Zurcher on Steoriods: An equilibrium life-cycle model of commuting, residential and work location choice

Christian L. Carstensen [1] Fedor Iskhakov [2] Maria J. Hansen [1]
John Rust [3] Bertel Schjerning [1]

[1]University of Copenhagen

[2]Australian National University

[3]Georgetown University

Center for Market Studies and Spatial Economics
St. Petersburg
January 29th, 2020

Overview of Rust (1987)

This is a path-breaking paper that introduces a methodology to estimate a single-agent dynamic discrete choice models.

Main contributions

1. Development and implementation of Nested Fixed Point Algorithm
2. Formulation of assumptions, that makes dynamic discrete choice models tractable.
3. Bottom-up approach: Micro-aggregated demand for durable assets
4. An illustrative application in a simple model of engine replacement.
5. The first researcher to obtain ML estimates of discrete choice dynamic programming models

Policy experiments:

- How does changes in replacement cost affect
  - the distribution of mileage
  - the equilibrium demand for engines

Structural Estimation in Microeconomics

Many methods for solving Dynamic Discrete Choice Models

- Rust (1987): MLE using Nested-Fixed Point Algorithm (NFXP)
- Hotz and Miller (1993): CCP estimator - (two step estimator)
- Keane and Wolpin (1994): Simulation and interpolation
- Rust (1997): Randomization algorithm (breaks curse of dimensionality)
- Aguirregabiria and Mira (2002): Nested Pseudo Likelihood (NPL).
- Norets (2009): Bayesian Estimation
- Su and Judd (2012): MLE using constrained optimization (MPEC)
- .... and a bunch of methods for solving games (NPL, NPEX, NFXP, CCP, BBL types)

NFXP is still the Swiss-army knife for structural estimation of dynamic structural choice models

Who cares about Harold Zurcher?

- Occupational Choice (Kaeane and Wolpin, JPE 1997)
- Retirement (Rust and Phelan, ECMA 1997)
- Brand choice and advertising (Erdem and Keane, MaScience 1996)
- Choice of college major (Arcidiacono, JoE 2004)
- Individual migration decisions (Kennan and Walker, ECMA 2011)
- High school attendance and work decisions (Eckstein and Wolpin, ECMA 1999)
- Sales and dynamics of consumer inventory behavior (Hendel and Nevo, ECMA 2006)
- Advertising, learning, and consumer choice in experience good markets (Ackerberg, IER 2003)
- Route choice models (Fosgerau et al, Transp. Res. B)
- Fertility and labor supply decisions (Francesconi, JoLE 2002)
- Car ownership, type choice and use (Gillingham et al, WP)
- Residential and Work-location choice (Buchinsky et al, ECMA 2015)
- ...and many more (2021 cites, Jan 2020)

Our paper

Goal: Analyze the consequences of increased commute cost and (eventually) the effects of infrastructure investments that reduces them.

We develop a dynamic equilibrium model that *simultaneously* tracks the following mechanisms

- Choice of residence, work location, housing demand, commuting
- ... allowing life-cycle aspects to play a key role.
- Trade of houses and sorting heterogeneous consumers.
- Equilibrium price mechanism in the housing market

We structurally estimate this model

- Danish register data: Lots heterogeneity and rich dynamics.
- Model fits key aspects of data well

We simulate the effects on house prices, job mobility, residential sorting and commuting in **two counterfactual equilibria:**

1. Increased supply of housing in the center of Copenhagen
2. Increased cost of commuting.

A brief history of related location choice models I

- Until recently, literature on household location decisions have mainly applied *static models* and focused on *either* work or home location choices.

- Kennan and Walker (ECMA, 2011)
  - First to estimate dynamic spatial model of location choice ... but restrict households to live and work in the same location
  - Focus on expected income changes as the main driver of migration decisions
- Bayer et al. (ECMA, 2015)
  - develop a dynamic model of residential choice and computationally light sequential estimation method
  - Focus is on estimating the marginal willingness to pay for several non-marketed amenities.
  - Ignores life-cycle aspects and do not solve the model for counterfactual simulations.

A brief history of related location choice models II

- Buchinsky et al. (ECMA, 2014)
  - The first to model the *joint residential and job location choices* in a structural, dynamic model
  - also includes occupational choice (unemployed, blue and white collar).
  - Data on immigrants to Israel from the Soviet Union.
  - Exogenous house prices.
- Oswald (2019, QE):
  - Life-cycle model of consumption, housing choice and migration in the presence of aggregate and regional income and price shocks.
  - Finds that migration elasticities vary substantially between renters and owners.
  - Does not distinguish between home and job location.

Data and Descriptives

Population density in Denmark, 2016

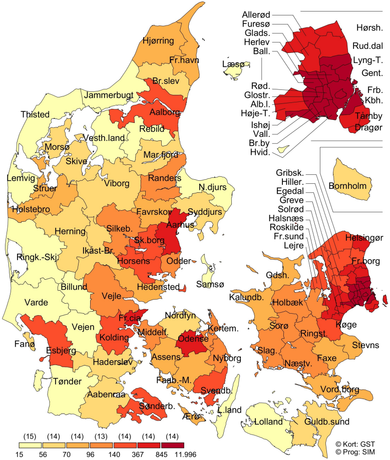
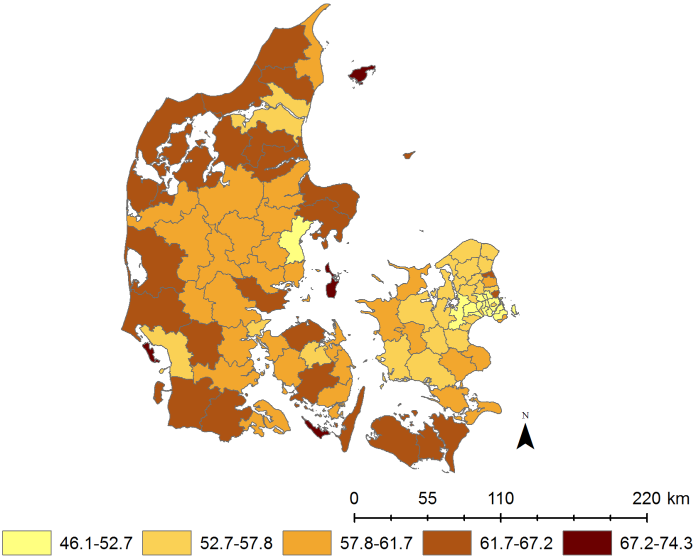
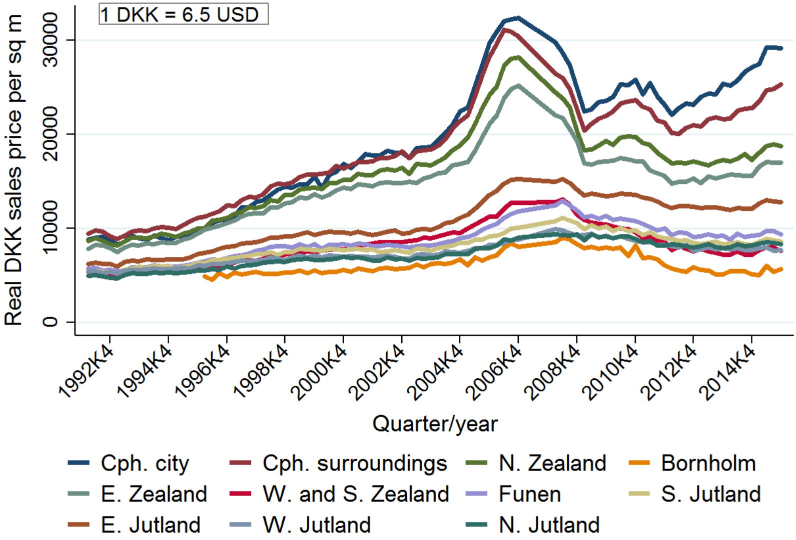


Figure: Population density (inhab/sq km) January 1st 2016 by municipality

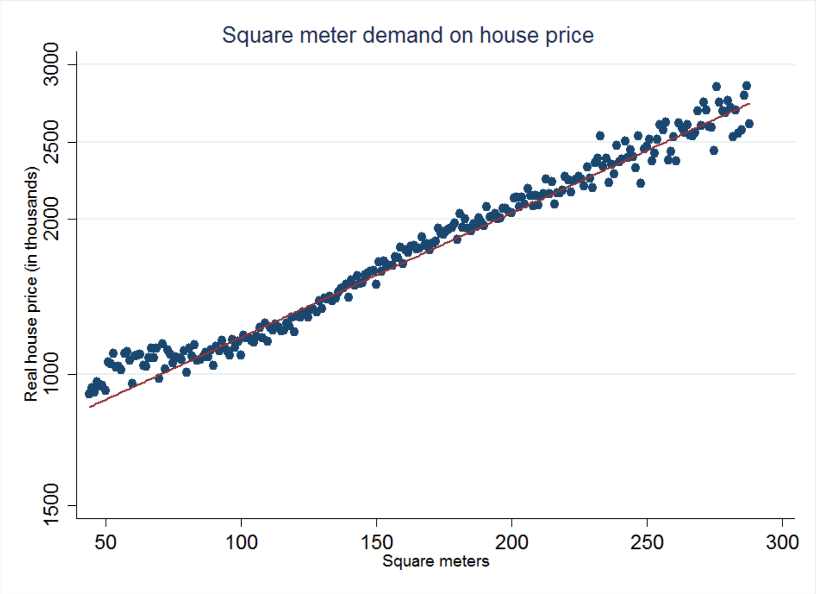
Avg. living area per person by municipality 1992-2015



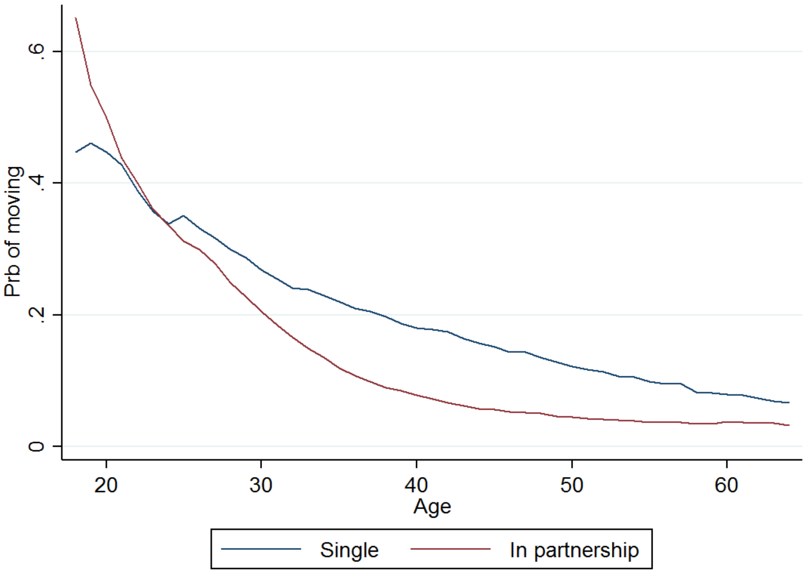
Sales price per sqm by province over time



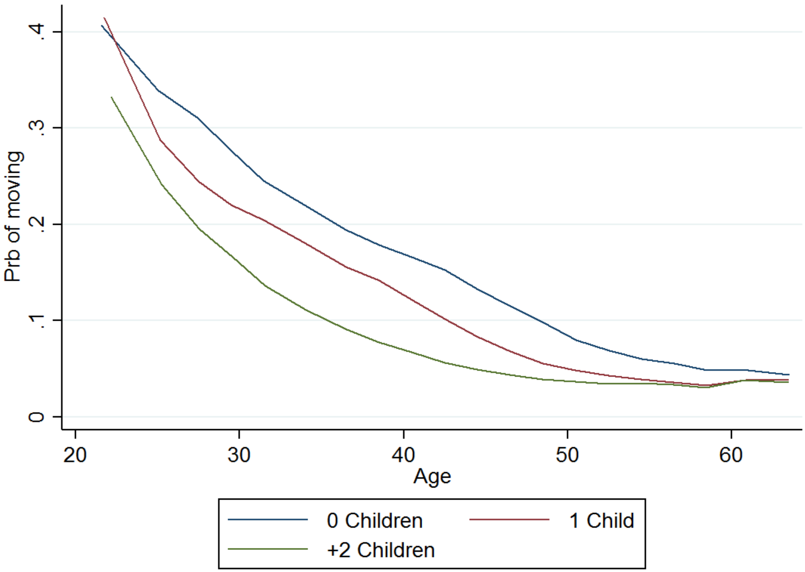
Prices linear in size



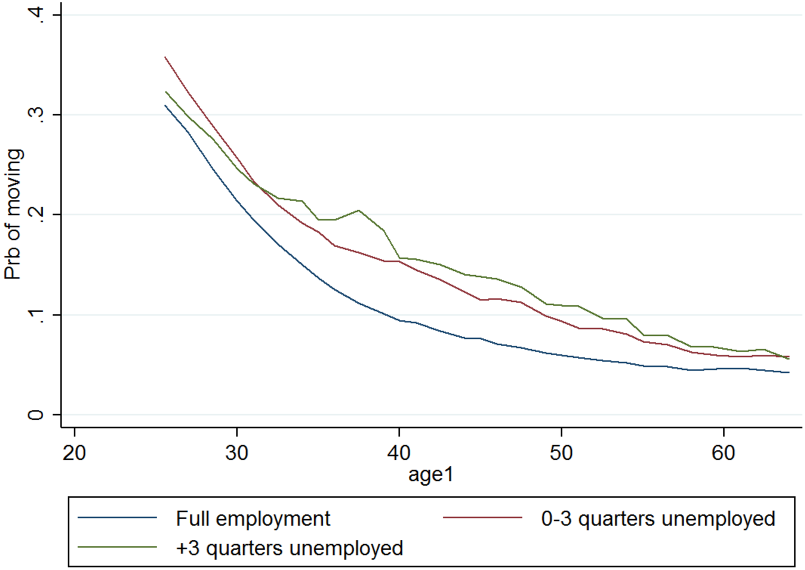
Move propensity by age and partnership status



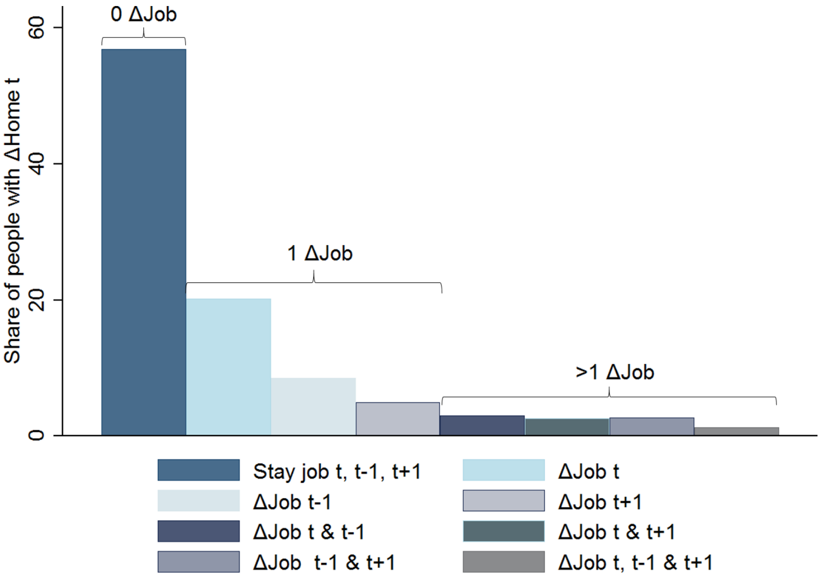
Move propensity by age and number of children



Probability of moving by age labor market status



Probability of moving home by job transition status



Key data features to capture

- Increased urbanization $\rightarrow$ regional price dispersion between urban and rural areas
- Simultaneity in residential and work location choices.
- Residential mobility
  - Falls with age, moving distance, when full-time employed, when children arrive, and when there is more than one person in the house hold (dual earner coordination problem)
- Job mobility
  - Falls with age, increase with education
- Housing demand
  - Increase with family size and income, but falls with prices
  - Trade-off between location and size of house
  - House prices are almost perfectly linear in house size

The Choice Model

Model overview

Model: Dynamic, finite-horizon, quasi-linear utility (later: overlapping generations).

States:
- **Household time-varying states:** Age, Marital status, Children status, Residential location, Work location
- **Household types:** Education type

Utility: Derived from

- Commuting (costly)
- Housing (costly, but deliver house services)
- Consumption of outside good

Decisions:
- Residential location, Work location (job search) $\rightarrow$ Commute distance
- House size demand (in m^2)

Choices and state transitions

Decision variables

Symbol	Description	Possible Values
d_t^{rl}	Residential location decision	one of the R zones
d_t^{wl}	Work location decision (search region)	one of the R zones or -1 for unemployment
h_t	House size decision	Continuous choice (in m^2)

State transitions

Symbol	State	Values		Transition
age_{it}	Age of individual	$t = 18, \ldots, 100$		deterministic
rl_{it}	Residential location	one of the R zones		deterministic by choice
wl_{it}	Work location	one of the R zones or -1 for unemployment		stochastic by state/choice
ms_{it}	Marital status	0	single	$\pi_{ms}(ms_{it+1} \mid ms_{it}, age_{it})$
		1	married/cohabitating	
cs_{it}	Children status	0	no children	$\pi_{cs}(cs_{it+1} \mid cs_{it}, age_{it})$
		1	children living at home	
es_{it}	Education status	0	Less than medium cycle	$\pi_{es}(es_{it+1} \mid es_{it}, age_{it})$
		1	Medium cycle (BA)	
		2	Long cycle(master/PhD)	

The Choice Model

$$\max_{\{d_{it}^{rl}, d_{it}^{wl}, h_{it}\}_{t=t_0}^{T}} \sum_{t=t_0}^{T} \rho_t(x_{it}, \tau)^t \mathbb{E}_t \Big[u_t(x_{it}, wl_{it}, rl_{it}, d'_{it}, \mathrm{h}_{it}, \tau_i) \Big]$$

$d'_{it} = (d_{it}^{rl}, d_{it}^{wl})$ *discrete* residential/work location decision

h_{it} *continuous* house size decision (in m^2)

t_0, T household age ($t_0 = 20$, $T = 85$)

$\rho_t(x_{it}, \tau)$ survival probability times discount factor ($\beta = 0.95$)

$u_t(\cdot)$ instantaneous utility function

(rl_{it}, wl_{it}) residential/work location in beginning of t, i.e "decision-outcome" component of state $s_{it} = (x_{it}, rl_{it}, wl_{it})$

x_{it} time varying individual states, $x_{it} = (ms_{it}, cs_{it})$

τ_i fixed types (educational status)

Sequential choice of work and residence locations

- States: Each period t starts off with a given work and residence locations, and other variables x_t forming the vector of state variables $s_t = (wl_t, rl_t, x_t)$.
- Sequential choice: Individuals make their work and residential location choices sequentially but instantaneously (at start of period t)
  1. Intended work location $d_t^{wl} \rightarrow$ *realized* work location wl_{t+1}.
  2. Residence location d_t^{wl} is then chosen *conditional* on wl_{t+1}, i.e. the realization of the their employment search
  3. Given new residence and job location, household determines optimal house size depending on their own characteristics and the chosen region of residence.
- Consumption (including housing) is enjoyed for the rest of the period, and the process transitions to the next period.

Possible transitions in the job search process

Work location transition probabilities when searching job in region d_t^w

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

For individual that chooses to search for a job in some new location $d_t^w \neq wl_t$, there are three possible outcomes:

1. job search is successful and individual receive a job offer in this location
2. individual does not get a job offer, but is able to keep existing job
3. individual's job search is unsuccessful and they are laid off from their current job.

Possible transitions in the job search process

Work location transition probabilities when not searching job
(i.e. $\pi_t^n(wl_t, wl_t, x_t) = 0$)

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

For currently unemployed individuals
(always possible to stay unemployed i.e. $\pi_t^k(\emptyset, x_t) = 1$)

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

Voluntary unemployment: If the individual choose to stop working, $d_t^w = \emptyset$, then $\pi_t^n(\emptyset, wl_t, x_t) = 1$, i.e. there is "perfect control" over the decision to stop working so that $w_{t+1} = \emptyset$ with probability 1.

Simplified Notation

- To make notation less cluttered we will drop time subscripts on the state and decision variables and write utility as $u_t(x, d, d', \tau)$.

Utility specification

$$u_t(x, d, d^{rl}, d^{wl}, h, \tau) =$$

$$\kappa(s) \underbrace{\left[y(s) - \psi_{uc} p^h(d^{rl}) h - c^m(d, d', x, \tau) \right]}_{\text{consumption}} + \underbrace{\phi_{h1}(s) h + \phi_{h2} h^2}_{\text{utility housing}}$$

$$+ \text{commute cost}(x, d') + \text{ammeneties}(d^{rl}) - c^u(d, d', x, \tau)$$

$\kappa(s)$ Constant marginal utility of money. Macro and income dependent. (no consumption/savings)

$p_h(d^{rl})$ price per m^2 in chosen region

$\phi_{h1}(s), \phi_{h2}$ utility from housing

commute cost depends travel-time between d^{rl} and d^{wl} and on states such as children, marital status

ammenties depends on regional attributes of residential location, d^{rl}

c^m, c^u Monetary and utility cost of moving job and residence (we may restrict dependence on d to simplify solution)

Simplifying the model

To simplify computations, we exploit:

House size: No dynamic implications of housing demand $\Rightarrow$ static choice.

Property: Continuation value depends only on the *destination* region of work and residence
(not the current region)

A: Separate static continuous choice

Next period value function $V_{t+1}(x, d, \epsilon, \tau)$ is independent of h, therefore the Bellman equation simplifies to

$$V_t(x, d, \epsilon, \tau) = \max_{d^{rl}, d^{wl}} \Big\{ \max_h \Big[u_t(x, d, d^{rl}, d^{wl}, h, \tau) \Big] + \epsilon(d') + \\ + \rho_t(x, \tau) EV_{t+1}(x, d^{rl}, d^{wl}, \tau) \Big] \Big\}$$

$$\frac{\partial u_t(\cdot)}{\partial h} = \phi_{h1}(s) + 2\phi_{h2} h - \kappa(s) p^h(d^{rl}) = 0 \quad \Rightarrow$$

$$h_t^* = \frac{\phi_{h1}(s) - \kappa(s) p^h(d^{rl})}{-2\phi_{h2}}$$

- Substituting expression optimal housing demand into the utility function defined above, we obtain the *indirect utility function* $u(wl, rl, wl', rl', x_t)$.
- Pure discrete choice model conditional on housing demand.

B: Reformulate in terms of expected value function

Next period value function $V_{t+1}(x, d', \epsilon, \tau)$ is only dependent of the chosen location d' and not on current location d, therefore solving the Bellman equation in terms of expected value function has much lower dimensionality

- Assume that there are N discrete possible values for x and so for each type τ $p_t(x'|x, \tau)$ is an $N \times N$ transition probability matrix.
- Evaluation of $v_t(x, d, d', \tau)$ for a single type τ at all possible values of (x, d, d') requires ND^2 operations.
- Calculation of $EV_{t+1}(x, d)$ for all possible values of (x, d) requires $(ND)^2$ operations
- Total work per iteration (value of t) is proportional to $ND^2 + (ND)^2$ which is dominated by the $(ND)^2$ term when D is large.

Recursive formulation of Bellman equations

Residence location choice probabilities logit formulas

$$P^r(d^r|wl, rl, wl', x) = \frac{\exp\{[u(wl, rl, wl', d^r, x) + \beta EV(wl', d^r, x)]/\sigma_r\}}{\sum_{d^r} \exp\{[u(wl, rl, wl', d^r, x) + \beta EV(wl', d^r, x)]/\sigma_r\}}.$$

Where EV is the expected value function

$$EV(wl', rl', x) = \sum_{x'} \pi^x(x, x') EV^w(wl', rl', x').$$

and $EV^w(wl, rl, x)$ be the ex ante expected value for an individual who has not learned the work location shocks $\{\epsilon^w(d^w)\}$

$$EV^w(wl, rl, x) = \sigma_w \log \left(\sum_{d^w} \exp \left\{ \sum_{wl} \pi(d^w, wl, x, wl) EV^r(wl, rl, wl, x)/\sigma_w \right\} \right).$$

Recursive formulation of Bellman equations

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

$$P_t^w(d^w|wl,rl,x) = \frac{\exp\{v^w(wl,rl,x,d^w)/\sigma_w\}}{\sum_{d^w}\exp\{v_t^w(wl,rl,x,d^w)/\sigma_w\}}.$$

Where $v^w(wl,rl,x,d^w)$ is the expected choice-specific value corresponding to the particular choice of job location d^w.

$$v^w(wl,rl,x,d^w) = \sum_{wl'} \pi_t(d^w,wl,x,wl')EV^r(wl,rl,wl',x).$$

and where EV^r is the *ex ante* expected value conditional on the employment location outcome wl' is given by the usual log-sum formula

$$EV^r(wl,rl,wl',x) = \sigma_r \log\left(\sum_{d^r}\exp\{[u(wl,rl,wl',d^r,x)+\beta EV(wl',d^r,x)]/\sigma_r\}\right).$$

Structural Estimation

Estimation strategy

We estimate the model sequentially in three separate steps:

1. Estimate the parameters governing the wage equations and transition probabilities of the children state
2. Estimate a reduced form housing demand equation
3. Estimate the remaining structural parameters by maximum likelihood applying the parameters obtained in 1) and 2).

The MLE is obtained as

$$\hat{\theta} = \mathrm{argmax}_{\theta} \frac{1}{N} \sum_{i} \sum_{t} \{ \log P_t^r(rl_{it+1} | wl_{it}, rl_{it}, wl_{it+1}, x_{it}; \theta) + \log \sum_{d^w} P_t^w(d^w | wl_{it}, rl_{it}, x_{it}; \theta) \pi_t(d^w, wl_{it}, x_{it}, wl_{it+1}; \theta) \}, \tag{1}$$

where N is the number of individuals.

- We proceed in the spirit of the Nested Fixed Point (NFXP) algorithm by Rust 1987 and solve the model via backwards induction for each evaluation of the likelihood function.

Functional forms, parameter estimates and model fit

Housing demand

Quadratic utility of housing

$$u_m + u_h = \kappa(inc_t)\big(inc_t - hcost_t\big) + \Phi(x_t)h_{t+1} + \frac{1}{2}\phi_{h2}\, h_{t+1}^2,$$

where $\phi_{h2} < 0$ (diminishing returns to house size)

Marginal utility of money depends on income

$$\kappa(inc_t) = \kappa_0 + \kappa_y inc_t.$$

$\Phi(x_t)$ allows for heterogeneity in marginal utility of housing

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

Housing costs are given by

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

Implied housing demand (linear regression).

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

Parameter estimates

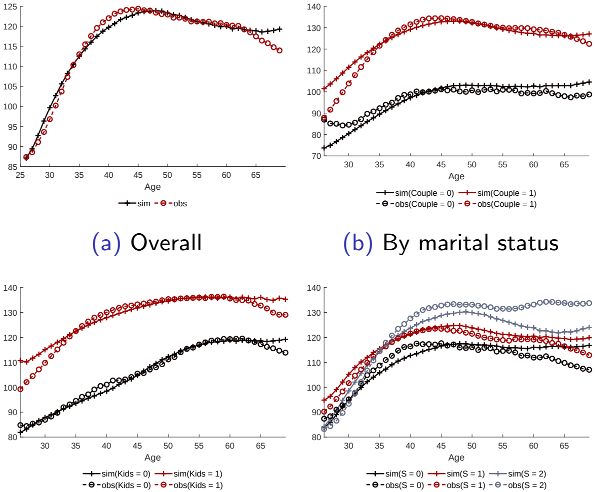
Table: First Stage Parameter Estimates, Housing Demand

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

Table: Curvature Parameter Housing Demand

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

Model fit: housing size over the life cycle



(a) Overall

(b) By marital status

(c) By children

(d) By schooling

Regional amenities

Amenities of regions come as a bundle of attributes, W, that each contribute to the experienced utility of a region.

$$amenities(d_{it}) = \boldsymbol{\alpha}' W(d_{it}^{rl}) \tag{2}$$

- Here, $W(d_{it}^{rl})$ is a column of attributes for region rl and $\boldsymbol{\alpha}$ is a vector of coefficients to those attributes.
- We can include a rich set of amenities almost without any additional computational cost associated with solving the model.
- This may require more parameters as number of amenities increase, but the number of parameters in the chosen specification is independent of the number of regions

Parameter estimates

Table: Regional Amenities

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

Job arrival and dismissal

Probability of getting a new job

$$\pi_t^n(d^w, wl, x) = \Bigg[1 + exp\Bigg(- \Big(\beta_0^{\pi(n)} + \beta_a^{\pi(n)} age + \beta_{\o}^{\pi(n)} 1_{wl=\o} + \beta_{jobdens}^{\pi(new)} jobdens(d^w) + \sum_{k=1}^{2} \big(\beta_s^{\pi(new)}(k) 1_{edu=k}\big)\Big)\Bigg)\Bigg]^{-1}, \tag{3}$$

Probability of keeping current job

$$\pi^k(wl, x) = \Bigg[1 + exp\Bigg(- \Big(\beta_0^{\pi(k)} + \beta_a^{\pi(k)} age + \sum_{k=1}^{2} \big(\beta_s^{\pi(k)}(k)\{1_{edu=k}\big)\Big)\Bigg)\Bigg]^{-1}. \tag{4}$$

Finally, we allow for disutility of work u_w though the fixed constant, c_{work}, which is applied whesn $wl' \neq \o$.

Parameter estimates

Table: Job arrival and dismissal

	Coeff. Est.	Standard Err.	Z-stat.
Probability of keeping job: $\pi_t^k(wl_t, x_t; \beta^k)$			
Const., $\beta_0^{\pi(keep)}$	2.2226	0.04122	53.9
Age, $\beta_a^{\pi(keep)}$	0.0384	0.00098	39.0
Schooling, $\beta_s^{\pi(keep)}$ (1)	0.8267	0.02178	38.0
Schooling, $\beta_s^{\pi(keep)}$ (2)	0.5677	0.01633	34.8
Probability of new job: $\pi_t^n(d_t^w, wl_t, x_t : \beta^n)$			
Const., $\beta_0^{\pi(new)}$	-0.2453	0.00617	-39.7
Age, $\beta_a^{\pi(new)}$	-0.0624	0.00014	-457.6
Schooling, $\beta_s^{\pi(new)}$ (1)	0.1455	0.00347	41.9
Schooling, $\beta_s^{\pi(new)}$ (2)	0.2580	0.00375	68.8
Job density $\beta_{jobdensity}^{\pi(new)}$	2.9591	0.00700	422.7
Prev. unempl., $\beta_{unemp}^{\pi(new)}$	1.2326	0.00337	365.6
Disutil. of work, c_{work}	2.2163	0.00189	1175.6

Utility cost of moving, $c^u(d, d', x, \tau)$

The utility cost of moving residence

$$c^u(d, d', x, \tau) =$$

$$1_{\{rl \neq rl'\}}\left[\gamma_0 + \gamma_a age + \gamma_{ms} ms + \gamma_{cs} cs + \sum_{k=1}^{2} \phi_{s,k} 1_{\{edu_t = k\}}\right],$$

Parameter estimates

Table: Utility Cost of Moving Residence

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

Commute cost

The commuting cost between d^{rl} and d^{wl}

$$c^{commute}(d', x, \tau) = \eta^1 f^{tt}(d^{rl}, d^{wl}) + \eta^{cs} cs + \eta^{ms} ms$$

where

- The $f^{tt}(\cdot)$ function denotes the travel-time by car between the argument locations.
- Commute cost are zero, when unemployed, i.e. $d^{rl} = -1$

Commute cost are assumed to be a function of

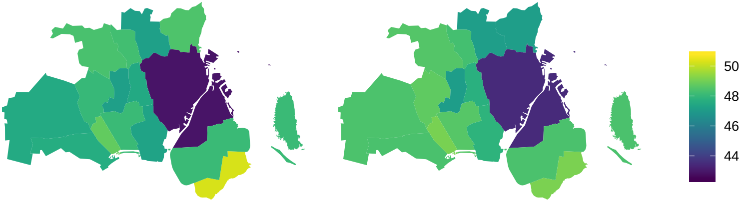
- the traveltime between the two destinations
- personal characteristics we see affecting commuting decisions (children status presumably affects marginal utility of leisure, marital status due to coordination/specialization effects)

Parameter estimates

Table: Commute Cost

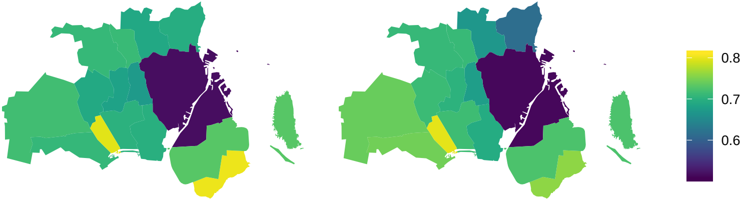
	Coeff. Estimates	Standard Error	Z-statistic
Cost of travel time, η_{ttime}	0.2369	0.00118	200.8

Model fit: residential sorting (age)



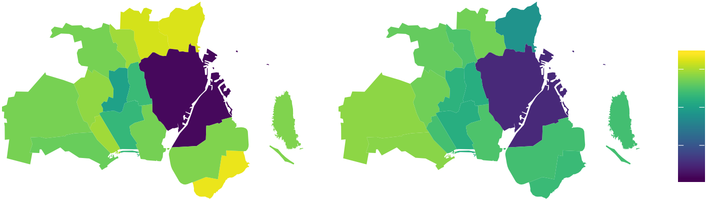
(a) Age (obs.)   (b) Age (sim.)

Model fit: residential sorting (couples)



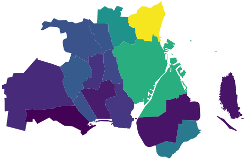
(a) Couples (obs.) (b) Couples (sim.)

Model fit: residential sorting (children)



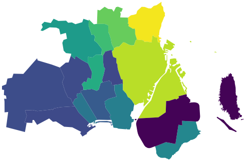
(a) Children (obs.) (b) Children (sim.)

Model fit: sorting of highly educated



(a) by home (obs.)

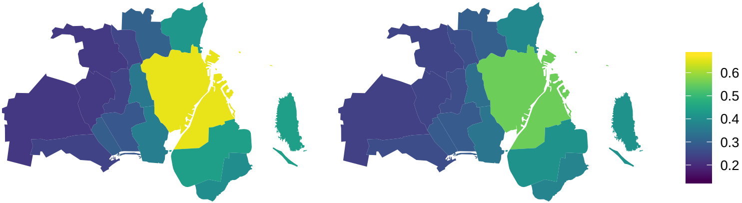
(b) by home (sim.)



(c) by work (obs.)

(d) by work (sim.)

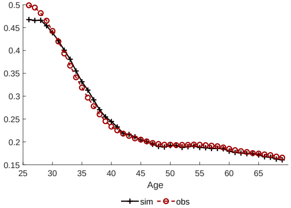
Model fit: working in Copenhagen by residenital location



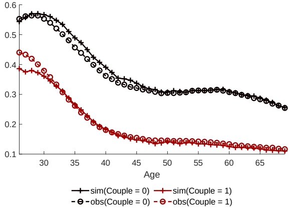
(a) observed

(b) Simualted

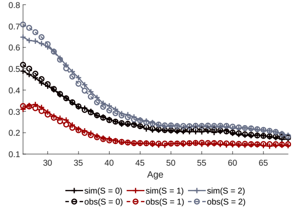
Model fit: share living in Copenhagen over the life cycle



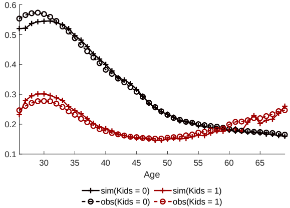
(a) Overall



(b) By marital status

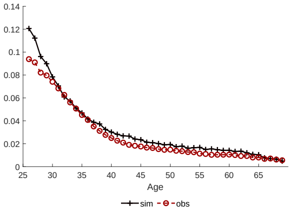


(c) By schooling

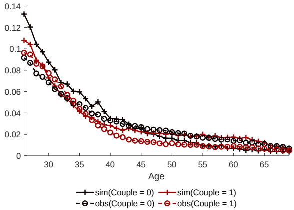


(d) By children

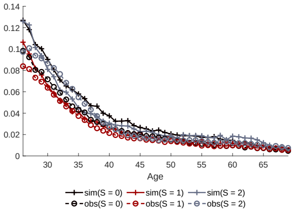
Model fit: moving residential location



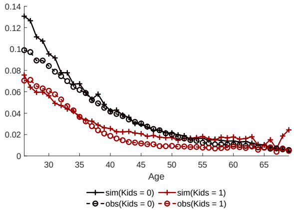
(a) Overall



(b) By marital status

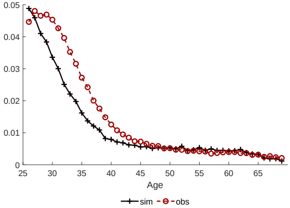


(c) By schooling

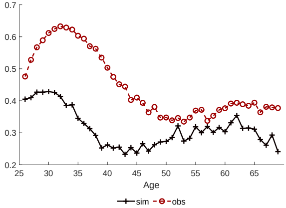


(d) By children

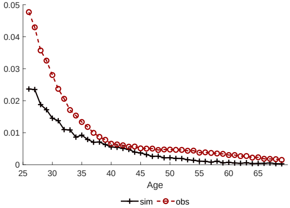
Model fit: moving residence from and to Copenhagen



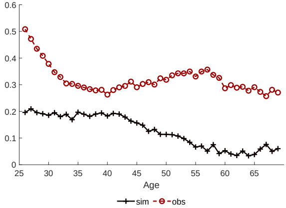
(a) From Cph. (all)



(b) From Cph. (movers)

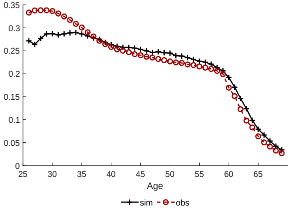


(c) To Cph. (all)

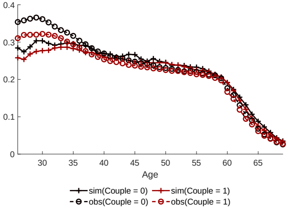


(d) To Cph (movers)

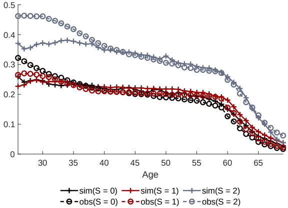
Model fit: share working in Copenhagen



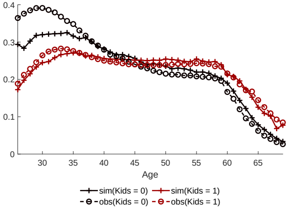
(a) Overall



(b) By marital status

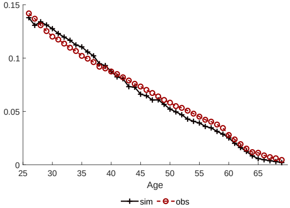


(c) By schooling

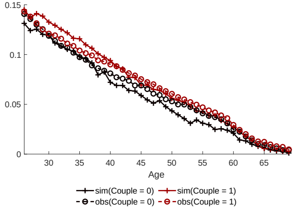


(d) By children

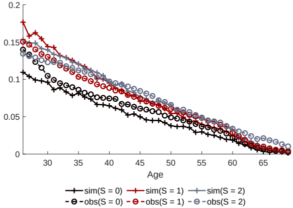
Model fit: moving work location



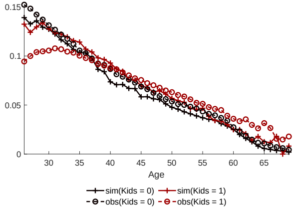
(a) Overall



(b) By marital status

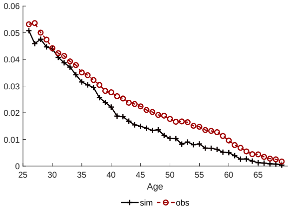


(c) By schooling

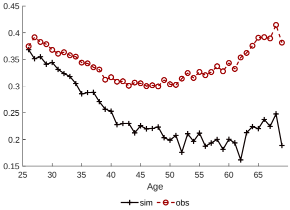


(d) By children

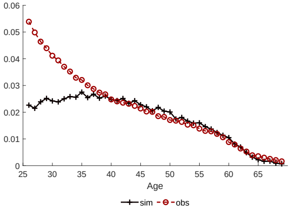
Model fit: moving work location from and to Copenhagen



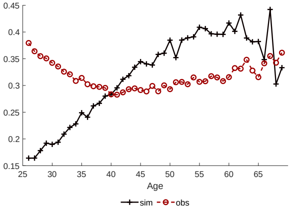
(a) Move away from Cph. (all)



(b) Move away from Cph. (movers)

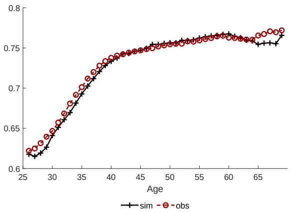


(c) Move to Cph. (all)

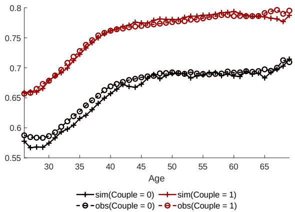


(d) Move to Cph. (movers)

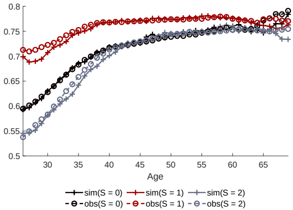
Model fit: commute time (hours)



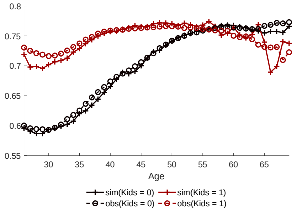
(a) Overall



(b) By marital status



(c) By schooling



(d) By Children

Equilibrium at the Housing Market

Demand and supply of housing

- **Equilibrium prices,** P^h: adjust so that the total expected demand $D_t(rl, P^h)$ for housing measured in square meters equals the (inelastic) supply $S_t(rl)$ in each residential region.
- **Housing supply** $S_t(rl)$: micro aggregated observed square meters of housing h_{it} for people who live in region $rl_{it} = rl$ at the beginning of each period t

$$S_t(rl) = \sum_{i=1}^{N} h_{it} 1(rl_{it} = rl)(1 - ms_{it}/2)$$

- **Expected housing demand** $D_t(rl, P^h)$: population average of housing demand weighted by choice probabilities of either staying or moving to region rl at the end of period t.

$$D_t(rl, P^h) = \sum_{i=1}^{N} h(rl, x_{it}; P^h(rl))\Pi_t(rl|wl_{it+1}, rl_{it}, x_{it}; P^h)(1-ms_{it}/2),$$

where $\Pi_t(rl|wl_{it+1}, rl_{it}, x_{it}; P^h)$ is the choice probability that individuals in state $s_{it} = (wl_{it+1}, rl_{it}, x_{it})$ chooses to live in region rl, given the vector of regional house prices, P^h

Equilibrium house prices

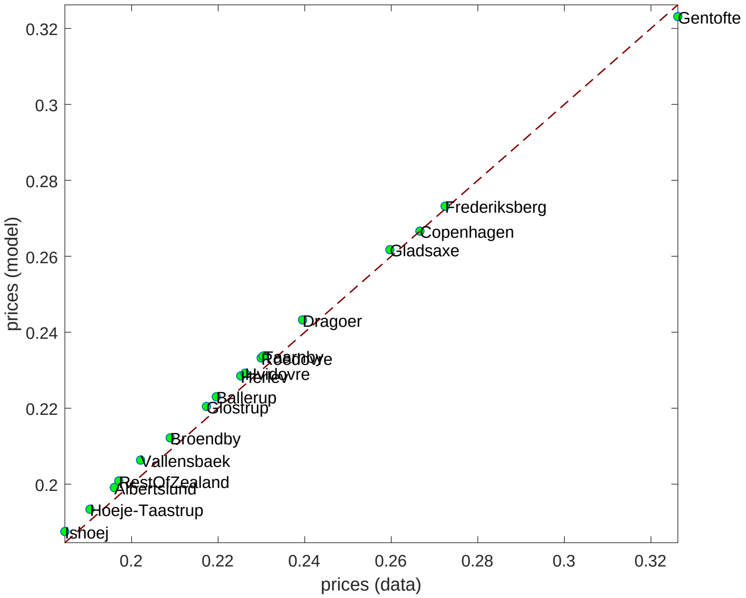
To compute the housing market equilibrium, P^h is set to solve

$$\begin{aligned} S_t(1) &= D_t(1, P^h) \\ &\vdots \\ S_t(R) &= D_t(R, P^h) \end{aligned}$$

where

- $P^h = (P^h(1), .., P^h(R))$ s the R-dimensional vector of regional square meter prices in each residential region $rl = \{1, \ldots R\}$
- $S_t(rl)$ the inelastic, exogenously fixed supply of total square meters of housing in region rl
- $D_t(rl, P^h)$ is the demand for available square meters of housing in region rl
- We can easily solve the R equilibrium equations with R unknowns using Newtons Method.

Model fit: Equilibrium housing demand and house prices



Model fit: Summary

The estimated model closely tracks observed life-cycle patterns and captures key aspects of worker mobility and residential sorting:

- House prices are almost perfectly linear in house size
- Regional prices/m^2 decreases in distance to high job-density areas
- Housing demand increases with family size and income, ...but falls with prices
- Trade-off between house size and location near jobs/amenities: ...households commute longer as they grow and demand more space.
- Residential sorting of highly educated households and into urban areas, but monocentric city model is poor approximation
- Residential mobility falls with age, moving distance, when children arrive, when married, when employed.
- Worker mobility falls with age, increase with education
- Clear simultaneity in residential and work location choices
- Equilibrium prices closely tracks the observed price ranking and overall price levels across regions.

Counterfactual Simulation

Counterfactual equilibrium simulation

Counterfactual I:

5 pct. higher housing supply in central Copenhagen

- Urbanization increase: more individuals move towards Copenhagen and away from rest of Zealand and the surroundings of the Greater Copenhagen Area
- Prices fall in all regions, but most in Copenhagen
- Sorting based on income change. Average income increase in fx i Copenhagen.

Counterfactual II:

50 pct. higher commute cost

- Commute time drops for workers
- Increase in i non-employment, especially in more remote residential regions
- Equilibrium prices falls in rural areas (rest of Zealand)

Counter-factual I: Increased housing supply

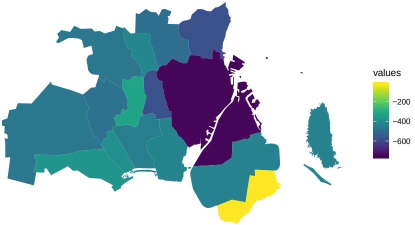
5 pct. increase in housing supply in Copenhagen/Frederiksberg

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

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

Counter-factual I: Increased housing supply

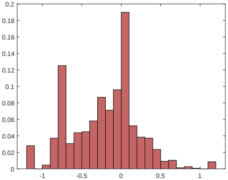
Simulated change in price (DKK/sqm)



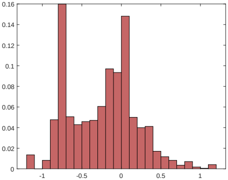
Note: Numbers are computed by subtracting baseline from counterfactual.

Counter-factual II: 50 pct. increase in commute cost

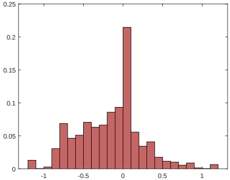
Distribution of simulated change in commute time in $t = 2$ in Counterfactual II (hours)



(a) Low education



(b) Medium education



(c) High education

Counter-factual II: 50 pct. increase in commute cost

Simulated relocations of work (wl) and home (rl) and price change

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

Counter-factual II: 50 pct. increase in commute cost

Simulated distribution of locations in $t = 2$ for $t = 0$ workers in Ishoej

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

Concluding remarks

We estimated a life-cycle equilibrium model that simultaneously tracks a large number of mechanisms:

- Residential choice, job search and work location outcomes (and resulting communing)
- Demand for square meters of housing
- Equilibrium house prices

We simulated the effects on house prices, job mobility, residential sorting and commuting in two counterfactual equilibria:

- Increased supply of housing in the center of Copenhagen
- Increased cost of commuting.

This is a first draft: We are currently working on further refining the model and simulating more counterfactuals.

Plan for future research I

- Include more heterogeneity in wage equation (persistence).
- Estimate choice model utilizing the panel data rather pooling the sample over time.
- Include more regions around Copenhagen area.
- Include more (time-varying) regional amenities
  ... and allow for taste variation.
- Add educational choice and retirement to the model.
- Model residential/job moves within regions
- Allow for more flexible substitution patterns than those implied by nested logit (GEV and GEM).
- Design counter factual analysis based on planned infrastructure investments (new metro).

Future research II

- Extend equilibrium to include the labor market.
- Allow for congestion feedback in equilibrium.
  ...residential/work location choices affects equilibrium congestion (fixed capacity in short run).
- Long run equilibrium with construction and labor demand.
- Add state variable and a dynamic housing demand choice subject adjustment costs.