity mostly affects highly educated workers. For this type of worker, the employment increases by 260 and the average wage increases by DKK 6,000. In turn, the total welfare of the highly educated workers increased by DKK 809 million. Unskilled and skilled workers experience a small decrease in their welfare due to lower wages and employment, whereas medium educated workers experience a small increase in their welfare. The change in total welfare for all worker types is DKK 791 million.

	Employment	Avg. wages	Avg. WW	Total WW
1. Unskilled labor	-0.01	-0.01	-0.01	-6
2. Skilled labor	-0.01	-0.03	-0.03	-28
3. Medium education	0.01	0.07	0.02	17
4. Long education	0.26	6.15	2.20	809
All	0.25	0.83	0.24	791

Table 10: Changes in total welfare of workers due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level skill occupations. Changes in employment are measured in 1,000. Changes in average wages and average in welfare of the worker are measured in DKK 1,000, whereas changes in total welfare of the worker are measured in DKK 1,000,000.

Table 11 shows that mainly firms in the manufacturing industry are affected by the increased productivity. In this industry, employment increases by 800 and average wages increase by nearly DKK 3,000. Despite the higher wages, the increased productivity increases total welfare of the firms in the manufacturing industry by DKK 288 million. Due to the higher wages, employment in the remaining industries is reduced. Further, the total welfare is similar reduced in these industries. However, when summing over all industries, the total welfare of the firms is increased by DKK 168 million.

	Employment	Avg. wages	Avg. WF	Total WF
1 Manufacturing	0.83	2.93	0.24	288
2 FIRE	-0.15	0.60	-0.04	-42
3 Public Services	-0.19	0.29	-0.06	-49
4 Other Services	-0.23	0.24	-0.08	-29
All	0.25	0.83	0.01	168

Table 11: Changes in total welfare of firms due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level skill occupations. Changes in employment are measured in 1,000. Changes in average wages and average welfare of the firm are measured in DKK 1,000, whereas changes in total total welfare of the firm are measured in DKK 1,000,000.

The government's budget is improved by DKK 1,212 million, see Table 12. This is mostly driven by higher tax payments of highly educated workers which constitutes DKK 1,158 million. However, the expenditures to unemployment benefits is reduced by DKK 34 million. which is especially a result of the increased employment among highly educated workers.

	Avg. taxes	Avg. benefits	Revenue	Expenditure	Total GB
1. Unskilled labor	0.00	0.00	1	1	-0
2. Skilled labor	-0.01	0.00	-6	2	-7
3. Medium education	0.03	-0.00	25	-1	26
4. Long education	3.15	-0.10	1,158	-35	1,193
All	0.35	-0.01	1,178	-34	1,212

Table 12: Changes in welfare of the government sector due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level skill occupations. Changes in average tax payments of workers and average benefit payments to workers are measured in DKK 1,000, whereas changes to Government budget (GB), expenditure, and net revenue are measured in DKK 1,000,000.

Finally, the total welfare is increased across workers, firms and the government sector. Hence, the welfare has increased by DKK 2,171 million in total, see Table 13.

Welfare of the workers	791
Welfare of the firms	168
Government budget	1,212
Total Welfare	2,171

Table 13: Changes in total welfare due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level skill occupations. Changes in welfare are measured in DKK 1,000,000.

7.2 Reduction in bottom tax rate

In this sub-section we study the counterfactual reduction of the bottom tax rate by 1 percentage point. Before turning to the welfare analysis note that taxes distort the matching of workers and firms, as it makes it more costly for firms to compensate workers for matching with them. Hence, under high taxes workers may to a higher degree choose the matches they like rather than the most productive matches. Dupuy *et al.* (2020) refer to this as matching distortion, but since they do not include unmatched agents, they cannot analyze matching distortions along the extensive margin, which is the change in the number of matches.

As the unmatched workers have a relatively low income compared to matched workers, the reduction in the bottom tax rate leads to a relative decrease in the post-tax income of unmatched workers. As a result, the supply of workers increases, which drives post-tax wages down and the number of matched workers up for most matches. For some low-paid matches, the reduction in the tax rate reduces the post-wage income relative to the remaining matches. For these matches the supply of workers decreases which drives post-tax wages up and the number of matches down. However, Table 14 shows that the average wages are increased for all types of workers. This is because the dominating effect is that matches with higher wages experience a larger tax-reduction. Hence, the supply of workers will increase the most for these matches, which implies that the composition of matched workers to a higher degree consists of workers in high-paying matches. In total employment increases by 1,350 and the average wage increases by DKK 1,480. As a result the total welfare of the workers increases by DKK 6,135 million. For comparison, the Danish Ministry of Taxation assumes that a reduction of the bottom tax rate by 1 percentage point increases the employment by only 250 individuals. (see DMT (2019)).[12] Compared to the computational procedure of the Ministry of Taxation the employment effect is relatively large.

	Employment	Avg. wages	Avg. WW	Total WW
1. Unskilled labor	0.09	0.37	0.98	1,162
2. Skilled labor	0.53	3.65	1.96	2,002
3. Medium education	0.38	0.51	2.02	1,582
4. Long education	0.34	0.06	3.78	1,389
All	1.35	1.48	1.83	6,135

Table 14: Changes in welfare of the workers due to a reduction in bottom tax rate of 1 percentage point. Changes in employment are measured in 1,000. Changes in average wages and average welfare of the workers are measured in DKK 1,000, whereas total changes in welfare of the workers are measured in DKK 1,000,000.

For industries with a sufficiently high wage level at the starting point, the tax reduction leads to falling post-tax wages and increased employment, which increase the welfare of the firms. However, as wages are relatively low for the *Public service* industry, the tax reduction tends to increase the wages and reduce the employment, which leads to a decrease in the total welfare of DKK 95 million. in this industry, see Table 15. However, the total welfare of the firms across industries is increased by DKK 272 million.

[12]The Danish Ministry of Taxation assumes more specifically that a 0.08 pct. point reduction in the bottom tax rate increases the employment by 20 individuals. The increased employment of 250 individuals is computed from this relationship.

	Employment	Avg. wages	Avg. WF	Total WF
1 Manufacturing	0.21	0.85	0.06	73
2 FIRE	0.68	-1.88	0.18	188
3 Public Services	-0.38	1.31	-0.12	-95
4 Other Services	0.83	2.54	0.29	106
All	1.35	1.48	0.10	272

Table 15: Changes in welfare of the firms due to a reduction in bottom tax rate of 1 percentage point. Changes in average wages and average welfare of the firms are measured in DKK 1,000, whereas total changes in welfare of the firms are measured in DKK 1,000,000.

The contribution of each worker type to the government finances have decreased, see Table 16. This is because the additional tax revenue from the higher average wages and the saved expenditures due to lower unemployment cannot offset the reduced tax revenue from the reduction in the bottom tax rate. Hence, the government budget is reduced by DKK 4,853 million.

	Avg. taxes	Avg. benefits	Revenue	Expenditure	Total GB
1. Unskilled labor	-1.03	-0.01	-1,218	-13	-1,205
2. Skilled labor	-0.74	-0.07	-752	-71	-680
3. Medium education	-2.05	-0.07	-1,607	-52	-1,555
4. Long education	-3.97	-0.13	-1,458	-46	-1,412
All	-1.50	-0.05	-5,035	-182	-4,853

Table 16: Changes in welfare of the government sector due to a reduction in bottom tax rate of 1 percentage point. Changes in average tax of workers and average benefit payments to workers are measured in DKK 1,000, whereas changes to Government revenue, expenditure and net revenue are measured in DKK 1,000,000.

Even though the public finances have decreased, the total welfare has increased by DKK 1,554 million from this counterfactual tax reduction, due to the large increase in welfare of the workers, see Table 17.

Welfare of the worker	6,135
Welfare of the firm	272
Government budget	-4,853
Total Welfare	1,554

Table 17: Changes in total welfare due to a reduction in bottom tax rate of 1 percentage point. Changes in welfare are measured in DKK 1,000,000.

8 Conclusion

This paper proposes an empirical framework for multidimensional matching where the behavior of participants is described by perturbed utility models. We show that a unique pair of the equilibrium wages and matching distribution exists based on the convexity of indirect utility and profit functions. Our result does not rely on the gross substitutability condition.

For our empirical application, we specify the perturbation function in terms of the similarity function and propose a maximum likelihood estimator. As a proof of concept, we apply our estimator to aggregate Danish data and find that workers with different education levels can be complements. As our framework allows the total employment to be determined as an equilibrium outcome, our welfare analysis includes changes along the extensive margin of the participants.

We highlight three directions for future work. The first direction is to extend our framework in order to allow the attributes of the participants to be continuously distributed as suggested by Dupuy and Galichon (2014). This would make the modeling of the matching process more realistic, as workers and firms are likely to consider a variety of attributes when matching.

Our analysis is restricted to a frictionless labor market. However, many questions about matching between workers and firms have been raised for markets with matching frictions. Hence, as a second direction of future research, introducing matching friction into the model in order to extend the analysis of the determinants of unemployment would be another interesting issue. For instance, search costs could be introduced in an optimal sequential search setting similar to Moraga-González *et al.* (2021).

Finally, much of the empirical two-sided sorting literature uses the nested CES production function in order to analyze substitution in firms' demand for heterogeneous workers (see, e.g., Ottaviano and Peri (2012), Diamond (2016), and Piyapromdee (2021)). Following Anderson *et al.* (1988) and Verboven (1996) for the (nested) CES and (nested) logit, linking the similarity functions and the CES-type functions will broaden substitution patterns further by accommodating complementarities.

Our model is open to many other economic applications beyond the labor market. Complementarities are used interchangeably with other terms such as team chemistry, synergy, peer effect and are prevalent throughout firms, sports teams, and classrooms. Further, the generalization to multidimensional matching could be assigned to many-to-one or many-to-many matching markets, for example, players and sports teams or a network of firms.

References

ALLEN, R. and REHBECK, J. (2019a). Identification with additively separable heterogeneity. *Econometrica*, **87** (3), 1021–1054.

— and — (2019b). Revealed stochastic choice with attributes. *SSRN Electronic Journal*.

ANDERSON, S. P., DE PALMA, A. and THISSE, J.-F. (1988). The ces and the logit: Two related models of heterogeneity. *Regional Science and Urban Economics*, **18** (1), 155–164.

AZEVEDO, E. M. and LESHNO, J. D. (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy*, **124** (5), 1235–1268.

BECKER, G. S. (1973). A theory of marriage: Part i. *Journal of Political Economy*, **81** (4), 813–846.

BERRY, S., GANDHI, A. and HAILE, P. (2013). Connected substitutes and invertibility of demand. *Econometrica*, **81** (5), 2087–2111.

BOJILOV, R. and GALICHON, A. (2016). Matching in closed-form: equilibrium, identification, and comparative statics. *Economic theory*, **61** (4), 587–609.

BORJAS, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics*, **118** (4), 1335–1374.

CARD, D. (2005). Is the new immigration really so bad? *The Economic journal (London)*, **115** (507), 300–323.

CHIAPPORI, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy*, **100** (3), 437–467.

—, IYIGUN, M. and WEISS, Y. (2009). Investment in schooling and the marriage market. *American Economic Review*, **99** (5), 1689–1713.

CHOO, E. and SIOW, A. (2006). Who marries whom and why. *Journal of Political Economy*, **114** (1), 175–201.

DIAMOND, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, **106** (3), 479–524.

DMT (2019). Skatteoekonomisk redegoerelse 2019. *Danish Ministry of Taxation*.

DUPUY, A. and GALICHON, A. (2014). Personality traits and the marriage market. *The Journal of political economy*, **122** (6), 1271–1319.

— and — (2022). A note on the estimation of job amenities and labor productivity. *Quantitative Economics*, **13** (1), 153–177.

—, —, JAFFE, S. and KOMINERS, S. D. (2020). Taxation in matching markets. *International Economic Review*, **61** (4), 1591–1634.

EECKHOUT, J. and KIRCHER, P. (2018). Assortative matching with large firms. *Econometrica*, **86** (1), 85–132.

FOSGERAU, M., KRISTENSEN, D. and NIELSEN, N. (2021a). Estimation of perturbed utility models using demand inversion.

—, MELO, E., DE PALMA, A. and SHUM, M. (2020). Discrete choice and rational inattention: A general equivalence result. *International Economic Review*, **61** (4), 1569–1589.

—, MONARDO, J. and PALMA, A. D. (2021b). The inverse product differentiation logit model.

— and NIELSEN, N. (2021). Similarity, perturbed utility and discrete choice.

FOX, J. T. (2018). Estimating matching games with transfers. *Quantitative economics*, **9** (1), 1–38.

—, YANG, C. and HSU, D. H. (2018). Unobserved heterogeneity in matching games. *Journal of Political Economy*, **126** (4), 1339–1373.

FUDENBERG, D., IIJIMA, R. and STRZALECKI, T. (2015). Stochastic choice and revealed perturbed utility. *Econometrica*, **83**, 2371–2409.

GALICHON, A., KOMINERS, S. D. and WEBER, S. (2019). Costly concessions: An empirical framework for matching with imperfectly transferable utility. *Journal of Political Economy*, **127** (6), 2875–2925.

— and SALANIE, B. (2021). Cupid's invisible hand: Social surplus and identification in matching models. *The Review of Economic Studies*, **In press**.

HOFBAUER, J. and SANDHOLM, W. (2002). On the global convergence of stochastic fictitious play. *Econometrica*, **70** (6), 2265–2294.

— and Sandholm, W. H. (2007). Stable games. In *2007 46th IEEE Conference on Decision and Control*, pp. 3416–3421.

Hummels, D., Jørgensen, R., Munch, J. and Xiang, C. (2014). The wage effects of offshoring: Evidence from danish matched worker-firm data. *American Economic Review*, **104** (6), 1597–1629.

Kelso, A. S. and Crawford, V. P. (1982). Job matching, coalition formation, and gross substitutes. *Econometrica*, **50** (6), 1483–1504.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics, Academic Press New York*, pp. 105–142.

Moraga-González, J. L., Sándor, Z. and Wildenbeest, M. R. (2021). Consumer search and prices in the automobile market, CEPR Discussion Paper No. DP16574.

Ottaviano, G. I. P. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, **10** (1), 152–197.

Piyapromdee, S. (2021). The impact of immigration on wages, internal migration, and welfare. *The Review of Economic Studies*, **88** (1), 406–453.

Rheinboldt, W. C. (1970). On m-functions and their application to nonlinear gauss-seidel iterations and to network flows. *Journal of Mathematical Analysis and Applications*, **32** (2), 274–307.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.

— and Wets, R. J. B. (1998). *Variational Analysis*. Springer Berlin Heidelberg.

Traiberman, S. (2019). Occupations and import competition: Evidence from denmark. *American Economic Review*, **109** (12), 4260–4301.

Verboven, F. (1996). The nested logit model and representative consumer theory. *Economics Letters*, **50** (1), 57–63.

A Derivation of fixed point

This section derives the fixed point, characterizing the optimal choice probabilities of the workers. Since the derivation of the fixed point of the optimal choice probabilities of the firms is identical, this is left out. The first order condition of workers problem is

$$U_x\left(W_{x\cdot\cdot}\right) + H\left(p^x\right) - c \cdot 1_{|\mathcal{Y}||\mathcal{Z}|+1} = 0,$$

where c is some constant. Exponentiate and multiply by the the vector of choice probabilities to have

$$p^x \circ e^{U_x(W_{x\cdot\cdot})+H(p^x)} = p^x e^c,$$

and divide by the sum of both sides

$$\frac{p^x \circ e^{U_x(W_{x\cdot\cdot})+H(p^x)}}{\left\langle p^x, e^{U_x(W_{x\cdot\cdot})+H(p^x)}\right\rangle} = \frac{p^x e^c}{\left\langle p^x, 1_{|\mathcal{Y}||\mathcal{Z}|+1} e^c\right\rangle} = \frac{p^x e^c}{e^c} = p^x.$$

B Contraction mapping

As suggested by Fosgerau and Nielsen (2021) the following iterative algorithms are applied in order to obtain the optimal choice probabilities of workers and firms

$$p^{x,m+1} = \frac{p^{x,m} \circ e^{[U_x(W_{x\cdot\cdot})+H_x(p^{x,m})]/\xi}}{\left\langle p^{x,m}, e^{[U_x(W_{x\cdot\cdot})+H_x(p^{x,m})]/\xi}\right\rangle} \to p^x, \quad q^{y,m+1} = \frac{q^{y,m} \circ e^{[V_y(W_{\cdot y\cdot})+M_y(q^{y,m})]/\psi}}{\left\langle q^{y,m}, e^{[V_y(W_{\cdot y\cdot})+M_y(q^{y,m})]/\psi}\right\rangle} \to q^y, \tag{16}$$

where

$$\xi = \max_k \sum \Gamma^w |\eta_k|, \quad \psi = \max_l \sum \Gamma^f |\lambda_l|.$$

Fosgerau and Nielsen (2021) show that Eq. (16) are contraction mappings. Hence, for any initial guess of $p^{x,0} \in \triangle^{|\mathcal{Y}||\mathcal{Z}|+1}$ and $q^{y,0} \in \triangle^{|\mathcal{X}||\mathcal{Z}|+1}$ Eq. (16) will eventually converge to the true choice probabilities of workers and firms (p^x, q^y).

C Proofs of the results in the main text

Proof of Proposition 1. We first note that $\mu^* \in \text{Int}\left(\triangle^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|+|\mathcal{X}|}\right)$ under Assumption 2.

Assume (μ^*, W^*) is an aggregate equilibrium outcome. Then, a generalized Williams-Daly-Zachary theorem implies that

$$\mu^*_{xyz} = r_x \partial G^*_x \left(U_x\left(W^*_{x\cdot\cdot}, w_0\right)\right) / \partial u_{xyz\cdot} = \partial \bar{F}^*_y \left(W^*_{\cdot y\cdot}\right) / \partial w_{xyz}.$$

Conversely, it follows with Assumption 2 that $\mu^*_{\cdot|x}$ and $\mu^*_{\cdot y\cdot}$ are the strategies that maximize workers' and firms' problems with wages $W^*_{x\cdot\cdot}$ and $W^*_{\cdot y\cdot}$, respectively. □

Proof of Lemma 1. Since $-\bar{F}^*(W)$ is convex in W, Theorem 12.17 in Rockafellar and Wets (1998) implies that $-d\bar{F}^*(W)/dW$ is monotone and maximal monotone. It is suffice to show that G_x^* is also monotone in $W_{x\cdots}$ convex in $W_{x\cdot\cdot}$ for all x. We first note that G_x^* is convex in $U_x(W_{x\cdot\cdot}) = \alpha_x W_{x\cdot\cdot} + \beta_{x\cdot\cdot}$, where $\beta_{x\cdot\cdot} = (\beta_{xyz})_{(y,z)\in\mathcal{Y}\times\mathcal{Z}}$. Then $dG_x^*(U_x(W_{x\cdot\cdot}))/dU_x$ is monotone in U_x, and is monotone in $W_{x\cdot\cdot}$ by Exercise 12.4.(b) and (d) in Rockafellar and Wets (1998). It follows from Exercise 12.4.(c) in Rockafellar and Wets (1998) that $-D(W) = -d\bar{F}^*(W)/dW + dG^*(U(W,w_0))/dU$ is also monotone and maximal monotone in W. □

Proof of Theorem 1. We employ Theorem 12.51 in Rockafellar and Wets (1998) to show that $D^{-1}(0)$ is nonempty and bounded. We first note that $D(W)$ is continuous in W, which implies that $-D(W)$ is maximal monotone. The remaining condition for Theorem is the existence of $\tilde{W} \in \mathbb{R}^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|}$ with $\left\langle D\left(\tilde{W}\right), W\right\rangle < 0$ for each nonzero $W \in \mathbb{R}^{|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|}$. Let $\mathcal{D}_1 := \{(x,y,z) : w_{xyz} \leq 0\}$ and $\mathcal{D}_2 := \{(x,y,z) : w_{xyz} > 0\}$. We define a sequence of wages, $\left\{\tilde{W}^k = \left(\tilde{w}^k_{xyz}\right)_{(x,y,z)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}}\right\}_{k=1,2,\ldots}$, by

$$\tilde{w}^k_{xyz} = \begin{cases} k, & (x,y,z) \in \mathcal{D}_1, \\ -k, & (x,y,z) \in \mathcal{D}_2. \end{cases}$$

Then, as $k \to \infty$,

$$\begin{aligned} \partial\bar{F}^*\left(\tilde{W}^k\right)/\partial w_{xyz} &\to 0, & (x,y,z) \in \mathcal{D}_1, \\ \partial G^*\left(U\left(\tilde{W}^k, w_0\right)\right)/u_{xyz} &\to 0, & (x,y,z) \in \mathcal{D}_2. \end{aligned}$$

Thus, for k large enough, $\left\langle D\left(\tilde{W}\right), W\right\rangle < 0$.

For the uniqueness, it is suffice to show the strict monotonicity of G_x^* or $-\bar{F}_y^*$ because the strict monotonicity of $-D$ implies that the solution set can have no more than one element. We note from Assumption 2 that $\bar{F}(W)$ is essentially smooth. Then Corollary 26.4.1 in Rockafellar (1970) implies that $-\bar{F}^*(W)$ is strictly convex, in turn, it is strictly monotone. □

Proof of Corollary 1. The result follows directly from combining Proposition 1 and Theorem 1. □

D Identification

Before moving on the estimation, we discuss identification issues. The parameter to be identified is $\left(\alpha_\ell, \gamma_\ell, \beta_{\ell tk}, \theta_{\ell tk}, \bar{\eta}, \bar{\lambda}\right)$. The optimality conditions yield

$$
\begin{aligned}
\frac{\partial G_x\left(\mu^t_{\cdot|x}\right)}{\partial \mu_{yz|x}} &= (1-\eta_{\mathcal{Y}}-\eta_{\mathcal{Z}})\log \mu^t_{yz|x} + \eta_{\mathcal{Y}} \log\left(\sum_{y'} \mu^t_{yz|x}\right) + \eta_{\mathcal{Z}} \log\left(\sum_{z'} \mu^t_{yz'|x}\right) + 1 \\
&= \alpha_x\left(w^t_{xyz} - \tau_t\left(w^t_{xyz}\right)\right) + \beta_{xyz}, \\
\frac{\partial G_x\left(\mu^t_{\cdot|x}\right)}{\partial \mu_{0|x}} &= \log \mu^t_{0|x} + 1 = \alpha_x\left(w^t_0 - \tau\left(w^t_0\right)\right) + \beta_{x0}; \\
\frac{\partial \bar{F}_y\left(\mu^t_{\cdot y \cdot}\right)}{\partial \mu_{xyz}} &= (1-\lambda_{\mathcal{X}}-\lambda_{\mathcal{Z}})\log \mu^t_{xyz} + \lambda_{\mathcal{X}} \log\left(\sum_{x'} \mu^t_{x'yz}\right) + \lambda_{\mathcal{Z}} \log\left(\sum_{z'} \mu^t_{xyz'}\right) + 1 \\
&= \pi_{xyz} - w^t_{xyz} \\
\frac{\partial \bar{F}_y\left(\mu^t_{\cdot y \cdot}\right)}{\partial \mu_{y0}} &= \log\left(\mu^t_{y0}\right) + 1 = \pi_{y0}.
\end{aligned}
$$

α_x and π_{y0} are easily identified with $\partial G_x\left(\mu^t_{\cdot|x}\right)/\partial \mu_{0|x}$ and $\partial \bar{F}_y\left(\mu^t_{\cdot y \cdot}\right)/\partial \mu_{y0}$, respectively. Then $(\eta_{\mathcal{Y}}, \eta_{\mathcal{Z}})$ is also identified with $\left\{\partial G_x\left(\mu^1_{\cdot|x}\right)/\partial \mu_{yz|x} - \partial G_x\left(\mu^2_{\cdot|x}\right)/\partial \mu_{yz|x}\right\}_{x,y,z}$, and β_{xyz} is in turn identified by $\left\{G_x\left(\mu^t_{\cdot|x}\right)/\partial \mu_{yz|x}\right\}$. For the firm side, $(\lambda_{\mathcal{X}}, \lambda_{\mathcal{Z}})$ is identified by $\left\{\partial \bar{F}_y\left(\mu^t_{\cdot y \cdot}\right)/\partial \mu_{xyz}\right\}_z$, and π_{xyz} is in turn identified.

E Counterfactual study of increased productivity

	y=1	y=2	y=3	y=4
$z = 1, x = 1$	-0.00	0.00	0.00	0.00
$z = 2, x = 1$	-0.00	0.00	0.00	0.00
$z = 3, x = 1$	-0.00	0.00	0.00	0.00
$z = 4, x = 1$	-0.06	0.00	0.00	0.00
$z = 5, x = 1$	-0.02	0.00	0.00	0.00
$z = 6, x = 1$	-0.01	0.00	0.00	0.00
$z = 1, x = 2$	-0.01	0.00	0.00	0.00
$z = 2, x = 2$	-0.00	0.00	0.00	0.00
$z = 3, x = 2$	-0.01	0.00	0.00	0.00
$z = 4, x = 2$	-0.11	0.00	0.00	0.00
$z = 5, x = 2$	-0.01	0.00	0.00	0.00
$z = 6, x = 2$	-0.01	0.00	0.00	0.00
$z = 1, x = 3$	-0.01	0.00	0.00	0.00
$z = 2, x = 3$	-0.01	0.00	0.00	0.00
$z = 3, x = 3$	-0.01	0.00	0.00	0.00
$z = 4, x = 3$	-0.01	0.00	0.00	0.00
$z = 5, x = 3$	-0.00	0.00	0.00	0.00
$z = 6, x = 3$	-0.00	0.00	0.00	0.00
$z = 1, x = 4$	-0.00	0.00	0.00	0.00
$z = 2, x = 4$	2.21	0.00	0.00	0.00
$z = 3, x = 4$	-0.00	0.00	0.00	0.00
$z = 4, x = 4$	-0.00	0.00	0.00	0.00
$z = 5, x = 4$	-0.00	0.00	0.00	0.00
$z = 6, x = 4$	-0.00	0.00	0.00	0.00

(a) Changes to excess demand, 1,000

	y=1	y=2	y=3	y=4
$z = 1, x = 1$	-0.00	0.00	0.00	0.00
$z = 2, x = 1$	-0.00	0.00	0.00	0.00
$z = 3, x = 1$	-0.00	0.00	0.00	0.00
$z = 4, x = 1$	-0.02	0.00	0.00	0.01
$z = 5, x = 1$	-0.00	0.00	0.00	0.00
$z = 6, x = 1$	-0.00	0.00	0.00	0.00
$z = 1, x = 2$	-0.00	0.00	0.00	0.00
$z = 2, x = 2$	-0.00	0.00	0.00	0.00
$z = 3, x = 2$	-0.00	0.00	0.00	0.00
$z = 4, x = 2$	-0.03	0.00	0.01	0.01
$z = 5, x = 2$	-0.00	0.00	0.00	0.00
$z = 6, x = 2$	-0.00	0.00	0.00	0.00
$z = 1, x = 3$	-0.00	0.00	0.00	0.00
$z = 2, x = 3$	-0.00	0.00	0.01	0.00
$z = 3, x = 3$	-0.00	0.00	0.00	0.00
$z = 4, x = 3$	-0.00	0.00	0.00	0.00
$z = 5, x = 3$	-0.00	0.00	0.00	0.00
$z = 6, x = 3$	-0.00	0.00	0.00	0.00
$z = 1, x = 4$	-0.01	-0.02	-0.01	-0.03
$z = 2, x = 4$	0.94	-0.12	-0.14	-0.18
$z = 3, x = 4$	-0.01	-0.02	-0.01	-0.03
$z = 4, x = 4$	-0.01	-0.01	-0.01	-0.03
$z = 5, x = 4$	-0.00	-0.00	-0.00	-0.01
$z = 6, x = 4$	-0.01	-0.01	-0.00	-0.01

(b) Changes to employment, 1,000

	y=1	y=2	y=3	y=4
$z = 1, x = 1$	-0.16	0.03	0.03	0.06
$z = 2, x = 1$	-0.15	0.03	0.02	0.06
$z = 3, x = 1$	-0.13	0.03	0.03	0.05
$z = 4, x = 1$	-0.09	0.02	0.03	0.03
$z = 5, x = 1$	-0.13	0.03	0.03	0.03
$z = 6, x = 1$	-0.13	0.03	0.03	0.03
$z = 1, x = 2$	-0.17	0.01	0.01	0.04
$z = 2, x = 2$	-0.15	0.01	0.01	0.03
$z = 3, x = 2$	-0.17	0.01	0.01	0.04
$z = 4, x = 2$	-0.10	0.01	0.01	0.02
$z = 5, x = 2$	-0.15	0.01	0.01	0.02
$z = 6, x = 2$	-0.17	0.01	0.01	0.03
$z = 1, x = 3$	-0.14	0.04	0.04	0.07
$z = 2, x = 3$	-0.14	0.04	0.03	0.07
$z = 3, x = 3$	-0.12	0.04	0.04	0.06
$z = 4, x = 3$	-0.14	0.04	0.04	0.07
$z = 5, x = 3$	-0.13	0.03	0.04	0.06
$z = 6, x = 3$	-0.14	0.04	0.04	0.07
$z = 1, x = 4$	2.58	2.76	2.72	2.79
$z = 2, x = 4$	56.30	2.71	1.63	2.73
$z = 3, x = 4$	2.33	2.66	2.32	2.45
$z = 4, x = 4$	2.38	2.55	2.48	2.72
$z = 5, x = 4$	2.16	1.65	2.32	2.47
$z = 6, x = 4$	2.57	2.76	2.77	2.79

(c) Changes to wages, DKK 1,000

Table 18: Changes to excess demand, equilibrium employment and equilibrium wages due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level occupations.

	Employment	Avg. wages	Avg. WW	Total WW
1. unskilled labor	0.00	0.00	0.00	2
2. skilled labor	-0.01	-0.02	-0.01	-15
3. medium education	0.00	0.01	0.01	9
4. long education	0.27	6.34	2.36	868
All	0.27	0.84	0.26	864

Table 19: Changes in total welfare of workers due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level occupations. Changes in employment are measured in 1,000. Changes in average wages and average in welfare of the worker are measured in DKK 1,000, whereas changes in total welfare of the worker are measured in DKK 1,000,000.

	Employment	Avg. wages	Avg. WF	Total WF
1 Manufacturing	0.82	2.91	0.29	342
2 FIRE	-0.16	0.64	-0.05	-53
3 Public Services	-0.15	0.26	-0.06	-44
4 Other Services	-0.25	0.26	-0.10	-37
All	0.27	0.84	0.02	207

Table 20: Changes in total welfare of firms due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level occupations. Changes in employment are measured in 1,000. Changes in average wages and average welfare of the firm are measured in DKK 1,000, whereas changes in total total welfare of the firm are measured in DKK 1,000,000.

	Avg. taxes	Avg. benefits	Revenue	Expenditure	Total GB
1. unskilled labor	-0.00	-0.00	-1	-0	-0
2. skilled labor	-0.00	0.00	-3	1	-4
3. medium education	0.00	-0.00	2	-1	2
4. long education	3.01	-0.10	1,105	-36	1,141
All	0.33	-0.01	1,103	-36	1,139

Table 21: Changes in welfare of the government sector due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level occupations. Changes in average tax payments of workers and average benefit payments to workers are measured in DKK 1,000, whereas changes to Government budget (GB), expenditure and net revenue are measured in DKK 1,000,000.

Welfare of the workers	864
Welfare of the firm	207
Government budget	1,139
Total Welfare	2,211

Table 22: Changes in total welfare due to an increase in productivity for highly educated workers employed in the manufacturing sector performing high level occupations. Changes in welfare are measured in DKK 1,000,000.

This page has been left blank intentionally.

Chapter 2

A perturbed spatial equilibrium model

A perturbed spatial equilibrium model*

Esben Scriver Andersen†

February 1, 2022

Abstract

This paper proposes an empirical spatial equilibrium framework for the housing and labor market, where the spatial substitution patterns of housing prices and wages can be analyzed. Unlike the existing literature, the proposed model introduces adjustment costs in the housing size decision and describes the behaviour of the households by the perturbed utility model, in order to allow for a rich and meaningful parameterization of spatial substitution patterns. As a proof of concept, the model is estimated on high-quality Danish administrative data, and the importance of distinguishing between adjustments in the demand for housing, along the intensive and extensive margins is illustrated by a counterfactual analysis of an increase in the supply of housing in the capital municipality, *Copenhagen*.

Keywords: Spatial sorting, residential location, work location, equilibrium model, perturbed utility, similarity.

*This research is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 740369). The author thanks Mogens Fosgerau and Bertel Schjerning for comments.

†Department of Economics, University of Copenhagen, Denmark, esben.scriver.andersen@econ.ku.dk.

1 Introduction

This paper proposes an empirical spatial equilibrium model for the housing and labor market, where the spatial substitution patterns of housing prices and wages can be analyzed. Unlike the existing literature, the proposed model introduces adjustment costs in the housing size decision and describes the behavior of the households by the perturbed utility model, in order to allow for a rich and meaningful parameterization of the spatial substitution patterns. Substitution patterns are important for the determination of equilibrium housing prices, since housing prices are not only affected by the quality of the location, but also by the availability of close substitutes. As a simple illustrative example, consider the case where the availability of locations with a view over the ocean is scarce, whereas the availability of locations with a view over grass fields is abundant. If these two attributes are close substitutes, then housing prices for locations with an ocean view will be relatively low. By contrast, if these two attributes are not close substitutes, the housing prices for locations with an ocean view will be relatively high.

Most of the existing empirical literature analyzes the residential choice through the additive random utility model of McFadden (1973, 1978), where choices are deterministic, as the decision maker - after observing the random taste-shock - chooses the residence with the highest payoff.[1] Conversely, choices in the perturbed utility model are stochastic, as the decision maker chooses a random alternative based on an optimal vector of choice probabilities that maximizes the decision maker's perturbed utility, which is given by the sum of the expected payoffs and the perturbation function. The latter term can be interpreted as representing taste-for-variety, and the specification of this perturbation function has important implications for the substitution patterns in the model. The perturbed utility model allows for more flexible substitution patterns than the additive random utility model, as it can allow different alternatives to be complements, which the random utility model rules out. Allen and Rehbeck (2019); Fosgerau et al. (2021); Fosgerau and Nielsen (2021) all showed that some perturbed utility models can allow for complementarity.[2]

Inspired by Buchinsky et al. (2014) and Carstensen et al. (2020), this paper proposes an empirical framework, where a population of heterogeneous households choose where to work, where to reside and the size of their houses. The households can - when deciding on their residential location and the work locations of their household members - choose between a set of locations that differ with respect to their spatial location. Households and household

[1] Bayer et al. (2004) and Bayer et al. (2007) both use this framework to study residential sorting.

[2] This paper is one of the first empirical applications of the perturbed utility model. Three notable examples of empirical applications are Fosgerau et al. (2021, 2022); Andersen and Lee (2022), who study consumption of cereal, route choices and matching between workers and firms.

members move and commute costly among locations.[3] The demand for housing and supply of labor is determined by the choices of the households. By contrast, the supply of housing and the demand for labor is assumed to be inelastic in the short run. A competitive housing and labor market determines the square meter prices and wages of each location.

This paper contribute to the existing literature on equilibrium models with moving costs by allowing both wages and square meter prices to be determined in equilibrium. Further, the existing literature assumes that households can change their housing size at no extra cost. By introducing adjustments costs to the housing size decision the proposed model discards this assumption, and illustrates that this has important implications for the spatial price substitution patterns, which is the focus of this paper. In order to allow for a rich parameterization of the substitution patterns, the perturbation function is specified in terms of the similarity function, as proposed by Fosgerau and Nielsen (2021). As the degree of substitution between different locations is important for the determination of equilibrium housing prices, a rich parameterization of the perturbation function allow substitution patterns to be determined by the data instead of being determined by the chosen functional form of the perturbation function. Further, the similarity function gives a meaningful interpretation of the substitution patterns in the model, as locations with similar characteristics are closer substitutes. For instance, such characteristics could be the type of view each residential location provides. In this case, the similarity function can accommodate that locations with a similar view are closer substitutes than locations with a different view.

This paper contributes to the existing literature on equilibrium models for housing by illustrating the importance of distinguishing between housing demand adjustments along the extensive and intensive margins when analyzing spatial price substitution patterns. Changes along the extensive margin reflect households that move from one location to another, and thereby reduce the demand for housing in the location the household moved away from, whereas the demand for housing is increased in the location the household moved to. By contrast, changes along the intensive margin reflect households that move within their current location, and thereby change the demand for housing within that location. Hence, changes along the intensive margin only affect the demand in one particular location and do not affect the demand for housing in any of the remaining locations. For instance, if the supply of housing increases in one location, this leads to a decrease in the square meter prices of that particular location, as demand and supply must equate. If the increase in demand for housing in this location only works through the intensive margin, the demand for housing

[3]Empirical papers, that considers sorting in both housing and labor market, include Kuminoff et al. (2012); Ahlfeldt et al. (2015); Monte et al. (2018). However, these papers assume frictionless markets. Bayer et al. (2016) and Kennan and Walker (2011) include switching costs, but only consider sorting in either housing or labor market.

in the remaining location will be unaffected, and as a result the square meter prices of the remaining locations will be unaffected as well. However, if part of the demand adjustment happens along the extensive margin, square meter prices of the remaining locations will be affected, as the demand in these locations have to adjust as well.

As a proof of concept, the proposed spatial equilibrium model is estimated on high-quality administrative register data for the entire population of Danish households from 2010 to 2013, where the choice set in the residential and work location problems consist of 92 locations given by a subset of the municipalities of Denmark. The equilibrium mechanisms of the estimated spatial equilibrium model are illustrated by analyzing a counterfactual increase of 1 percent in the supply of housing in the capital municipality, *Copenhagen*, that constitutes the most populous municipality. This counterfactual analysis shows that the square meter prices in *Copenhagen* decrease by 10.3 percent in the short run, and the average square meter prices decrease by 4.4 percent in the remaining municipalities of the Capital Region of Denmark that constitute the municipalities closest to *Copenhagen*. Conversely, if the counterfactual analysis is based on an estimated model where households can adjust their housing size at no cost and substitution patterns of the residential choice are determined by the independence of irrelevant alternatives property, the counterfactual increase in the housing supply of *Copenhagen* only leads to a 0.5 percent decrease in the square meter prices of these locations. This is because the demand adjustment only works through the intensive margin in this model. As a result, the square meter prices in the remaining locations are unaffected of the increased supply of housing.

Finally, the counterfactual analysis shows that average wages in *Copenhagen* in the short run are reduced for females and males, by 0.6 and 0.4 percent, respectively. This is because, as more households move to *Copenhagen*, the supply of labor increases, which in turn implies that wages must decrease in order to equate demand and supply of labor. The effect is smallest for males, as they, in line with existing literature, are more willing to commute in order to receive a higher wage than females. Locations sufficiently far from *Copenhagen* experience an increase in the average wages, as the supply of labor is reduced in these locations.

Organization of the paper. Section 2 introduces the perturbed utility model used to describe the behavior of the households and firms. This is followed by Section 3 that introduces the data used for estimation, and Section 4 sets up the proposed spatial equilibrium model. Section 5 describes the applied estimation procedure and Section 6 presents the estimation results. Section 7 conducts a counterfactual analysis of an increase in the housing supply of *Copenhagen*, and Section 8 concludes.

Some notation. Vectors are denoted $\bar{q}$. The probability simplex in the $J-1$ dimensional simplex is denoted $\Delta^J = \{\bar{q} \in \mathbf{R}_+^J : \sum_{j=1}^J q_j = 1\}$. The integers between A and B, where the end points are included, is denoted $\{A \,..\, B\} = \{A, A+1, ..., B\}$.

2 The perturbed utility model

This section introduces the perturbed utility model used to describe the joint residential and work location decision of the households.

Definition 1 ***(Perturbed utility model)*** *A decision maker (DM) faces a discrete choice problem with J different alternatives. The discrete choice, $d \in \{1, ..., J\}$, is stochastic with choice probabilities, $\bar{q}^* = (q_1^*, q_2^*, ..., q_J^*)$, given as the solution to the utility maximization problem*

$$\max_{\bar{q} \in \Delta^J} \left\{ \sum_{j=1}^{J} q_j v_j - Y(\bar{q}) \right\}, \tag{1}$$

where $v = (v_1, v_2, ..., v_J)$ is a vector of choice-specific payoffs and $Y : \mathbb{R}_{++}^J \to \mathbb{R}$ is the perturbation function.

If the perturbation function, Y, is strictly convex and satisfies that $\nabla_{\bar{q}} Y \to \infty$ as $q_j \to 0$ for any j, then there exists a unique optimal solution in the interior of the probability simplex. In this case, the perturbation function can be interpreted as representing taste-for-variety as it penalizes small choice probabilities.

Example 1 ***(Shannon entropy)*** *When the perturbation function is the negative Shannon entropy*

$$Y(\bar{q}) = \sum_{j=1}^{J} q_j \log q_j,$$

it can be shown that the solution to the DM's utility maximization problem is the multinomial logit choice probabilities

$$q_k^* = \frac{\exp(v_k)}{\sum_{j=1}^J \exp(v_j)}, \text{ for } k = 1, 2, ..., J.$$

For the logit choice probabilities, the substitution patterns are governed by the independence of irrelevant alternative property, and the corresponding expected utility of the DM is given by the well-known log-sum

$$EV = \log\left\{ \sum_{j=1}^{J} \exp\Big(v_j\Big) \right\}.$$

Fosgerau and Nielsen (2021) introduce a class of perturbation functions that provide a way to parameterize the degree of substitutability among groups of alternatives. Definition 2 is based on their work.

Definition 2 ***(Similarity function)*** *Let $\mathcal{C} = \{1, ..., C\}$ be a finite set of characteristics and let $\Psi = \{\psi\}_{c \in C, j \in J}$ be a $C \times J$ matrix with nonnegative entries ψ_{cj} and columns that sum to 1, $\sum_{c=1}^{C} \psi_{cj} = 1$. A similarity function $G : \mathbb{R}^J_{++} \to \mathbb{R}$ takes the form*

$$G(\bar{q}) = \sum_{j=1}^{J} q_j \cdot g_j(\bar{q}),$$
$$g_j(\bar{q}) = \left(1 - \sum_{c=1}^{C} \eta_c \psi_{cj}\right) \log q_j + \sum_{c=1}^{C} \eta_c \psi_{cj} \log\left(\sum_{k=1}^{K} \psi_{ck} q_k\right),$$

where $\bar{\eta} = (\eta_1, \eta_2, ..., \eta_C) \in \mathbb{R}^C$ is a vector of parameters associated with the C characteristics and satisfies $\sum_{c=1}^{C} \max(\eta_c, 0) \cdot \psi_{cj} < 1$ for all $j \in \{1 .. J\}$.

Fosgerau and Nielsen (2021) show that for $Y(\bar{q}) = G(\bar{q})$ the choice probabilities that maximize the DM's utility are determined by the unique fixed point equation

$$q_k^* = \frac{\exp\left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right)\right)}{\sum_{j=1}^{J} \exp\left(v_j + \log q_j^* - g_j\left(\bar{q}^*\right)\right)}, \text{ for } k = 1, 2, ..., J. \tag{2}$$

As shown in Appendix A the corresponding expected utility can be shown to be given as

$$EV = \log\left\{\sum_{j=1}^{J} \exp\left(v_j + \log q_j^* - g_j\left(\bar{q}^*\right)\right)\right\}. \tag{3}$$

Note that the similarity function reduces to the negative Shannon entropy if $\eta_c = 0$ for all $c \in \{1 .. C\}$, which implies that the optimal strategy is given by the multinomial logit choice probabilities of the logit model. Further, the class of similarity models nests the nested logit and inverse product differentiation logit model of Fosgerau et al. (2021) as special cases.

For each alternative, the parameter ψ_{cj} describes the degree to which alternative j is associated with characteristic c. Options that lie close together in the space of characteristics are considered to be more similar than alternatives far from each other. Higher positive values of η_c result in stronger substitution among similar alternatives relative to the logit model, while negative values result in weaker substitution, and even complementarity when η_c is sufficiently large negative.

3 Data

This section outlines the data sources and the construction of the data set used for estimation. The section ends by showing some descriptive statistics for the housing and labor market of Denmark.

The geographical units of this paper - the locations where households can choose to reside and work - are given by the municipalities of Denmark. Denmark consists of 98 municipalities that belong to five different regions (see Figure 1). As *Frederiksberg* is surrounded by *Copenhagen*, these two municipalities are treated as a single location throughout this paper, and this location will be referred to as *Copenhagen*. Further, the municipalities *Bornholm*, *Ærø*, *Fanø*, *Samsø* and *Furesø* are ignored, as these municipalities are all relatively small islands and not well-connected to the remaining municipalities. Hence, in the proposed model, households can only choose to reside or work in 92 different locations.

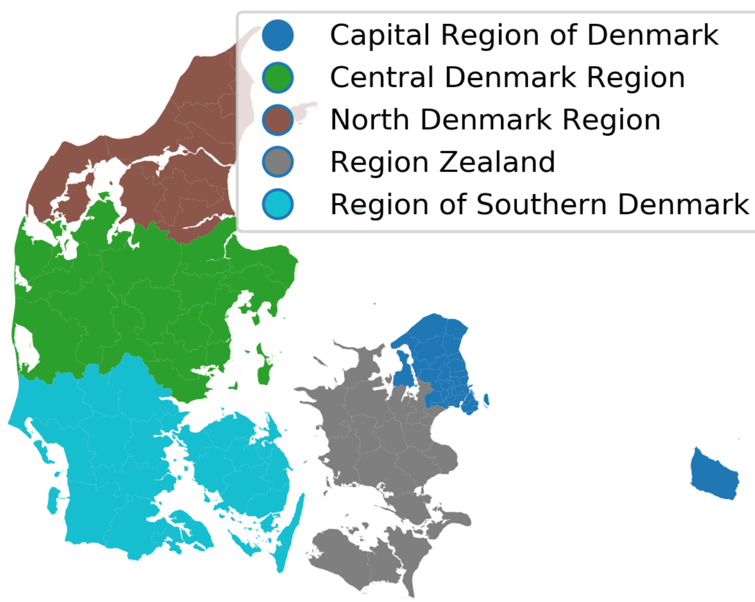


Figure 1: Regions and municipalities of Denmark.

The data are drawn from several administrative registers that cover the entire population of Denmark. Since all individuals are associated with a unique identifier, each individual can be followed over time, and information from the different registers can be linked. The main data source is the population register, which each year records all individuals who are officially registered with an address in Denmark on January 1st. This register contains information on demographics, address of residency, and a household identifier such that all household members can be linked. Statistics Denmark defines households as individuals who reside together at the same address. Households can consist of singles or couples with and without children.

Based on the household's address, it is possible to link information on the dwelling. The key variable *housing size* can be found in the dwelling register for each household. Another key variable is *square meter prices*, which is based on the register for sales of real property that contains information on sales prices for every property sold in a given year.

The income register contains detailed income history on, for example, pension benefits, transfer payments, and wages. Further, this register also classifies each individual by their attachment to the labor market in a given year based on their income history. Information on workplace is recorded on a monthly basis, and as a default the workplace in January is linked to each individual. Finally, the attainment register gathers information about the highest level of completed education for each individual.

The constructed data set contains all households where both members are 25 years of age or older with an address in one of the 92 considered municipalities in a given year, and where at least one of the household members had an address in one of the 92 municipalities the year before. The data set contains 10 million households or 15 million individuals. Table 1 shows some descriptive statistics for the data set.

Individuals who do not work and who, based on their labor market attachment, are defined as retired in the income register are in this paper defined as *outside the labor force*.[4] Hence, this group includes non-working individuals who receive public pension, early retirement pay, and early retirement pension. This groups constitutes 31 percent of the population. Individuals who are not classified as *outside the labor force* are classified into four mutually exclusive groups based on their highest attained education level. This classification is based on the International Standard Classification of Education (ISCED) two-digit code. *Unskilled* household members include individuals with high school completion or less (ISCED code 10-20 or missing); *Skilled* household members include individuals with vocational education (ISCED code 30); *Medium educated* household members include individuals with short- or medium-cycle higher education and bachelors (ISCED code 35-60); *Highly educated* household members include individuals with a masters or PhD degree (ISCED code 70-80). *Highly educated* household members is the smallest group. This group constitutes 7 percent

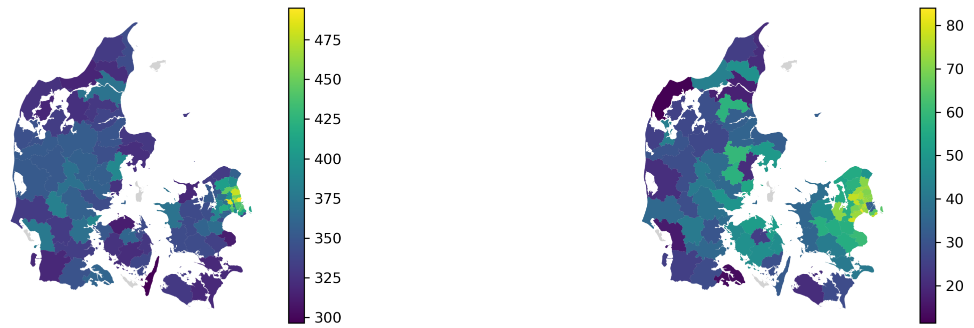
[4]The definition of individuals *outside the labor force* is based on the variable *BESKST02* that describes each individual's labor market attachment.

of the population and *Medium educated* individuals constitutes around 17 percent of the population. *Skilled* and *Unskilled* individuals constitute roughly 26 and 20 percent of the population, respectively. Figures 10 and 2a in Appendix B show that the municipalities are heterogeneous with respect to their composition of the four educational groups and the average wages differ across educational groups and municipalities. Hence, the proposed model describes the population of households by these four educational groups.

	Data set	
Households	10,036,748	
Individuals	15,007,505	
	(1) mean	(2) std.
Wage (DKK 1,000)	235.004	358.437
1 if working	0.615	0.487
1 if *commuting**	0.252	0.434
1 if unskilled**	0.193	0.394
1 if skilled**	0.260	0.439
1 if medium educ.**	0.169	0.375
1 if highly educ.**	0.071	0.256
1 if *outside the labor force*	0.308	0.462
1 if couple	0.495	0.500
1 if moved address	0.100	0.299
1 if moved municipality	0.029	0.168
Sales	130,467	
Housing size (m^2)	124.588	44.854
Sales prices (DKK 1,000)	2,030.264	1,483.385
Sales prices (DKK 1,000/m^2)	16.808	14.777

Table 1: Descriptive statistics. The samples include individuals of age 25 older from 2010 to 2013. Individuals who reside or work in *Bornholm*, *Ærø*, *Fanø*, *Samsø* or *Furesø* are excluded. All monetary units are measured in 2018 prices. **Commuting* refer to individuals who work in another municipality than they reside in. **Individuals *outside the labor force* are not included.

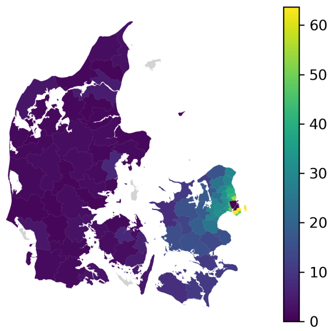
Table 1 shows that 61 percent of the population are working and 25 percent of the employed individuals commute to another municipality than the one they reside in. Given that a relatively large fraction of the employed individuals commutes, this favors that the residential and work location choices should be modelled jointly. Figure 2 illustrates the relation between average wages and share of commuters. The figures show that the average wages are relatively high in the four biggest cities (*Copenhagen*,*Odense*,*Aarhus*, and *Aalborg*) and the municipalities north of *Copenhagen* and that the share of commuters is relatively high in the surrounding municipalities. This shows that individuals residing outside high-paying areas often choose to commute to these areas instead of moving to these areas.



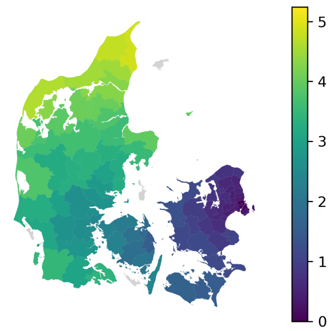
(a) Average wages **(b)** Share of commuters

Figure 2: Average wages (DKK 1,000) and share of employed individuals who commute to another municipality (percent). Average wages are based on the municipality the individual work. *Commuters* are defined as individuals who work in a different municipality than the reside in.

Commuting time is another key variable in the presented model, as household members are expected to receive compensation for longer commuting time in both the housing and the labor market (see Rosen (1986) and Zax (1991)). This paper relies on data from *Google Maps* to measure commuting time. The commuting time between any pair of municipalities is calculated as the travel time by car between the centroids of the municipalities. Figure 3 illustrates the share of commuters that commute to Copenhagen and the commuting times to *Copenhagen*. Commuters are defined as individuals who work in another municipality than the one they reside in. The figures show, as expected, that the shares of commuters who commute to *Copenhagen* are negatively correlated with commuting time to *Copenhagen*. Similar figures are shown in Appendix C for *Odense*, *Aarhus*, and *Aalborg*, which all show similar patterns. Figure 3 shows that the share of commuters to *Copenhagen* reduces drastically with commuting time, and very few individuals commute to *Copenhagen* from municipalities with commuting time longer than one hour.



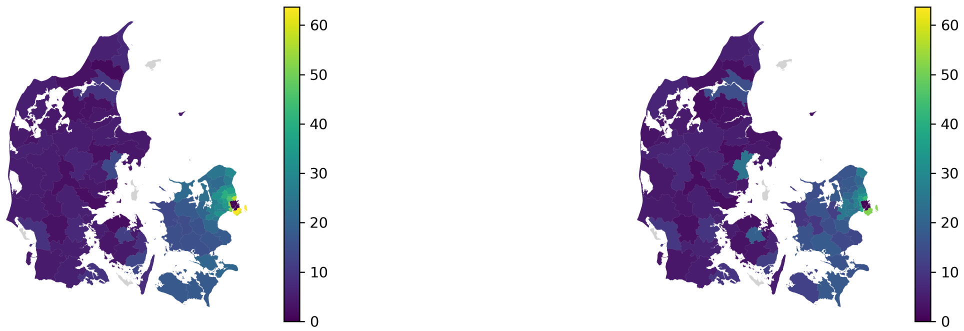
(a) Share of *commuters* that commute to Copenhagen



(b) Commuting time to Copenhagen

Figure 3: Share of *commuters* that commute to Copenhagen (percent) and travel time to Copenhagen (hours), 2010-2013. *Commuters* are defined as individuals who work in a different municipality than they reside in. By definition the share of *commuters* that commutes from *Copenhagen* to *Copenhagen* is zero. Commuting time is based on data from *Google Maps*.

Note from Table 1 that only 10 percent of the households moved address and only 3 percent moved to another municipality. As a small fraction of the households move, any realistic modelling of the residential location choice needs to incorporate frictions, such as moving costs. Further, any model should also be able to capture that households tend to move inside the municipality in which they currently reside, and if they move municipality, they tend to move to locations close by. The latter is illustrated for *Copenhagen* in Figure 4. Figure 4b shows for all municipalities the probability of moving to *Copenhagen* conditioned on moving municipality. The figure shows that the probability of moving to *Copenhagen* is higher for municipalities close to *Copenhagen*. However, the figure also shows that this probability is relatively big for the three second largest municipalities (*Odense*, *Aarhus*, and *Aalborg*).



(a) Share of *movers* that moves from Copenhagen **(b)** Share of *movers* that move to Copenhagen

Figure 4: Share of *movers* that move from Copenhagen (percent) and share of *movers* that move to Copenhagen (percent), 2010-2013. *Movers* are defined as households that move to another municipality.

Figure 5 illustrates the average square meter prices and housing size across municipalities. The figure does not indicate a clear correlation between square meter prices and housing size. However, it should be noted that the average housing size is relatively low for the four biggest municipalities (*Copenhagen*, *Odense*, *Aarhus*, and *Aalborg*), where the square meter prices are relatively high.

(a) Average housing size **(b)** Average square meters

Figure 5: Average housing size (m^2) and average square meter prices of sold properties (DKK 1,000/m^2).

4 Model

Inspired by Carstensen et al. (2020), this section sets up a spatial equilibrium model for the housing and labor market based on the joint residential and work location decision of a population of heterogeneous households. In the presented model, the economy consists of H_t households at time t and N different geographical locations that differ with respect to their spatial position. The households consist of single females, single males, and couples. It is assumed that households are not forward-looking and have static expectations, which implies that a static discrete choice model can be used to describe the behavior of the households.[5] Further, when faced with a discrete choice, it is assumed that the behavior of households can be described by the perturbed utility model from the previous section, where the perturbation function is specified in terms of the similarity function.

At each point in time, the household chooses its residential location, its housing size and the work location of each of its household members based on its residential and housing size choices in last period. Household members not participating in the labor force receive a fixed income and do not face a work location choice. Finally, it is assumed that in each time period the household makes its decisions sequentially in three stages.

In the first stage, household i chooses its residential location, $d^R_{it} \in \{1 \,..\, N\}$. If the household chooses to move to another location than its previous location, $d^R_{it} \neq d^R_{it-1}$, it pays a moving cost. Secondly, the household chooses its housing size $h_{it} \in \mathbf{R}$, where it is assumed that for any given square meter price there exists an optimal housing size. If the household in the previous stage chose to incur the moving costs and moves to another location, $d^R_{it} \neq d^R_{it-1}$, it will with certainty choose the optimal housing size by assumption. In contrast, if the household chose to stay, $d^R_{it} = d^R_{it-1}$, it faces the binary choice, $d^B_{it} \in (0, 1)$, of keeping its previous housing size, $d^B_{it} = 0$, or paying the moving costs and updating to the optimal housing size, $d^B_{it} = 1$. In the final third stage, each household member chooses its work location, $d^g_{it} \in \{0 \,..\, N\}$, based on the household's chosen residential location choice, where the household member takes into account the offered wages and commuting time. Let superscript $g = \{F, M\}$ denote whether the household member is female (F) or male (M). For $d^g_{it} = 0$ the household member chooses the outside option of not working and for $d^g_{it} \in \{1 \,..\, N\}$ it chooses to work in one of the N locations.

Since the household decides sequentially, the household - when deciding on its residential location - does not know the outcomes of the work location and housing size decisions. Hence, the model is solved by backward induction, such that each household member first solves the work location problem and the binary housing size problem - for any given residential choice - and then finally solves the residential location choice problem, where it bases its decision on the expected utility of the work location and housing size decision.

The supply of housing is fixed across locations, whereas the demand for workers in each

[5] This is in contrast to Carstensen et al. (2020) who describe the behaviour of the households by a dynamic discrete choice model.

location is described by a representative firm that uses the different types of household members as inputs in its production. The equilibrium can is determined by a set of market clearing conditions such that the demand for housing and labor equals the supply. The market clearing conditions are defined at the end of this section after the choice problems of the households and firms are presented.

4.1 Work location

Let $c_i \in \{S, C\}$ denote the civil status of the household. For $c_i = S$ the household consists of a single female or single male household member and for $c_i = C$ the household is a couple that consists of household members of each gender. Further, let $f_i \in \{0 .. E\}$ denote the educational type of the female household member and let $m_i \in \{0 .. E\}$ denote the educational type of the male household member. $f_i = m_i = 0$ refers to a female or male household members outside the labor force, who does not face a work location decision.

Let subscript s denote the residential location of the household. Then household member g chooses its work location based on the vector of choice probabilities, $\bar{q}_{ist}^{g*} = \left(q_{is0t}^{g*}, ..., q_{isNt}^{g*}\right)^\top \in \Delta^{N+1}$, that maximizes its perturbed utility

$$\bar{q}_{ist}^{g*} = \underset{\bar{q}_{ist}^{g} \in \Delta^{N+1}}{\arg\max} \left\{ \sum_{j=0}^{N} q_{isjt}^{g} \cdot u_{isjt}^{g} - Y^g\left(\bar{q}_{ist}^{g}\right) \right\}, \tag{4}$$

where $Y^g : \mathbb{R}_{++}^{N+1} \to \mathbb{R}$ is the perturbation function of household member g and u_{isjt}^{g} is the choice-specific utility of not working or working in location $j = 1, ..., N$. For the female household member the choice-specific utilities are given as

$$\begin{aligned} u_{isjt}^{F} &= \alpha_{c_i} \cdot w_{f_ijt}^{F}, && \text{for } j = 0, \\ u_{isjt}^{F} &= \alpha_{c_i} \cdot w_{f_ijt}^{F} + \lambda^F \cdot dist(s,j) + \mu_j^F, && \text{for } j = 1, ..., N, \end{aligned} \tag{5}$$

where α_{c_i} is the utility of money for household i, $w_{f_i0t}^{F}$ is some exogenous benefit payment the female household member receives when unemployed, and w_{f_ijt} for $j = 1, ..., N$ is the offered wage at location j for the female household member of type f_i. Further, $dist(s, j)$ refers to the travel time between location s and j, and λ^F is the marginal cost of commuting for female household members. The parameter μ_j^F captures the location-specific utility of working. The choice-specific utilities of the male household member are given by similar expressions

$$\begin{aligned} u_{isjt}^{M} &= \alpha_{c_i} \cdot w_{m_ijt}^{M}, && \text{for } j = 0, \\ u_{isjt}^{M} &= \alpha_{c_i} \cdot w_{m_ijt}^{M} + \lambda^M \cdot dist(s,j) + \mu_j^M, && \text{for } j = 1, ..., N. \end{aligned} \tag{6}$$

To sum up, for a household member g in the labor force with the residential location s, the optimal choice probabilities, $\bar{q}^{g*}_{ist}$ and the corresponding expected utility, EV^g_{ist} of the work location decision can be found from Eq. (2) and Eq. (3), for a given vector of choice-specific utilities, $(u^g_{is0t}, ..., u^g_{isNt})^\top$, and for a given specification of the perturbation function, Y^g. Household members outside the labor force, $f_i = m_i = 0$, choose with certainty the outside option of not working by assumption, $\bar{q}^{g*}_{ist} = (1, 0, ..., 0)^\top$. Hence, their expected utility of any given residential choice is $EV^g_{ist} = \alpha_{c_i} \cdot w^g_{0t}$. Where w^g_{0t} is a fixed income that household members outside the labor force receive.

4.2 Housing size

Recall, that subscript s denotes the the household's chosen residential location, and that any household that moves to another residential location will with certainty choose the optimal housing size, $h^*_i(p_{st})$, where p_{st} denotes the square meter price of the chosen location. Hence, in this case the vector of choice probabilities of this binary decision is $\bar{q}^{B*}_{ist} = (0, 1)^\top$ by assumption.

In contrast, a household that chooses to stay at its residential location faces a binary housing size choice of keeping its housing size or updating to the optimal housing size. This household makes its binary decision based on the vector of choice probabilities, $\bar{q}^{B*}_{ist} = \left(q^{B*}_{is0t}, q^{B*}_{is1t}\right)^\top \in \Delta^2$, that maximizes its perturbed utility

$$\bar{q}^{B*}_{ist} = \underset{\bar{q}^B_{ist} \in \Delta^2}{\arg\max} \left\{ \sum_{b=0}^{1} q^B_{isbt} \cdot u^B_{isbt} - Y^B\left(\bar{q}^B_{ist}\right) \right\}, \tag{7}$$

where $Y^B : \mathbb{R}^2_{++} \to \mathbb{R}$ is the perturbation function of the household and u^B_{isbt} is the choice specific value, when residing in location s with housing size h_{it-1}. The choice-specific payoff is

$$\begin{aligned} u^B_{is0t} &= -\alpha_{c_i} \cdot \phi \cdot p_{st} \cdot h_{it-1} + \beta_{1,c_i} \cdot h_{it-1} + \tfrac{1}{2} \cdot \beta_2 \cdot h^2_{it-1}, \\ u^B_{is1t} &= -\alpha_{c_i} \cdot \phi \cdot p_{st} \cdot h_{it} + \beta_{1,c_i} \cdot h_{it} + \tfrac{1}{2} \cdot \beta_2 \cdot h^2_{it} - \rho^B, \end{aligned} \tag{8}$$

where ρ is the moving costs and ϕ is the annual user cost of housing. Hence, $\phi \cdot p_{st} \cdot h_{it-1}$ is the annual housing cost when keeping the housing size. In the spirit of Carstensen et al. (2020), the utility of housing is quadratic in housing size. As a result, the optimal housing size can be found from the first-order condition

$$\frac{\partial u^B_{ir1t}}{\partial h_{it}} = 0 \Leftrightarrow h^*_i(p_{st}) = -\tfrac{\beta_{1,c_i}}{\beta_2} + \tfrac{\alpha_{c_i}}{\beta_2} \cdot \phi \cdot p_{st}. \tag{9}$$

This implies that if the marginal utility of money is positive, $\alpha_{c_i} > 0$, and the utility of housing is concave, $\beta_2 < 0$, then the optimal housing size decreases linearly in the square meter price of the chosen location. Hence, households of the same type will choose the same housing size if they move to the same location. In turn, the utility of updating housing size can be rewritten as

$$u^B_{is1t} = -\tfrac{1}{2} \cdot \beta_2 \cdot h^*_i(p_{st})^2 - \rho^B. \tag{10}$$

Due to the inclusion of moving costs - when the household chooses to update its housing size - the housing size choice problem must be solved for any value of h_{it-1}. However, as this is computationally costly, it is assumed that h_{it-1} is common for all households of the same type that reside in the same location. This is consistent with the long-run equilibrium of the model, as in the long run all households that move will choose the same housing size as the households - of the same type - that already reside in that location. It is shown in Section 7 that it is important to obtain realistic results to include cost when the household changes its housing size.

For a household that chooses to stay at its residential location, the optimal choice probabilities, $\bar{q}^B_{ist}$, and the expected utility, EV^B_{ist}, of the binary housing size decision can be found from Eq. (2) and Eq. (3), respectively, for a given the vector of choice-specific utilities, $(u^B_{is0t}, u^B_{is1t})^\top$, and for a given specification of the perturbation function, Y^B. For households that choose to move to another location, the choice probabilities are $\bar{q}^{B*}_{ist} = (0,1)^\top$, and the expected utility is simply given by the utility of having the corresponding optimal housing size of the chosen residential location, $EV^B_{ist} = u^B_{is1t}$.

4.3 Residential location

For simplicity let subscript r denote the previous chosen residential location of the household. The individual household chooses its residential location based on the vector of choice probabilities, $\bar{q}^{R*}_{irt} = \left(q^{R*}_{ir1t}, ..., q^{R*}_{irNt}\right)^\top \in \Delta^N$, that maximizes its perturbed utility

$$\bar{q}^{R*}_{irt} = \underset{\bar{q}^R_{irt} \in \Delta^N}{\arg\max} \left\{ \sum_{s=1}^{N} q^R_{irst} \cdot u^R_{irst} - Y^R\Big(\bar{q}^R_{irt}\Big) \right\}, \tag{11}$$

where $Y^R : \mathbb{R}^N_{++} \to \mathbb{R}$ is the perturbation function of the residential problem and u^R_{irst} is the choice specific value of the residential location s when residing in location r.

As the outcomes of the work location and housing size decisions are unknown, the household bases its decision on the expected utilities of these decisions. For a couple, the expected choice-specific utilities are

$$u^R_{irst} = \kappa_s + \rho^D \cdot dist(r,s) + EU^B_{ist} + EU^F_{ist} + EU^M_{ist}, \quad \text{for } s = 1, ..., N, \tag{12}$$

where κ_s is the utility of residing in location s and ρ^D is the cost of moving from location r to s. This term is included in order to capture that most households, when moving, move to a location close to their previous location. EU^B_{ist} is the expected utility of the binary housing size decision when location s is chosen. Similarly, (EU^F_{ist}, EU^M_{ist}) are the expected utilities of the female and male household members' work location decision when the household chooses location s. For a single female household or single male household, the expected utility of the work location decision of the opposite gender drops out.

Given the vector of choice-specific utilities, $(u^R_{is1t}, ..., u^R_{isNt})^\top$, and for a given specification of the perturbation function, Y^R, the optimal choice-probabilities, $\bar{q}^R_{ist}$, can be found from Eq. (2).

4.4 Firms' choices

The representative firm in location j uses the different household members as input in it production. However, the production is independent of the residential location of the hired household members, and the representative firm is assumed to have an inelastic aggregate demand for labor given by Q_{jt}. As shown by Andersen and Lee (2022), the firm's profit maximization problem can be written as a perturbed utility maximization problem, where the firm chooses the composition of household members, $\bar{q}^{C*}_{jt} = (q^{C*}_{j1Ft}, ..., q^{C*}_{jEFt}, q^{C}_{j1Mt}, ..., q^{C*}_{jEMt})^\top$, in order to maximize its profit

$$\bar{q}^{C*}_{jt} = \underset{\bar{q}^C_{jt} \in \Delta^{2 \cdot E}}{\arg\max} \left\{ \sum_{e=1}^{E} \sum_{g \in (F,M)} q^C_{jegt} \cdot v^C_{jegt} - Y^C\left(\bar{q}^C_{jt}\right) \right\} \cdot Q_{jt}. \tag{13}$$

$Y^C : \mathbb{R}^{2 \cdot E} \to \mathbb{R}$ is the perturbation function of the firm that represents the non-linear part of the profit function, and v^C_{jegt} is the choice-specific net revenue from hiring a household member of type e and gender g, and constitutes the linear part of the profit function. The choice-specific net revenue for hiring a female or male household member is given as

$$\begin{aligned} v^C_{jeFt} &= \pi^F_{jet} - w^F_{jet}, \quad \text{for } e = 1, ..., E, \\ v^C_{jeMt} &= \pi^M_{jet} - w^M_{jet}, \quad \text{for } e = 1, ..., E. \end{aligned} \tag{14}$$

$(\pi^F_{jet}, \pi^M_{jet})$ represent the productivity of a female or male household member of educational type e employed in the representative firm at location j, and (w^F_{jet}, w^M_{jet}) represent the associated wage for the female and male household member that the firm must pay when hiring.

As the optimal composition of household members is independent of Q_{jt}, the representative firm's optimal composition of household members, $\bar{q}_{jt}^{C*}$, can be found from Eq. (2) for a given vector of choice-specific net revenues, $(v_{j1Ft}^C, ..., v_{jEFt}^C, v_{j1Mt}^C, ..., v_{jEMt}^C)^\top$, and for a given specification of the perturbation function, Y^C.

4.5 Perturbation functions

Recall that the perturbation functions are specified in terms of the similarity function defined in Section 2. In order to reduce the computational complexity, the similarity functions for the work location and binary housing size problem of the households, (Y^B, Y^F, Y^M), are given by the Shannon entropy. Similarly, the similarity function of the firm, Y^C, is given by the Shannon entropy. Hence, for these decision problems the optimal choice probabilities or optimal composition of household members and the corresponding expected utilities and profits simplify to the multinomial logit choice probabilities and the *log-sum*, respectively. This implies that the substitution patterns between the alternatives of these decision problems are governed by the IIA property.

As this paper focuses on the spatial substitution patterns, the similarity function for the residential location problem, Y^R, is given a rich parameterization in order to allow substitution among the residential locations to be determined by the data. Hence, it is assumed that the locations can belong to two nests to a varying degree. The degree by which the locations belongs to the first nest is described by the vector $\bar{\omega} = (\omega_1, ..., \omega_N)^\top$. As a result, the degree by which the locations belong to the second nest is given by $(1-\omega_1, ..., 1-\omega_N)^\top$. The nesting parameter that determines the degree of substitution inside each nest is assumed to be identical for both nests, $\eta_1 = \eta_2 = \eta$.

$$
\begin{aligned}
Y^R\left(\bar{q}_{irt}^R\right) &= \sum_{s=1}^{J} q_{irst}^R \cdot g_s^R\left(\bar{q}_{irt}^R\right), \\
g_s^R\left(\bar{q}_{irt}^R\right) &= \left(1-\eta\right) \log q_{irst}^R + \eta \cdot \omega_s \log\left(\sum_{k=1}^{K} \omega_k \cdot q_{irkt}^R\right) \\
&\quad + \eta \cdot (1-\omega_s) \log\left(\sum_{k=1}^{K} (1-\omega_k) \cdot q_{irkt}^R\right).
\end{aligned}
$$

As described by Fosgerau and Nielsen (2021), the parameter η describes the degree of of substitution inside the nests relative to the logit model. For $\eta = 0$ the function reduces to $y_s^R\left(\bar{q}_{irt}^R\right) = \log q_{irst}^R$ and substitution is described by the IIA property. For $\eta > 0$, locations with similar sized ω_s are closer substitutes than for the logit model. In contrast, for $\eta < 0$, locations with similar sized ω_s are less substitutes than for the logit model. For $\eta > 0$, the model is closely related to the generalized nested logit model by Wen and Koppelman (2001). Further, for $\eta > 0$, the similarity model reduces to the nested logit model if ω_s equals 0 or 1 for $s = 1, ..., N$, as each location in this case only belongs to either the first or the second nest.

4.6 Equilibrium

In equilibrium, the demand for housing must equal the supply of housing across all locations, and the supply of labor must equal the demand for labor across all locations and types of household members in the labor force. This paper takes a short run perspective and assumes a fixed supply of housing and fixed aggregated demand for labor. Thus the longer run dynamics, where the supply of housing and the aggregated demand for labor adjust to changes in square meter prices and wages, are ignored.

All households face the same vector of square meter prices, $\bar{p}_t = (p_{1t}, ..., p_{Nt})^\top$, but the different types of household members face different offered wages. Hence, the household member of gender g and type e faces the vector of wages $\bar{w}^g_{et} = (w^g_{e1t}, ..., w^g_{eNt})^\top$. For later use, define the stacked matrix $W_t = (\bar{w}^F_{1t}, ..., \bar{w}^F_{E-1,t}, \bar{w}^M_{1t}, ..., \bar{w}^M_{E-1,t})$ that contains all the vectors of wages.

The demand for housing in location $s = 1, ..., N$ is given in terms of the housing sizes, the choice probabilities of the residential location problem, and the choice probabilities of the binary choice

$$D^R_{st}(\bar{p}_t, W_t) = \sum_{i=1}^{H_t} q^{R*}_{irst} \cdot \left(q^{B*}_{is0t} \cdot h_{it-1} + q^{B*}_{is1t} \cdot h^*_i(p_{st}) \right), \tag{15}$$

where $q^{B*}_{is0t} = 0$ and $q^{B*}_{is1t} = 1$ for $s \neq r$. Since utility increases with consumption of the outside good, households prefer to reside in less expensive locations. In turn, the demand for housing decreases with square meter prices. In contrast, the presence of commuting costs implies that the demand for housing increases with the offered wages of that location, as households prefer to reside close to locations that offer higher wages.

The supply of female or male labor of type $e = 1, .., E-1$ in location $s = 1, ..., N$ is given in terms of the choice probabilities of the residential and work location problems

$$S^F_{set}(\bar{p}_t, W_t) = \sum_{i=1}^{H} q^{R*}_{ijst} \cdot q^{F*}_{ijst} \cdot \mathbb{1}(f_i = e), \tag{16}$$

$$S^M_{set}(\bar{p}_t, W_t) = \sum_{i=1}^{H} q^{R*}_{ijst} \cdot q^{M*}_{ijst} \cdot \mathbb{1}(m_i = e). \tag{17}$$

The resulting supply of labor at a given location therefore decreases with square meter prices and increases with the offered wages, since households prefer to reside close to where they work and receive high wages.

The demand for female and male workers of each type, (D^F_{set}, D^M_{set}) is given by the firm's choice probabilities

$$\begin{aligned} D^F_{set}(W^*_t) &= Q_{st} \cdot q^{C*}_{seFt}(W^*_t), \\ D^M_{set}(W^*_t) &= Q_{st} \cdot q^{C*}_{seMt}(W^*_t). \end{aligned} \tag{18}$$

Note that the aggregated demand for labor in location s is fixed at Q_{jt}, but the firm can adjust its composition of household members, $(q^{C*}_{seFt}, q^{C*}_{seMt})$.

As the supply of square meters, S^R_{st}, is fixed across locations, the square meter prices and wages, $(\bar{p}_t, W_t)$, must adjust such that demand equals supply for type $e = 1, ..., E-1$ and location $s = 1, ..., N$

$$\begin{aligned} D^R_{st}(\bar{p}^*_t, W^*_t) &= S^R_{st}, \\ D^F_{set}(W^*_t) &= S^F_{set}(\bar{p}^*_t, W^*_t), \\ D^M_{set}(W^*_t) &= S^M_{set}(\bar{p}^*_t, W^*_t). \end{aligned} \tag{19}$$

Given the decentralized nature of the housing and labor market, square meter prices and wages are assumed to adjust in order to clear the markets. If, for example, the supply of housing exceeds the demand in a given location, square meter prices must decrease in order for the demand to increase. As prices fall, more households will be willing to move to this particular location, which in turn leads to an increase in the supply of workers in that location. As a result, wages will also have to decrease in order to clear the labor market. In contrast if the supply of labor exceeds the demand in a given location, wages of that location must decrease. As a consequence, the demand for housing in that location will decrease, which leads to a decrease in the square meter price of that location.

5 Estimation

This section describes the estimation procedure used when estimating the structural parameters of the proposed model. The user cost parameter, ϕ, is set to 0.05. Hence, the remaining structural parameters to be estimated are denoted by the vector $\theta = (\bar{\alpha}, \bar{\beta}_1, \beta_2, \bar{\kappa}, \bar{\rho}, \lambda^F, \lambda^M, \bar{\mu}^F, \bar{\mu}^M, \eta, \bar{\omega}, \bar{\pi}^F, \bar{\pi}^M)$. The structural parameters can be divided into two sets $\theta = (\theta_1, \theta_2)$, where $\theta_1 = (\bar{\alpha}, \bar{\beta}_1, \beta_2, \bar{\kappa}, \bar{\rho}, \lambda^F, \lambda^M, \bar{\mu}^F, \bar{\mu}^M, \eta, \bar{\omega})$ describes the preferences of the households and $\theta_2 = (\bar{\pi}^F, \bar{\pi}^M)$ describes the productivity of the firms. The two vectors of parameters, (θ_1, θ_2), are estimated independently of each other.

In order to reduce the computational complexity, is θ_1 estimated in two steps, as proposed by Carstensen et al. (2020). The first step is based on Eq. (9), that describes the optimal housing size. The reduced form parameters $\gamma_{1,c_i} = -\beta_{1,c_i}/\beta_2$ and $\gamma_{2,c_i} = \alpha_{c_i}/\beta_2$ can be estimated from the following equation by simple OLS

$$\widetilde{h}_{it} = \sum_{k\in\{S,C\}} \left\{\gamma_{1,k} + \gamma_{2,k}\cdot\phi\cdot\widetilde{p}_{it}\right\}\mathbb{1}(c_i = k) + \varepsilon_{it}, \text{ for } i = 1, ..., H_t, \tag{20}$$

for $t = 1, ..., T$. The variable $\widetilde{h}_{it}$ is the observed housing size, $\widetilde{p}_{it}$ is the observed square meter price and ε_{it} is assumed to be a random measurement error. Recall that $c_i \in \{S, C\}$ indicates whether household i consists of a single or a couple.

Given the reduced form estimates, $(\hat{\gamma}_{1,S}, \hat{\gamma}_{1,C}, \hat{\gamma}_{2,S}, \hat{\gamma}_{2,C})$, the corresponding structural parameters $(\alpha_S, \alpha_C, \beta_{1,S}, \beta_{1,C})$ can be expressed in terms of β_2

$$\begin{aligned} \beta_{1,c_i} &= -\hat{\gamma}_{1,c_i}\cdot\beta_2, \quad \text{for } c_i = S, C, \\ \alpha_{c_i} &= \hat{\gamma}_{2,c_i}\cdot\beta_2, \qquad \text{for } c_i = S, C. \end{aligned} \tag{21}$$

As a result, these structural parameters can be concentrated out of the likelihood function, such that the dimensionality of the maximum likelihood problem solved in the second step of the estimation procedure is reduced. The likelihood contribution of household i - that previously resided in location d^R_{it-1} with housing size h_{it-1} - can be written as the product of the choice probabilities of the household's residential location decision, the choice probabilities of the households binary choice, and the choice probabilities of the work location decision of the household members in the labor force

$$Pr(d^R_{it}, d^B_{it}, d^F_{it}, d^M_{it}|d^R_{it-1}, h_{it-1}, \theta_1) = q^{R*}_{i,d^R_{it-1},d^R_{it},t}(h_{it-1}, \theta_1)\cdot q^{B*}_{i,d^R_{it},d^B_{it},t}(h_{it-1}, \theta_1)\prod_g q^{g*}_{i,d^R_{it},d^g_{it},t}(\theta_1), \tag{22}$$

where the product is over the individual household members of household i. It is assumed that households that move to another address choose to update their previous chosen housing size to the optimal housing size, $d^B_{it} = 1$.

The maximum likelihood estimator is the vector $\hat{\theta}$ that maximizes the log-likelihood function, which due to Eq. (22) can be expressed by the sum of the log-likelihood function of the residential choices, $L_R(\theta)$, the log-likelihood function of the binary choices, $L_B(\theta)$, and the log-likelihood function of the work location choices, $L_W(\theta)$,

$$\hat{\theta}_1 = \arg\max_{\theta_1}\ L(\theta_1) = L_R(\theta_1) + L_B(\theta_1) + L_W(\theta_1), \tag{23}$$

respectively given as

$$\begin{aligned}
L_R(\theta_1) &= \sum_{t=1}^{T}\sum_{i=1}^{H_t} \log q^{R*}_{i,d^R_{it-1},d^R_{it},t}(h_{it-1},\theta_1), \\
L_B(\theta_1) &= \sum_{t=1}^{T}\sum_{i=1}^{H_t} \log q^{B*}_{i,d^R_{it},d^B_{it},t}(h_{it-1},\theta_1), \\
L_W(\theta_1) &= \sum_{t=1}^{T}\sum_{i=1}^{H_t}\sum_{g} \log q^{g*}_{i,d^R_{it},d^g_{it},t}(\theta_1).
\end{aligned} \tag{24}$$

The remaining structural parameters describing the productivity of the representative firms, θ_2, are estimated non-parametrically for each time period, $t = 1,...,T$, from the observed choice probabilities and average wages

$$\begin{aligned}
\hat{\pi}^F_{jet} &= \log(\tilde{q}^{C*}_{jeFt}) - \log(\tilde{q}^{C*}_{j1Ft}) + \tilde{w}^F_{jet} - \tilde{w}^F_{j1t}, \text{ for } j = 2,...,E, \\
\hat{\pi}^M_{jet} &= \log(\tilde{q}^{C*}_{jeMt}) - \log(\tilde{q}^{C*}_{j1Ft}) + \tilde{w}^M_{jet} - \tilde{w}^F_{j1t}, \text{ for } j = 1,...,E,
\end{aligned} \tag{25}$$

where π^F_{j1t} is normalized to zero across representative firms and time periods in order to identify the productivity parameters.

6 Results

This section presents the estimation results of the first and second step of the estimation procedure described in Section 5. In order to compare the quantitative implications on the counterfactual analysis in Section 7 of the assumptions behind the proposed model, three different models are estimated for the households choices, $(d^R_{it}, d^B_{it}, d^F_{it}, d^M_{it})$.

The proposed model in Section 4 is referred to as *model* {*3*}. *Model* {*2*} is similar, except that η is restricted to be zero. This implies that the perturbation function of the residential choice problem is specified in terms of the Shannon entropy, and the resulting choice probabilities are given by the multinomial logit choice probabilities. Hence, for this model, the substitution patterns are solely described by the IIA property.

For *model* {*1*}, η is restricted to be zero as well, but the households in this model can, similarly to Carstensen et al. (2020), adjust their housing size at no extra cost. Hence, households in this model do not face the binary housing size decision, as they will always choose the optimal housing size. As a result, when estimating this model, the information on households that move inside their current municipality is ignored.

6.1 Estimation Results

Table 2 shows the reduced-form estimates of the first step of the estimation procedure, on which the estimation of the structural parameters of all the three models are based. This first step is based on observed housing sizes and observed square meter prices for 130,000 property sales across all the included municipalities in the period 2010 to 2013.

The table shows that the estimates are statistically significant at a 99% significance level and the estimates have the expected sign. However, the R^2 is relatively low, $R^2 = 0.064$. Hence, the model can only explain 6 percent of the variation in housing sizes of the sold properties. The estimated constants $(\hat{\gamma}_{1,S}, \hat{\gamma}_{1,C})$ are both positive, and since $\hat{\gamma}_{1,S} < \hat{\gamma}_{1,C}$ couples have a higher baseline demand for square meters. Further, the price coefficients $(\hat{\gamma}_{2,S}, \hat{\gamma}_{2,C})$ are both negative, which implies that the demand for square meters decreases with square meter prices as expected. Since $\hat{\gamma}_{2,S} < \hat{\gamma}_{2,C}$, the demand of couples is less price sensitive compared with the demand of singles. The estimates imply that the demand for housing is higher for couples compared with singles, for any given square meter price.

	coef.	s.e.	t-stat.
$\hat{\gamma}_{1,S}$	1.388	0.004	335.970
$\hat{\gamma}_{1,C}$	1.505	0.004	367.998
$\hat{\gamma}_{2,S}$	-2.498	0.045	-55.967
$\hat{\gamma}_{2,C}$	-2.359	0.045	-52.828

Table 2: Reduced form estimates are based on $130,467$ property sales from 2010 to 2013. $R^2 = 0.0643$.

Table 3 shows the estimated structural key parameters and log-likelihood values of the models. It should be stressed that *model* {*1*} is not nested in any of the other models, which implies that the log-likelihood values of this model cannot be compared with the other models. In contrast, *model* {*2*} is nested in *model* {*3*}. Hence, the log-likelihood values are comparable for these two models. The table shows that the log-likelihood value increases when η is not restricted to zero, but estimated freely.[6] However, *model* {*3*} contains 372 parameters to be estimated, whereas *model* {*2*} only contains 280 parameters. A comparison of the AIC scores of the two models clearly favors *model* {*3*}, since $P = \exp\left((52,361,218 - 53,593,792)/2\right) \approx 0$. The estimated value of η is 0.55 for this model, which implies that the substitution patterns between the different residential locations are not governed by the IIA property of the logit model. Hence, municipalities with

[6] As the null hypothesis, $\omega_j = 0$, is on the boundary of the parameter set, a likelihood ratio test is not applicable, as the test-statistic is not χ^2-distributed.

similar ω_j are closer substitutes compared with the logit model.

The estimates for $\bar{\omega}$ are shown in Figure 20 in Appendix E. The figure shows that municipalities closest to *Copenhagen* mostly belong to the second nest, whereas the municipalities on the mainland mostly belong to the first nest. As a result, the municipalities closest to *Copenhagen* are closer substitutes compared with the logit model.

	model {1}	model {2}	model {3}
$\hat{\alpha}_S$	5.7983	5.8333	5.6367
	(0.0004)	(0.0004)	(0.0004)
$\hat{\alpha}_C$	5.4764	5.5094	5.3238
	(0.0003)	(0.0004)	(0.0004)
$\hat{\beta}_{1,S}$	3.2214	3.2408	3.1316
	(0.0001)	(0.0001)	(0.0002)
$\hat{\beta}_{1,C}$	3.4937	3.5147	3.3963
	(0.0001)	(0.0001)	(0.0002)
$\hat{\beta}_2$	-2.3215	-2.3355	-2.2568
	(0.0001)	(0.0001)	(0.0001)
$\hat{\rho}^B$	-6.2420	-3.2792	-2.6984
	(0.0002)	(0.0002)	(0.0002)
$\hat{\rho}^D$	-1.0989	-4.2789	-0.8221
	(0.0002)	(0.0002)	(0.0002)
$\hat{\lambda}^F$	-4.6085	-4.6010	-4.5827
	(0.0001)	(0.0001)	(0.0002)
$\hat{\lambda}^M$	-3.6665	-3.6625	-3.6403
	(0.0001)	(0.0001)	(0.0001)
$\hat{\eta}$	0	0	0.5515
	(-)	(-)	(0.0001)
$L(\hat{\theta}_1)$	-23,632,958	-26,796,616	-26,180,237
$L_R(\hat{\theta}_1)$	-2,482,564	-2,987,275	-2,481,013
$L_B(\hat{\theta}_1)$	-	-2,658,572	-2,537,002
$L_J(\hat{\theta}_1)$	-21,150,394	-21,150,768	-21,162,221
Parameters	280	280	372
AIC	47,266,476	53,593,792	52,361,218

Table 3: Structural parameters are estimated by maximum likelihood based on the choices of 10,186,442 households. Commuting time is measured in hours. All monetary units are measured in DKK 1,000,000. Housing size is measured in 100 m^2.

The estimated productivity parameters of the firms, $(\hat{\bar{\pi}}^F, \hat{\bar{\pi}}^M)$, are difficult to interpret, as the reference alternative differs across municipalities. Hence, the productivity parameters can only be compared within municipalities and not across municipalities. The average estimated productivity parameters of the representative firms in the different municipalities are shown in Table 11 in Appendix E.

Since utility is linear in income, utilities can for comparison be transformed into monetary units. Table 4 shows the average estimated (annual) value of housing and moving. The table shows that the average estimated value of housing is DKK 4,700 per square meter across models. The identical estimates reflect that the monetary value of housing only depends on the reduced form ratios - which are identical across models - and not the structural parameter.[7] As the average user cost per square meter is DKK 900, the average estimated value of housing is relatively big.

	model {1}	model {2}	model {3}
Housing (DKK 1,000/m^2)	4.7	4.7	4.7
Moving* (DKK 1,000)	0	-29.0	-24.7
Moving municipality** (DKK 1,000)	-65.0	-68.2	-32.2

Table 4: Average estimated annual monetary value of attributes. *refers to the costs of moving inside the municipality in which the household currently resides. **refers to the costs of moving to another municipality than the municipality in which the household currently resides. The average estimated annual monetary values of moving address or municipality are conditioned on moving inside the current municipality or moving to another municipality, respectively. For comparison, the moving costs are turned into a infinite annuity based on a 5 percent interest rate.

The average estimated costs of moving differs across the models. For *model {1}* households can freely move inside their current municipality. Hence, the average costs of moving address is zero for this model. In contrast, the average costs of moving to another municipality is DKK 65,000 for this model. Recall that moving costs include financial and none-financial costs of moving. However, this is a very high estimate for moving costs. It is a general result of the literature on dynamic discrete choice models that the estimated switching costs are upward biased, as all relevant unobserved heterogeneity cannot be included in the model, see, for example, Honoré and Kyriazidou (2000). For that reason, most of the existing literature does not report the average estimated value of moving. Due to the inclusion of the additional parameters of the similarity function, $(\eta, \bar{\omega})$, the average estimated moving costs of moving municipality is considerably lower (DKK 32,000), as the additional parameters control for unobserved heterogeneity.

[7] Recall that utility of housing is $\beta_{1,c_i} \cdot h_{it} + \beta_{2,c_i} \cdot h_{it}^2$. Dividing by α_{c_i} it can be shown that the monetary value of housing equals $-\gamma_{1,c_i} \cdot \gamma_{2,c_i}^{-1} \cdot h_{it} + \gamma_{2,c_i}^{-1} \cdot h_{it}^2$.

Table 5 shows the average estimated cost of commuting one hour. Note that monetary costs, lost time and discomfort are included in the commuting costs. Across models, the estimated commuting costs for females and males are DKK 1,900 and 1,500, respectively. These estimates are much higher than those found in the empirical literature on commuting costs[8]. Similar to the high estimates of moving costs, this is a consequence of unoberserved heterogeneity. However, the relative difference in commuting costs between females and males is in line with Le Barbanchon et al. (2020), who find that females value commute around 20 percent more than males do.

	model {1}	model {2}	model {3}
Females (DKK 1,000)	-1.9	-1.8	-1.9
Males (DKK 1,000)	-1.5	-1.5	-1.5

Table 5: Average monetary value of commuting one hour. The average monetary values of commuting for e.g. females is an average of of the term $(\hat{\lambda}^F/\hat{\alpha}_i)/(2\cdot 224)$ for females that commute. 224 refers the number of yearly workdays and household members are assumed to commute back and forth between the residential and work location each day.

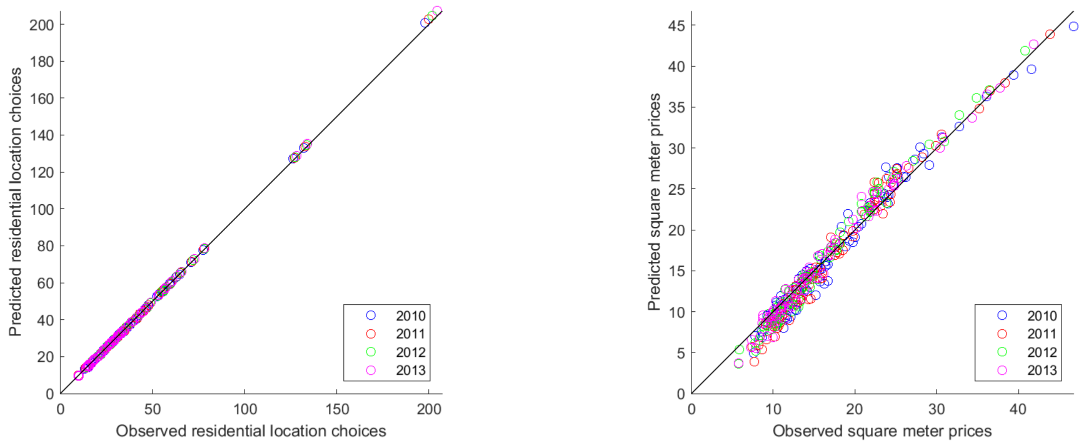
6.2 Model fit

This subsection illustrates the model fit of the estimated spatial equilibrium model, where households' behavior is described by *model {3}*. Hence, the equilibrium concept of Subsection 4.6 is applied to the estimated model, and predicted square meter prices and wages are given by the equilibrium square meter prices and wages such that markets clears.

Figure 6a compares the predicted and observed number of inhabitants across municipalities. Since households rarely move and the proposed model includes moving costs, the model does as expected a relatively good job in predicting the number of inhabitants. Hence, most of the observations are located on the 45-degree line.

The model does a slightly worse job in predicting the observed square meter prices, as illustrated by Figure 6b. This is because square meter prices are determined by the marginal buyer, which implies that small initial differences between supply and demand for housing can result in relatively large adjustments of square meter prices. Differences between demand and supply in the labor market can magnify the adjustment of the square meter prices due to spill-over effects.

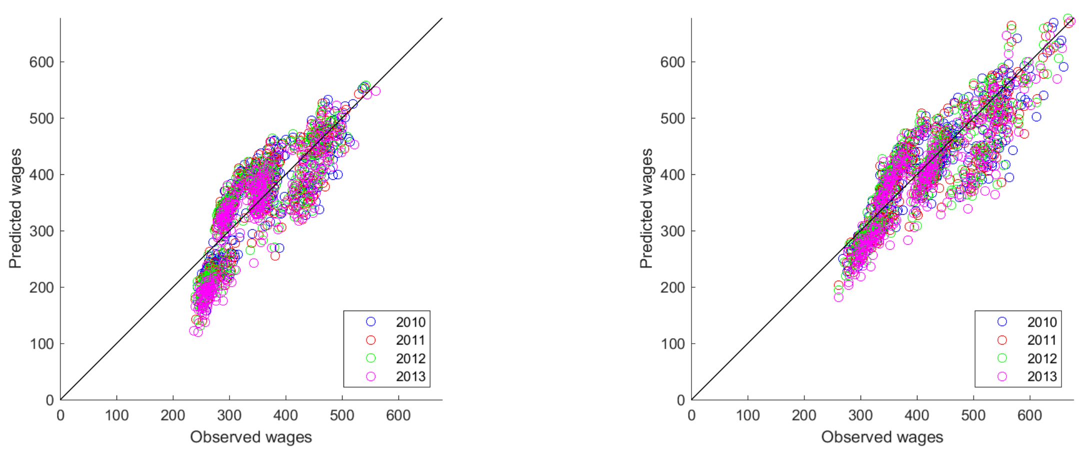
[8] E.g Van Ommeren and Fosgerau (2009) find that, on average, workers' marginal costs of one hour of commuting are about EUR 17 (roughly DKK 160 in 2018 prices). It should be noted that the estimates in Table 5 is based on gross wages. The average Danish tax rate is close to 50 percent.



(a) Observed and predicted number of inhabitants. **(b)** Observed and predicted square meter prices.

Figure 6: *Model {3}*: Observed and predicted number of inhabitants (1,000) and square meter prices (DKK 1,000).

In contrast to the residential location decision, the work location decision does not include switching costs and since commuting to another municipality is more prevalent in the data than moving to a another municipality, it is more difficult to predict the labor supply than the housing demand. Hence, the model does a relatively worse job in predicting the observed wages for females and males. This is illustrated by Figures 7a and 7b.



(a) Observed and predicted wages for females. **(b)** Observed and predicted wages for males.

Figure 7: *Model {3}*: Observed and predicted wages (DKK 1,000).

7 Counterfactual study

Based on the estimated models presented in the previous section, this section conducts a counterfactual study of an increase of 1 percent in the supply of housing in *Copenhagen*. The counterfactual analysis has two purposes. First, the analysis illustrates the consequences of relaxing the assumption by Carstensen et al. (2020) that households can adjust their demand for housing at no extra cost. Second, it illustrates the importance of spatial substitution patterns.

Before turning to the results, it should be noted that the demand for housing can change along the extensive and intensive margins. Changes along the extensive margin reflect households that move from one municipality to another, and thereby increase the demand for housing in the municipality they move to and reduce the demand for housing in the municipality they move away from. In contrast, changes along the intensive margin reflect households that move inside their current municipality, and thereby only affect the demand for housing in that particular municipality. To illustrate the importance of distinguishing between changes along these two margins, consider the case where the supply of housing in a single municipality increases. The square meter price of that municipality must decrease in order for the demand and supply of housing to clear. If the demand only adjusts along the intensive margin, this implies that the demand for housing in the remaining municipalities will be unaffected, as the demand and supply in the starting point equates. Hence, the square meter prices of these municipalities will be left unaffected.

Table 6 shows, for each of the three estimated models, the corresponding changes in the number of inhabitants and square meter prices for *Copenhagen* due to the counterfactual increase in the supply of housing in this location. At the bottom of the table, the total adjustment of the demand for housing in *Copenhagen* decomposed into the percentage change along the intensive and extensive margin is shown.

	model {1}	model {2}	model {3}
Inhabitants (1,000)	0.0	2.3	2.5
Inhabitants (percent)	0.0	0.5	0.5
Avg. sqm. prices (DKK 1,000)	-0.1	-3.4	-2.6
Avg. sqm. prices (percent)	-0.5	-13.5	-10.3
Intensive margin (percent)	98.8	40.0	43.8
Extensive margin (percent)	1.2	60.0	56.2

Table 6: Changes in equilibrium outcomes for *Copenhagen* due to a 1 percent increase in housing supply in *Copenhagen*. The adjustment in demand for housing in *Copenhagen* is decomposed into changes along the intensive and extensive margins.

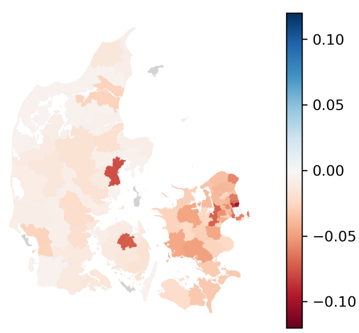
The table shows for *model {1}* that the intensive margin accounts for 98.8 percent of the total adjustment of the demand for housing in *Copenhagen*. This is due to the high estimated cost of moving to another municipality and due to the assumption that households can change their housing size at no extra cost. Due to the high moving costs, the square meter prices must decrease a lot to make it attractive for households to move to *Copenhagen*. However, since households that already reside in *Copenhagen* can increase their housing size at no extra cost, a small reduction in the square meter price is sufficient for demand and supply to equate. As a result, the increase in the number of inhabitants in *Copenhagen* is below 100, and the 1 percent increase in the supply of housing in *Copenhagen* only reduces the average square meter price by 0.5 percent (corresponding to a decrease of DKK 100). Table 8 and Figure 18 in Appendix D show that the number of inhabitants and average square meter prices for the remaining municipalities are unaffected for this model, as very few households move to *Copenhagen*.

Table 6 shows for *model {2}* that the intensive margin only constitutes 40.0 percent of the total adjustment. This is because households in this model can only adjust their housing size at a cost. As a result, the square meter price of *Copenhagen* must decrease by a relatively large fraction for the demand and supply of housing to clear. This implies - despite the high moving costs - that a substantial number of households in the remaining municipalities are willing to move to *Copenhagen*. As some of the adjustment works through the extensive margin, the square meter prices are reduced in the remaining municipalities (see Table 9 in Appendix D). For *model {2}*, the population in *Copenhagen* increases by 0.5 percent (or 2,300 inhabitants) and the average square meter price of this location decreases by 13.5 percent (corresponding to DKK 3,400). The large decrease in the square meter price of *Copenhagen* is a result of the high estimates of the moving costs parameters.

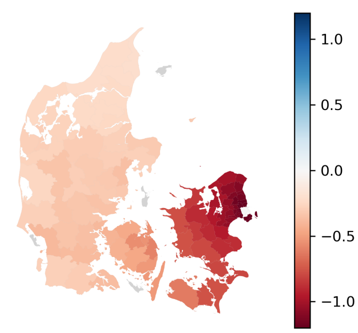
Finally, Table 6 shows that the changes along the intensive margins are of similar size for *model {3}* and *model {2}*. Hence, the changes in the number of inhabitants and square meter prices for *Copenhagen* are relatively similar for these models. For *model {2}*, the population in *Copenhagen* increases by 0.5 percent (or 2,500 inhabitants) and the average square meter price of this location decreases by 10.3 percent (corresponding to DKK 2,600).

However, the changes in the counterfactual equilibrium outcomes for the remaining municipalities differ for these two models. For *model {2}* and *model {3}*, tables 9 and 10 in Appendix D show the counterfactual changes for the remaining municipalities, which are grouped by the five regions (as given by Figure 1). The tables show that the average square meter prices in the *Southern, Central and North Regions of Denmark* are more affected in *model {3}* compared to *model {2}*. These regions constitute the municipalities located furthest away from *Copenhagen*, and the fact that the average square meter prices are more affected for these municipalities in *model* {*3*} compared to *model* {*2*} is due to the lower estimated moving costs. This is especially due to the lower estimate of the moving cost parameter $\hat{\rho}^D$, as this implies that the costs of moving from any of these municipalities to *Copenhagen* are relatively low.

Figure 8 shows the counterfactual changes in the number of inhabitants and square meter prices across municipalities for *model* {*3*}. As mentioned above, the square meter prices are mostly affected in the municipalities close to *Copenhagen* due to the relatively high estimates of $\hat{\rho}^D$. Further, the figure shows that the counterfactual change in the supply of housing mostly affects the number of inhabitants in the municipalities on Zealand, and especially for the municipalities located closest to *Copenhagen*. However, this figure shows that the number of inhabitants in the three second largest municipalities, *Odense*, *Aarhus*, and *Aalborg*, are also affected by the counterfactual change in the supply of housing. Hence, the changes in the number of inhabitants as illustrated by this figure mimic the moving patterns described Figure 4b in Section 3 relatively well. In contrast, *model* {*2*} predicts that the number of inhabitants in these municipalities are almost unaffected (see Table 19a in Appendix D).



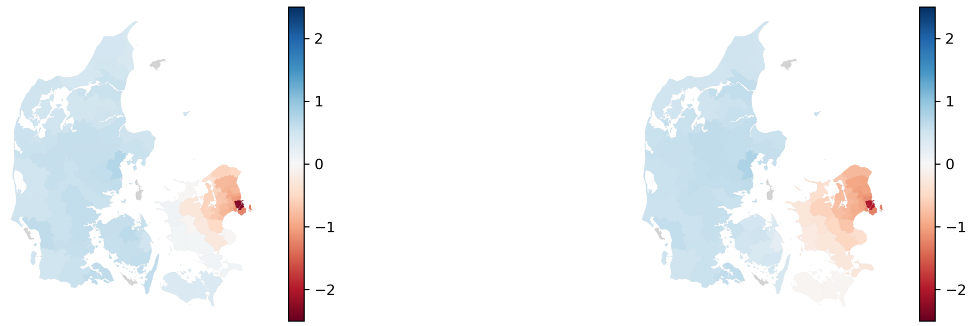
(a) Changes in inhabitants.



(b) Changes in square meter prices.

Figure 8: Model {3}: Counterfactual changes in the number of inhabitants (1,000) and square meter prices (DKK 1,000). For illustrative reasons, changes in *Copenhagen* are not included in the figures.

Figure 9 shows the counterfactual changes to the average wages for females and males. The figures show that the average wages decrease the most for *Copenhagen* and for the municipalities closest to this location, as the supply of labor increases for these municipalities. For the municipalities sufficiently far from *Copenhagen* the average wages increase as the supply of labor is reduced. Notice that the average wages for municipalities far from *Copenhagen* decrease more for males than females. This is because males commute longer than females due to the lower estimated commuting costs.



(a) Changes in average wages for females

(b) Changes in average wages for males

Figure 9: Model {3}: Counterfactual changes in average wages for females and males (DKK 1,000). Changes in average wages are based on the municipality the household members work in.

Table 7 summarizes the counterfactual changes in the average wages for the different regions, and shows, for example, that the average wages for males and females in *Copenhagen* decrease by 0.6 and 0.4 percent, respectively (corresponding to a decrease of DKK 2,100 and DKK 1,800).

	Copenhagen	Capital*	Zealand	Southern Denmark	Central Denmark	North Denmark
Females:						
Avg. wages (DKK 1,000)	-2.1	-0.8	-0.1	0.5	0.5	0.4
Avg. wages (percent)	-0.6	-0.2	-0.0	0.2	0.2	0.1
Males:						
Avg. wages (DKK 1,000)	-1.8	-0.9	-0.3	0.5	0.6	0.5
Avg. wages (percent)	-0.4	-0.2	-0.1	0.1	0.1	0.1

Table 7: Model {3}: Changes in equilibrium wages for females and males due to a 1 percent increase in housing supply in *Copenhagen*. *Changes in wages of *Copenhagen* are not included. Average wages are based on the municipality that the household members work in.

It should be stressed that the mentioned upward bias of the estimated moving costs implies that the counterfactual predictions overestimate the decreases in the square meter prices. For the same reasons, assuming households can adjust their housing size at no extra cost leads to a downward bias in the counterfactual predictions of the decreases in square meter prices. Similarly, the upward bias of the estimated commuting cost parameters implies that the absolute changes in wages are overestimated.

8 Conclusion

This paper proposes a spatial equilibrium model for the housing and labor market, where square meter prices and wages are determined in equilibrium. In contrast to the existing literature households can only costly adjust their housing size, and the behavior of the households are described by the perturbed utility model, which allows for a rich parameterization of the substitution patterns. As a proof of concept, the model is estimated on high-quality Danish administrative data. A counterfactual analysis of an 1 percent increase in the supply of housing in *Copenhagen* shows that the proposed model improves the realism of the price substitution patterns of housing for two reasons. First, when households freely can adjust their housing size most of the adjustment in the demand for housing only works through the intensive margin, which implies that square meter prices of the remaining municipalities are unaffected. In contrast, when households only costly can adjust their housing size, the a large fraction of the adjustment in the demand for housing works through the intensive margin, which implies that the square meter prices of the remaining municipalities are affected. Second, a rich parameterization of the perturbation function of the residential location problem implies that the counterfactual decreases in the square meter prices are largest for municipalities, where a large share of households move to *Copenhagen*, as these municipalities are close substitutes to *Copenhagen*. In contrast, when the perturbation function is given by the Shannon entropy, the counterfactual decreases in the square meter prices do not match this observed moving pattern.

I now highlight three directions for future work. A first direction is to extent the static perturbed utility model to a dynamic perturbed utility model in order to allow the households to be forward-looking - which the moving costs of the model inherently imply - but maintain the more general substitution patterns of the model compared to the random utility model. Andersen (2022) extents the proposed model of this paper to a dynamic spatial equilibrium model. However, due to the computational burden of solving the dynamic model, he only considers a choice set constituting of 24 locations.

As household members tend to stay at their current job, a second direction for future research is to incorporate frictions into the work location decision. This could be modelled by introducing switching costs into the work location decision similar to the moving costs in the residential location decision. However, this will greatly increase the number of different households in the model, as most types of households consist of two household members, which both need to be described by their previous work location. This extension will increase the realism of the model, as it induces households to move to locations close to their current work location.

Finally, a serious problem of the proposed empirical framework is that observed wages and square meter prices are equilibrium outcomes and therefore endogenous variables. Hence, the estimates are likely to be biased, as unobserved heterogeneity may affect both the choices of the households and the square meter prices and offered wages. Hence, a third direction of future research is to apply instrumental variable techniques in order to control for endogeneity. This is complicated, as there exists no generally accepted instrument for housing prices.

References

G. M. Ahlfeldt, S. J. Redding, D. M. Sturm, and N. Wolf. The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189, 2015. ISSN 0012-9682.

R. Allen and J. Rehbeck. Identification With Additively Separable Heterogeneity. *Econometrica*, 87(3):1021–1054, May 2019.

E. S. Andersen. A dynamic spatial equilibrium model for the housing and labor market. *Working paper*, 2022.

E. S. Andersen and Y. J. Lee. Multidimensional Matching and Labor Market Complementarity. *Working paper*, 2022.

P. Bayer, R. McMillan, and K. Rueben. An equilibrium model of sorting in an urban housing market. *National Bureau of Economic Research*, 2004.

P. Bayer, F. Ferreira, and R. McMillan. A unified framework for measuring preferences for schools and neighborhoods. *The Journal of political economy*, 115(4):588–638, 2007. ISSN 0022-3808.

P. Bayer, R. McMillan, A. Murphy, and C. Timmins. A dynamic model of demand for houses and neighborhoods. *Econometrica*, 84(3):893–942, 2016. ISSN 0012-9682.

M. Buchinsky, C. Gotlibovski, and O. Lifshitz. Residential Location, Work Location, and Labor Market Outcomes of Immigrants in Israel. *Econometrica*, 82(3):995–1054, May 2014.

C. L. Carstensen, F. Iskhakov, M. Juul-Hansen, J. Rust, and B. Schjerning. A dynamic equilibrium model of commuting, residential and work location choices. *Working paper*, 2020.

M. Fosgerau and N. Nielsen. Similarity, perturbed utility and discrete choice. *Working paper*, 2021.

M. Fosgerau, J. Monardo, and A. de Palma. The Inverse Product Differentiation Logit Model. *SSRN working Papers*, 2021.

M. Fosgerau, M. Paulsen, and T. K. Rasmussen. A perturbed utility route choice model. *Transportation Research Part C: Emerging Technologies*, 136:103514, 2022.

B. E. Honoré and E. Kyriazidou. Panel data discrete choice models with lagged dependent variables. *Econometrica*, 68(4):839–874, 2000.

J. Kennan and J. R. Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251, 2011. ISSN 0012-9682.

N. V. Kuminoff, C. Parmeter, D. Phaneuf, J. Pope, C. Timmins, and E. V. Partial identification of preferences in a dual-market sorting equilibrium, 2012.

T. Le Barbanchon, R. Rathelot, and A. Roulet. Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, 136(1):381–426, 2020.

D. McFadden. Conditional logit analysis of qualitative choice behaviour. *Frontiers in Econometrics*, pages 105–142, 1973.

D. McFadden. Conditional modeling the choice of residential location. *in eds. Karlquist, A., et al., Spatial Interaction Theory and Planning Models, Elsevier North-Holland, New York*, 1978.

F. Monte, S. J. Redding, and E. Rossi-Hansberg. Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–90, December 2018.

S. Rosen. Chapter 12 the theory of equalizing differences. In *Handbook of Labor Economics*, volume 1, pages 641–692. Elsevier, 1986.

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

C.-H. Wen and F. S. Koppelman. The generalized nested logit model. *Transportation Research Part B: Methodological*, 35(7):627–641, 2001.

J. S. Zax. Compensation for commutes in labor and housing markets. *Journal of Urban Economics*, 30(2):192–207, 1991.

A Appendix

To ease notation define the function $K_j : \mathbb{R}^J_{++} \rightarrow \mathbb{R}$

$$K_j(\bar{q}) = -\sum_{c=1}^{C} \eta_c \psi_{cj} \log q_j + \sum_{c=1}^{C} \eta_c \psi_{cj} \log \Bigg(\sum_{k=1}^{K} \psi_{ck} q_k, \Bigg). \tag{26}$$

Insert the optimal choice probabilities, $\bar{q}^*$, into the function g_j (given in Eq. (**??**)) that describes the similarity function, G

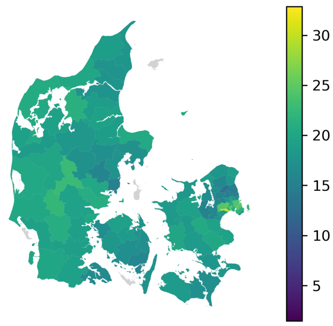
$$\begin{aligned}
g_j(\bar{q}^*) &= \Bigg(1 - \sum_{c=1}^{C} \eta_c \psi_{cj}\Bigg) \log q_j^* + \sum_{c=1}^{C} \eta_c \psi_{cj} \log \Bigg(\sum_{k=1}^{K} \psi_{ck} q_k^* \Bigg), \\
&= \log q_j^* + K_j(\bar{q}^*), \\
&= \Bigg(v_j + \log q_j^* - g_j(\bar{q}^*) - \log \Bigg\{ \sum_{k=1}^{J} \exp \Bigg(v_k + \log q_k^* - g_k\Big(\bar{q}^*\Big)\Bigg)\Bigg\}\Bigg) + K_j(\bar{q}^*), \\
&= \Bigg(v_j + \log q_j^* - \Bigg(\log q_j^* + K_j(\bar{q}^*)\Bigg) - \log \Bigg\{ \sum_{k=1}^{J} \exp \Bigg(v_k + \log q_k^* - g_k\Big(\bar{q}^*\Big)\Bigg)\Bigg\}\Bigg) + K_j(\bar{q}^*), \\
&= v_j - \log \Bigg\{ \sum_{k=1}^{J} \exp \Bigg(v_k + \log q_k^* - g_k\Big(\bar{q}^*\Big)\Bigg)\Bigg\},
\end{aligned}$$

where the expression for the optimal choice probabilities (given by Eq. 2) are inserted in the third line and $g_j(\bar{q}^*)$ is inserted in the fourth line. Insert this expression and the optimal choice probabilities into Eq. (1) that describe the expected utility prior to the realized choice

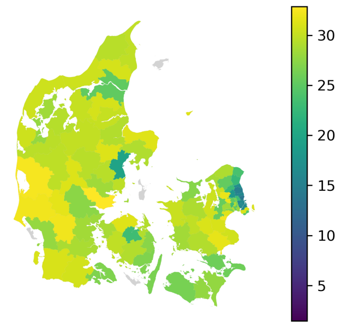
$$\begin{aligned}
EV &= \sum_j v_j \cdot q_j^* - G(\bar{q}^*), \\
&= \sum_j v_j \cdot q_j^* - \sum_j q_j^* \cdot g_j(\bar{q}^*), \\
&= \sum_j v_j \cdot q_j^* - \sum_j q_j^* \cdot \Bigg(v_j - \log \Bigg\{ \sum_{k=1}^{J} \exp \Bigg(v_k + \log q_k^* - g_k\Big(\bar{q}^*\Big)\Bigg)\Bigg\}, \\
&= \sum_j q_j^* \cdot \log \Bigg\{ \sum_{k=1}^{J} \exp \Bigg(v_k + \log q_k^* - g_k\Big(\bar{q}^*\Big)\Bigg)\Bigg\}, \\
&= \log \Bigg\{ \sum_{k=1}^{J} \exp \Bigg(v_k + \log q_k^* - g_k\Big(\bar{q}^*\Big)\Bigg)\Bigg\}.
\end{aligned}$$

This ends the proof.

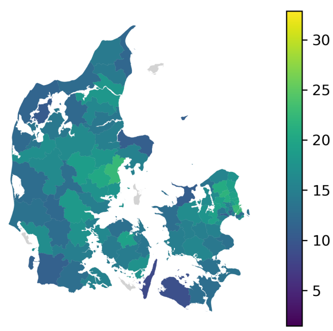
B Appendix



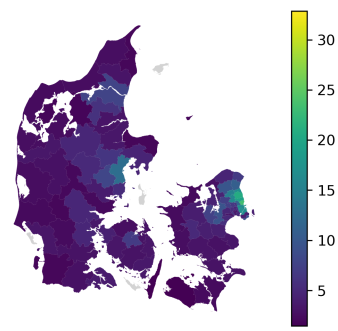
(a) Shares of unskilled.



(b) Shares of skilled.

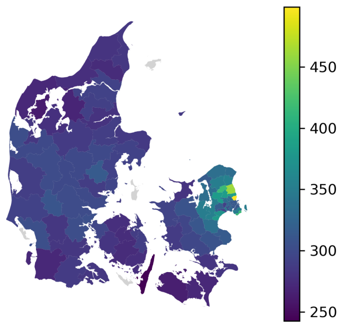


(c) Shares of medium educated.

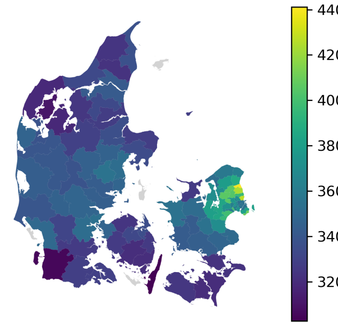


(d) Shares of highly educated.

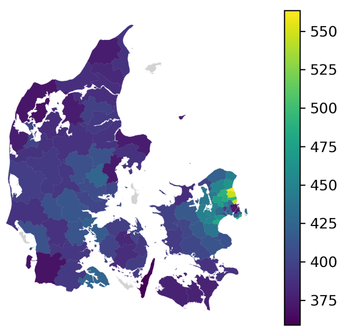
Figure 10: Educational composition of municipalities (percent). The share are based on the residential location of the individuals.



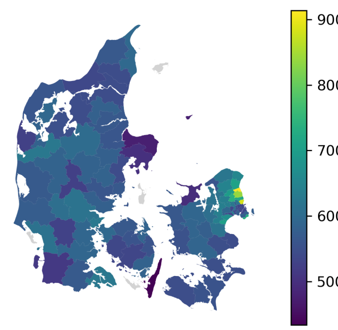
(a) Average wages of unskilled.



(b) Average wages of skilled.



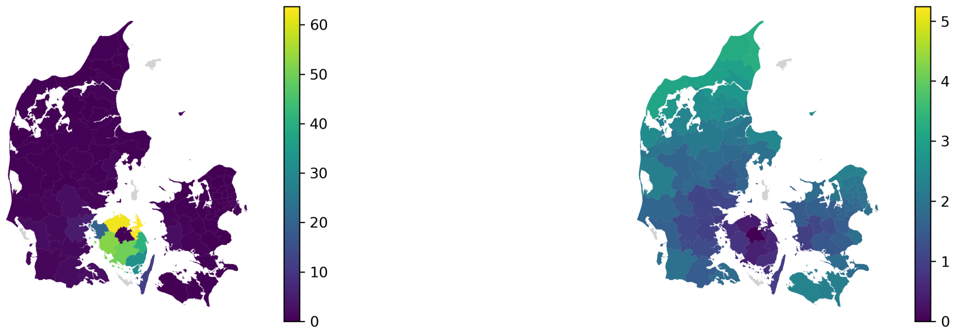
(c) Average wages of medium educated.



(d) Average wages of highly educated.

Figure 11: Average wages across educational groups (DKK 1,000). Average wages are based on the municipality that the individuals work in.

C Appendix



(a) Share of *commuters* that commute to Odense

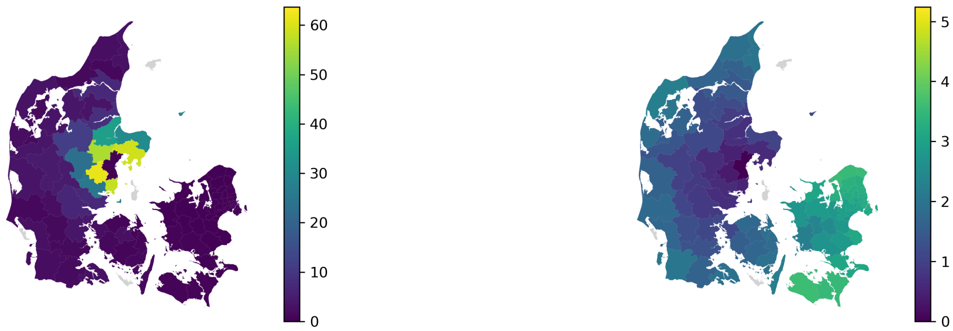
(b) Commuting time to Odense

Figure 12: Share of *commuters* that commute to Odense (percent) and travel time to Odense (hours), 2010-2013. *Commuters* are defined as individuals who work in a different municipality than they reside in. Commuting time is based on data from *Google Maps*.

(a) Share of *movers* that moves from Odense

(b) Share of *movers* that move to Odense

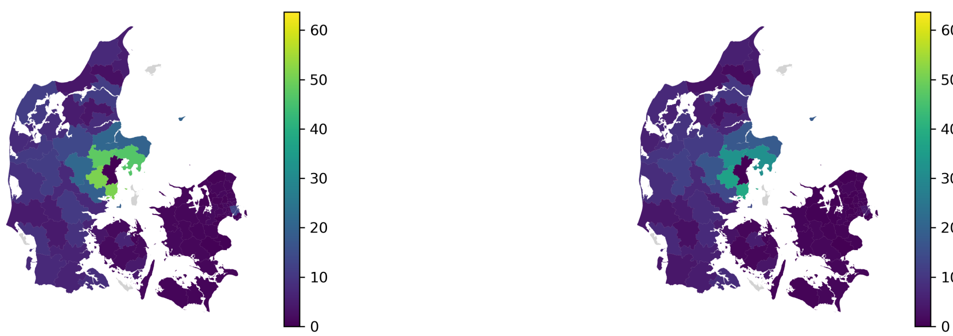
Figure 13: Share of *movers* that move from Odense (percent) and share of *movers* that move to Odense (percent), 2010-2013. *Movers* are defined as households that move to another municipality.



(a) Share of *commuters* that commute to Aarhus

(b) Commuting time to Aarhus

Figure 14: Share of *commuters* that commute to Aarhus (percent) and travel time to Aarhus (hours), 2010-2013. *Commuters* are defined as individuals who work in a different municipality than they reside in. Commuting time is based on data from *Google Maps*.



(a) Share of *movers* that moves from Aarhus

(b) Share of *movers* that move to Aarhus

Figure 15: Share of *movers* that move from Aarhus (percent) and share of *movers* that move to Aarhus (percent), 2010-2013. *Movers* are defined as households that move to another municipality.

(a) Share of *commuters* that commute to Aalborg

(b) Commuting time to Aalborg

Figure 16: Share of *commuters* that commute to Aalborg (percent) and travel time to Aalborg (hours), 2010-2013. *Commuters* are defined as individuals who work in a different municipality than they reside in. Commuting time is based on data from *Google Maps*.

(a) Share of *movers* that moves from Aalborg

(b) Share of *movers* that move to Aalborg

Figure 17: Share of *movers* that move from Aalborg (percent) and share of *movers* that move to Aalborg (percent), 2010-2013. *Movers* are defined as households that move to another municipality.

D Appendix

	Capital*	Zealand	Southern Denmark	Central Denmark	North Denmark
Inhabitants (1,000)	-0.0	-0.0	-0.0	-0.0	-0.0
Inhabitants (percent)	-0.0	-0.0	-0.0	-0.0	-0.0
Avg. sqm. prices (DKK 1,000)	-0.0	-0.0	-0.0	-0.0	-0.0
Avg. sqm. prices (percent)	-0.0	-0.0	-0.0	-0.0	-0.0

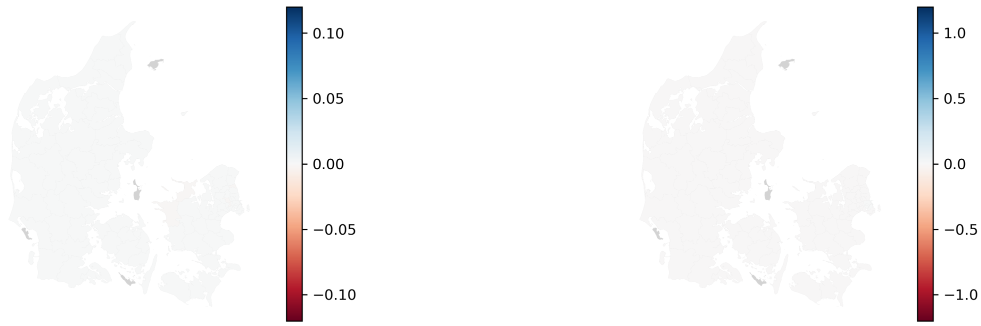
Table 8: Model {1}: Changes in equilibrium outcomes due to a 1 percent increase in housing supply in *Copenhagen* for each of the five regions that constitute the remaining municipalities of Denmark. *Changes in the equilibrium outcomes of *Copenhagen* are not included in this table.

	Capital*	Zealand	Southern Denmark	Central Denmark	North Denmark
Inhabitants (1,000)	-1.8	-0.5	-0.0	-0.0	-0.0
Inhabitants (percent)	-0.3	-0.1	-0.0	-0.0	-0.0
Avg. sqm. prices (DKK 1,000)	-1.0	-0.6	-0.1	-0.0	-0.0
Avg. sqm. prices (percent)	-4.0	-4.7	-0.5	-0.1	-0.0

Table 9: Model {2}: See footnote of Table 8.

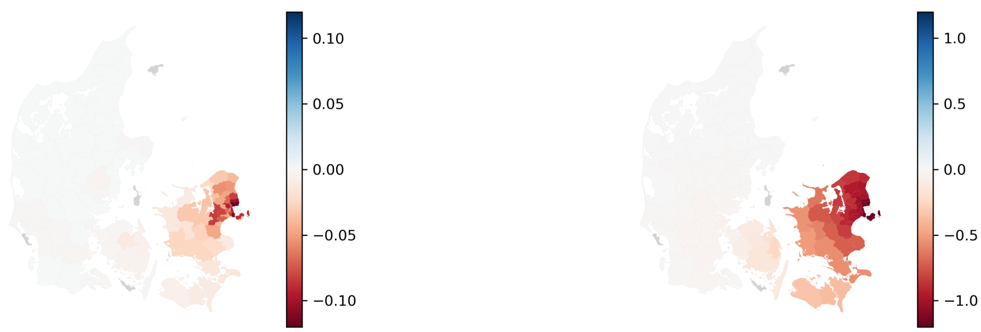
	Capital*	Zealand	Southern Denmark	Central Denmark	North Denmark
Inhabitants (1,000)	-1.2	-0.6	-0.3	-0.2	-0.1
Inhabitants (percent)	-0.2	-0.1	-0.0	-0.0	-0.0
Avg. sqm. prices (DKK 1,000)	-1.1	-0.9	-0.4	-0.3	-0.2
Avg. sqm. prices (percent)	-4.4	-6.2	-3.3	-2.1	-2.1

Table 10: Model {3}: See footnote of Table 8.



(a) Changes in inhabitants. **(b)** Changes in square meter prices.

Figure 18: Model {1}: Changes in inhabitants (1,000) and square meter prices (DKK 1,000). For illustrative reasons, changes in *Copenhagen* are not included in the figures.



(a) Changes in inhabitants. **(b)** Changes in square meter prices.

Figure 19: Model {2}: Changes in inhabitants (1,000) and square meter prices (DKK 1,000). For illustrative reasons, changes in *Copenhagen* are not included in the figures.

E Appendix D

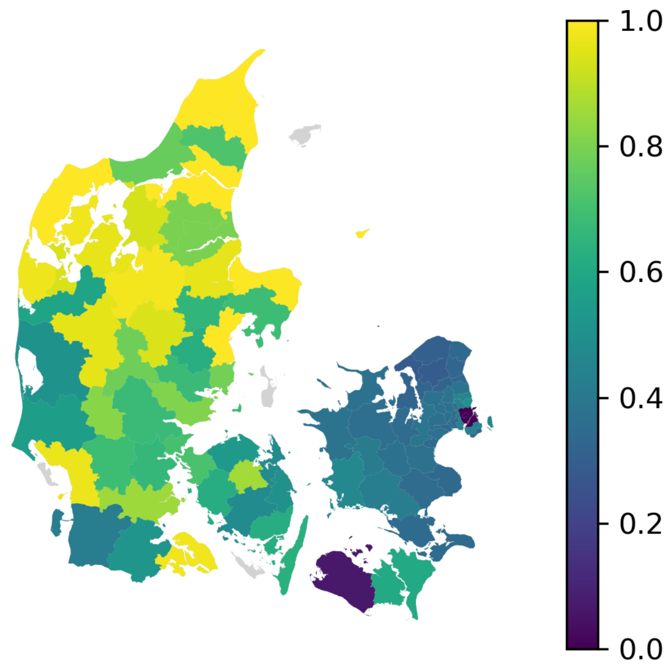


Figure 20: Estimates for $\bar{\omega}_j$.

Table 11 shows the estimated average productivity parameters across municipalities for unskilled, skilled, medium educated and highly educated household members. Across representative municipalities unskilled female household members are used as a reference group. For the representative firm in municipality j, the average productivity for the educational group e is given as

$$\frac{\sum_{t=1}^{T}\sum_{g\in(F,M)}\hat{\pi}^{g}_{jet}\cdot Q_{jt}\cdot q^{g}_{jegt}}{\sum_{t=1}^{T}\sum_{g\in(F,M)}Q_{jt}\cdot q^{g}_{jegt}}.$$

	Unskilled	Skilled	Medium educ.	Highly educ.
Copenhagen	0.1912	0.3903	0.6318	0.5068
Ballerup	0.3607	1.4204	1.0451	0.2290
Brøndby	0.8309	1.5716	0.6800	-0.9891
Dragør	0.1326	1.0281	0.2024	-2.0870
Gentofte	0.0582	0.5020	1.0086	0.7803

	Unskilled	Skilled	Medium educ.	Highly educ.
Gladsaxe	0.3227	1.2049	1.4090	0.6734
Glostrup	0.4598	1.6155	1.0675	-0.0083
Herlev	0.3013	1.5693	1.2706	-0.1894
Albertslund	0.7941	1.6515	0.6269	-1.6281
Hvidovre	0.5912	1.3889	0.7368	-1.0561
Høje-Taastrup	0.4512	1.3821	0.1288	-1.4068
Lyngby-Taarbæk	0.2156	0.8423	1.0910	0.8988
Rødovre	0.5558	1.6286	0.3115	-2.1104
Ishøj	0.4915	1.0947	-0.1477	-2.2373
Tårnby	0.5485	0.7701	-0.7879	-3.3561
Vallensbæk	0.1690	1.0840	0.1453	-1.8355
Furesø	0.2727	1.3592	0.9245	-0.8385
Allerød	0.3560	1.2769	1.1285	-0.2668
Fredensborg	0.1268	0.9127	0.5894	-0.8767
Helsingør	0.1396	0.8764	0.2880	-1.7152
Hillerød	0.2032	1.5447	1.2195	0.1522
Hørsholm	0.0529	0.9397	0.9173	0.1245
Rudersdal	0.2316	0.9374	1.0389	0.5790
Egedal	0.3141	1.5359	0.8114	-1.2627
Frederikssund	0.3075	1.4192	0.3854	-2.2596
Greve	0.4937	1.4425	0.0734	-2.4492
Køge	0.4081	1.5150	0.2680	-1.7618
Halsnæs	0.1725	1.0674	-0.2971	-2.8401
Roskilde	0.2873	1.5676	0.8935	-0.1132
Solrød	0.1353	1.2799	0.1029	-2.1991
Gribskov	0.2116	1.1553	-0.0978	-2.3948
Odsherred	0.1830	1.2226	-0.3600	-3.0657
Holbæk	0.2425	1.3396	0.3026	-2.0838
Faxe	0.2705	1.3529	-0.1142	-2.9937
Kalundborg	0.3090	1.6897	0.4198	-2.0962
Ringsted	0.4968	1.6485	0.3112	-2.3450
Slagelse	0.2312	1.3975	0.2603	-2.2970
Stevns	0.3421	1.3850	-0.2353	-3.1762
Sorø	0.3103	1.7169	0.4470	-1.1323
Lejre	0.2208	1.4892	0.4187	-2.2611

	Unskilled	Skilled	Medium educ.	Highly educ.
Lolland	0.1698	1.3427	-0.3340	-3.4050
Næstved	0.2210	1.4906	0.4784	-1.8754
Guldborgsund	0.2580	1.6448	0.2476	-2.3945
Vordingborg	0.2093	1.2816	0.0058	-2.7181
Middelfart	0.2351	1.5515	0.4323	-2.2451
Assens	0.2355	1.4342	0.2474	-2.9009
Faaborg-Midtfyn	0.3261	1.5225	0.3364	-2.8370
Kerteminde	0.5600	1.8108	0.2596	-3.2373
Nyborg	0.2414	1.3373	0.4695	-2.8730
Odense	0.1771	1.3028	0.7937	-0.9259
Svendborg	0.1091	1.5416	0.7466	-1.7963
Nordfyns	0.2549	1.3749	0.0335	-3.2779
Langeland	0.2779	1.4648	-0.2729	-3.9435
Haderslev	0.3053	1.4598	0.1761	-2.2737
Billund	0.1228	0.9357	-0.3065	-2.5180
Sønderborg	0.1337	1.5095	0.6690	-1.8661
Tønder	0.2057	1.3799	-0.2725	-3.1927
Esbjerg	0.3362	1.5375	0.6254	-2.1719
Varde	0.3205	1.3921	-0.2184	-3.4123
Vejen	0.4554	1.5184	-0.0454	-3.1326
Aabenraa	0.2649	1.5804	0.1630	-2.5545
Fredericia	0.4351	1.5440	0.6769	-1.8001
Horsens	0.2652	1.3034	0.3757	-2.0943
Kolding	0.2764	1.4641	0.5347	-1.8890
Vejle	0.2941	1.3230	0.4577	-1.7726
Herning	0.4617	1.5000	0.4511	-2.3055
Holstebro	0.3423	1.5651	0.4375	-2.2116
Lemvig	0.3762	1.5091	-0.0765	-2.3576
Struer	0.3266	1.1998	0.0849	-2.7293
Syddjurs	0.2779	1.4477	0.2259	-2.2840
Norddjurs	0.3802	1.5739	-0.1057	-2.7426
Favrskov	0.3069	1.5886	0.6950	-2.0316
Odder	0.0556	1.5832	0.4712	-2.0725
Randers	0.2636	1.4832	0.4514	-1.7099
Silkeborg	0.2761	1.6400	0.7824	-1.4976

	Unskilled	Skilled	Medium educ.	Highly educ.
Skanderborg	0.2170	1.3286	0.7462	-2.0116
Aarhus	0.3464	1.3942	1.3133	0.2242
Ikast-Brande	0.2740	1.2855	0.2431	-2.5217
Ringkøbing-Skjern	0.3526	1.5590	-0.0544	-2.9423
Hedensted	0.3883	1.3083	0.1466	-3.4618
Morsø	0.4897	1.5487	-0.2610	-3.3905
Skive	0.3251	1.4381	0.1268	-2.8948
Thisted	0.3337	1.3832	-0.2996	-3.1365
Viborg	0.3465	1.5692	0.5875	-1.1603
Brønderslev	0.3225	1.7239	0.3747	-2.4649
Frederikshavn	0.2449	1.5260	-0.1902	-3.1967
Vesthimmerlands	0.4030	1.3335	-0.1338	-3.2124
Rebild	0.3469	1.5192	0.0660	-2.5715
Mariagerfjord	0.3676	1.5034	0.2243	-2.6544
Jammerbugt	0.3006	1.2205	-0.2475	-3.3939
Aalborg	0.3216	1.5575	0.8391	-0.6382
Hjørring	0.2886	1.5052	0.1942	-2.2817

Table 11: Average estimated productivity parameters. Unskilled female workers are used as reference group in each location. Wages are measured in DKK 1,000,000 in 2018 prices.

This page has been left blank intentionally.

Chapter 3

A dynamic spatial equilibrium model for the housing and labor market

A dynamic spatial equilibrium model for the housing and labor market*

Esben Scriver Andersen†

February 1, 2022

Abstract

This paper proposes a dynamic spatial equilibrium model based on the joint residential and work location decision of heterogeneous households. The model allows square meter prices and wages to be determined in equilibrium. The discrete choices of the households are described by the perturbed utility model, which includes a larger class of models than the random utility model and allows for a rich and meaningful parameterization of the spatial substitution patterns between the locations. As proof of concept, the model is estimated based on high-quality administrative data for the entire population of Danish households from 2010 to 2013. Based on the estimated model, this paper finds strong responses in square meter prices and wages following a counterfactual increase in the supply of housing of 1 percent in *Copenhagen*. For this location the average square meter price and wages decrease by 10.8 and 0.2 percent, respectively.

Keywords: Spatial sorting, residential location, work location, dynamic discrete choice model, equilibrium model, perturbed utility, similarity.

*This research is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 740369). The author thanks Mogens Fosgerau and Bertel Schjerning for comments.

†Department of Economics, University of Copenhagen, Denmark, esben.scriver.andersen@econ.ku.dk.

1 Introduction

Due to large transaction costs, households' location decision is inherently dynamic in nature, as these costs induce households to be forward-looking. However, most studies of the housing market have adopted the static approach, that ignores any forward-looking motive. Further, since at least Roback (1982) it has been well understood that the spatial distribution of wages and housing prices is determined simultaneously in equilibrium. However, the existing literature on dynamic equilibrium models of the housing market has until now ignored this fact by assuming exogenous wages. In contrast, this paper sets up a dynamic spatial equilibrium model where both housing prices and wages are determined in equilibrium.

Unlike the existing literature on dynamic discrete choice models, the proposed model does not describe the discrete choices of the households by the random utility model of McFadden (1973, 1978). Instead, households' choices are described by a wider class of perturbed utility models that allows for a rich and meaningful parameterization of the spatial substitution patterns between the residential locations. A rich parameterization of substitution pattern between the set of residential locations allows the data to determine the substitution patterns, which is important as the housing prices are not only affected by the quality of the locations, but also by the availability of close substitutes. As a simple illustrative example, consider the case where the availability of locations with a view over the ocean is scarce, whereas the availability of locations with view over grass fields is abundant. If these two attributes are close substitutes, then housing prices for locations with an ocean view will be relative low. In contrast, if these two attributes are not close substitutes, then the housing prices for locations with an ocean view will be relative high.[1]

In the perturbed utility model, the decision-maker faces the discrete choice of choosing between a set of alternatives and makes a random decision based on a probability distribution given by the vector of choice probabilities that maximizes the decision-maker's perturbed utility. This perturbed utility consists of a term that is linear in the choice probabilities, and the perturbation function that is concave in the choice probabilities. As the perturbation function penalizes small choice probabilities, it represents taste-for-variety. This model can accommodate more flexible substitution patterns, and can even allow different alternatives to be complements, which the random utility model rules out. Allen and Rehbeck (2019); Fosgerau et al. (2021); Fosgerau and Nielsen (2021) all showed that some perturbed utility models can allow for complementarity. As suggested by Fosgerau and Nielsen (2021), the perturbation function for the residential choice is given by the similarity function, which allows locations with similar characteristics to be closer substitutes. For instance, the simi-

[1] This example was originally proposed by Andersen (2022).

larity function can allow locations close to each other to be closer substitutes than locations far apart. To the best of my knowledge, this paper is the first to set up and solve a dynamic discrete choice problem based on perturbed utility. The solution is given in terms of the expected value function, which can be expressed as a fix point equation, and it is shown that this equation defines a contraction mapping. Hence, there exists a unique solution to the fixed point, which can be obtained through iterations of the fixed point equation.[2]

In the proposed dynamic spatial equilibrium model, forward-looking households and household members sort themselves into the set of available residential and work locations based on housing prices, offered wages and commuting distances. The model is inspired by Carstensen et al. (2020), who set up a dynamic discrete choice model of the joint residential and work location decision. As they include a continuous housing size decision, they are unlike the existing literature on dynamic equilibrium models able to identify housing prices from a set of market clearing conditions. Besides describing the discrete choices by the perturbed utility model, this paper makes two notable extensions to this type of dynamic equilibrium models. First, by introducing an additional set of market-clearing conditions, wages can be determined in equilibrium. Second, this paper relaxes the assumption that households can costlessly adjust their housing size.[3] Hence, in contrast to Carstensen et al. (2020), the households not only face a cost when moving to another residential location, they also face a cost when adjusting their current housing size.[4] Further, to set up estimate a dynamic spatial equilibrium model for both the combined housing and labor market.

Andersen (2022) shows that this has important implications for the adjustment to equilibrium. To see this, note that changes in the demand for housing at a given location can be due to changes along the extensive and the intensive margin. Changes along the extensive margin reflect households that move to another location and thereby increase the demand for housing in that location and reduce the demand for housing in the location they move away from. Changes along the intensive margin reflect households that choose to stay in their current location, but change their demand for housing. Hence, in contrast to changes along the extensive margin, changes along the intensive margin do not have spill-over effects on the demand for housing in the remaining locations. Similarly to Carstensen et al. (2020), Andersen (2022) analyzes a counterfactual increase in the supply of housing in *Copenhagen*, and shows that ignoring adjustment costs in the presence of large moving costs leads the

[2]Until now, the only empirical applications of the perturbed utility model study static decisions. Three notable examples are Fosgerau et al. (2021, 2022); Andersen and Lee (2022), who study the consumption of cereals, route choices and matching between workers and firms.

[3]Andersen (2022) considers similar extensions but uses a static perturbed utility model to describe the discrete choices.

[4]In contrast to Carstensen et al. (2020), this paper ignores any life-cycle motives and any stochasticity in the work location choice is ignored. Hence, in this paper unemployment is assumed to be voluntary.

model to predict that 99 percent of the adjustment in the housing demand is along the intensive margin. As a result, square meter prices adjust only in *Copenhagen*. In contrast, when adjustment costs are included in the model, the demand adjustment is approximately evenly distributed along the intensive and extensive margins. As a relative big fraction of the adjustment in the demand for housing in *Copenhagen* happens along the extensive margin, square meter prices in the remaining municipalities are affected, which is not the case when adjustment costs are ignored.

As a proof of concept, the proposed model is estimated on high-quality administrative register data for the entire population of Danish households from 2010 to 2013. The equilibrium mechanisms of the model are illustrated by analyzing a counterfactual increase of 1 percent in the supply of housing in *Copenhagen*. The counterfactual analysis shows that the square meter prices of *Copenhagen* decrease by DKK 2,700 and DKK 1,300 on average. for the remaining *Capital municipalities*, respectively (this corresponds to a decrease of 10.8 and 4.8 percent, respectively). The analysis shows that 60 and 40 percent of the demand adjustment works through the intensive and extensive margin, respectively. The increased housing supply further decreases the average wages by DKK 700 and DKK 500 (0.2 and 0.1 percent) in *Copenhagen* and the remaining *Capital municipalities* as the supply of labor increases in the *Capital municipalities*, since more households move to this area. In contrast, the average wages for the remaining municipalities increases by DKK 200 (0.1 percent) as the supply of labor is reduced in this area.

Related literature. Spatial equilibrium models date back to Rosen (1974) and Roback (1982), but they were made popular by Bayer et al. (2004, 2007), who combined the discrete choice framework of McFadden (1973, 1978) with a set of market clearing conditions to study residential sorting and estimate households' preferences for neighbors' socio-demographics. More recently, Diamond (2016) and Piyapromdee (2021) studied the skill sorting and immigration in an extended model framework, where workers sort across locations based on wages and housing prices. Even though these papers use data on finely-detailed spatial units, little attention was paid to the spatial linkages between these units. In contrast, Ahlfeldt et al. (2015) and Monte et al. (2018) emphasize the spatial location of the location units, as they model the joint residential and work location decision. Note that this strand of the literature uses a static approach that focuses on long-run equilibrium outcomes as individuals are implicitly assumed to behave myopically and any forward-looking motive is therefore ignored.

However, individuals face large financial and non-financial costs when moving. As a result, individuals rarely move. Due to the presence of moving costs, the location choice is inherently a dynamic choice. However, the estimation of dynamic models comes with substantial computational costs, and often requires very rich data. Hence, empirical applications of dynamic models of the location choice are limited. Kennan and Walker (2011) set up a dynamic discrete choice model based on the empirical framework of Hotz and Miller (1993) to analyze how income differences across spatial units effect the migration decision over the life cycle. Bishop and Murphy (2011) and Bayer et al. (2016) use a similar framework to estimate the willingness to pay for non-market amenities, such as neighborhood air pollution, violent crime, and racial composition. Buchinsky et al. (2014) were the first to set up and estimate a dynamic discrete choice model where the spatial location is taken seriously, as they model the joint residential and work location choices of immigrants and spouses. These papers consider a partial equilibrium model in which housing prices and/or wages are taken to be exogenous. Carstensen et al. (2020) extended the model of Buchinsky et al. (2014) to a general equilibrium model, where square meter prices of housing can be determined in equilibrium. They then evaluate the changes in equilibrium outcomes due to a local increase in the supply of square meters. Andersen (2022) uses a static version, to analyze the spatial substitution patterns of both housing prices and wages of a similar counterfactual change.

Organization of the paper. Section 2 introduces the perturbed utility model that is used to describe the discrete choices of the households. The Danish administrative register data used for estimation is then described in Section 3. Section 4 sets up the economic model for the joint residential and work location decision and specifies the equilibrium concepts of the model and Section 5 sets up the maximum likelihood estimator used to estimate the structural parameters of the model. Section 6 shows the estimation results and the model fit. Section 7 analyzes the changes to the equilibrium outcomes due to a counterfactual increase in the housing supply of the municipalities of *Copenhagen* and Section 8 concludes.

Some notation. Vectors are denoted $\bar{q}$. The probability simplex in the $J-1$ dimensional simplex is denoted $\Delta^J = \{\bar{q} \in \mathbf{R}_+^J : \sum_{j=1}^J q_j = 1\}$. The integers between A and B, where the end points are included, is denoted $\{A \,..\, B\} = \{A, A+1, ..., B\}$.

2 The perturbed utility model

This section builds on Andersen (2022), when introducing the perturbed utility model in the static case. The section then shows how to set up a simple dynamic discrete choice problem based on perturbed utility. In Section 4, these perturbed utility models are used to describe the joint residential and work location decision of the households.

Definition 1 ***(Perturbed utility model)*** *A decision-maker (DM) faces a discrete choice problem with J different alternatives. The discrete choice, $d \in \{1, ..., J\}$, is stochastic with choice probabilities, $\bar{q}^* = (q_1^*, q_2^*, ..., q_J^*)$, given as the solution to the utility maximization problem*

$$\max_{\bar{q} \in \Delta^J} \left\{ \sum_{j=1}^{J} q_j v_j - Y(\bar{q}) \right\}, \tag{1}$$

where $v = (v_1, v_2, ..., v_J)$ is a vector of choice-specific payoffs and $Y : \mathbb{R}_{++}^J \to \mathbb{R}$ is the perturbation function.

If the perturbation function, Y, is strictly convex and satisfies that $\nabla_{\bar{q}} Y \to \infty$ as $q_j \to 0$ for any j, then there exists a unique optimal solution in the interior of the probability simplex. In this case, the perturbation function can be interpreted as representing taste-for-variety as it penalizes small choice probabilities.

Example 1 ***(Shannon entropy)*** *When the perturbation function is the negative Shannon entropy*

$$Y(\bar{q}) = \sum_{j=1}^{J} q_j \log q_j,$$

it can be shown that the solution to the DM's utility maximization problem is the multinomial logit choice probabilities

$$q_k^* = \frac{\exp\left(v_k\right)}{\sum_{j=1}^{J} \exp\left(v_j\right)}, \text{ for } k = 1, 2, ..., J.$$

For the logit choice probabilities, the substitution patterns are governed by the independent of irrelevant alternative property, and the corresponding expected utility of the DM is given by the well-known log-sum

$$EV = \log\left\{ \sum_{j=1}^{J} \exp\left(v_j\right) \right\}.$$

Fosgerau and Nielsen (2021) introduce a class of perturbation functions that provides a way to parameterize the degree of substitutability among groups of alternatives. Definition 2 is based on their work.

Definition 2 ***(Similarity function)*** *Let $\mathcal{C} = \{1, 2, ..., C\}$ be a finite set of characteristics and let $\Psi = \{\psi\}_{c \in C, j \in J}$ be a $C \times J$ matrix with nonnegative entries ψ_{cj} and columns that* *sum to 1,* $\sum_{c=1}^{C} \psi_{cj} = 1$*. A similarity function* $G : \mathbb{R}_{++}^{J} \to \mathbb{R}$ *takes the form*

$$G(\bar{q}) = \sum_{j=1}^{J} q_j \cdot g_j(\bar{q}), \tag{2}$$

$$g_j(\bar{q}) = \left(1 - \sum_{c=1}^{C} \eta_c \psi_{cj}\right) \log q_j + \sum_{c=1}^{C} \eta_c \psi_{cj} \log\left(\sum_{k=1}^{K} \psi_{ck} q_k\right), \tag{3}$$

where $\bar{\eta} = (\eta_1, \eta_2, ..., \eta_C) \in \mathbb{R}^C$ *is a vector of parameters associated with the* C *characteristics and satisfies* $\sum_{c=1}^{C} \max(\eta_c, 0) \cdot \psi_{cj} < 1$ *for all* $j \in \{1 .. J\}$*.*

Fosgerau and Nielsen (2021) show that for $Y(\bar{q}) = G(\bar{q})$ the choice probabilities that maximize the DM's utility are determined by the unique fixed point equation

$$q_k^* = \frac{\exp\left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right)\right)}{\sum_{j=1}^{J} \exp\left(v_j + \log q_j^* - g_j\left(\bar{q}^*\right)\right)}, \text{ for } k = 1, 2, ..., J. \tag{4}$$

As shown in Appendix A the corresponding expected utility can be shown to be given as

$$EV = \log\left\{\sum_{j=1}^{J} \exp\left(v_j + \log q_j^* - g_j\left(\bar{q}^*\right)\right)\right\}. \tag{5}$$

Note that the similarity function reduces to the negative Shannon entropy if $\eta_c = 0$ for all $c \in \{1 .. C\}$, which implies that the optimal strategy is given by the multinomial logit choice probabilities of the logit model. Further, the class of similarity models nests the nested logit and inverse product differentiation logit model of Fosgerau and Nielsen (2021) as special cases.

For each alternative, the parameter ψ_{cj} describes the degree to which alternative j is associated with characteristic c. Options that lie close together in the space of characteristics are considered to be more similar than alternatives far from each other. This implies that higher positive values of η_c result in stronger substitution among similar alternatives relative to the logit model, while conversely negative values result in weaker substitution, and even complementarity when η_c is a sufficiently large negative.

2.1 Dynamic perturbed utility model

This subsection illustrates by example that perturbed utility can be used to describe a dynamic discrete choice problem. Let the DM base its sequence of discrete choices, $\{d_s\}_{s=t}^{\infty}$

on the corresponding sequence of choice probabilities $\{\bar{q}_s\}_{s=t}^{\infty}$ that maximizes its discounted expected perturbed utility over an infinite time of horizon

$$EV_t(d_{t-1}) = \max_{\{\bar{q}_s\}_{s=t}^{\infty} \in \Delta^J} E_t \Bigg[\sum_{s=t}^{\infty} \delta^{s-t} \Bigg(\sum_{j=1}^{J} q_{js} u_{js}(d_{s-1}) - Y(\bar{q}_s) \Bigg) | d_{t-1} \Bigg], \tag{6}$$

where the state variable in period t is assumed to be the choice in the previous period, d_{t-1}. δ is the discount factor and $u_{js}(d_{t-1})$ is the choice-specific instantaneous utility of the DM. Since the perturbed utility is additive across time periods, the maximization problem of Eq. (6) can be written in terms of a Bellman equation

$$\begin{aligned} EV_t(d_{t-1}) &= \max_{\bar{q}_t \in \Delta^J} \Bigg\{ \sum_{j=1}^{J} q_{jt} u_{jt}(d_{t-1}) - Y(\bar{q}_t) + \delta \sum_{j=1}^{J} q_{jt} EV_{t+1}(d_t = j) \Bigg\}, \\ &= \max_{\bar{q}_t \in \Delta^J} \Bigg\{ \sum_{j=1}^{J} q_{jt} v_{jt}(d_{t-1}) - Y(\bar{q}_t) \Bigg\}, \end{aligned}$$

where $EV_{t+1}(d_t = j)$ is the expected value function and $v_{jt}(d_{t-1})$ is the choice specific value of the current period

$$v_{jt}(d_{t-1}) = u_{jt}(d_{t-1}) + \delta EV_{t+1}(d_t = j). \tag{7}$$

This implies that the stated dynamic problem has a representation similar to Definition 1, that defines the perturbed utility model. Hence, for a given vector of choice-specific values the optimal choice probabilities and the corresponding expected utility are given by Eq. (4) and Eq. (5). As the DM has an infinite time horizon, the time subscript can be dropped from EV. By inserting the choice-specific values of Eq. (7) into the formula for the expected utility, the EV can be described by the fixed point

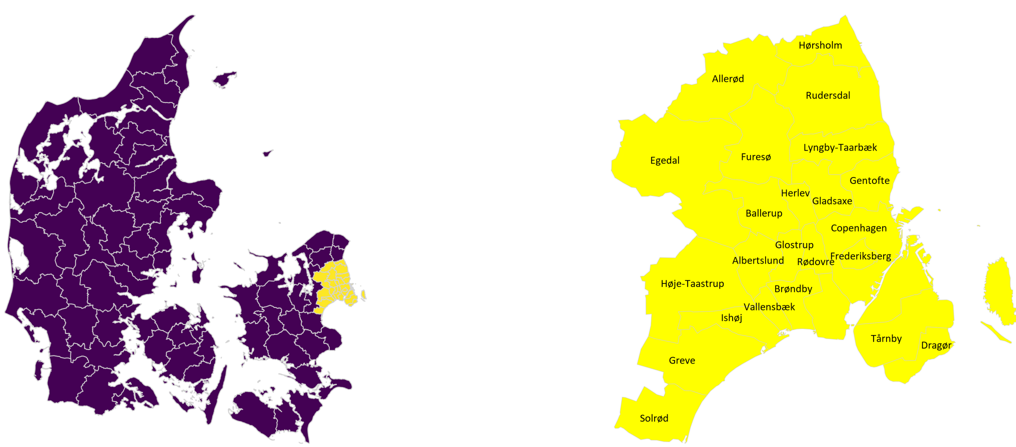
$$EV(d = k) = \log \Bigg\{ \sum_{j=1}^{J} \exp \Bigg(u_j(d = k) + \delta \cdot EV(d = j) + \log q_j^* - y_j\Big(\bar{q}^*\Big) \Bigg) \Bigg\}, \tag{8}$$

for $k = 1, 2, ..., J$. As shown in Appendix B Eq. (8) defines a contraction mapping for EV. This implies that the fixed point algorithm proposed by Rust (1987) can be used to efficiently solve for EV.

3 Data

This section outlines the data sources and the construction of the data set used.[5] The section ends by showing some descriptive statistics for the housing and labor market of the *Capital municipalities* of Denmark.

The geographical units of this paper - the locations where the households can choose to reside and work - are given by the municipalities of Denmark. However, for computational reasons, this paper focuses on the *Capital municipalities*.[6] Denmark consists of 98 municipalities, of which 24 municipalities constitute the *Capital municipalities*, see Figure 1. As *Frederiksberg* is surrounded by *Copenhagen* these two municipalities are treated as a single location throughout this paper, and this location will be referred to as *Copenhagen* in the remaining if this paper. Further, as this paper focuses on the *Capital municipalities*, the remaining 74 municipalities are combined into a single location. Hence, in the proposed model, households can only choose to reside or work in one of the 23 defined *Capital municipalities* or they can choose to reside or work outside the *Capital municipalities*.



(a) Municipalities of Denmark.

(b) Capital municipalities.

Figure 1: Municipalities and capital municipalities of Denmark.

The data are drawn from several administrative registers that cover the entire population

[5]The construction of the data set is similar to Andersen (2022).

[6]This definition of the *Capital municipalities* is based on Statistics Denmark's classification of municipality groups. Statistics Denmark classifies the municipalities into five groups based on accessibility to jobs and the number of inhabitants in the largest city in the municipality. The *Capital municipalities* constitute one of these five groups.

of Denmark. Since all individuals are associated with a unique identifier, each individual can be followed over time, and information from the different registers can be linked. The main data source is the population register, that each year records all individuals who are officially registered with an address in Denmark on January 1st. This register contains information on demographics, address of residency and a household identifier such that all household members can be linked. Statistics Denmark defines households as individuals that reside together at the same address. Households can consist of singles or couples with and without children.

Based on the household's address, it is possible to link information on the dwelling. The key variable *housing size* can be found in the dwelling register for each household. Another key variable is *square meter prices*, which is based on the register for sales of real property that contains information on sales prices for every property sold in a given year.

The income register contains detailed income history on e.g. pension benefits, public transfers and wages. Further, this register also classifies each individual after their attachment to the labor market in a given year based on their income history. Information on workplace is recorded on a monthly basis and as a default the workplace in January is linked to each individual. Based on the the residential address and workplace address in November, Statistics Denmark calculates the commute distance for employed individuals. Finally, the attainment register gathers information about the highest completed education for each individual.

The constructed data set contains all household members 25 years of age or older with an address in Denmark in a given year, and where at least one of the household members had an address in Denmark in the year before. The data set contains 10 million households or 15 million individuals. Table 1 shows some descriptive statistics for the data set. The column *All municipalities* shows the descriptive statistics for the entire data set, whereas the column *Capital municipalities* shows the descriptive statistics for all individuals residing in the *Capital municipalities*.

	All municipalities		Capital municipalities	
Households	10,186,442		2,809,618	
Individuals	15,232,708		3,993,647	
	(1) mean	(2) std.	(3) mean	(4) std.
Wage (DKK 1,000).	234.589	357.171	276.316	424.490
1 if working	0.615	0.487	0.643	0.479
1 if working in Copenhagen	0.089	0.285	0.290	0.454
1 if *commuting**	0.251	0.434	0.385	0.487
1 if *commuting* to Copenhagen*	0.047	0.212	0.129	0.335
1 if unskilled**	0.193	0.394	0.212	0.409
1 if skilled**	0.260	0.439	0.206	0.404
1 if medium educ.**	0.169	0.375	0.193	0.394
1 if highly educ.**	0.070	0.255	0.133	0.339
1 if *outside the labor force*	0.308	0.462	0.257	0.437
1 if couple	0.495	0.500	0.421	0.494
1 if residing in Copenhagen	0.130	0.337	0.473	0.499
1 if moved address	0.100	0.300	0.108	0.310
1 if moved municipality	0.031	0.174	0.044	0.205
Property sales	132,473		40,462	
	(1) mean	(2) std.	(3) mean	(4) std.
Housing size (m^2)	124.550	44.855	112.450	46.943
Sales prices (DKK 1,000)	2,017.987	1,479.334	2,889.966	1,956.326
Sales prices (DKK 1,000/m^2)	16.711	14.706	25.180	8.946

Table 1: Descriptive statistics. The samples includes individuals of age 25 or older from 2010 to 2013. All monetary units are measured in 2018 prices. The top of the table describes the individuals of the samples, whereas the bottom describe property sales in the samples. **Commuting* refer to individuals who work in another municipality than they reside in. **Individuals *outside the labor force* are not included.

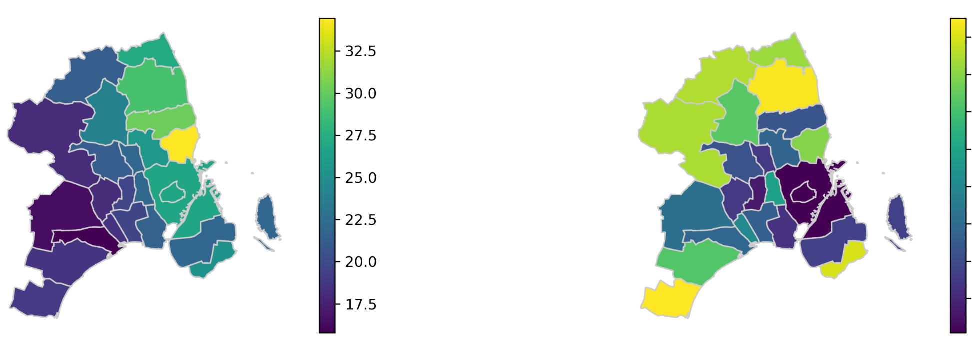
Individuals who do not work and the income register defines as retired based on their labor market attachment are defined as *outside the labor force*.[7] Hence, this group includes non-working individuals who receive public pension, early retirement pay and early retirement pension. This groups constitutes 31 percent of the population. Individuals that are

[7]The definition of individuals *outside the labor force* is based on the variable *BESKST02* that describes each individual's labor market attachment. Unemployed individuals that are not retired are included in the labor force.

not classified as *outside the labor force* are classified into four mutual excluding groups based on their highest attained education level. This classification is based on the International Standard Classification of Education (ISCED) two-digit code. *Unskilled* household members includes individuals with high school completion or less (ISCED code 10-20 or missing); *Skilled* household members includes individuals with vocational education (ISCED code 30); *Medium educated* household members includes individuals with short- or medium-cycle higher education and bachelors (ISCED code 35-60); *Highly educated* household members include individuals with a masters or PhD degree (ISCED code 70-80). *Highly educated* household members is the smallest group. This group constitutes 7 percent of the population and *Medium educated* individuals constitutes around 17 percent of the population. *Skilled* and *Unskilled* individuals constitute roughly 26 and 20 percent of the population, respectively.

For the *Capital municipalities* 64 percent of the population are working and 39 percent commute to another municipality than where they reside. Given that a relatively large fraction of the employed individuals commutes, this favors that the residential and work location choices should be modelled jointly. Further, note that 10 percent of the households changed address and only 2 percent moved to another capital municipality. Hence, any realistic modelling of the residential choice needs to incorporate frictions such as moving cost.

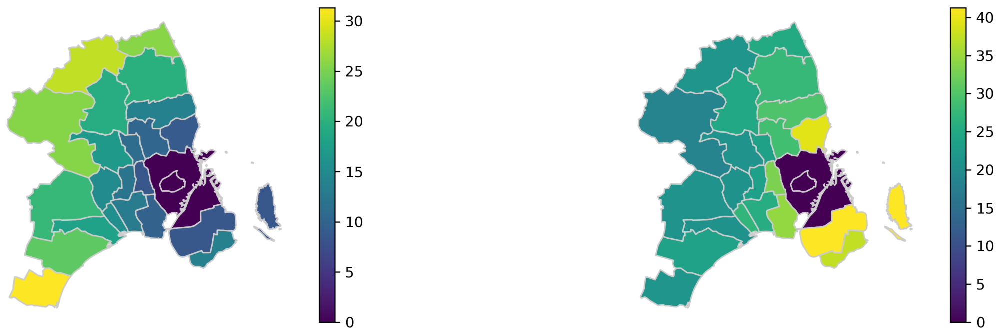
Figure 2 illustrates the average square meter prices of sold properties and housing size for the *Capital municipalities*. The figure does not indicate a clear correlation between square meter prices and housing size. It should be noted that the smallest average housing size is observed in *Copenhagen*, where the square meter prices are relative high.



(a) Average square meter prices (DKK 1,000/m^2). **(b)** Average housing size (m^2)

Figure 2: Average square meter prices of sold properties and average housing size of households, 2010-2013.

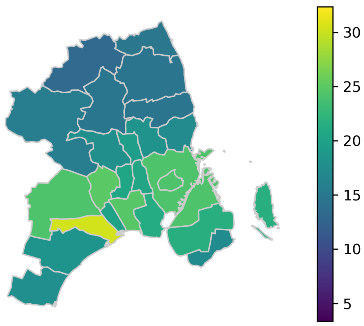
Figure 3 illustrates the average commuting distances to Copenhagen and the shares of employed individuals that commute to Copenhagen. The figures show, as expected, that the shares of commuters are negatively correlated with commuting distance to Copenhagen. A similar pattern recurs for the other *Capital municipalities*.



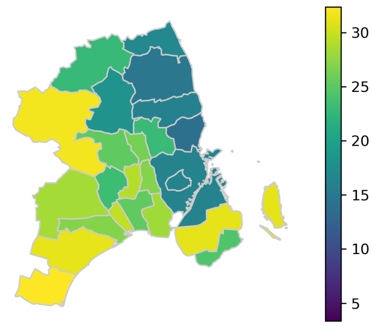
(a) Average commuting distance (km).   **(b)** Share that commutes (percent).

Figure 3: Average commuting distance to *Copenhagen* and share of employed individuals that commute to *Copenhagen*. The latter is based on the municipality the individuals reside in.

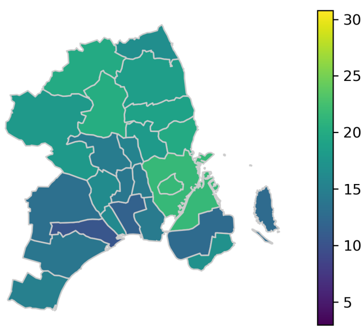
Figure 4 shows the educational composition of the different *Capital municipalities*. The figures show that the percentage share of the different educational groups varies across the municipalities. E.g skilled individuals are to a relative high degree concentrated in the municipalities south-west of Copenhagen, whereas highly educated individuals are more concentrated in Copenhagen and the municipalities north of Copenhagen. The average wages for the different educational groups can be found in Appendix C. These figures show that the wages offered vary over the educational groups and across municipalities. Across educational groups, wages offered are highest for the municipalities north of Copenhagen, where the square meter prices are highest as well. This indicates that educational level is an important source of heterogeneity to include when modeling the residential choice, and square meter prices and wages influence each other.



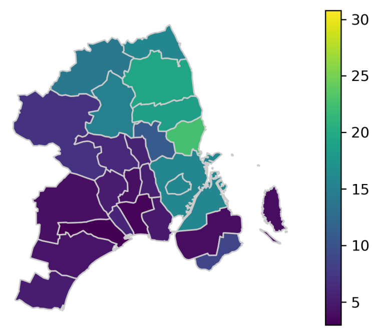
(a) Shares of unskilled (percent).



(b) Shares of skilled (percent).



(c) Shares of medium educated (percent).



(d) Shares of highly educated (percent).

Figure 4: Educational composition of the *Capital municipalities*. The share are based on the municipality the individuals reside in.

4 Model

Based on the joint residential and work location decision of the households, this section sets up a dynamic spatial equilibrium model for the housing and labor market for the *Capital municipalities* of Denmark. In this model, the economy consists of H_t different households at time t and at $N+1$ different geographical locations, that differ with respect to their spatial position, square meter prices of housing, offered wages etc. N refers to the number of *Capital municipalities* and in addition the households can choose to reside and/or work outside this region. The households consist of single females, single males and couples. When facing a discrete choice, it is assumed that household behavior can be described by the perturbed utility model discussed in Section 2.

At each point in time, the individual household chooses its residential location, its housing size and the work location of each of its members. Household members not participating in the labor force receive a fixed income and do not face a work location choice.

As the wealth accumulation motive is ignored in this model, the forward-looking behavior is only induced by the moving costs. The households have static expectations about the future square meter prices and wages and have an infinite time horizon. Finally, it is assumed that in each time period the households make the decisions sequentially in two stages.

First, the household chooses its residential location and housing size based on its previous residential location and housing size choices. If the household decides to move to another residential location or adjust its housing size, it incurs a moving cost. Secondly, each household member chooses its work location based on the chosen residential location choice. When choosing work location, each household member only takes into account the offered wages and the commute distances.

The model is solved by backward induction, such that each household member first solves the work location problem for any given residential choice, and then solves the residential location choice problem, where the member takes into account the expected utility of the work location choice for any chosen residential location choice.

The remainder of this section is organized as follows. Subsection 4.1 introduces the choices, specifies the instantaneous utility function on which the household bases its decisions and derives the optimal housing size. Subsections 4.2 and 4.3 set up and solve the dynamic discrete work and residential location problem, respectively. Subsection 4.4 specifies the perturbation functions and Subsection 4.5 explains the equilibrium concept of the model.

4.1 Instantaneous utility

Let $d_{it}^R \in \{0 \,..\, N+1\}$ denote the residential location choice of household i at time t. For $d_{it}^R = 0$ the household moves out of the *Capital municipalities* and for $d_{it}^R \in \{1 \,..\, N\}$ the household chooses to move to one of the N *Capital municipalities*. In both cases the household chooses a new housing size, $h_{it} \in \mathbb{R}_{++}$, as well. For $d_{it}^R = N+1$ the household stays at its previous chosen location, $n_{it-1} \in \{0 \,..\, N\}$, and keeps its housing size, $h_{it} = h_{it-1}$. The chosen residential location is then denoted by $n_{it} \in \{0 \,..\, N\}$.

Let superscript $g \in \{F, M\}$ denote the gender of the household member and let $d_{it}^g \in \{0 \,..\, N+1\}$ denote the work decision of the female (F) or male (M) household member. For $d_{it}^g = N+1$ the household member chooses the outside option of not working and for $d_{it}^g \in \{0 \,..\, N\}$ the household member chooses to work in one of the $N+1$ geographical locations.

It is assumed that any household member belongs to one of E different educational types. Hence, let $f_i \in \{0 .. E\}$ denote the type of the female member of household i, and let $m_i \in \{1 .. E\}$ denote the type of the male member. For both genders $f_i = m_i = 0$ refers to household members not in the labor force. Further, $c_i = \{S, C\}$ denotes the civil status of the household. For $c_i = S$, the household is characterized as a single female or single male, and for $c_i = C$, the household is characterized as a couple.

In the spirit of Carstensen et al. (2020) the instantaneous choice-specific utility for a couple that moves, $d_{it}^R \in \{0 .. N\}$, is

$$\begin{aligned} u_i(d_{it}^R, n_{it}, h_{it}, d_{it}^F, d_{it}^M) = \alpha_{c_i} \cdot & \left(w_{f_i}^F(d_{it}^F) + w_{m_i}^M(d_{it}^M) - \phi \cdot p_t(n_{it}) \cdot h_{it} \right) \\ & + \beta_{1,c_i} \cdot h_{it} + \tfrac{1}{2} \cdot \beta_2 \cdot h_{it}^2 + \kappa(n_{it}) + \rho_1 + \rho_2 \cdot \mathbb{I}(n_{it} \neq n_{it-1}) \\ & + \lambda^F \cdot dist(d_{it}^F, n_{it}) + \lambda^M \cdot dist(d_{it}^M, n_{it}) \\ & + \pi^F(d_{it}^F) + \pi^M(d_{it}^M), \end{aligned} \tag{9}$$

and the instantaneous choice-specific utility of a couple that stays, $d_{it}^R = N + 1$, is

$$\begin{aligned} u_i(d_{it}^R, n_{it-1}, h_{it-1}, d_{it}^F, d_{it}^M) = \alpha_{c_i} \cdot & \left(w_{f_i}^F(d_{it}^F) + w_{m_i}^M(d_{it}^M) - \phi \cdot p_t(n_{it-1}) \cdot h_{it-1} \right) \\ & + \beta_{1,c_i} \cdot h_{it-1} + \tfrac{1}{2} \cdot \beta_2 \cdot h_{it-1}^2 + \kappa(n_{it-1}) \\ & + \lambda^F \cdot dist(d_{it}^F, n_{it-1}) + \lambda^M \cdot dist(d_{it}^M, n_{it-1}) \\ & + \pi^F(d_{it}^F) + \pi^M(d_{it}^M). \end{aligned} \tag{10}$$

The instantaneous choice-specific utilities are given by similar expressions for households of single females or single males, but the terms related to a household member of the opposite sex drops out of Eq. (9) and (10).

The income of the household is given by $w_{f_i}(d_{it}^F) + w_{m_i}(d_{it}^M)$ and depends on the chosen work location. Household members not in the labor force do not face a work location choice and receive a fixed exogenous income. Let $p_t(n_{it})$ denote the square meter price of the chosen residential location, and let ϕ denote the annual user cost. Hence, $\phi \cdot p_t(n_{it}) \cdot h_{it}$ constitutes the housing expenses of the household. In turn, α_{c_i} can be interpreted as the marginal utility of consumption of the outside good. Utility is quadratic in square meters of housing, and the parameters (β_{1,c_i}, β_2) describe the curvature of this second order polynomial function. The parameter $\kappa(n_{it})$ represents the utility of residing in a given location and $\rho_1 + \rho_2 \cdot \mathbb{I}(n_{it} \neq n_{it-1}) < 0$ represents the financial and non-financial moving costs that for $\rho_2 \neq 0$ vary for households that move within their previously chosen residential location and households that move to another location. The commuting distance between the residential and workplace location is given by $dist(d_{it}^g, n_{it})$, and λ^g represents the marginal costs of increasing the commuting distance for the household member. Household members that choose not to work do not face any commuting costs, as $dist(0, n_{it}) = 0$. Finally, the parameter $\pi^g(d^g_{it})$ represents the (dis-)utility of working in a given location.

Since utility is linear in consumption of the outside good, this implies that the work location decisions of the household members are independent of each other, and the derived utility from this decision only depends on the offered wage and commute distance to the chosen work location. Let subscript $j = n_{it}$ denote the realized residential choice of the household and subscript k denote the work location choice. Then the choice-specific utilities of the female work location problem are given as

$$
\begin{aligned}
u^F_{ijkt} &= \alpha_{c_i} \cdot w^F_{f_ikt} + \lambda^F \cdot dist(j,k) + \pi^F_k, \quad && \text{for } k = 0, ..., N, \\
u^F_{ijkt} &= \alpha_{c_i} \cdot w^F_{f_ikt}, && \text{for } k = N+1,
\end{aligned}
\tag{11}
$$

where $w^F_{f_ikt}$ is the offered wage to the female workers of type f_i in location $k = 0, ..., N$ and $w^F_{f_i,N+1,t}$ is some exogenous benefit payments the female household member receives when unemployed.

The choice-specific utilities of the male work location problem have similar specifications

$$
\begin{aligned}
u^M_{ijkt} &= \alpha_{c_i} \cdot w^M_{m_ikt} + \lambda^M \cdot dist(j,k) + \pi^M_k, \quad && \text{for } k = 0, ..., N, \\
u^M_{ijkt} &= \alpha_{c_i} \cdot w^M_{m_ikt}, && \text{for } k = N+1.
\end{aligned}
\tag{12}
$$

Let subscript $r = n_{it-1}$ denote the previously chosen residential location of the household and let subscript s denote the residential location choice. The expected choice-specific utilities of the residential location problem are

$$
\begin{aligned}
u^R_{irst}(h_{it-1}) &= -\alpha_{c_i} \cdot \phi \cdot p_{st} \cdot h_{it} + \beta_{1,c_i} \cdot h_{it} + \tfrac{1}{2} \cdot \beta_2 \cdot h^2_{it} \\
&\quad + \kappa_s + \rho_1 + \rho_2 \cdot \mathbb{I}(n_{it} \neq n_{it-1}) + EU^F_{ist} + EU^M_{ist}, \qquad && \text{for } s = 0, ..., N, \\
u^R_{irst}(h_{it-1}) &= -\alpha_{c_i} \cdot \phi \cdot p_{rt} \cdot h_{it-1} + \beta_{1,c_i} \cdot h_{it-1} + \tfrac{1}{2} \cdot \beta_2 \cdot h^2_{it-1} \\
&\quad + \kappa_r + EU^F_{irt} + EU^M_{irt}, && \text{for } s = N+1,
\end{aligned}
\tag{13}
$$

where (EU^F_{ist}, EU^M_{ist}) are the expected utilities of the female and male household members' work location choices if residential location s is chosen, and (EU^F_{irt}, EU^M_{irt}) are the expected utilities from staying at the current residential location, r. A closed form expression for expected utility of the work location decision is derived in next subsection.

Finally, as the household has static expectations, wealth accumulation is ignored and utility is linear in consumption of the outside good, and the resulting housing size problem is static. Hence, the optimal housing size - when moving - can be found from the first order condition

$$\frac{\partial u_{irst}^{R}(h_{it-1})}{\partial h_{it}} = 0 \Leftrightarrow h_i^*(p_{st}) = -\frac{\beta_{1,c_i}}{\beta_2} + \frac{\alpha_{c_i}}{\beta_2} \cdot \phi \cdot p_{st}. \tag{14}$$

This implies for $\alpha_{c_i} > 0$ and $\beta_2 < 0$ that optimal housing size decreases linearly with the square meter price of the chosen location. Hence, households of the same type will choose the same housing size if they move to the same location.

4.2 Work location problem

This subsection sets up and solves the work location choice problem of the female and male household member. Conditioned on the chosen residential location the work location choice can be seen as a static choice, as the household members can freely change their work location and any wealth accumulation motive is ignored.

The individual household member chooses its work location based on the vector of choice probabilities, $\bar{q}_{ijt}^{g*} = \left(q_{ij0t}^{g*}, ..., q_{ijN+1t}^{g*}\right)^{\top} \in \Delta^{N+2}$, that maximizes the member's perturbed utility

$$\bar{q}_{ijt}^{g*} = \underset{\bar{q}_{ijt}^{g} \in \Delta^{N+2}}{\arg\max} \left\{ \sum_{k=0}^{N+1} q_{ijkt}^{g} \cdot u_{ijkt}^{g} - Y^{g}\left(\bar{q}_{ijt}^{g}\right) \right\}, \tag{15}$$

where the choice-specific utility, u_{ijkt}^{g}, of working in location k, when residing in location j, is given by Eq. (11) or (12) for the female or male household members. In order to reduce the computational complexity of solving the model, the perturbation function, $Y^{g} : \mathbb{R}_{++}^{N+2} \rightarrow \mathbb{R}$, equals the negative Shannon entropy. As a result, the optimal choice probabilities are given by the logit choice probabilities

$$q_{ijkt}^{g*} = \frac{\exp\left(u_{ijkt}^{g}\right)}{\sum_{l=0}^{N} \exp\left(u_{ijlt}^{g}\right)}, \text{ for } k = 0, ..., N+1. \tag{16}$$

For a household member residing in location j, the expected perturbed utility from the work location decision is given by the logsum

$$EU_{ijt}^{g} = \log\left\{ \sum_{k=0}^{N+1} \exp\left(u_{ijkt}^{g}\right) \right\}. \tag{17}$$

Hence, the choice-specific utilities of the residential choice problem can be evaluated by inserting Eq. (17) into Eq. (13) of last subsection.

4.3 Residential location problem

This subsection sets up and solves the residential location problem. Due to the presence of moving costs, the realization of the residential location choice affects the future welfare of the household. Hence, the residential choice problem is a dynamic discrete choice problem, where the state variables at time t are given by the previously chosen residential location and housing size, (n_{it-1}, h_{it-1}). Let subscript $r = n_{it-1}$ denote the previously chosen residential location.

The individual household chooses its residential location based on the vector of choice probabilities, $\bar{q}^{R*}_{irt}(h_{it-1}) = \left(q^{R*}_{ir0t}(h_{it-1}), ..., q^{R*}_{ir,N+1,t}(h_{it-1})\right)^\top \in \Delta^{N+2}$, that maximizes its perturbed utility

$$\bar{q}^{R*}_{irt}(h_{it-1}) = \underset{\bar{q}^R_{irt} \in \Delta^{N+2}}{\arg\max} \left\{ \sum_{s=0}^{N+1} q^R_{irst} \cdot v^R_{irst}(h_{it-1}) - Y^R\Big(\bar{q}^R_{irt}\Big) \right\}, \tag{18}$$

where $v^R_{irst}(h_{it-1})$ is the choice-specific value from residing at location s when previously residing in location r with housing size h_{it-1}. The perturbation function, $Y^R : \mathbb{R}^{N+2}_{++} \to \mathbb{R}$, is specified in terms of the similarity function in order to allow for a flexible parameterization of the substitution patterns between the set of residential locations

$$Y^R\Big(\bar{q}^R_{irt}\Big) = \sum_{r=0}^{N+1} q^R_{irst} \cdot y^R_r\Big(\bar{q}^R_{irt}\Big), \tag{19}$$

where the exact specification of y^R_r is given in Subsection 4.4. In turn, the choice probabilities are given by the fixed point

$$q^{R*}_{irst}(h_{it-1}) = \frac{\exp\left(v^R_{irst}(h_{it-1}) + \log q^{R*}_{irst}(h_{it-1}) - y^R_s\Big(\bar{q}^{R*}_{irt}(h_{it-1})\Big)\right)}{\sum_{u=0}^{N+1} \exp\left(v^R_{irut}(h_{it-1}) + \log q^{R*}_{irut}(h_{it-1}) - y^R_u\Big(\bar{q}^{R*}_{irt}(h_{it-1})\Big)\right)}, \tag{20}$$

for $s = 0, ..., N+1$.[8] The corresponding expected utility of the residential decision is

$$EV^R_{irt}(h_{it-1}) = \log\left\{ \sum_{s=0}^{N+1} \exp\left(v^R_{irst}(h_{t-1}) + \log q^{R*}_{irst}(h_{t-1}) - y^R_s(\bar{q}^{R*}_{irt}(h_{t-1}))\right) \right\}, \tag{21}$$

[8] Fosgerau and Nielsen (2021) provide a contraction mapping that can be used to obtain the optimal choice probabilities.

for $r = 0, ..., N$. The choice-specific value function, $v^R_{irst}(h_{it-1})$, is given by the sum of the instantaneous choice-specific utility and the discounted expected value function

$$
\begin{aligned}
v^R_{irst}(h_{it-1}) &= u^R_{irst}(h_{it-1}) + \delta EV^R_{ist+1}(h^*_i(p_{st})), \quad \text{for } s = 0, ..., N, \\
v^R_{irst}(h_{it-1}) &= u^R_{irst}(h_{it-1}) + \delta EV^R_{irt+1}(h_{it-1}), \qquad \text{for } s = N+1,
\end{aligned}
\tag{22}
$$

where $EV^R_{ist+1}(h^*_i(p_{rt}))$ is the expected value from residing in location s with the optimal housing size and $EV^R_{irt+1}(h_{it-1})$ is the expected value from staying in the previously chosen location and keeping the housing size.

Since the household has an infinite time horizon, the time subscript on the expected value function, EV^R_{ir}, can be dropped. Similar to Subsection 2.1, a contraction mapping for EV^R_{ir} can be derived by plugging the expression for the choice-specific value function, Eq. (22), into the formula for the expected-choice specific value function, Eq. (21). However, due to the additional choice and state variable, the expression is more complicated. The obtained contraction mapping is given by Eq. (23) and (24) stated below, where Eq. (23) is the expected value of residing in location $r = 0, ..., N$ with the optimal housing size

$$
\begin{aligned}
EV^R_{ir}(h^*_i(p_{rt})) = \log \Bigg\{ &\sum_{s=0}^{N} \exp \Bigg(u^R_{irst}(h^*_i(p_{st})) \\
&+ \delta EV^R_{is}(h^*_i(p_{st})) + \log q^{R*}_{irst}(h^*_i(p_{st})) - y^R_s\Big(\bar{q}^{R*}_{irt}(h^*_i(p_{st}))\Big)\Bigg) \\
&+ \exp \Bigg(u^R_{ir,N+1,t}(h^*_i(p_{rt})) \\
&+ \delta EV^R_{ir}(h^*_i(p_{rt})) + \log q^{R*}_{ir,N+1,t}(h^*_i(p_{rt})) - y^R_{N+1}\Big(\bar{q}^{R*}_{irt}(h^*_i(p_{rt}))\Big)\Bigg)\Bigg\},
\end{aligned}
\tag{23}
$$

and Eq. (24) is the expected value of staying in the previously chosen location and keep the housing size

$$
\begin{aligned}
EV^R_{i,n_{it-1}}(h_{it-1}) = \log \Bigg\{ &\sum_{s=0}^{N} \exp \Bigg(u^R_{i,n_{it-1},st}(h_{it-1}) \\
&+ \delta EV^R_{is}(h_{it-1}) + \log q^R_{i,n_{it-1},st}(h_{it-1}) - y^R_s\Big(\bar{q}^R_{i,n_{it-1},t}(h_{it-1})\Big)\Bigg) \\
&+ \exp \Bigg(u^R_{i,n_{it-1},N+1,t}(h_{it-1}) \\
&+ \delta EV^R_{i,n_{it-1}}(h_{it-1}) + \log q^R_{i,n_{it-1},N+1,t}(h_{it-1}) - y^R_{N+1}\Big(\bar{q}^R_{i,n_{it-1},t}(h_{it-1})\Big)\Bigg)\Bigg\}.
\end{aligned}
\tag{24}
$$

In principle, the residential choice problem must be solved for any value of h_{it-1}. However, as this is computationally costly, it is assumed that h_{it-1} is common for all households of the same type that reside in the same location. This is consistent with the long-run equilibrium of the model.[9] As a result, the number of state and choice combinations of the joint residential and work location decision is proportional to $2(N+1)(N+2)$. As the number of dimensions is quadratic in the number of geographical locations, the model is subject to a curse of dimensionality. Note that the combination of a static solution to the work location decision and the sequential decision process reduces the number of state and choice combinations. If the household had to choose its residential and work location simultaneously rather than sequentially, the number of combinations would be proportional to $(N+1)(N+2)^2$.

4.4 Perturbation function

In order to break down the IIA property, the perturbation function of the residential choice problem is given by the similarity function, $Y^R : \mathbb{R}^{N+2}_{++} \to \mathbb{R}$. Recall that Y^R is described by the vector function $y^R_s : \mathbb{R}^{N+2}_{++} \to \mathbb{R}$ (see Eq. 19). It is assumed that $\{\psi_{cj}\}_{c\in\{0\,..\,2\},j\in\{0\,..\,N\}}$ is a $3 \times N$ matrix that describes to which degree the $N+1$ residential locations belong to three different nests, and the associated parameters are identical across nests, $\eta_0 = \eta_1 = \eta_2 = \eta$. Hence, the function y^R_s for $s = 0, ..., N$ is given as

$$y^R_s\left(\bar{q}^R_{irt}\right) = (1-\eta)\log q^R_{irst} + \eta \sum_{c=0}^{2} \psi_{cs} \log\left(q^R_{ir,N+1,t} \cdot \psi_{cn_{it-1}} + \sum_{u=0}^{N} q^R_{irut} \cdot \psi_{cu}\right),$$

and for $s = N+1$ this function is given as

$$y^R_{N+1}\left(\bar{q}^R_{irt}\right) = (1-\eta)\log q^R_{ir,N+1,t} + \eta \sum_{c=0}^{2} \psi_{cn_{it-1}} \log\left(q^R_{ir,N+1,t} \cdot \psi_{cn_{it-1}} + \sum_{u=0}^{N} q^R_{irut} \cdot \psi_{cu}\right).$$

It is further assumed that the location outside the *Capital municipalities* belongs to its own nest, $c = 0$, which implies that $\psi_{00} = 1$ and $\psi_{0j} = 0$ for $j = 1, ..., N$. Further, $\psi_{01} = \psi_{02} = 0$ as the outside location does not belong to the remaining two nests.

In contrast, the *Capital municipalities* can, to a varying degree, belong to the remaining nests. Define the vector of structural nesting parameters, $\bar{\omega} = (\omega_1, \omega_2, ..., \omega_N)$, where $\omega_j \in [0,1]$ for all $j = 1, ..., N$. The remaining elements of ψ_{cj} are given in terms of this vector. More specifically $\psi_{1j} = \omega_j$ and $\psi_{2j} = 1 - \omega_j$ for $j = 1, ..., N$.

[9] Andersen (2022) shows that it is important to relax the assumption that a household can costlessly adjust their housing size. The proposed model is a simple way to introduce adjustment costs, and implies that the proposed model should be interpreted as a representative agent model.

$(\eta, \bar{\omega})$ are all parameters to be estimated. This rich parameterization of the function y_s^R is chosen in order to allow the substitution patterns to be determined by the data instead of a predetermined functional form.

$\bar{\omega}$ describes to which degree the residential locations belong to the nests, and η describes the degree of substitution inside the two nests relative to the logit model. For $\eta = 0$, the function reduces to $y_s^R\left(\bar{q}_{irt}^R\right) = \log q_{irst}^R$ and substitution is described by the IIA property. For $\eta > 0$, locations with similar sized ω_s are closer substitutes than for the logit model. In contrast for $\eta < 0$, locations with similar sized ω_s are less substitutes than for the logit model. For $\eta > 0$, the model is closely related to the generalized nested logit model by Wen and Koppelman (2001). Further, for $\eta > 0$, the similarity model reduces to the nested logit model if ω_s equals 0 or 1 for $s = 1, ..., N$, as each location in this case only belongs to either the first or the second nest.

4.5 Equilibrium

In equilibrium, the demand for housing must equal the supply of housing across all locations, and the supply of labor must equal the demand for labor across all locations and types of household members in the labor force. This paper takes a short-run perspective and assumes a fixed supply of housing and a fixed demand for labor. Thus, the longer run dynamics, where the supply of housing and the demand for labor adjust to changes in square meter prices and wages, are ignored.

All households face the same vector of square meter prices, $\bar{p}_t = (p_{0t}, ..., p_{Nt})^\top$, but the different types of household members face different offered wages. Hence, the household member of gender g and type e faces the vector of wages $\bar{w}_{et}^g = (w_{e0t}^g, ..., w_{eNt}^g)^\top$. For later use, define the stacked matrix $W_t = (\bar{w}_{1t}^F, ..., \bar{w}_{E,t}^F, \bar{w}_{1t}^M, ..., \bar{w}_{E,t}^M)$, that contains all the vectors of wages.

The demand for housing in location $s = 0, ..., N$ is given in terms of the housing sizes and the choice probabilities of the residential location problem

$$D_{st}^R(\bar{p}_t, W_t) = \sum_{i=1}^{H_t} \left\{ q_{irst}^{R*} \cdot h_i^*(p_{st}) + q_{ir,N+1,t}^{R*} \cdot h_{it-1} \cdot \mathbb{1}(n_{it-1} = s) \right\}. \tag{25}$$

Since utility is increasing in consumption of the outside good, households prefer to reside in less expensive locations. In turn, the demand for housing is decreasing in the square meter prices. In contrast, the presence of commuting costs implies that the demand for housing is increasing in the offered wages of that location, as households prefer to reside close to locations that offer higher wages.

The supply of female or male labor of type $e = 1, .., E$ in location $s = 0, ..., N$ is given in terms of the choice probabilities of the residential and work location problems

$$S^F_{set}(\bar{p}_t, W_t) = \sum_{i=1}^{H} \left(q^{R*}_{ijst} + q^{R*}_{ij,N+1,t} \cdot \mathbb{1}(n_{it-1} = s) \right) \cdot q^{F*}_{ijst} \cdot \mathbb{1}(f_i = e), \tag{26}$$

$$S^M_{set}(\bar{p}_t, W_t) = \sum_{i=1}^{H} \left(q^{R*}_{ijst} + q^{R*}_{ij,N+1,t} \cdot \mathbb{1}(n_{it-1} = s) \right) \cdot q^{M*}_{ijst} \cdot \mathbb{1}(m_i = e). \tag{27}$$

The supply of labor at a given location is therefore decreasing in square meter prices and increasing in the offered wages, since households prefer to reside close to where they work and receive high wages.

As the supply of square meters, S^R_{st}, and the demand for female and male workers of each type, (D^F_{set}, D^M_{set}), are fixed across locations, the square meter prices and wages, $(\bar{p}_t, W_t)$, must adjust such that demand equals supply for type $e = 1, ..., E$ and location $s = 0, ..., N$

$$\begin{aligned} D^R_{st}(\bar{p}^*_t, W^*_t) &= S^R_{st}, \\ D^F_{set} &= S^F_{set}(\bar{p}^*_t, W^*_t), \\ D^M_{set} &= S^M_{set}(\bar{p}^*_t, W^*_t). \end{aligned} \tag{28}$$

Given the decentralized nature of the housing and labor market, square meter prices and wages are assumed to adjust in order to clear the markets. If e.g. the supply of housing exceeds the demand in a given location, square meter prices must decrease in order for the demand to increase. As prices fall, more households will be willing to move to this particular location, which in turn leads to an increase in the supply of workers in that location. As a result, wages will also have to decrease in order to clear the labor market. In contrast, if the supply of labor exceeds the demand in a given location, wages of that location must decrease. As a consequence the demand for housing in that location will decrease, which leads to a decrease in the square meter price of that location.

5 Estimation

This section describes the estimation procedure used to estimate the structural parameters of the proposed model and builds on Andersen (2022). As is standard in the literature, the discount factor, δ, is not estimated but set to 0.95 and the user cost parameter, ϕ, is set to 0.05.[10] The remaining structural parameters to be estimated are denoted by the vector $\theta =$

[10] Kennan and Walker (2011); Bishop and Murphy (2011); Bayer et al. (2016); Carstensen et al. (2020) all set the discount factor to 0.95.

$(\alpha_S, \alpha_C, \beta_{1,S}, \beta_{1,C}, \beta_2, \bar{\kappa}, \rho_1, \rho_2, \lambda^F, \lambda^M, \bar{\pi}^F, \bar{\pi}^M, \eta, \bar{\omega})^\top$. In order to reduce the computational complexity of the estimation procedure, the structural parameters are estimated in two steps, as proposed by Carstensen et al. (2020).

Based on Eq. (14) that describes the optimal housing size the reduced form parameters $\gamma_{1,c_i} = -\beta_{1,c_i}/\beta_2$ and $\gamma_{2,c_i} = \alpha_{c_i}/\beta_2$ can be estimated in a first step based on the following equation by simple OLS

$$\widetilde{h}_{it} = \sum_{k \in \{S,C\}} \left\{ \gamma_{1,k} + \gamma_{2,k} \cdot \phi \cdot \widetilde{p}_{it} \right\} \mathbb{1}(c_i = k) + \varepsilon_{it}, \text{ for } i = 1, ..., H_t, \tag{29}$$

and for $t = 1, ..., T$. The variable $\widetilde{h}_{it}$ is the observed housing size, $\widetilde{p}_{it}$ is the observed square meter price and ε_{it} is assumed to be a random measurement error. Recall that $c_i \in \{S, C\}$ indicates whether household i consists of a single or a couple.

Given the reduced form estimates, $(\hat{\gamma}_{1,S}, \hat{\gamma}_{1,C}, \hat{\gamma}_{2,S}, \hat{\gamma}_{2,C})$, the corresponding structural parameters $(\alpha_S, \alpha_C, \beta_{1,S}, \beta_{1,C})$ can be expressed in terms of β_2

$$\begin{aligned} \beta_{1,c_i} &= -\hat{\gamma}_{1,c_i} \cdot \beta_2, \quad && \text{for } c_i = S, C, \\ \alpha_{c_i} &= \hat{\gamma}_{2,c_i} \cdot \beta_2, \quad && \text{for } c_i = S, C. \end{aligned} \tag{30}$$

As a result, these structural parameters can be concentrated out of the likelihood function, such that the dimensionality of the maximum likelihood problem solved in the second step of the estimation procedure is reduced. The likelihood contribution of household i - that previously resided in location n_{it-1} with housing size h_{it-1} - can be written as the product of the choice probabilities of the household's residential location decision and the choice probabilities of the work location decision by the household members in the labor force

$$Pr(d_{it}^R, d_{it}^F, d_{it}^M | n_{it-1}, h_{it-1}, \theta) = q_{i,n_{it-1},d_{it}^R,t}^{R*}(h_{it-1}, \theta) \prod_g q_{i,n_{it},d_{it}^g,t}^{g*}(\theta), \tag{31}$$

where the product is over the individual household members of household i. The maximum likelihood estimator is the vector $\hat{\theta}$ that maximizes the log-likelihood function, which due to Eq. (31) can be expressed by the sum of the log-likelihood function of the residential choices, $L_R(\theta)$, and the log-likelihood function of the work location choices, $L_W(\theta)$,

$$\hat{\theta} = \arg\max_{\theta} \; L(\theta) = L_R(\theta) + L_W(\theta), \tag{32}$$

respectively given as

$$L_R(\theta) = \sum_{t=1}^{T} \sum_{i=1}^{H_t} \log q^{R*}_{i,n_{it-1},d^R_{it},t}(h_{it-1},\theta),$$
$$L_W(\theta) = \sum_{t=1}^{T} \sum_{i=1}^{H_t} \sum_{g} \log q^{g*}_{i,n_{it},d^g_{it},t}(\theta). \tag{33}$$

This second step of the estimation procedure is based on observed choices, average square meter prices across locations and time, average wages for the different educational groups across locations and time and average commuting distance for all combinations of residential and workplace location.

6 Results

This section presents the estimation results of the first and second step of the estimation procedure described in Section 5. For comparison, the model is estimated both when the perturbation function for the residential choice problem is specified in terms of the negative Shannon entropy and in terms of the similarity function. Recall that when the perturbation function is specified in terms of the negative Shannon entropy, the choice probabilities reduce to the logit choice probabilities that suffer from the IIA property. In contrast, the choice probabilities based on the similarity function allow for more complex substitution patterns, as it allows similar residential locations to be closer substitutes. For a comparison of the two estimated models, the estimated monetary values of different attributes are calculated afterwards. The section ends by illustrating the fit of the preferred model.

6.1 Estimation Results

Table 2 shows the reduced-form estimates of the first step of the estimation procedure. This estimation step is based on observed housing sizes and observed square meter prices for 132,000 property sales across all the municipalities of Denmark in the period 2010 to 2013.

The table shows that all estimates are statistically significant at a 99% significance level and the estimates have the expected sign. However, the R^2 is relatively low, $R^2 = 0.063$. Hence, the model can only explain 6 percent of the variation in housing sizes of the sold properties. The estimated constants $(\hat{\gamma}_{1,S}, \hat{\gamma}_{1,C})$ are both positive, and since $\hat{\gamma}_{1,S} < \hat{\gamma}_{1,C}$ couples have a higher baseline demand for square meters. Further, the price coefficients $(\hat{\gamma}_{2,S}, \hat{\gamma}_{2,C})$ are both negative, which implies that the demand for square meters is decreasing in square meter prices as expected. Since $\hat{\gamma}_{2,S} < \hat{\gamma}_{2,C}$, the demand of couples is less price sensitive compared to the demand of singles. The estimates imply that the demand for housing is higher for couples compared to singles, for any given square meter price.

	Reduced form estimates:		
	(1) coef.	(2) s.e.	(3) t-stat.
$\hat{\gamma}_{1,S}$	1.374	0.005	305.130
$\hat{\gamma}_{1,C}$	1.481	0.004	334.640
$\hat{\gamma}_{2,S}$	-2.372	0.050	-47.439
$\hat{\gamma}_{2,C}$	-2.107	0.050	-42.294

Table 2: Reduced form parameters. Estimation is based on 132,473 property sales. All monetary units are measured in 1,000,000 DKK. Housing size is measured in 100 m^2. Standard errors are based on the sum of squared residuals.

The structural parameters are estimated in the second step by maximum likelihood based on the choices of 10 million households from 2010 to 2013. Table 3 shows the estimated key parameters for the two different specifications of the perturbation function. The estimated model, based on the Shannon entropy, is referred to as the *Logit model* and the estimated model, based on the similarity function, is referred to as the *Similarity model*. For both models, all the parameters have the expected signs and are significant on a 99 significance level. However, the table shows that the likelihood is higher for the *Similarity model* compared to the *Logit model*. This is primarily driven by a better fit of the residential location choices.[11] It should be noted that the *Similarity model* includes 23 additional parameters. A comparison of the AIC scores of the two models clearly favors the *Similarity model*, since $P = \exp((32,912,326 - 33,030,294)/2) \approx 0$. It should be noted that the estimation of the *Similarity model* turned out to be sensitive to the initial values. Hence, the reported estimates for the *Similarity model* are obtained from a sequence of different starting values.

[11] As the null hypothesis, $\omega_j = 0$, is on the boundary of the parameter set, a likelihood ratio test is not applicable, as the test-statistic is not χ^2-distributed.

	Structural estimates:	
	(1) Logit model	(2) Similarity model
$\hat{\alpha}_S$	3.9979	5.1505
	(0.0004)	(0.0005)
$\hat{\alpha}_C$	3.5522	4.5764
	(0.0003)	(0.0004)
$\hat{\beta}_{1,S}$	2.3167	2.9847
	(0.0001)	(0.0002)
$\hat{\beta}_{1,C}$	2.4959	3.2154
	(0.0002)	(0.0002)
$\hat{\beta}_2$	-1.6856	-2.1716
	(0.0001)	(0.0001)
$\hat{\rho}_1$	-2.3566	-2.7925
	(0.0001)	(0.0002)
$\hat{\rho}_2$	-3.8714	-4.0418
	(0.0002)	(0.0002)
$\hat{\pi}_F$	-5.2375	-5.6423
	(0.0001)	(0.0001)
$\hat{\pi}_M$	-4.3712	-4.5636
	(0.0001)	(0.0001)
$\hat{\eta}$	0	-0.1828
	(-)	(0.0001)
$L(\hat{\theta})$	-16,515,094	-16,456,087
$L_R(\hat{\theta})$	-4,094,294	-4,053,706
$L_J(\hat{\theta})$	-12,420,800	-12,402,381
Parameters	53	76
AIC	33,030,294	32,912,326

Table 3: Structural parameters are estimated by maximum likelihood based on the choices of 10,186,442 households. Commuting distance is measured in 100 km. All monetary units are measured in 1,000,000 DKK. Housing size is measured in 100 m^2.

The applied specification of the similarity function allows the degree to which the *Capital municipalities* belongs to the two nests to be estimated instead of being pre-specified. Figure 5 illustrates the estimates of $\bar{\omega}$ that describes the degree to which the different municipalities belong to the first nest. To ensure identification, the nesting parameter characterizing *Copenhagen* is set to zero, $\omega_1 = 0$. For all the municipalities, the corresponding parameter estimate is either 1 or 0. The figure shows that the municipalities closest to *Copenhagen* belong to the first nest, and the municipalities in the outskirts of the region belong to the second nest. Hence, the municipalities closest to *Copenhagen* are in the same nest and are relatively close substitutes compared to the municipalities in the outskirt. Note that, as $\eta < 0$, the model differs from the nested logit model, as the residential locations are substitutes to a lesser degree.

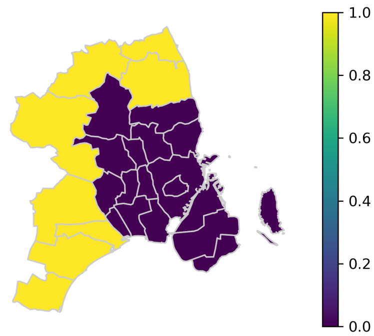


Figure 5: Estimated similarity parameters, $\hat{\hat{\omega}}$.

6.2 Estimated values of attributes

Since the utility scale differs for the two estimated models, the estimated structural parameters are not comparable. However, as utility is assumed to be linear in consumption of the outside good, the implied utilities can for comparison be transformed into monetary units. Table 4 shows the estimated average annual value of different attributes of the households' choices. As estimation is based on gross wages, the reported estimates reflect pre-tax values. As the average tax rate in Denmark is approximately 50 %, the estimated value can be divided by two get a rough estimate for the net values.

The table shows that the estimated average value of housing is DKK 3,700 per square meter for both models. The identical estimates reflect that the monetary value of housing only depends on the reduced form ratios - which are identical for both models - and not the structural parameter.[12] As the calculated average user costs are DKK 900 per square meter, the estimated average values of housing are relative large. Further, for both models the estimated average monetary costs of moving are very large. E.g. for moving residential municipality the estimated costs are DKK 80,800 for the *Logit model* and DKK 68,800 for the *Similarity model*. A general result of the literature on dynamic discrete choice models

[12] Recall that utility of housing is $\beta_{1,c_i} \cdot h_{it} + \beta_{2,c_i} \cdot h_{it}^2$. Hence, the monetary value of housing equals $-\gamma_{1,c_i} \cdot \gamma_{2,c_i}^{-1} \cdot h_{it} + \gamma_{2,c_i}^{-1} \cdot h_{it}^2$.

is that the estimated switching costs are upward biased, unless all relevant unobserved heterogeneity is included in the model, see e.g. Honoré and Kyriazidou (2000).[13] Hence, it is not surprising that the estimated moving costs are lower for the *Similarity model*, as the additional parameters capture some of this unobserved heterogeneity.

Further, Table 4 shows that the estimated marginal costs of commuting an additional kilometer to the work location for females and males are respectively DKK 12,000 and DKK 9,700. These costs include financial costs, travel time and discomfort. The estimates indicate that men are more willing to commute than women, as men on average only need to receive an additional annual income of DKK 9,700 for commuting an additional kilometer, whereas women require DKK 12,000. Note that the cost of commuting is on average 24 percent higher for females than males. This is in line with the existing literature. E.g. Le Barbanchon et al. (2020) find that females value commute around 20 percent more than males.[14] The large estimates for commuting costs are due to unobserved heterogeneity in the work location choice. As the household members are assumed to freely choose their work location, this parameters are primarily identified by differences in commuting distances and differences in wages. However, if household members cannot freely choose their work location they will tend to search for job where the probability of finding a job is highest, and where the cost of searching is lowest. Hence, if commuting distances and wages are correlated with these unobserved variables, the estimated commuting costs will be biased.

Compared to the existing literature, the estimated commuting costs are very high. In proportion to the average yearly income, females or males have to be compensated by 3.6 or 2.6 percent of yearly income in order to commute an additional kilometer. In contrast, based on administrative data on Danish firms that relocate, Mulalic et al. (2014) find that a 1 km increase in commuting distance induces a wage increase of only 0.15 percent three years after the relocation.

In order to interpret the implications of the estimated commuting costs, the average hourly cost of commuting is calculated by assuming 224 annual work days and a commuting speed of 33.8 km per hour.[15] Based on these rough calculations, the average hourly costs of commuting for females and males are DKK 450 and DKK 350 per hour post-tax, respectively. For a comparison Van Ommeren and Fosgerau (2009) find that, on average, workers' marginal costs of one hour of commuting are about EUR 17 (roughly DKK 160 in 2018 prices) and Small et al. (2005) estimate the value of time to be USD 21.46 per hour (roughly DKK 230 in

[13] For that reason, most of the existing literature does not report the estimated value of moving.

[14] Van Ommeren and Fosgerau (2009) also find that the marginal costs of commuting are larger for women than for men, but the difference is insignificant.

[15] According to the Danish commuting survey the average speed when commuting in the *Capital region* is 33.8 kilometers per hour.

2018 prices). Hence, the reported estimates of this paper are roughly 50-180 percent higher than the reported estimates in the existing literature.

	Average annual monetary value:	
	(1) Logit model	(2) Similarity model
Housing (DKK 1,000/m^2)	3.7	3.7
Moving address (DKK 1,000)	-30.5	-28.1
Moving municipality (DKK 1,000)	-80.8	-68.8
Females commuting (DKK 1,000 /km)	-14.3	-12.0
Males commuting (DKK 1,000 /km)	-12.0	-9.7

Table 4: All monetary units are measured in 2018 prices. The average annual value of commuting is conditioned on working. The average annual monetary value of moving is conditioned on moving and is transformed into an infinite annuity based on a 5 percent interest rate for comparison.

Figure 6 shows for the *similarity model* the estimated average monetary values of residing in the different *capital municipalities*. The estimated monetary values are measured relative to *Copenhagen*. The figure shows that the monetary values are positive for the municipalities north of *Copenhagen*, where square meter prices are high. E.g. the monetary value of residing in *Rudersdal* is DKK 22,000 annually. In contrast, the monetary values are negative for the municipalities south-west of *Copenhagen*. E.g. the monetary cost of residing in *Høje-Taastrup* is DKK 15,000 annually.

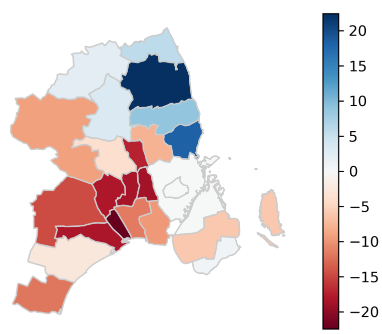


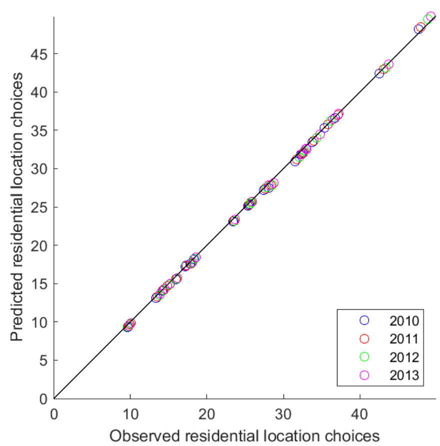
Figure 6: Average estimated annual monetary value of residing in the *Capital municipalities* (DKK 1,000). Estimates are based on $\hat{\hat{\kappa}}$ of the *Similarity model*.

6.3 Model fit

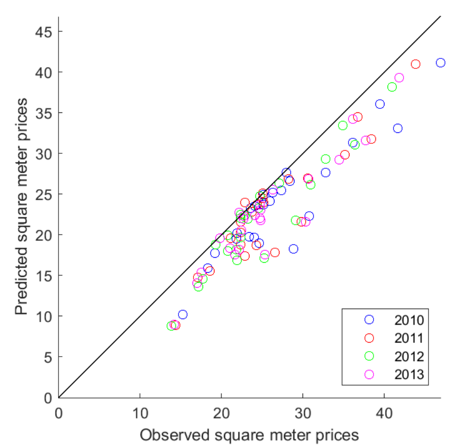
This subsection illustrates the model fit of the *Similarity model* by comparing the model's predictions with the observed outcomes. Recall that the observed average square meter prices and observed average wages are used when estimating the model. However, given the estimated structural parameters the model-consistent predictions for the different equilibrium outcomes can be solved by applying the equilibrium concept of Subsection 4.5.

Figure 7a illustrates the observed and predicted number of inhabitants in the different municipalities. As the figure shows, the predicted numbers of inhabitants match the observed numbers across the municipalities closely. This is not surprising, since most households do not move and the model includes moving costs (ρ_1, ρ_2).

In contrast, the model generally underestimates the square meter prices, as illustrated by Figure 7b. Equilibrium square meter prices reflect the marginal buyer's willingness to pay, and this is a more difficult equilibrium outcome to predict, as small differences between supply and demand for square meters can lead to large changes in square meter prices.



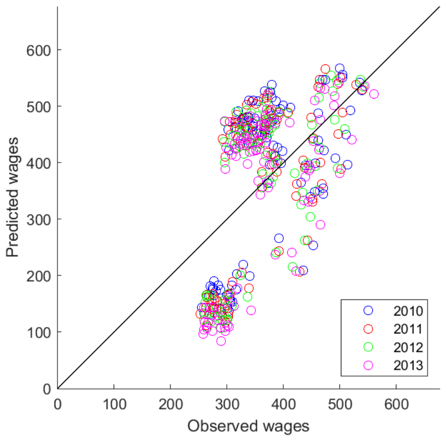
(a) Observed and predicted number of inhabitants



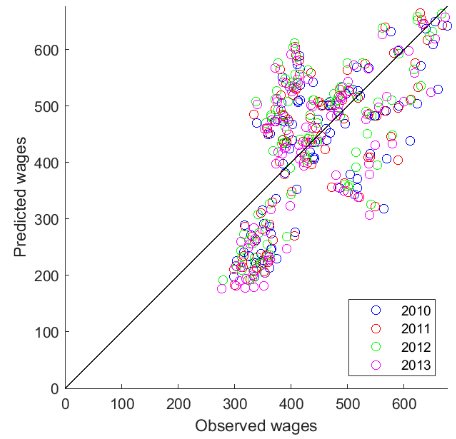
(b) Observed and predicted square meter prices

Figure 7: Observed and predicted inhabitants (1,000) and square meter prices (DKK 1,000). For illustrative purpose the number of inhabitants in Copenhagen is not included in panel (a).

Figure 8 illustrates the observed and predicted wages for females and males across municipalities. The figures show that the model is worse at predicting wages than predicting the square meter prices. Further, the figures show that the predicted wages are more spread than the observed wages. This is especially the case for female workers, where the observed wages are more compactly distributed. The relative poor fit of the observed wages is because the model is relatively bad at predicting the work location choice based on the observed wages. Hence, wages have to adjust a lot in order to clear the labor market. As the labor and housing markets are interlinked, this implies that the square meter prices have to adjust. Hence, the fit of the square meter prices would be better if wages were taken as exogenous.



(a) Observed and predicted wages for females.



(b) Observed and predicted wages for males.

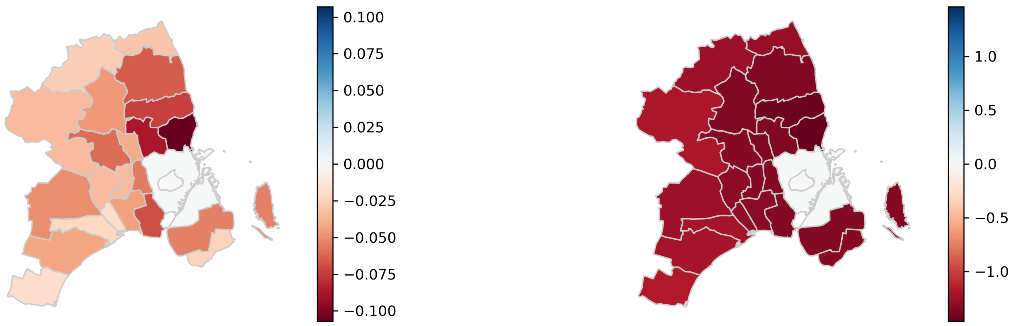
Figure 8: Observed and predicted wages (DKK 1,000).

7 Counterfactual study

This section studies the adjustments of the equilibrium outcomes of the estimated *Similarity model* due to a counterfactual increase of 1 percent in the supply of square meters in *Copenhagen*.

Before turning to the analysis of the equilibrium adjustment, note that changes in the demand for square meters can be due to changes along the extensive and the intensive margin. Changes along the extensive margin reflect households that move from one municipality to another, and thereby decrease the demand for square meters in the municipality they move away from and increase the demand in the municipality they move to. In contrast, changes along the intensive margin reflect households that move inside the municipality in which they already reside, and thereby change their demand for square meters in that municipality. Hence, changes along the intensive margin do not have spill-over effects on the demand for square meters in any other municipality.

Figure 9 illustrates the changes in the number of inhabitants and changes in the square meter prices across the *Capital municipalities* due to the counterfactual increase in the supply of housing in *Copenhagen*. The figures show that the number of inhabitants decreases across all other municipalities than *Copenhagen* and the square meter prices decrease across all *Capital municipalities*. However, as shown later, the expected decreases in the square meter prices are largest for *Copenhagen*.

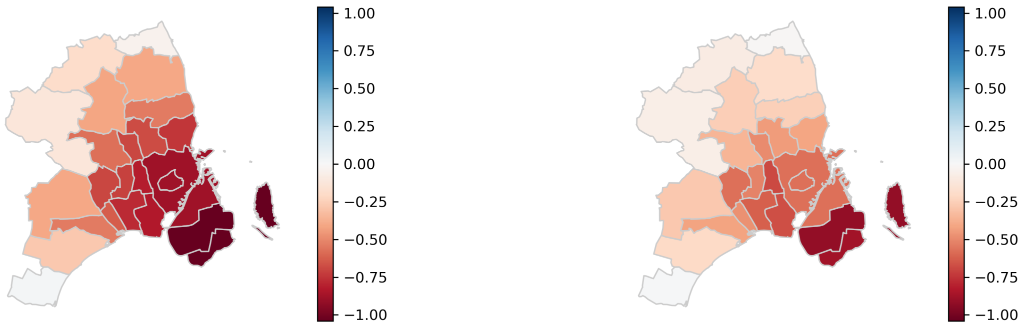


(a) Changes in inhabitants (1,000). **(b)** Changes in square meter prices (DKK 1,000).

Figure 9: Changes in square meter prices and inhabitants. For illustrative purpose the changes in inhabitants of Copenhagen and Frederiksberg is ignored in panel (a)

The changes in the equilibrium outcomes can be summarized as follows. The counterfactual increase in the supply of square meters in *Copenhagen* leads to an excess supply of square meters in this location. In order to equate demand and supply, the square meter prices have to decrease for this location. As some of the increased demand propagates through the extensive margin, this leads to excess supply of square meters in the remaining municipalities as well. In turn the square meter prices decrease for all municipalities. Figure 9b shows that the square meter prices of the municipalities that belong to the same nest as *Copenhagen* are more affected by the counterfactual increase in supply, since these are closer substitutes.

As more households move to *Copenhagen*, this leads to excess supply of labor in *Copenhagen* and the municipalities closest to this location. Hence, wages will decrease for these municipalities. For municipalities sufficiently far from *Copenhagen*, this will in contrast lead to excess demand for labor. As a result, wages will increase for these municipalities. This is illustrated by Figure 10, that shows the counterfactual changes in the average wages for females and males. As the estimated costs of commuting are largest for female, the effects on wages are largest for this group.



(a) Changes in average wages for females **(b)** Changes in average wages for males

Figure 10: Changes in average wages (DKK 1,000).

Table 5 summarizes the counterfactual changes to different equilibrium outcomes. The table shows that the counterfactual increase in the housing supply of 1 percent implies that the number of inhabitants in *Copenhagen* increases by 0.6 percent, corresponding to 2,600 individuals. In contrast, the number of inhabitants decreases by 1,000 for the remaining *Capital municipalities* and by 1,600 for the outside location. Further, the counterfactual increase in the supply of housing leads to a large decrease in the square meter prices in the *Copenhagen*. For this location, the square meter prices decrease by DKK 2,700, corresponding to a decrease of 10.8 percent. For the remaining areas, the average square meter prices decrease by DKK 1,300 and DKK 700, corresponding to a decrease of 4.8 percent and 5.2 percent, respectively. The large decrease in average square meter prices is a result of the large estimates of the moving cost parameters, as prices have to decrease a lot in order for households to move or adjust their housing size.

	Copenhagen and Frederiksberg	Other Capital municipalities	Outside Capital municipalities
Inhabitants (1,000)	2.6	-1.0	-1.6
Inhabitants (percent)	0.6	-0.2	-0.1
Avg. sqm. prices (DKK 1,000)	-2.7	-1.3	-0.7
Avg. sqm. prices (percent)	-10.8	-4.8	-5.2
Avg. wages (DKK 1,000)	-0.7	-0.5	0.2
Avg. wages (percent)	-0.2	-0.1	0.1

Table 5: Changes in equilibrium outcomes due to a 1 percent increase in housing supply in *Copenhagen*.

The counterfactual change in the supply of housing also affects the labor market. Due to the higher supply of labor in the *Capital municipalities*, the average wages decrease by DKK 700 and DKK 500 in *Copenhagen* and in the remaining *Capital municipalities*, respectively. In contrast, the average wages increase by DKK 200 in the outside location due to a lower supply of labor in this location.

Finally, Table 6 decomposes the adjustment in the demand for square meters in *Copenhagen* into adjustments along the extensive and intensive margins. The table shows that 60.4 percent of the increase in the demand for square meters in *Copenhagen* is along the intensive margin and the remaining 39.6 percent is along the extensive margin. This result illustrates that when households cannot costlessly adjust their demand for housing, a substantial fraction of the adjustment will propagate along the extensive margin. However, as it is less costly to move inside a municipality, a larger fraction of the adjustment is along the intensive margin. In contrast, if households can costlessly adjust their housing demand, adjustment will mostly propagate along the intensive margin, and as a result the increased supply of square meters in one location will in this case have a much lower spill-over effect on the square meter prices of the other locations.

	Adjustment in demand
Intensive margin (percent)	60.4
Extensive margin (percent)	39.6

Table 6: Decomposition of the adjustment of the demand for square meters in *Copenhagen* into changes along the extensive and intensive margin.

8 Conclusion

This paper is the first paper to propose a dynamic discrete choice model in which the behavior of the decision-makers is described by the perturbed utility model. Based on the joint residential and workplace location decision of heterogeneous households, this paper proposes an empirical spatial equilibrium model for the housing and labor market of the *Capital municipalities* of Denmark, where square meter prices and wages are determined in equilibrium, such that markets clear.

As a proof of concept, the structural parameters of the proposed model are estimated in two steps by maximum likelihood based on the observed choices of the entire population of Denmark from 2010 to 2013. Based on the estimate, a counterfactual analysis is conducted of equilibrium adjustment of the model due to a 1 percent increase in the housing supply of *Copenhagen and Frederiksberg*. This analysis shows that the equilibrium square meter prices and wages respond strongly to the counterfactual change in the supply of housing.

I now highlight three directions for future work. The first direction is to include additional dynamics into the model. One problem of the proposed model is the maintained assumption of static expectations with respect to square meter prices and wages. This leads initially prices and wages to respond too strongly. Incorporating forward-looking expectations into the model would eliminate this issue. In addition, the model should include financial wealth accumulation similar to Bayer et al. (2016), as a household may anticipate selling their house at some point in the future and thus explicitly consider the expected evolution of housing prices when deciding where and when to purchase or sell a house.

As household members tend to stay at their current job, a second direction for future research is to incorporate frictions into the work location decision. This could be modelled by introducing job-change costs into the work location decision similar to the moving costs in the residential location decision. However, this will greatly increase the state space of the model, as most types of household consist of two household members. For these households, a state variable that describes their previous work location has to be included for each household member. However, this will increase the realism of the model, as it will induce households to move to locations close to their current work location.

Finally, in the presented model the supply of housing and demand for labor are assumed to be exogenous given in the short run. However, even in the short run the supply side for housing and demand side for labor should be allowed to adjust to price and wage changes. As adjustments in these quantities most realistic are costly a dynamic model is the obviously starting point to analyze responses of demand and supply to changes in house prices and wages. By allowing the supply of housing and demand for labor to adjust due to changes in house prices and wages will most likely reduce the obtained counterfactual equilibrium changes in house prices and wages, as the equilibrium quantities now will adjust and therefor house prices and wages do not have to absorb the entire equilibrium adjustment.

The dynamic discrete choice model based on the perturbed utility is open to many other economic applications, where forward looking behaviour and a flexible parameterization of the substitution patterns are considered to be important. As the existing literature on static perturbed utility models has shown, alternatives are even allowed to be complements, which is ruled out in the random utility model.

References

G. M. Ahlfeldt, S. J. Redding, D. M. Sturm, and N. Wolf. The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189, 2015. ISSN 0012-9682.

R. Allen and J. Rehbeck. Identification With Additively Separable Heterogeneity. *Econometrica*, 87(3):1021–1054, May 2019.

E. S. Andersen. A perturbed spatial equilibrium model. *Working paper*, 2022.

E. S. Andersen and Y. J. Lee. Multidimensional Matching and Labor Market Complementarity. *Working paper*, 2022.

P. Bayer, R. McMillan, and K. Rueben. An equilibrium model of sorting in an urban housing market. *NBER Working Paper Series*, 2004.

P. Bayer, F. Ferreira, and R. McMillan. A unified framework for measuring preferences for schools and neighborhoods. *The Journal of political economy*, 115(4):588–638, 2007. ISSN 0022-3808.

P. Bayer, R. McMillan, A. Murphy, and C. Timmins. A dynamic model of demand for houses and neighborhoods. *Econometrica*, 84(3):893–942, 2016. ISSN 0012-9682.

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

M. Buchinsky, C. Gotlibovski, and O. Lifshitz. Residential Location, Work Location, and Labor Market Outcomes of Immigrants in Israel. *Econometrica*, 82(3):995–1054, May 2014.

C. L. Carstensen, F. Iskhakov, M. Juul-Hansen, J. Rust, and B. Schjerning. A dynamic equilibrium model of commuting, residential and work location choices. *Working paper*, 2020.

R. Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524, March 2016.

M. Fosgerau and N. Nielsen. Similarity, perturbed utility and discrete choice. *Working paper*, 2021.

M. Fosgerau, J. Monardo, and A. de Palma. The Inverse Product Differentiation Logit Model. *SSRN working Papers*, 2021.

M. Fosgerau, M. Paulsen, and T. K. Rasmussen. A perturbed utility route choice model. *Transportation Research Part C: Emerging Technologies*, 136:103514, 2022.

B. E. Honoré and E. Kyriazidou. Panel data discrete choice models with lagged dependent variables. *Econometrica*, 68(4):839–874, 2000.

V. J. Hotz and R. A. Miller. Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529, 1993. ISSN 00346527, 1467937X.

J. Kennan and J. R. Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251, 2011. ISSN 0012-9682.

T. Le Barbanchon, R. Rathelot, and A. Roulet. Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, 136(1):381–426, 2020.

D. McFadden. Conditional logit analysis of qualitative choice behaviour. *Frontiers in Econometrics*, pages 105–142, 1973.

D. McFadden. Conditional modeling the choice of residential location. *in eds. Karlquist, A., et al., Spatial Interaction Theory and Planning Models, Elsevier North-Holland, New York*, 1978.

F. Monte, S. J. Redding, and E. Rossi-Hansberg. Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–90, December 2018.

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

S. Piyapromdee. The impact of immigration on wages, internal migration, and welfare. *Review of Economic Studies*, 88(1):406–453, 2021.

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

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

J. Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5), 1987.

K. A. Small, C. Winston, and J. Yan. Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4):1367–1382, 2005.

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

C.-H. Wen and F. S. Koppelman. The generalized nested logit model. *Transportation Research Part B: Methodological*, 35(7):627–641, 2001.

A Appendix

To ease notation define the function $K_j : \mathbb{R}^J_{++} \rightarrow \mathbb{R}$

$$K_j(\bar{q}) = -\sum_{c=1}^{C} \eta_c \psi_{cj} \log q_j + \sum_{c=1}^{C} \eta_c \psi_{cj} \log \left(\sum_{k=1}^{K} \psi_{ck} q_k, \right). \tag{34}$$

Insert the optimal choice probabilities, $\bar{q}^*$, into the function g_j (given in Eq. (2)) that describes the similarity function, G

$$\begin{aligned}
g_j(\bar{q}^*) &= \left(1 - \sum_{c=1}^{C} \eta_c \psi_{cj}\right) \log q_j^* + \sum_{c=1}^{C} \eta_c \psi_{cj} \log \left(\sum_{k=1}^{K} \psi_{ck} q_k^* \right), \\
&= \log q_j^* + K_j(\bar{q}^*), \\
&= \left(v_j + \log q_j^* - g_j(\bar{q}^*) - \log \left\{ \sum_{k=1}^{J} \exp \left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right) \right) \right\} \right) + K_j(\bar{q}^*), \\
&= \left(v_j + \log q_j^* - \left(\log q_j^* + K_j(\bar{q}^*) \right) - \log \left\{ \sum_{k=1}^{J} \exp \left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right) \right) \right\} \right) + K_j(\bar{q}^*), \\
&= v_j - \log \left\{ \sum_{k=1}^{J} \exp \left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right) \right) \right\},
\end{aligned}$$

where the expression for the optimal choice probabilities (given by Eq. 4) are inserted in the third line and $g_j(\bar{q}^*)$ is inserted in the fourth line. Insert this expression and the optimal choice probabilities into Eq. (1) that describe the expected utility prior to the realized choice

$$\begin{aligned}
EV &= \sum_j v_j \cdot q_j^* - G(\bar{q}^*), \\
&= \sum_j v_j \cdot q_j^* - \sum_j q_j^* \cdot g_j(\bar{q}^*), \\
&= \sum_j v_j \cdot q_j^* - \sum_j q_j^* \cdot \left(v_j - \log \left\{ \sum_{k=1}^{J} \exp \left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right) \right) \right\} \right), \\
&= \sum_j q_j^* \cdot \log \left\{ \sum_{k=1}^{J} \exp \left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right) \right) \right\}, \\
&= \log \left\{ \sum_{k=1}^{J} \exp \left(v_k + \log q_k^* - g_k\left(\bar{q}^*\right) \right) \right\}.
\end{aligned}$$

This ends the proof. The proof is similar to Andersen (2022).

B Appendix

For ease of notation let $EV_k = EV(d = k)$ denotes the expected value of the DM that in previous period chose alternative k, and define the vector of expected values $\bar{EV} = (EV_1, ..., EV_J)^\top$. Lastly, define the function $f_k : \mathbb{R}^J \to \mathbb{R}$ as the expected value function

$$f_k(\bar{EV}) = \log\left\{\sum_{j=1}^{J} \exp\left(v_{kj}(EV_j) + \log q^*_{kj} - y_j\left(\bar{q}^*_k\right)\right)\right\}, \text{ for } k = 1, ..., J, \tag{35}$$

where $v_{kj}(EV_j) = u_{kj} + \delta \cdot EV_j$ is the expected payoff from choosing alternative j given alternative k was chosen in last period. For later use differentiate Eq. (35) with respect to EV_j and apply the envelope theorem

$$\nabla_{EV_j} f_k(\bar{EV}) = \frac{\exp\left(v_{kj} + \log q^*_{kj} - g_j\left(\bar{q}^*_k\right)\right)}{\sum_{m=1}^{J} \exp\left(v_{km} + \log q^*_{km} - g_m\left(\bar{q}^*_k\right)\right)} \cdot \delta = \delta \cdot \bar{q}^*_{kj}(\bar{EV}). \tag{36}$$

Next, define $\bar{Q}^*(\bar{EV}) = (\bar{q}^*_1, ..., \bar{q}^*_J)$ as the stacked matrix of choice probabilities, and based on Eq. (35) and (36) define the vector function $f(\bar{EV}) = \left(f_1(\bar{EV}), ..., f_J(\bar{EV})\right)^\top$ and the $(J \times J)$ matrix $\nabla f(\bar{EV}) = \delta \cdot \bar{Q}^*(\bar{EV})$ containing its derivatives.

The mean value theorem implies

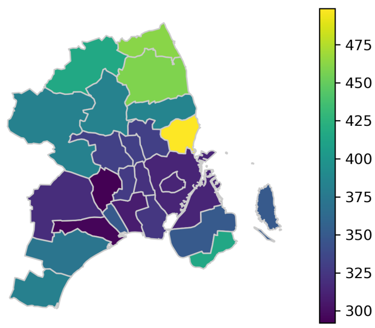
$$f(\bar{EV}_1) - f(\bar{EV}_0) = \nabla f(\bar{EV}) \cdot (\bar{EV}_1 - \bar{EV}_0),$$

for some $\bar{EV} \in (\bar{EV}_0, \bar{EV}_1)$. Eq. (35) can be shown to define a contraction, by the use of the infinity norm as distance metric

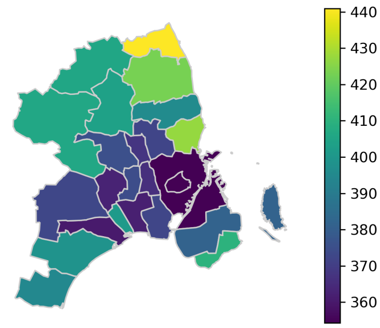
$$\begin{aligned}
\|f(\bar{EV}_1) - f(\bar{EV}_0)\|_\infty &= \|\nabla f(\bar{EV}) \cdot (\bar{EV}_1 - \bar{EV}_0)\|_\infty, \\
&\leq \|\delta \cdot \bar{Q}^*(\bar{EV})\|_\infty \cdot \|(\bar{EV}_1 - \bar{EV}_0)\|_\infty, \\
&\leq \delta \cdot \|\bar{EV}_1 - \bar{EV}_0\|_\infty.
\end{aligned}$$

Where it has been used that $\|\bar{Q}^*(\bar{EV})\|_\infty = 1$, since every row of $\bar{Q}^*(\bar{EV})$ sums to unity. This ends the proof. As a result for any $\bar{EV}_0$ the sequence of iterates $\bar{EV}_0, f(\bar{EV}_0), f(f(\bar{EV}_0)), ...$ converges to a unique fixed point of f.

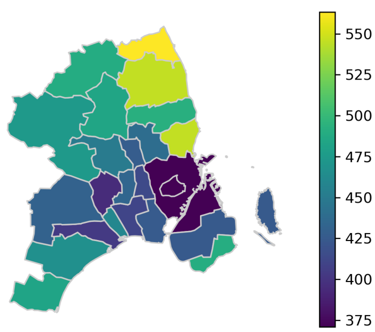
C Appendix



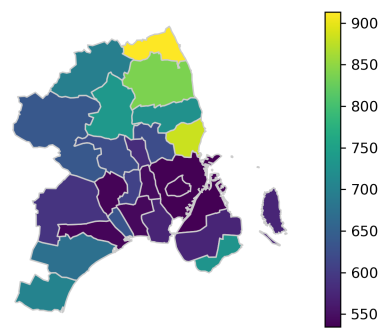
(a) Average wages for unskilled (DKK 1,000).



(b) Average wages for unskilled (DKK 1,000).



(c) Average wages for medium educated (DKK 1,000).



(d) Average wages for highly educated (DKK 1,000).

Figure 11: Average wages for educational groups based on the municipality the individuals reside.