Location patterns in the greater Copenhagen area

Evidence from a residential sorting model

Ismir Mulalic

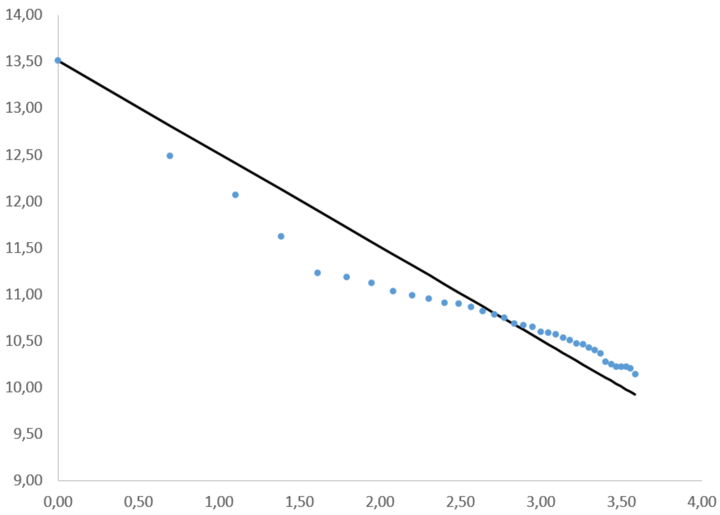
DTU Management Engineering & Kraks Fond Institute for Urban Economic Research

March 15

Good urban policy is vitally important

- The importance of cities in shaping the lives of billions makes undestanding how they work imortant.
- Do you remember Tikal, Babel, Ctesiphon, Mohenjo-daro, Mesa, Verde, Ani, Thebes, Vijayanagar, Persepolis, Palenque, Petra, Angkor, Carthage, Troy?
- Denmark:
  - Ribe (704–710)
  - *Hedeby (808)*
  - Århus (948)

Zipf's law


14,00
13,50
13,00
12,50
12,00
11,50
11,00
10,50
10,00
9,50
9,00
0,00
0,50
1,00
1,50
2,00
2,50
3,00
3,50
4,00

Smart City

- Cities are "*organised complexity*" (Jane Jacobs)
- Agglomeration:
  - It was essential overcome the "*tyrany of distance*"
  - Technological progress, social interaction and the excange of ideas.
- Cities have evolved rather like natural systems:
  - inovations introduced in one city, and if they worked (market places, public spaces, mass transit systems), they spred,
  - it yhey stopped being useful (e.g. city walls) they disappeared.
- Cities: labor market, consumption (amenities) and transport (derived demand).

The choices of residential location and car ownership are most likely interrelated

- Public and private transport are substitutes $\Leftrightarrow$ households make a choice which type of transport to use.
- The attractiveness of owning a car is related to the residential location:
  - The presence of many amenities at walking distance decreases the value of owning a car: the share of car-owners is lower in urban than in rural areas (Dargay (TRPE, 2002)).
  - Choice of a rural area implies in many cases the necessity to own a car.
  - Living in or close to city centers implies cruising for parking and parking fees, while accessibility of public transport is often much better.

This project

Structural approach

Horizontal (logit-based) Equilibrium Sorting Model (ESM)

Choice alternatives are combinations of:

- Geographical zone
- House type (single family – multifamily)
- Car ownership (0,1,2)

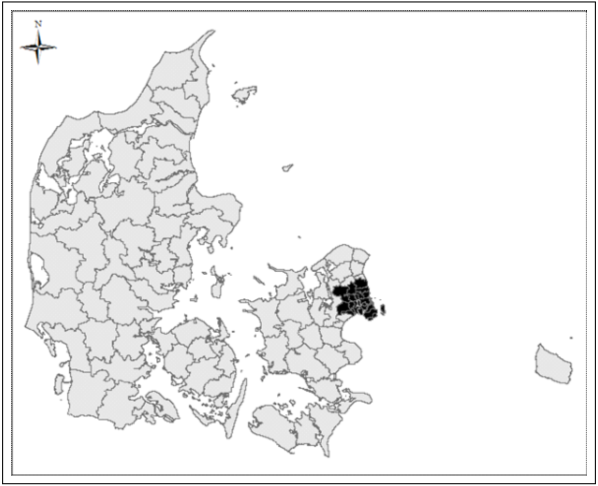
Distinction between single and dual earner households

Estimated for Greater Copenhagen Area (GCA)
Simulation of the impact of an extension of the metro network

This presentation

1. The study area
2. The model
3. Estimation results
4. Simulation

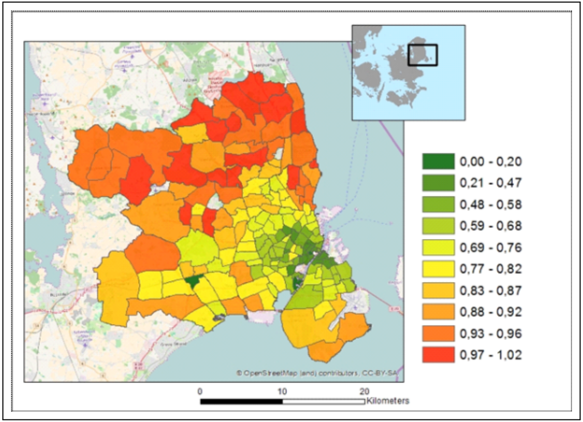
The Greater Copenhagen Area (GCA)



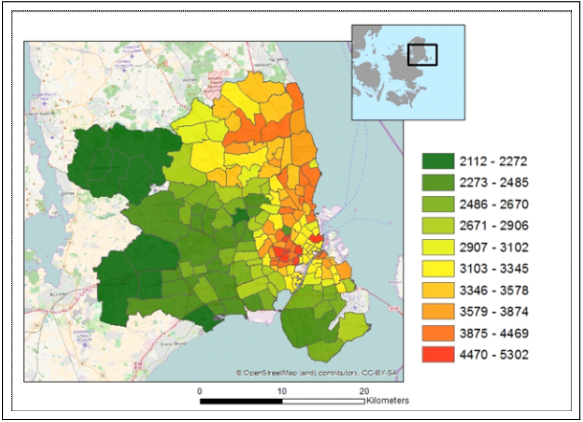
The Greater Copenhagen Area (GCA)

- Copenhagen (the capital city of Denmark) is the centre of the GCA.
- The GCA is the political, administrative, and educational centre of Denmark.
- The GCA accounts for more than 40% of Denmark's GDP, 1.6 mio. people (app. one third of Danish population), and 1 million workplaces.

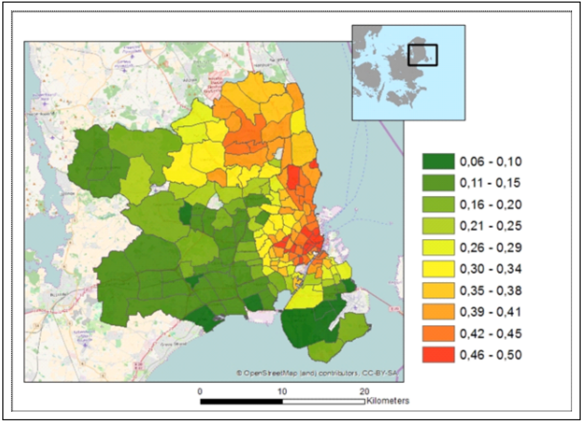
Car ownership (number of cars per household)



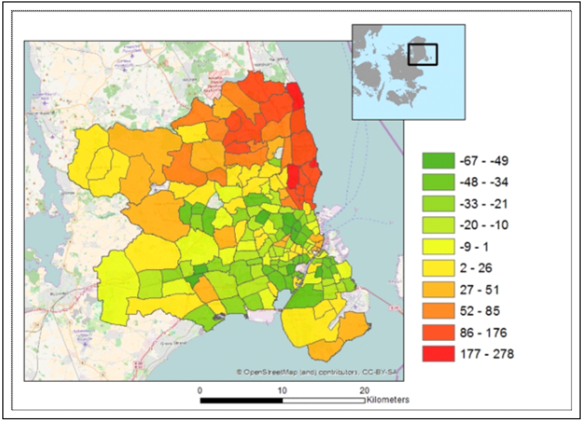
Std. housing price (1000 DKK)



Higher educated (share)



Households income (dev. from the average income)



Preliminaries

- We estimate a version of "*horisontal*" sorting model of the type proposed by Bayer and Timmins (2007 EJ) and Kuminof et. al. (2013 JEL).
  - The choice alternatives are combinations of residential areas and car ownership.
- The model includes residential area characteristics.
- The methodology we use is based on Berry et al. (Econometrica, 1995) and Bayer et al. (JPE, 2007):
  - *basically a logit model of the Berry-Levinsohn-Pakes type (BLP).*

A discrete choice model and its implications for car ownership

- We consider households who derive utility from housing, owning a car, local amenities and a composite that represents all other consumption goods.
- A household considers living in a residential area with and without having a car and chooses the alternative that offers the highest utility.
- Car ownership is included as a simple indicator that takes on the dichotomous values of 0 and 1 $\Rightarrow$ *we ignore the heterogeneity of cars in the interest of focusing on the interaction between the availability of public transport and car ownership*.

Housing services

- Housing services are available at a given price per unit that is specific for the residential area.
- The number of units consumed is determined by choosing from the stock or adjusting an existing house (Muth (1969), Epple and Platt (JUE, 1998), Rouwendal (1998)).
  - this allows the researchers to abstract from heterogeneity in the housing stock.
- The neglect of the durable aspects of housing may be problematic if quality differences are substantial $\Rightarrow$ *we distinguish between single and multifamily housing*.

Our model

- The utilities (area $a = 1 \dots n$, house type $h = s, m$, and car ownership $c = 0, 1$):

$$u^i_{a,h,c} = v^i_{a,h,c} + \varepsilon^i_{a,h,c}$$

- We assume that the random term $(\varepsilon^i_{a,h,c})$ are multivariate extreme value (MEV) distributed $\Rightarrow$ characterized by a generator function $G\left(e^{v^i_{a,h,c}}\right)$ where $e^{v^i_{a,h,c}}$ is the vector of the exponentiated deterministic parts of the utilities:

$$\pi^i_{a,h,c} = \frac{e^{v^i_{a,h,c}} G_{a,h,c}\left(e^{v^i}\right)}{G\left(e^{v^i_{a,h,c}}\right)}$$

- If $G\left(e^{v^i_{a,h,c}}\right) = \sum_a \sum_h \sum_c e^{v^i_{a,h,c}} \Rightarrow$ the choice probabilities are given by the multinomial logit model (MNL).

Car ownership

- The consumer will own a car if the maximum utility of the alternatives in which a car is owned exceeds the maximum utility of the alternatives in which no car is owned

$$U_1^i = \max\left\{u_{a,h,c}^i \middle| c = 1\right\} = \ln\left(\sum_a \sum_h e^{v_{a,h,1}^i}\right) + \varepsilon_1^i$$

  the random term ε_1^i iid Extreme Value Type I distributed.
- The probability of car ownership:

$$\pi_{c=1}^i = \frac{e^{\ln\left(\sum_a \sum_h e^{v_{a,h,1}^i}\right)}}{e^{\ln\left(\sum_a \sum_h e^{v_{a,h,1}^i}\right)} + e^{\ln\left(\sum_a \sum_h e^{v_{a,h,0}^i}\right)}}$$

Car ownership

- Our model differs from one in which we estimate car ownership *conditional on the choice of a residential area* and housing type:

$$\pi^i_{c=1} = \frac{e^{v^i_{a,h,1}}}{e^{v^i_{a,h,1}} + e^{v^i_{a,h,0}}}$$

  - compares the utility a household would be able to reach with and without owning a car in a given neighborhood.
- **Our model allows the consumer to choose a different neighborhood and housing type depending on whether a car will be owned.**

The impact of public transport on car ownership

- In our empirical model we use two variables: accessibility of jobs through public transport and accessibility of the metro network:
  - nonnegative impact on the utility of all choice alternatives
  - the impact on the utility of a given residential area and housing type without a car is at least as large as that on utility with a car
- The CV of car ownership will never increase when public transport improves $\Rightarrow$ **improving public transport will have a nonpositive impact on car ownership.**

The utility function

Combines car ownership and housing/area choice:

$$v^i_{a,h,c}(apt_a, amt_a, d_c, d_h, P_{h,a}, X_a; y^i, Z^i) =$$

$$\underbrace{\alpha^i_1 apt_a + \alpha^i_2 amt_a + \alpha^i_3 d_c +}_{\text{public transport and car ownership}}$$

$$\underbrace{\beta^i_1 d_h + \beta^i_2 P_{h,a} + \beta^i_3 X_a +}_{\text{neighborhood amenities}}$$

$$\underbrace{\left(\gamma^i_1 apt_a + \gamma^i_2 amt_a + \gamma^i_3 d_h + \gamma^i_4 X_a\right) d_c +}_{\text{cross effects}}$$

$$\underbrace{\xi_{a,h,c}}_{\text{unobserved characteristics}}$$

In practice we do not use the full specification.

The coefficients are individual-specific

- We specify α, β, γ as

$$\alpha_{i,k} = \widetilde{\alpha}_i^0 + \widetilde{\alpha}_i^1 \ln\left(y^i\right) + \sum_{l=1}^{L} \widetilde{\alpha}_i^{l+1} Z_l^i$$

  where Z_l^i is the value of the l'th characteristic of household i.
- The household characteristics are demeaned $\Rightarrow \widetilde{\alpha}_j^i$ is the average value of the coefficients $\widetilde{\alpha}_j^i$ in the population.

The annual register data

- We use a 20% sample of the GCA population living in owner-occupied housing.
- The estimation is based on the data derived from the administrative register data for *owner-occupiers* with residence in the GCA for the year 2008 spread over 166 zones (designed for the purpose of detailed traffic modeling).

Selection of sample

- We distinguish between living in a house or an apartment in the GCA.
- We also distinguish between being a car owner or not in both housing situations.
- We estimate two models:
    - one referring to the single earner households (66, 012 households and 538 alternatives), and
    - one referring to the dual earners households (87, 330 households and 636 alternatives).

The socioeconomic variables

1. Age and age squared,
2. Three dummy variables indicating highest education obtained,
3. Number of children in household,
4. Households income
5. Dummy variable indicating singles (single earner households).

Local amenities

1. Standardized house and apartment prices (from the two separated hedonic models, i.e. one for the houses and one for the apartments),
2. Employment access (using the number of the full time job equivalents for each zone and the travel time by public transport),
3. Proximity to the nearest metro station (km),
4. Number of conserved/protected buildings per sq.km.,
5. Distance to the CBD,
6. Share of higher educated population,
7. Share of social housing, and
8. Parking charging.

Unobserved characteristics of alternatives

- Ignoring the unobserved characteristics of the alternatives will not affect the model if it is uncorrelated with the X's.
- Housing prices and unobserved location characteristics are most likely correlated.
- A possible solution: Berry et al. (1995 Econometrica) proposed to estimate the model in two steps (BLP):
  1. Estimate the alternative specific constants (asc's) and household-specific parameters in the MNL model
  2. Use mean utility estimates from Step 1 and estimate mean household preference parameters in regression model (endogeneity)!

Endogeneity

House prices

- Demand is affected by unobserved char.
- Predicted prices in the absence of unobserved characteristics
- Use equilibrium condition on housing market, Bayer et al. (JPE, 2007):
  - Calculate prices that clear housing market at all locations.
  - Instruments are effectively functions of exogenous variables X and housing supply at each location.

Endogeneity

Share of higher educated

- Determined by choice behavior that we study
- Affected by unobserved char. of alternatives
- Instrument: private schools from late 19th cent.
  - correlated with current concentrations of higher educated,
  - arguably independent of unobserved characteristics that are currently important.

Endogeneity

Accessibility of employment

- Location choice of firms (in some industries) affected by location of workers, households demanding their products...
- Instrument: stations constructed before WWII
    - not constructed to serve commuters
    - often still important public transport ‘hubs’

Decomp. of the mean utilities (single wage-earners)

Example

	OLS		IV	
	Estimate	Std.err.	Estimate	Std.err.
Log(standarized house/apartment price)	-2.178	0.324	-3.032	0.517
Share of higher educated	1.874	0.532	3.130	1.043
Number of conserved/protected buildings per sq.km.	0.937	0.167	0.903	0.167
Proximity to the nearest metro station * nocar	0.454	0.207	0.547	0.230
Dummy variable indicating one car	0.960	0.227	0.889	0.304

Decomp. of the mean utilities (dual wage-earners)

Example

	OLS		IV	
	Estimate	Std.err.	Estimate	Std.err.
Log(standarized house/apartment price)	-2.320	0.361	-3.357	0.651
Share of higher educated	2.644	0.586	3.880	1.255
Number of conserved/protected buildings per sq.km.	0.897	0.159	0.848	0.161
Proximity to the nearest metro station * nocar	0.712	0.215	0.800	0.236
Dummy variable indicating one car	1.728	0.298	1.770	0.392
Dummy variable indicating two cars	1.033	0.327	0.912	0.444

Decomposition of the mean utilities

- For the alternatives in which no car is owned, accessibility to employment by public transport and proximity to a metro station are important.
- Ownership of a car makes a choice alternative more attractive.
- Houses are preferred to apartments and a higher housing price makes an alternative less attractive.
- The presence of higher educated households and monuments make a zone more attractive and the presence of social housing has a negative impact.
- The interactions of car and neighbourhood characteristics have no significant impact on the average household.
- Having one or two cars is better than having none, but one car is clearly the preferred situation.

Interaction parameter estimates (single wage-earners)

Example

	Proximity to the nearest metro station * nocar	Dummy variable indicating one car
Log(households income)	-0.062 (0.062)	0.501 (0.082)
Age	0.019 (0.009)	-0.033 (0.011)
Age squared/1000	-0.243 (0.096)	0.329 (0.120)
Number of children in household	-0.054 (0.042)	0.221 (0.053)
Medium education	-0.069 (0.057)	0.211 (0.078)
Higher education	0.016 (0.059)	0.030 (0.087)
Singles	-0.109 (0.068)	-0.830 (0.088)

Interaction parameter estimates (dual wage-earners)

Example

	Proximity to the nearest metro station * nocar
Log(households income)	-0.561 (0.106)
Age, head of the household	0.043 (0.049)
Age squared / 1000, head of the household	-0.439 (0.518)
Medium education, head of the household	0.139 (0.091)
Higher education, head of the household	0.317 (0.095)
Age, partner	0.039 (0.052)
Age squared / 1000, partner	-0.668 (0.588)
Medium education, partner	0.295 (0.088)
Higher education, partner	0.273 (0.097)
Number of children in household	-0.185 (0.039)

Interaction parameter estimates

- The results show the importance of household income
  - higher income households are less sensitive to the availability of public transport if no car is owned, but owning a car becomes much more attractive
  - the sensitivity to the housing price decreases, but the presence of higher educated is appreciated more
  - the combination of a single family house and a car gets more important with income.
- Accessibility to public transport as well as owning a car become less important with age (at a decreasing rate).
- Households with children have stronger preferences for cars and single family houses.
- The combination of children and living in an area with parking charges is unattractive.

Nesting structures and endogeneity

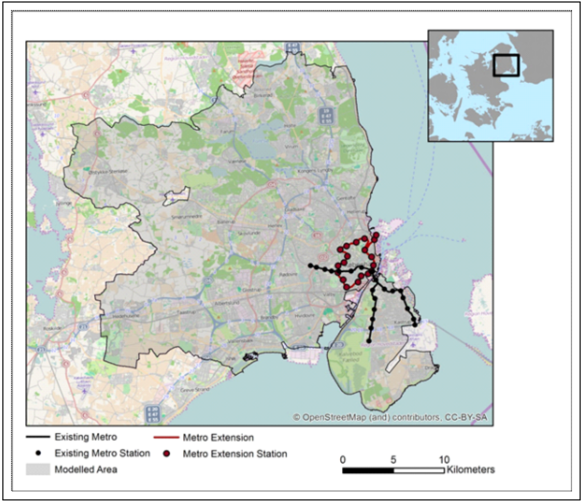
Nesting structures:

- Correlation between error terms ε seems a priori plausible
- Nesting possibilities:
  - Car ownership
  - Housing type
  - Zone
- Mixed logit allows for all simultaneously
  $u^i_{a,h,c} = v^i_{a,h,c} + \left(\theta^i_a + \mu^i_a + \sigma^i_a + \varepsilon^i_{a,h,c}\right)$
  - In principle we can let the "data speak"
- Endogeneity of house prices:
  - Instrument: share of divorced couples.
    The results remain robust.

Simulation study

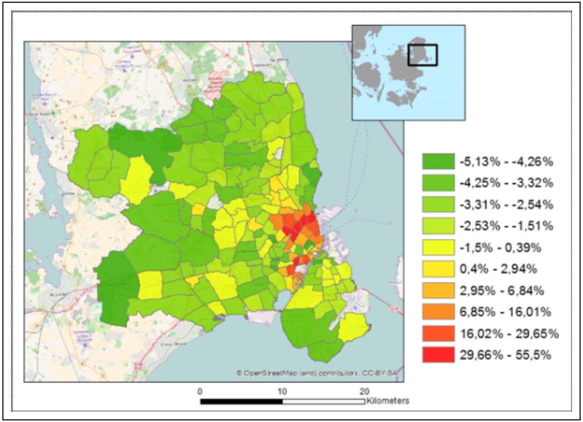
- Extension of metro network
- Impact:
  - Under ceteris paribus conditions
    - no mobility
  - With elastic housing supply
    - mobility but house prices are constant
  - With inelastic housing supply
    - prices equilibrate the market

The metro system extension in 2019



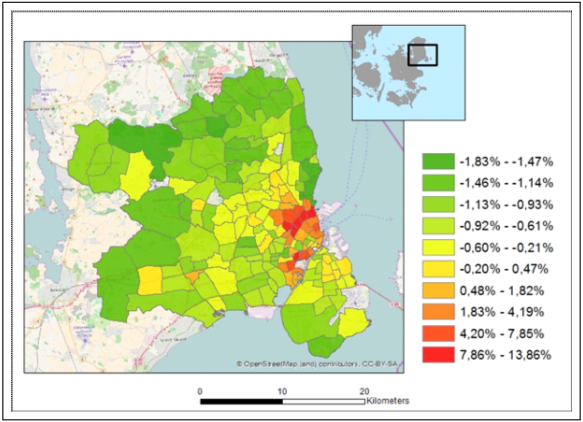
Results: with elastic housing supply

Households will tend to relocate closer to the CBD

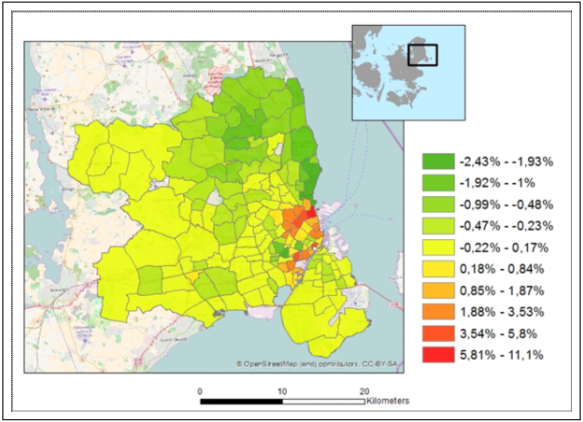


Results: with inelastic housing supply

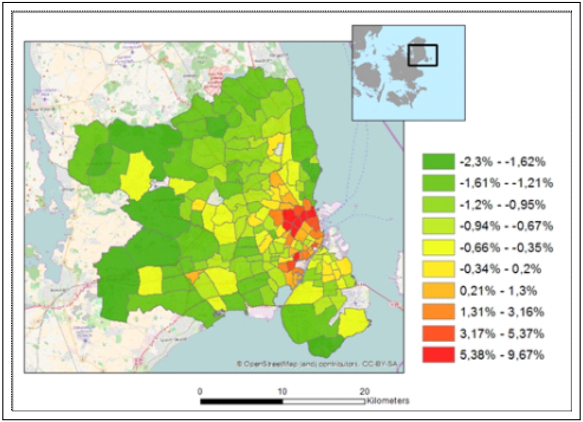
Prices equilibrate the market



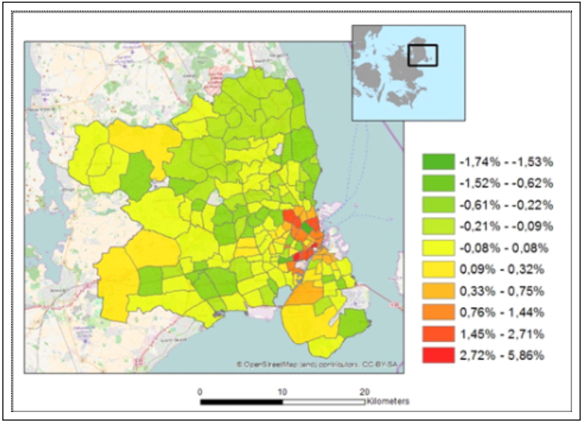
Change in household income



Pct. change in the share of higher educated



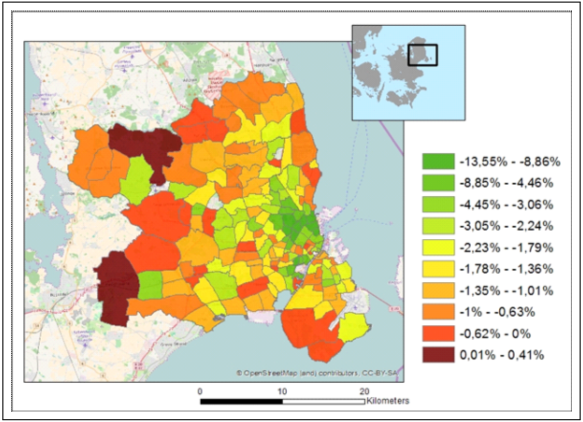
Pct. change in the number of households with children



Car ownership

	Reference scenario	Scenario 1	Scenario 2
		Fixed prices	Fixed supply
One car households	85,388	82,906	83,389
Two cars households	17,495	16,695	16,949
Total number of cars	120,378	116,295	117,287

Car ownership (percentage point change)



Compensating variations of the extension of the metro network

			[1] No mobility	[2] Elastic supply	[3] House prices adjust
	All	Av	11,062	12,026	11,899
Single		Av %	2.8	3.1	3.0
earner	Only affected	Av	33,753	34,386	24,324
	alternatives	Av %	8.6	8.7	6.2
	All	Av	13,271	13,669	13,012
		Av %	2,1	2.2	2.1
Dual	Dir. affected	Av	53,156	53,413	38,641
earners	(no car)	Av %	8.4	8.4	6.1
	Dir. affected	Av	12,019	12,412	3,518
	(one car)	Av %	1.9	2.0	0.6

Conclusion

- We developed a model for the joint choice of residential location and car ownership.
- Estimation results suggest a significant impact of metro network on attractiveness of zones and on car ownership.
- Simulations suggest a potentially large impact of the extension of metro network.
- Our results suggest that a place-based policy which focuses on areas close to attractive city centres will attract relatively wealthier households and most likely cause more segregation.