The Geography of Life

Gabriel M. Ahlfeldt[1] Ismir Mulalic[3] Caterina Soto Vieira[2] Daniel M. Sturm[2]

[1]Humboldt University
[2]LSE
[3]CBS

April 23, 2024

Berlin School of Economics Opportunity Day

Motivation

- Large literatures have investigated how **age and life events**, such as marriage, children or retirement, **shape economic decisions**:
  - Franco Modigliani's pioneering work introduced the idea that wages, consumption and savings are intimately linked to age.
  - Gary Becker's work portrays marriage and children as fundamental determinants of labour supply and time allocation.
  - These ideas have generated vast empirical literatures that show that age and life events profoundly shape **labour supply, wages and savings**.
- Despite this long tradition **we know surprisingly little about how age and life events shape location choices across space**.
- For **efficient planning**, we need to understand how **demographic trends** (again, fertility, single-hood) will affect where people live and work

This Paper

- This paper uses **newly assembled geocoded matched employer-employee-property data** for Copenhagen spanning 40 years
- We provide **reduced-form evidence** (stylized facts and event studies) on how location and housing consumption choices are affected by age and life events
- We develop a **quantitative spatial model** to show how these location choices are explained by housing expenditure shares, commuting costs and amenities.
- We use model **counterfactuals** to explore how demographic trends such as aging and fertility shape the spatial organisation of cities.

Related Literature

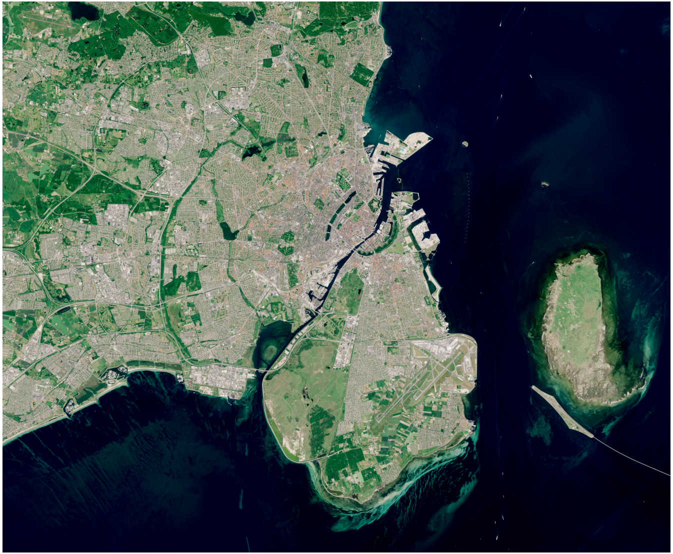
- **Effect of age on wages, income and savings**: Modigliani (1966), Mincer (1974), Meghir and Pistaferri (2011)
- **Effect of marriage and children on labour outcomes and consumption:** Becker (1973, 1974), Eckstein and Wolpin (1989), Blundell et al. (1994), Van Der Klaauw (1996), Adda et al. (2017), Kleven et al. (2018)
- **Quantitative models of cities**: Ahlfeldt et al. (2015), Allen et al. (2015), Monte et al. (2018), Heblich et al. (2020), Tsivanidis (2022), Miyauchi et al. (2022)
- **Age, Fertility and Location Choices**: Komissarova (2022), Moreno-Maldonado and Santamaría (2022), Coeurdacier et al. (2023), Albuoy and Faberman (2024)

Overview of the Presentation

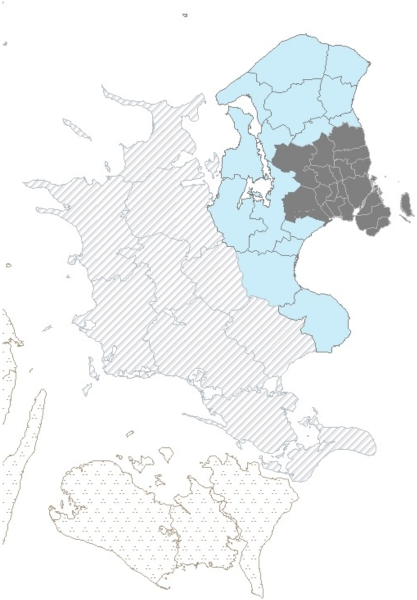
- Empirical Context and Data
- Stylized Facts on Age and Life Events
- Theoretical Model
- Quantification
- Counterfactuals
- Conclusion

Empirical Setting

A View from Space



Copenhagen Metro Area

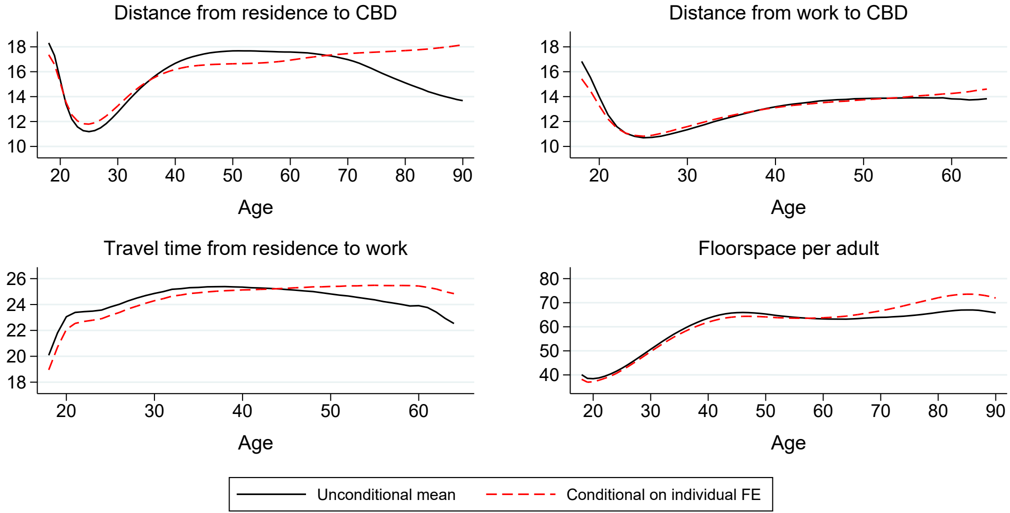


Data

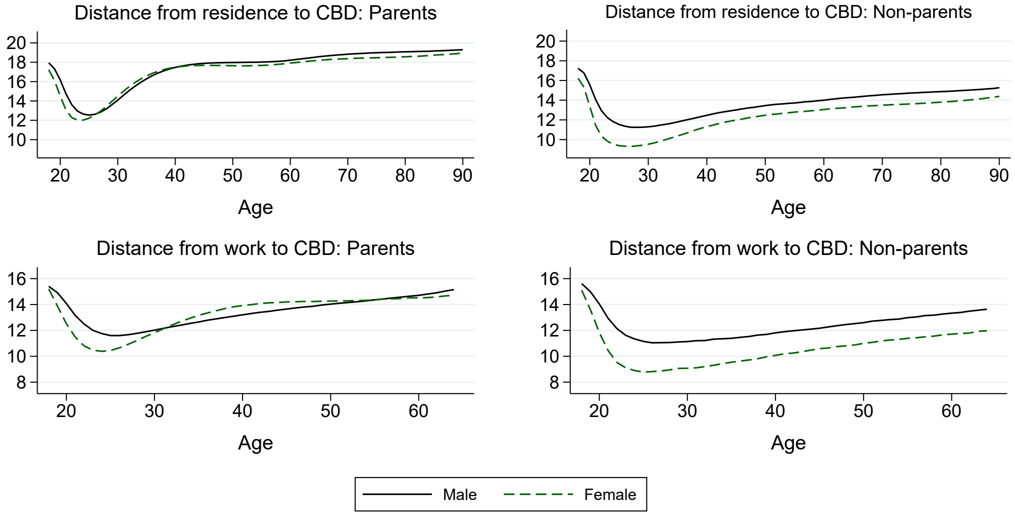
- Our main dataset is an annual population panel of both workers and the non-working population in the **Copenhagen metro area starting in 1983**.
- For each person we observe the following information:
  - Residence and workplace (if working) **location** in 100 x 100m grid cells.
  - Wage and non-wage **income and sector** of employment (if working).
  - Size and type of **residence** including estimates of the square meter price.
  - **Family status**, including number and age of children and marital status.
- We have the same data also for other parts of Denmark and see when people move away from or into Copenhagen.
- We combine this data with detailed information on the **geography** of Copenhagen including travel times by several different modes.

Stylized Facts: Age

Age(ing) in cities



Parents Versus Non-Parents and Gender Gaps



Stylized Facts: Life Events

Empirical Strategy Life Events

- We estimate **event-study regressions** to determine to what extent the patterns across age groups are caused by life events.
- We consider the following **life events**: cohabitation, children, separation, empty nesting, retirement, and death of the spouse (which can all repeat).
- For estimation we consider all life events that happen to **at least 2.5% of the people in our sample**, but here concentrate on the most frequent events.
- For efficiency, we run separate regressions for early and late life events (above and below median occurance at age 40). ▸ Early life ▸ Late life
- Empirically, the timing and sequence of life events varies a lot between different individuals. ▸ Frequency Early life ▸ Frequency Late life

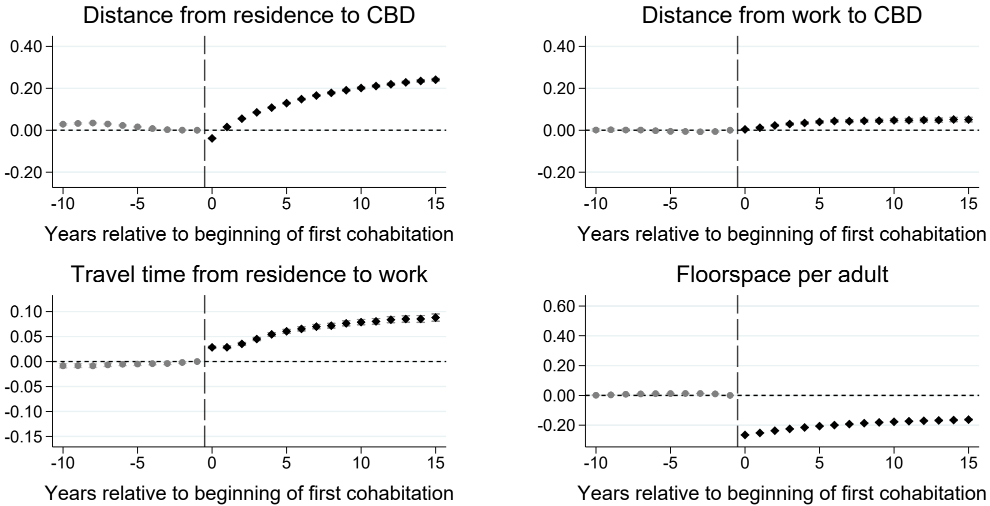
Regression Specification

- We estimate the following event-study regression using a variant of the imputation method (Borusyak, Jaravel and Spiess, 2024):

$$y_{it} = \alpha_i + \eta_s + \sum_{e \in \mathbb{E}} \sum_{\substack{h=-a \\ h \neq -1}}^{b} \beta_h^e \mathbb{1}[K_{it}^e = h] + \varepsilon_{i,t}$$

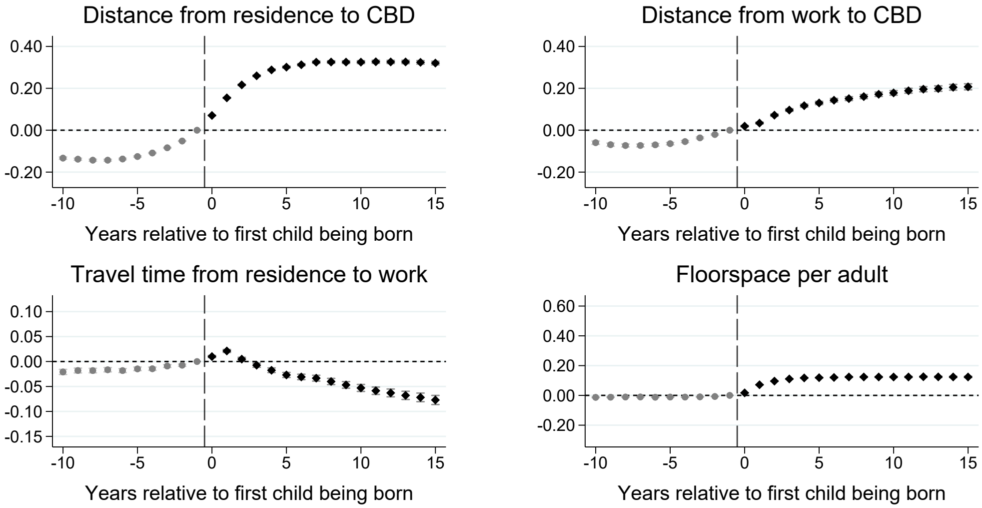
  - y_{it}: Outcome of worker i in year t
  - α_i, η_s: Individual (cohort) and age fixed effects
  - $K_{it}^e = t - E_i^e$ the difference between the current year (t) and the year in which individual i experiences event e (E_i^e), and $\mathbb{1}[K_{it}^e = h]$ is a dummy for difference h.
  - β_h^e: are the treatment effects of the a leads and b lags of life event $e \in \mathbb{E}$, where $\mathbb{E}$ can either be the early or late life events.
  - The regressions contain all leads and lags but the graphs show -10 to +15.
  - **One OLS regression with leads and lags** on imputed outcome to accommodate multiple treatments and avoid artificial jumps under pre-trends (Roth, 2024)

First Cohabitation



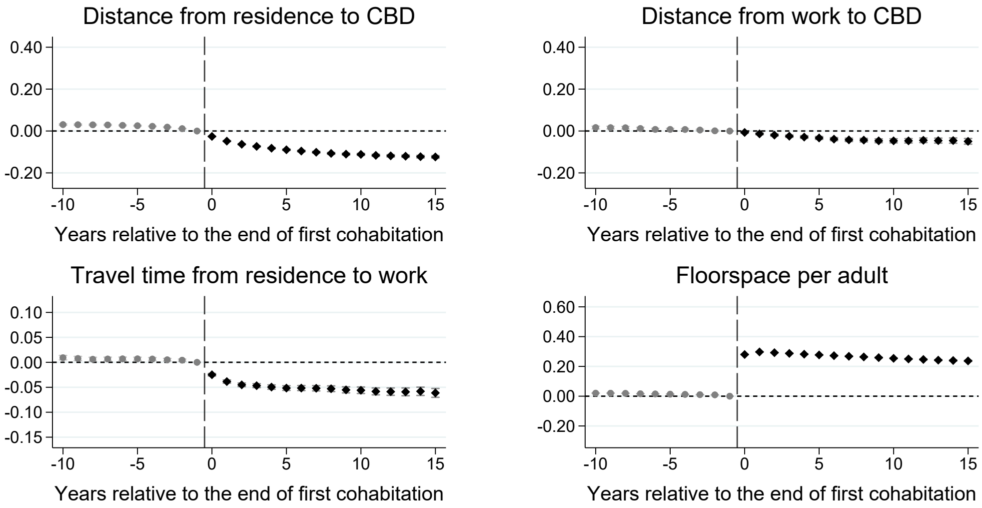
By Gender OLS Robustness

First Child

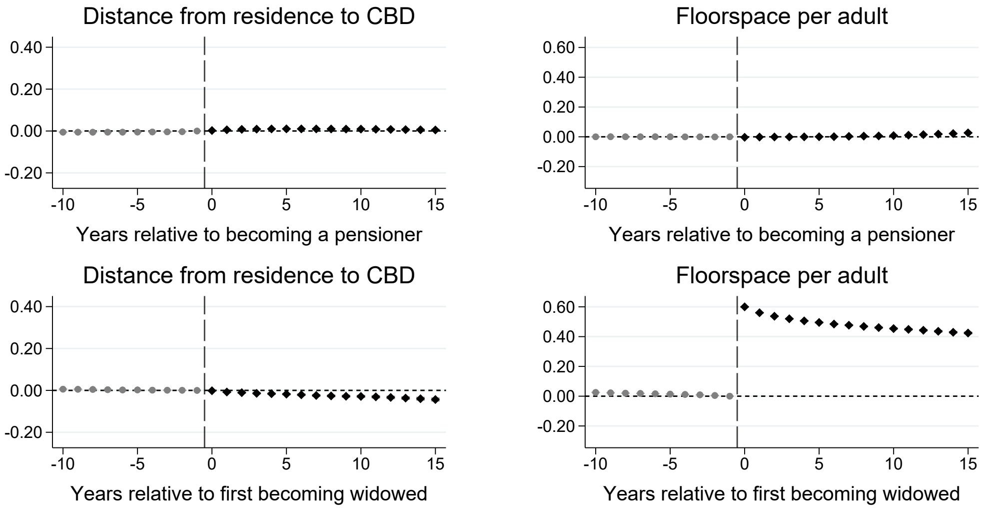


By Gender
Second Child

First Separation



Retirement and Death of Spouse



Theoretical Framework

Model Overview

- We develop a quantitative urban model in the tradition of Ahlfeldt et al. (2015) which differs from the exiting literature in three main ways:
  - Different **worker types** (low/high skilled and young/old)
  - **Non-working population** (pensioners and students)
  - Workers can have different **family types** (married, children etc.), which affect commuting costs, housing expenditure and preferences over amenities.
- The model is **static** and captures the steady-state distribution of different types of agents in space. We want to use the model to find
  - **unobserved amenity values** different groups assign to different locations
  - the counterfactual under a **different endowment with worker groups**

▸ Groups

Model Setup

- The city consists of locations that are connected by a transport technology.
- Workers and non-working agents consume a final good and floor space and value residential amenities depending on their family type f and occupation o.
- Workers choose where to live and work, while non-workers only do the first.
- Firms use labour and floor space as inputs to produce a freely tradable good.
- In production, **workers are perfectly substitutable across family type, but not across skill and age groups**.
- Floor space is produced using capital and land and optimally allocated.
- City is closed (worker group endowment is forcing variable in counterfactuals).
- All markets are competitive.

Preferences and Production

- Indirect utility of worker ω living in location n, working in location i, of occupation o and family type f is:

$$U_{ni}^{of}(\omega) = \frac{B_{ni}^{of} w_i^o z_{ni}^{of}(\omega)}{\kappa_{ni}^{of}(P_n)^{\alpha^{of}}(Q_n)^{1-\alpha^{of}}} \quad 0 < \alpha^{of} < 1. \tag{1}$$

- Indirect utility function of non-worker ρ of group r living in n is:

$$U_n^r(\rho) = \frac{B_n^r \bar{w}^r z_n^r(\rho)}{(P_n)^{\alpha^r}(Q_n)^{1-\alpha^r}} \quad 0 < \alpha^r < 1 \tag{2}$$

- Output (Y_i) in i is produced using all types of labour (L_i^o) and floor space (H_i):

$$Y_i = A_i \prod_{o\in\mathbb{O}} \left(\frac{L_i^o}{\beta_i^o}\right)^{\beta_i^o} \left(\frac{H_i}{\beta^H}\right)^{\beta^H}, \quad 0 < \beta_i^o, \beta^H < 1, \quad \sum_{o\in\mathbb{O}} \beta_i^o + \beta^H = 1 \tag{3}$$

Quantification

Estimation of Key Model Parameters

- We estimate relative housing expenditure shares (α^{of} and α^{r}) for all combinations of workers and non workers (17 groups). ▸ Details
- We estimate gravity commuting equations for all family and worker type combinations (12 groups) using PPML. ▸ Details
- We estimate Frechet shape parameters (ϵ^{of} and ϵ^{r}) for all worker and non-worker types. ▸ Details
- We calibrate location specific labour input shares (β_i^{o}) using the observed composition of employment across locations. ▸ Details
- We set the share of floor space in total production costs (β^{H}) to 0.15.

The Role of Residential Amenities

- The model inversion suggests that a substantial part of the variation in **location choices** is **due to differences in preferences over local amenities**.
  - **Singles and childless couples prefer dense areas** over more peripheral locations.
  - **Pensioners dislike dense areas** in the center of Copenhagen and prefer suburbs.
  - **High-skilled** have a stronger **preference for dense areas** than the low-skilled.

Counterfactuals

Counterfactuals

- We use model counterfactuals to explore how the striking differences in location preferences will reshape cities through demographic change.
- We consider three different model counterfactuals:
  1. An **increase** in the share of the **old** (40+) population by 10%.
  2. A **decrease** in the share of **families** with children until the Total Fertility Rate (TFR) reaches 1.
  3. An **increase** in the share of **single households** by 10%.
- Today we will focus on the first two counterfactuals.

Aging Counterfactual: Increase in the 40+ Population by 10%

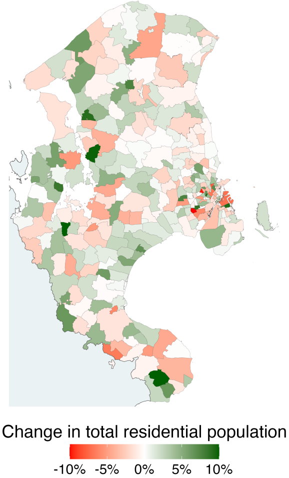


Figure: Residential population

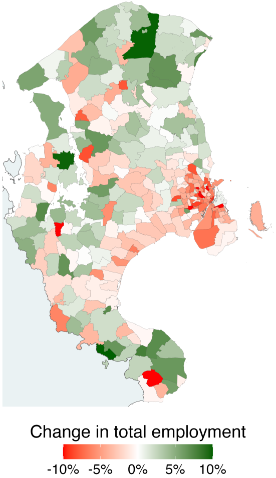


Figure: Employment

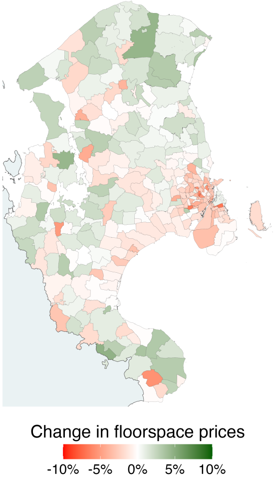


Figure: Residential prices

Lower Fertility Counterfactual: Fertility Drops to 1 Child per Women

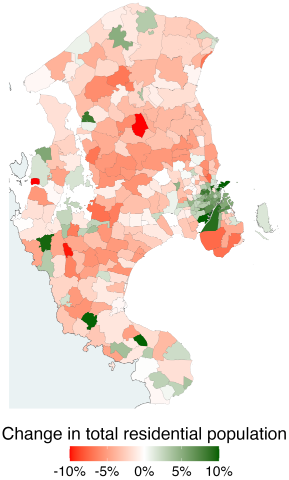


Figure: Residential population

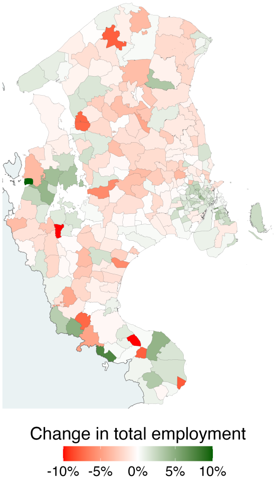


Figure: Employment

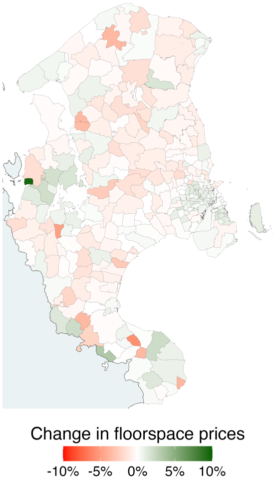


Figure: Residential prices

Conclusion

- This paper provides evidence that **age and life events** have a substantial effect on the **spatial sorting** of people across locations within cities.
- We use a quantitative spatial model estimate to uncover how the striking **sorting** of different groups in cities is **driven by amenity preferences**.
- We use **counterfactuals** to show how demographic changes that alter the **composition** of a cities population affect housing prices and sorting in cities.
- The results suggest that **demographic changes** such as fertility or aging can **change the geography of cities** substantially.

Appendix

Early Life Events

Table: Age Distribution of Early Life Events

Event	p10	p50	p90	Treated Individuals	Share of sample (%)
First Child	23	29	36	660503	22.3
Second Child	25	31	38	517545	17.5
Third Child	28	34	41	172159	5.8
First Cohabitation	21	26	41	870719	29.4
Second Cohabitation	25	32	51	498638	16.8
Third Cohabitation	28	37	54	145563	4.9
First Separation	22	31	55	804221	27.2
Second Separation	26	36	54	241615	8.2

Back

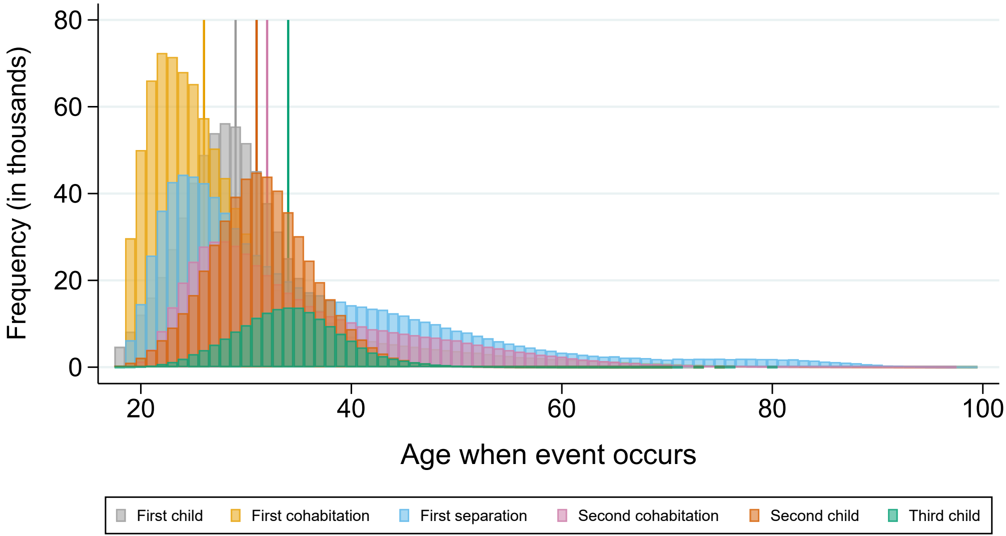
Late Life Events

Table: Age Distribution of Late Life Events

Event	p10	p50	p90	Treated Individuals	Share of sample (%)
Empty Nest	42	52	62	630665	21.3
Pension	49	62	67	671887	22.7
First Widowhood	52	70	84	201439	6.8
First Late-Life Separation	40	56	77	118477	4.0

▸ Back

Frequency of Early Life Events by Age

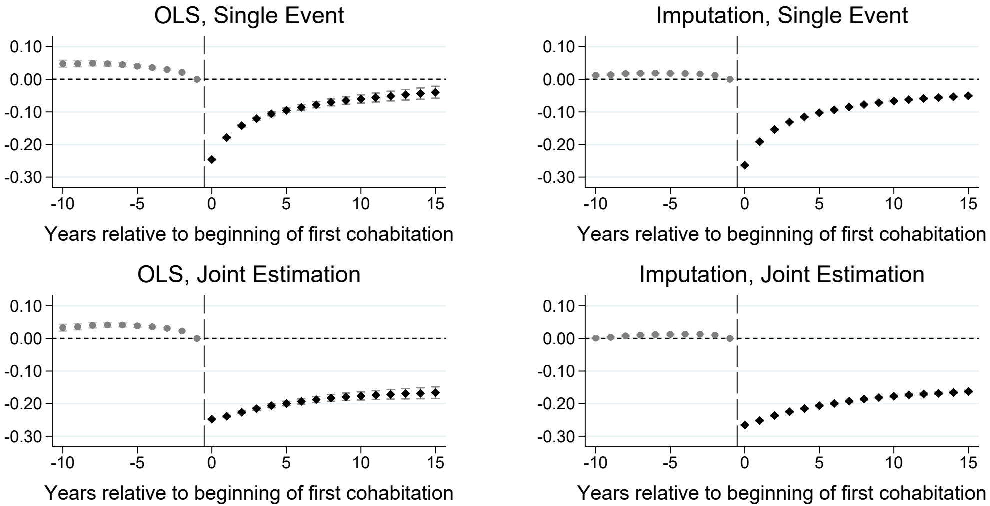


Back

Frequency of Late Life Events by Age

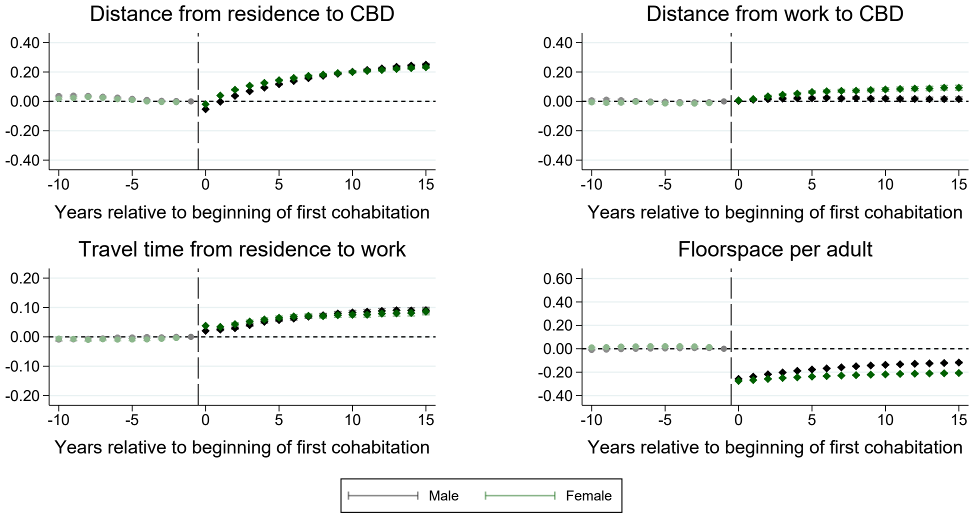
▸ Back

Impact on floor space per adult: Imputation versus OLS



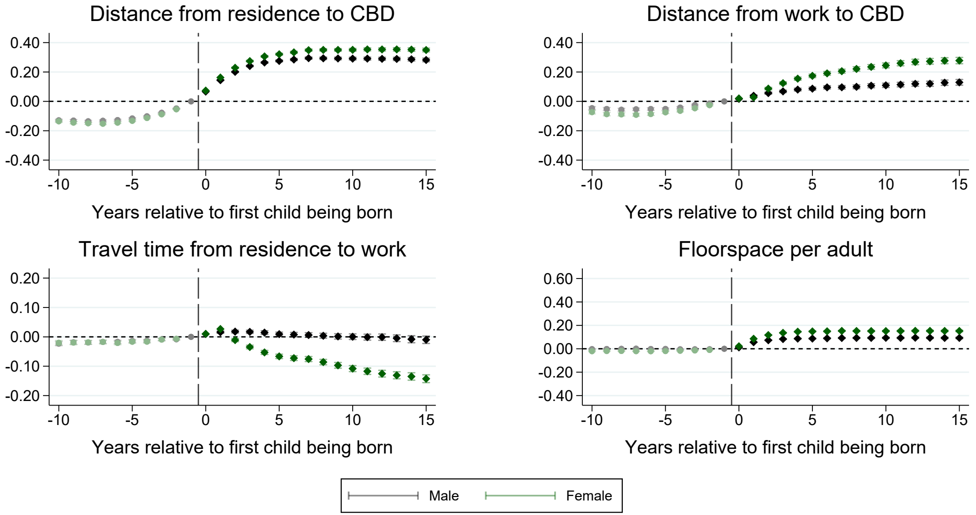
Back

First Cohabitation by Gender



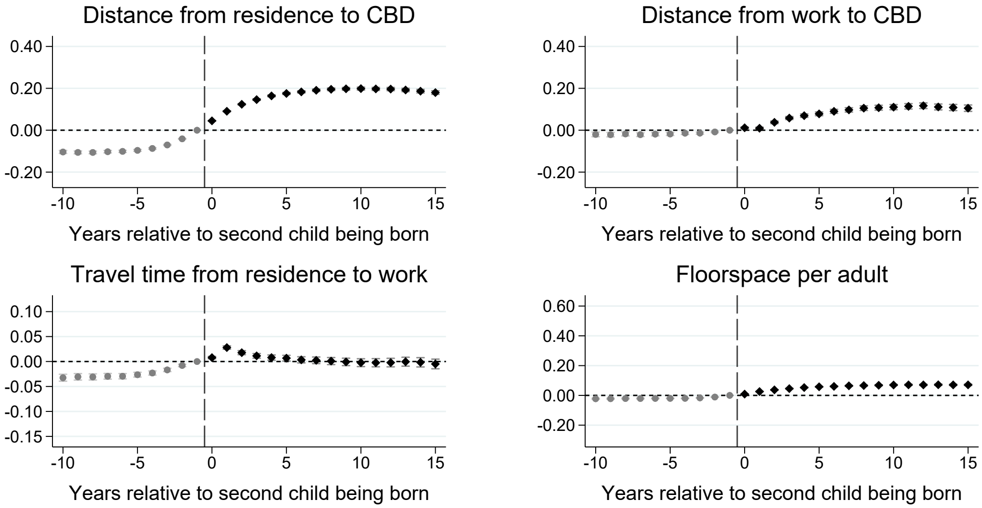
Back

First Child by Gender



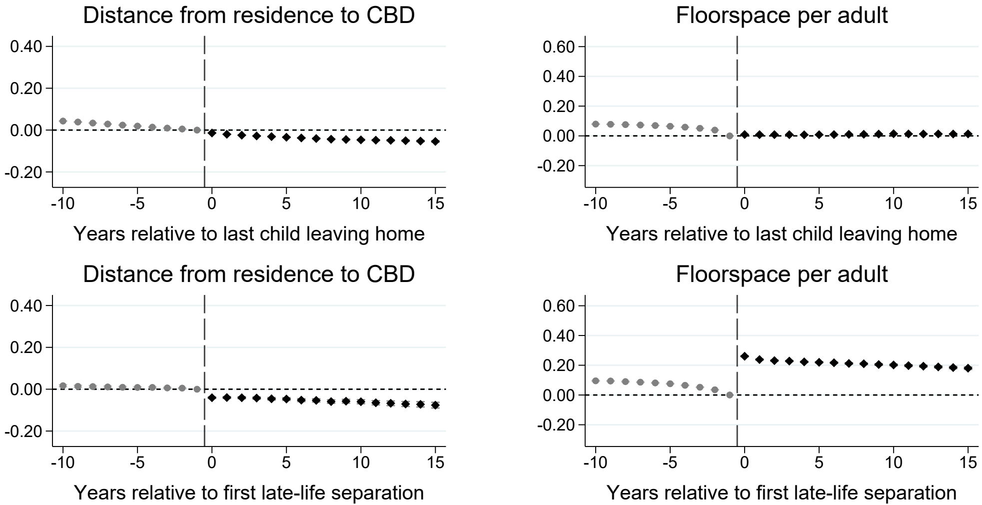
▸ Back

Second Child



Back

Empty Nest and Late Life Separation



▸ Back

Model Groups

Table: Overview of Model Groups

	Age	Skill	Family type
Non-workers	Students	-	Single
	Pensioners	LS, HS	Single, Cohabiting
Workers	Young worker	LS, HS	Single, Cohabiting, Cohabiting with Children
	Senior worker	LS, HS	Single, Cohabiting, Cohabiting with Children

Back

Calibration Details

- Housing expenditure shares per group are calibrated by group.
  - We calibrate the annual yield to target 30% aggregate housing expenditure, according to the Danish Sage's report of the Danish Economy in 2021.
  - We consider the residential floor space price index of Ahlfeldt, Heblich and Seidel, 2023 and annual net income to estimate the individual housing expenditure.
- The overall housing share in production is $\beta^H = 0.15$ following the report 'Produktivitet 2021' from The Danish Sages.
- The occupation-specific labour input shares (β_i^o) are obtained using the model implied wage bill shares by group for each location.

Back

Housing expenditure shares

Table: Estimating α_g

Group	α_g	m^2 quantity	m^2 price	Net income
Population	30.0 %	100.0	100.0	100.0
Student	39.4 %	79.7	113.9	64.2
Young, single, low-skill	34.8 %	91.9	100.2	73.9
Young, single, high-skill	33.1 %	102.9	124.6	108.3
Young, cohabiting, low-skill	27.3 %	82.9	94.0	83.6
Young, cohabiting, high-skill	26.3 %	91.6	118.2	120.0
Young, cohabiting with children, low-skill	25.7 %	95.4	84.9	93.4
Young, cohabiting with children, high-skill	25.8 %	108.6	107.2	137.4
Senior, single, low-skill	33.1 %	101.2	100.2	84.5
Senior, single, high-skill	32.8 %	120.9	117.0	129.7
Senior, cohabiting, low-skill	24.8 %	86.6	94.0	101.3
Senior, cohabiting, high-skill	23.7 %	101.3	110.7	149.6
Senior, cohabiting with children, low-skill	24.9 %	93.8	87.8	103.2
Senior, cohabiting with children, high-skill	24.0 %	113.8	108.0	174.3
Pensioner, single, low-skill	35.9 %	118.4	98.3	92.1
Pensioner, single, high-skill	32.4 %	135.4	113.6	147.8
Pensioner, cohabiting, low-skill	31.7 %	89.1	88.8	81.0
Pensioner, cohabiting, high-skill	26.3 %	109.5	107.2	154.3

▸ Calibration

Gravity Equations

- To obtain the semi-elasticity of commuting flows to commuting time, we estimate a gravity equation at the parish level.
- We take flows and travel times at a much granular level and aggregate them to the parish level. Flows are simply summed, while commuting times are the average of smaller unit travel times, weighted by commuting flows.
- This avoids the issue of granularity, while making use of rich commuting flows data.
- We estimate the following gravity equation for each worker group:

$$\ln \pi_{ni}^{of} = \nu^{of} \tau_{ni} + \phi_n + \phi_i + u_{ni}$$

▸ Calibration

Estimating epsilon

- Epsilon is estimated using real wages per worker group in each location, such that the two moment conditions are satisfied:

    -

- The epsilon of the non-working groups is the same as the closes working group to them in terms of age, skill and family type.

Calibration