Household's location choice, transport accessibility and car ownership

Ismir Mulalic

Technical University of Denmark

September 2015

Lecture Agenda

1. What is transport demand?
2. Residential sorting model (Bayer et al. (2007) and Kuminoff et al. (2012))
3. Location choice, transport accessibility and car ownership (Mulalic et al., 2015)
4. Case study (Mulalic et al., 2015)

Bayer, P. and C. Timmins. 2007. Estimating equilibrium models of sorting across locations. Economic Journal, vol. 117(518), pp. 353-374.

Kuminoff, N.V., V.K. Smith and C. Timmins. 2012. The new economics of equilibrium sorting and policy evaluation using housing markets. Journal of Economic Literature, 51(4), pp. 1007-1064.

Mulalic, I., N. Pilegaard and J. Rouwendal. 2015. Public transport and car ownership: evidence from the Copenhagen metropolitan area. Manuscript.

What is transport demand?

- The demand for transport is a **derived demand**
  - an economic term, which refers to demand for one good or service in one sector occurring as a result of demand from another.
- The users of transport are very often consuming the service not because they benefit from consumption directly (*except in cases such as pleasure cruises*), but because they wish to partake in other consumption elsewhere, e.g. housing market, labor market, etc.

The household's location decision

- According to economic theory, *local wages*, *housing costs*, and *the cost of commuting* (accessibility to employment) represent the three most important economic determinants of the household's location decisions.
- More recent literature acknowledges that location choices can also be affected by urban amenities (Brueckner et al. (1999 EER); Glaeser et al. (2001 JOG)).
- *Many of urban amenities are transport related, e.g. access to public transportation.*
- We focus on a residential sorting model that allows us to identify determinants of the household's location decisions and to compute the marginal willingness to pay for the considered urban amenities.

Preliminaries

- We focus on a "*horisontal*" residential sorting model of the type proposed by Bayer and Timmins (2007 EJ) and Kuminof et. al. (2013 JEL).
- 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).

The model

We adopt the multinomial logit (**MNL**) model:

$$P_{in} = \frac{e^{u_{in}}}{\sum_m e^{u_{im}}}$$

The indirect utility of household i from choosing location n is

$$u_{in} = \sum_{k=1}^{K} \alpha_{ik} X_{kn} + \varepsilon_{in}$$

where X_{kn} is the value of the k-th characteristic of alternative n, the α's are coefficients, and the ε's are random variables that are iid. extreme value type I distributed.

Coefficients are related to household characteristics

The coefficients α's are individual-specific:

$$\alpha_{ik} = \beta_{0k} + \sum_{l=1}^{L} \beta_{kl} \left(Z_{il} - \overline{Z}_l \right)$$

where Z_{il} is the value of the l-th characteristic of household i, $\overline{Z}_l$ the population mean of characteristic l, and the β's are coefficients.

The utility function

We can rewrite the utility function as:

$$u_{in} = \sum_{k=1}^{K} \beta_{0k} X_{kn} + \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \beta_{kl} \left(Z_{il} - \overline{Z}_l \right) \right) X_{kn} + \varepsilon_{in}$$

This is similar to the error component formulation of the mixed logit model because the second term on its right-hand side can be interpreted as resulting from a random draw of a household from the population.

The MNL model has two important shortcomings

1. The restrictive Independence of Irrelevant Alternatives (IIA) property (also known as binary independence or the independence axiom),
2. In its standard form, the MNL model has difficulties in dealing with unobserved characteristics of alternatives.

IIA

- The IIA property implies that two residential locations with the same market shares will have the same cross-price derivative with respect to any third location $\Rightarrow$ in contrast with intuition since locations with similar characteristics should be close substitutes.
- By allowing for interaction between household and location characteristics the IIA assumption is to some extent relaxed:
  - different household types will have different market shares and by that different substitution patterns
  - on the aggregate level the IIA property does not hold.

Unobserved characteristics of alternatives

- Ignoring the unobserved characteristics of the alternatives (ξ) will not affect the model if it is uncorrelated with the X's:

$$u_{in} = \sum_{k=1}^{K} \beta_{0k} X_{kn} + \xi_n + \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \beta_{kl} \left(Z_{i,l} - \overline{Z}_l \right) \right) X_{kn} + \varepsilon_{in}$$

- **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):

$$u_{in} = \delta_n + \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \beta_{kl} \left(Z_{il} - \overline{Z}_l \right) \right) X_{kn} + \varepsilon_{in}$$

$$\delta_n = \sum_{k=1}^{K} \beta_{0k} X_{kn} + \xi_n$$

Step 1

Estimate the alternative specific constants (asc's) and household-specific parameters in the MNL model:

$$u_{in} = \delta_n + \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \beta_{kl} \left(Z_{il} - \overline{Z}_l \right) \right) X_{kn} + \varepsilon_{in}$$

Problem: Large number of alternatives gives large number of parameters.
Solution: Iterate the following steps:

1. For fixed asc's estimate household-specific parameters.
2. For fixed household parameters use fixed point algorithm to calculate asc's (observed and predicted location shares must be equal).

Step 2

Use mean utility estimates $\hat{\delta}_j$ from Step 1 and estimate mean household preference parameters in regression model:

$$\hat{\delta}_n = \sum_{k=1}^{K} \bar{\beta}_k X_{kn} + \xi_n$$

Housing prices and unobserved location characteristics are correlated $\Rightarrow$ use instruments for housing prices:

- Exclusion type instruments, Ferreira (2010, J Public Econ)
- Use equilibrium condition on housing market, Bayer et al. (JPE, 2007):
  - Calculate prices that clear housing market at all locations when $\xi_n = 0$.
  - Instruments are effectively functions of exogenous variables X_{kj} and housing supply at each location.

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.

Literature

- The interaction between car ownership and public transport seems to have been neglected in recent decades.
- It has been addressed in an older literature, e.g. Goodwin (TRANSP, 1993):
  - road congestion and pollution by cars,
  - cities as green places (Kahn, 2006),
  - geographic literature on the impact of urban form and urban amenities on car ownership,
  - literature in economics looking at car ownership.

Preliminaries

- The choice alternatives are combinations of residential areas and car ownership.
- The model explains car ownership while paying special attention to its relationship with residential area characteristics.

A discrete choice model and its implications for car ownership

- A household considers living in a residential area with and without having a car and chooses the alternative that offers the highest utility.
- We consider households who derive utility from housing, owning a car, local amenities and a composite that represents all other consumption goods.
- 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)).
- 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*.

The choice alternatives

- Choice alternatives are defined by three variables: area ($a = 1 \ldots A$), house type ($h = s, m$), and car ownership ($c = 0, 1$).
- The utility of a choice alternative for household i

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

- The multinomial logit model (MNL)

$$\pi^i_{a,h,c} = \frac{e^{v^i_{a,h,c}}}{\sum_{a'} \sum_{h'} \sum_{c'} e^{v^i_{a',h',c'}}}$$

The model specification

The deterministic part (for single earner households) of the utility of a choice alternative is

$$\begin{aligned} v^i_{a,h,c} &= \alpha^i_1 opt_a + \alpha^i_2 amt_a + \alpha^i_3 d_c + \\ &\quad \beta^i_1 d_h + \beta^i_2 P_{h,a} + \beta^i_3 X_a + \\ &\quad \left(\gamma^i_1 opt_a + \gamma^i_2 amt_a + \gamma^i_3 d_h + \gamma^i_4 X_a\right) d_c + \xi_{a,h,c} \end{aligned}$$

where X_a is the value of the a-th characteristic of alternative a, the α, β, γ are coefficients, and the ξ reflects unobserved (by the researcher) characteristic of the alternative.

Coefficients are related to household characteristics

The coefficients α, β and γ are individual-specific:

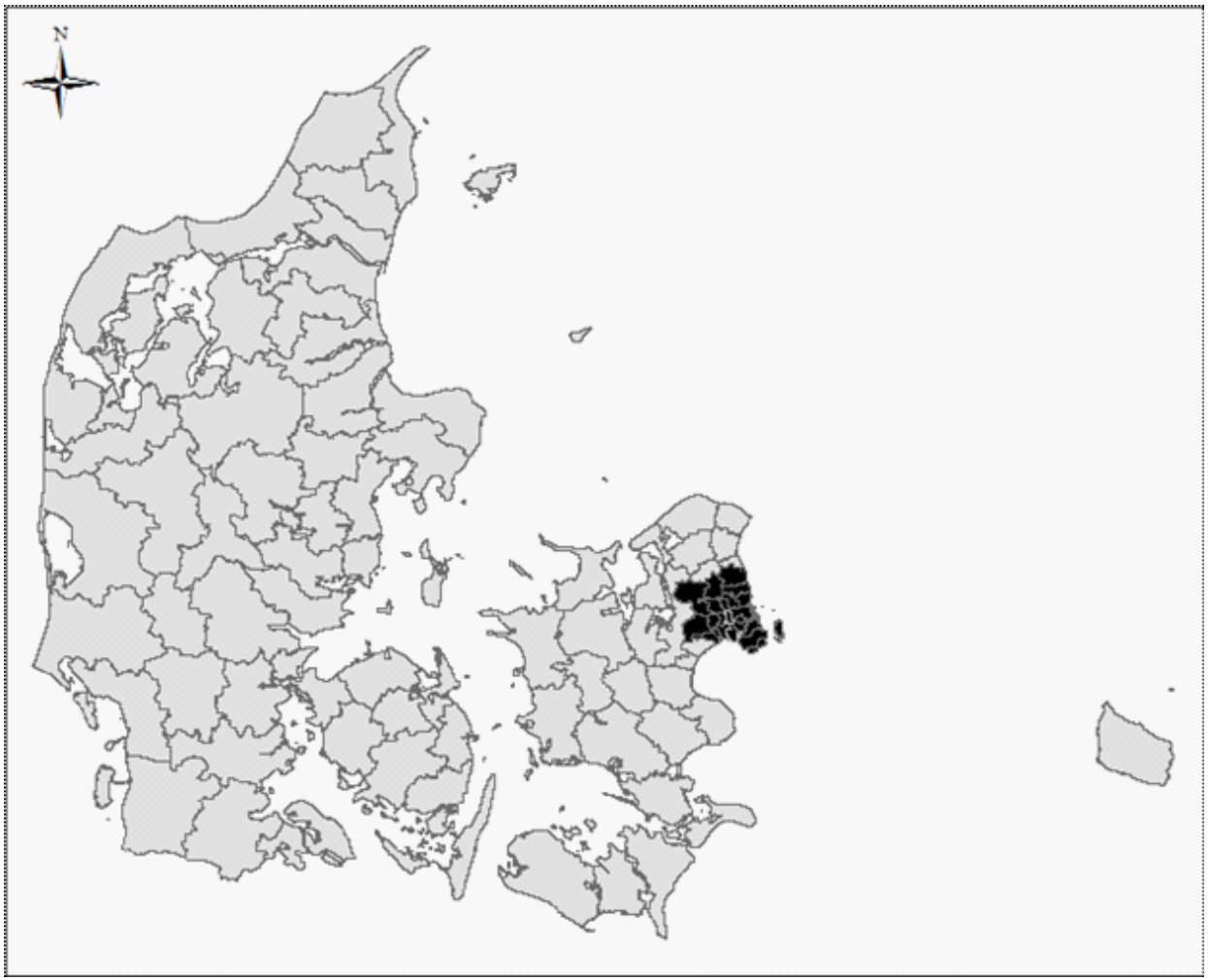
$$\alpha_j^i = \widetilde{\alpha}_j^0 + \widetilde{\alpha}_j^1 y^i + \sum_{l=1}^{L} \widetilde{\alpha}_j^{l+1} \left(Z_{il} - \overline{Z}_l \right)$$

where Z_{il} is the value of the l-th characteristic of household i, $\overline{Z}_l$ the population mean of characteristic l, and the α's are coefficients. We have and analogous expressions for the βs and γs.

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.

The Greater Copenhagen Area (GCA)



The annual register data

- We use a 20% sample of the GCA population living in owner-occupied housing.
- The estimation of the equilibrium sorting model 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 modelling).

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).

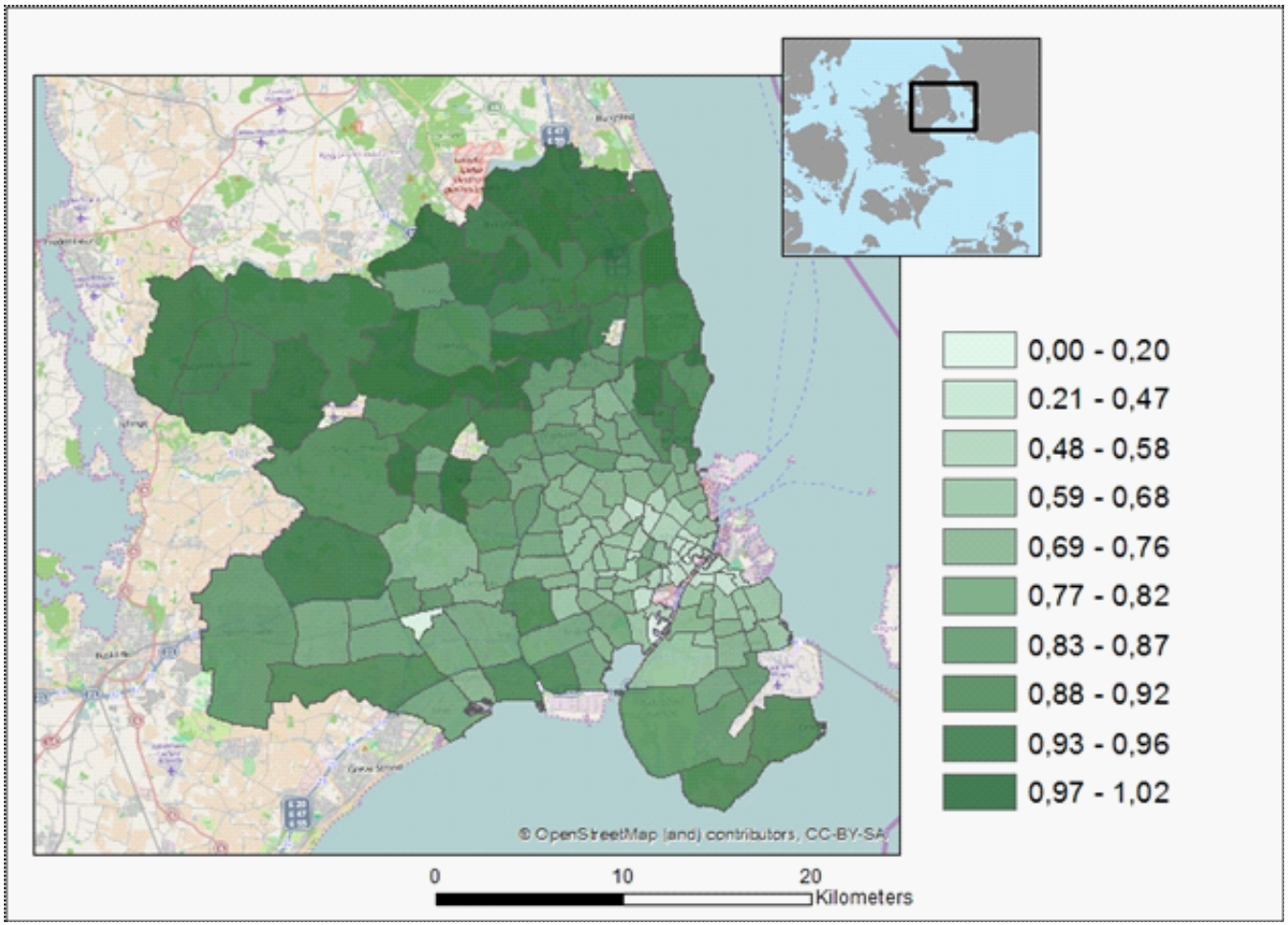
The socioeconomic variables

1. Age and age squared,
2. Three dummy variables indicating highest education obtained,
3. Number of children in household,
4. Households income.

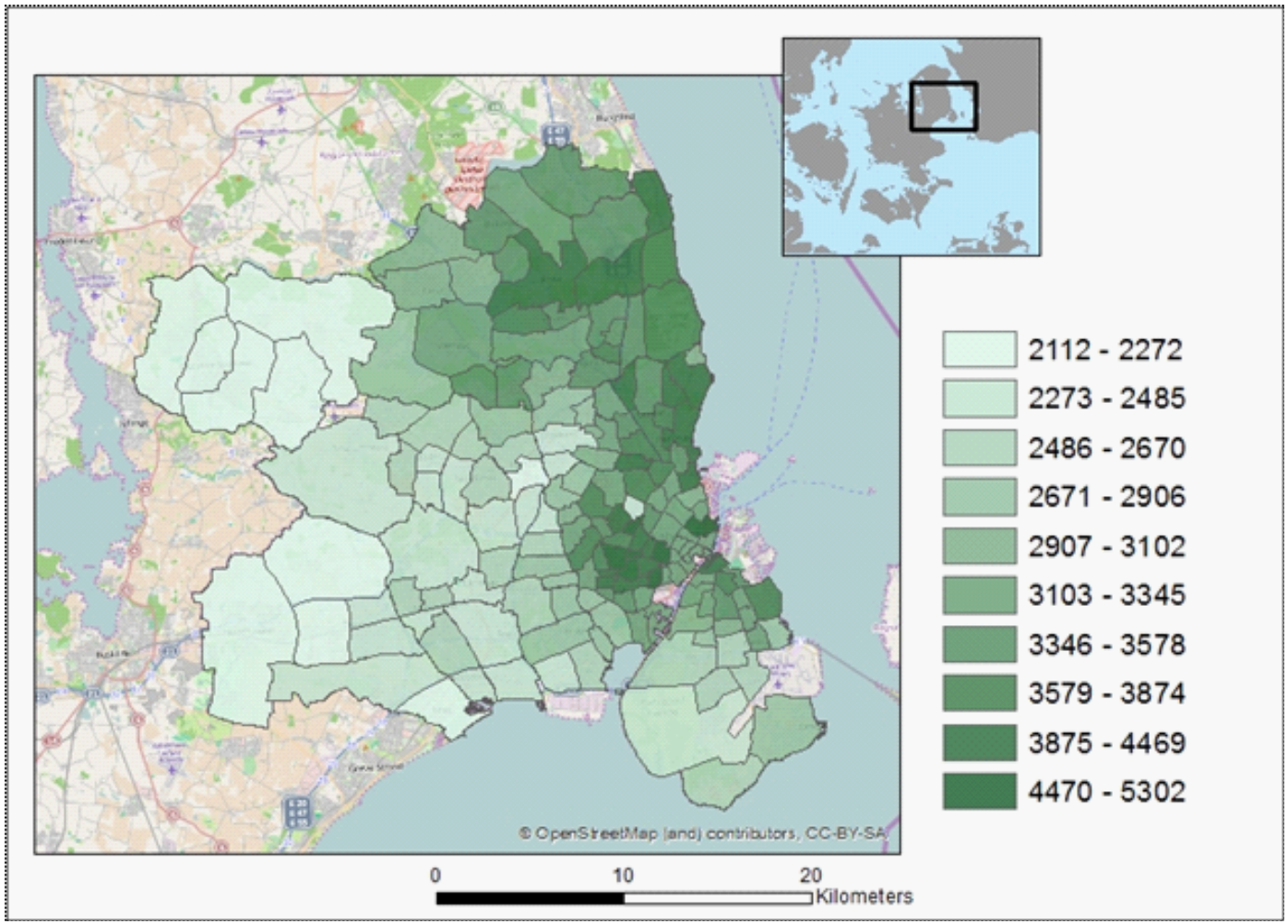
Local amenities

1. Standardized house/apartment price (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.

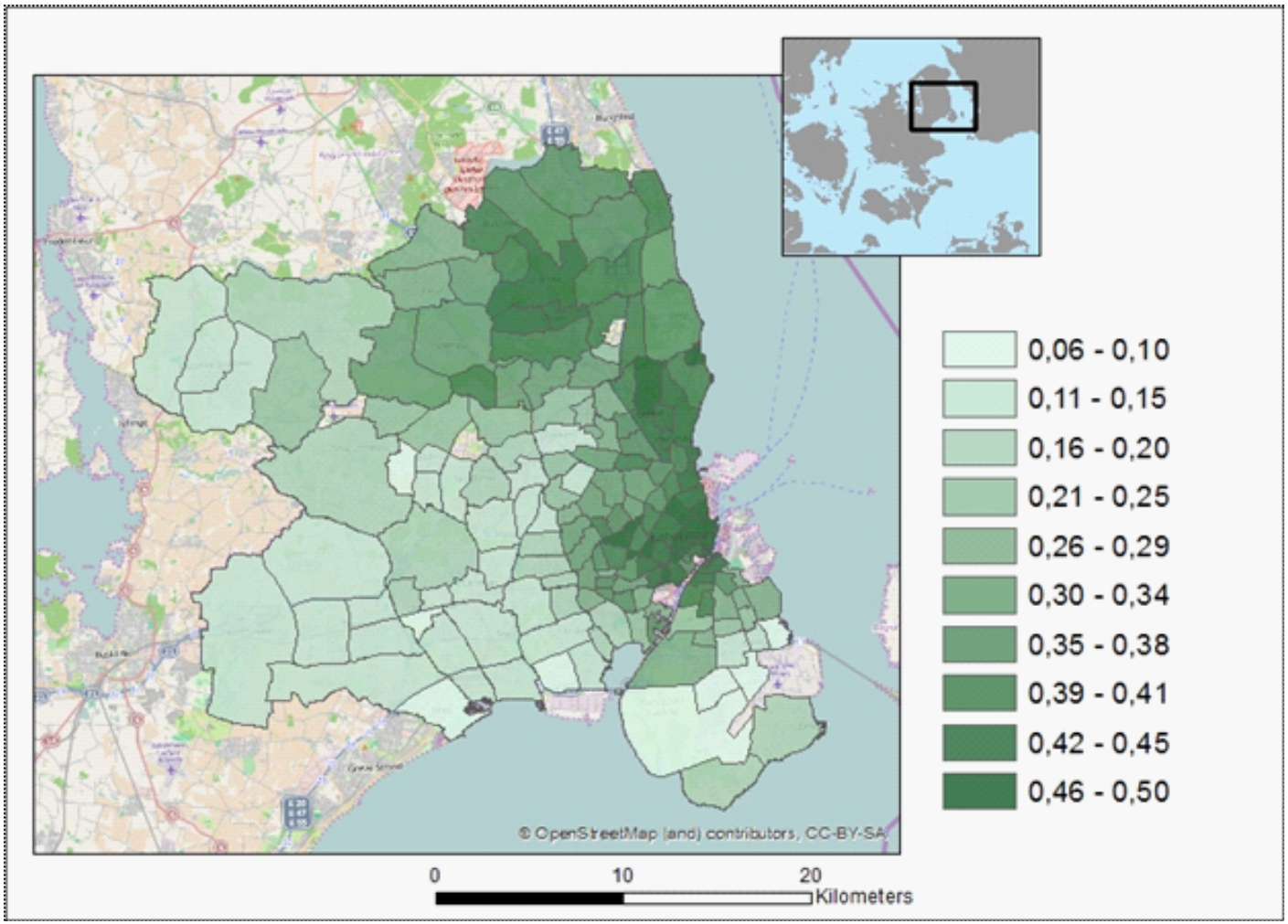
Car ownership (share)



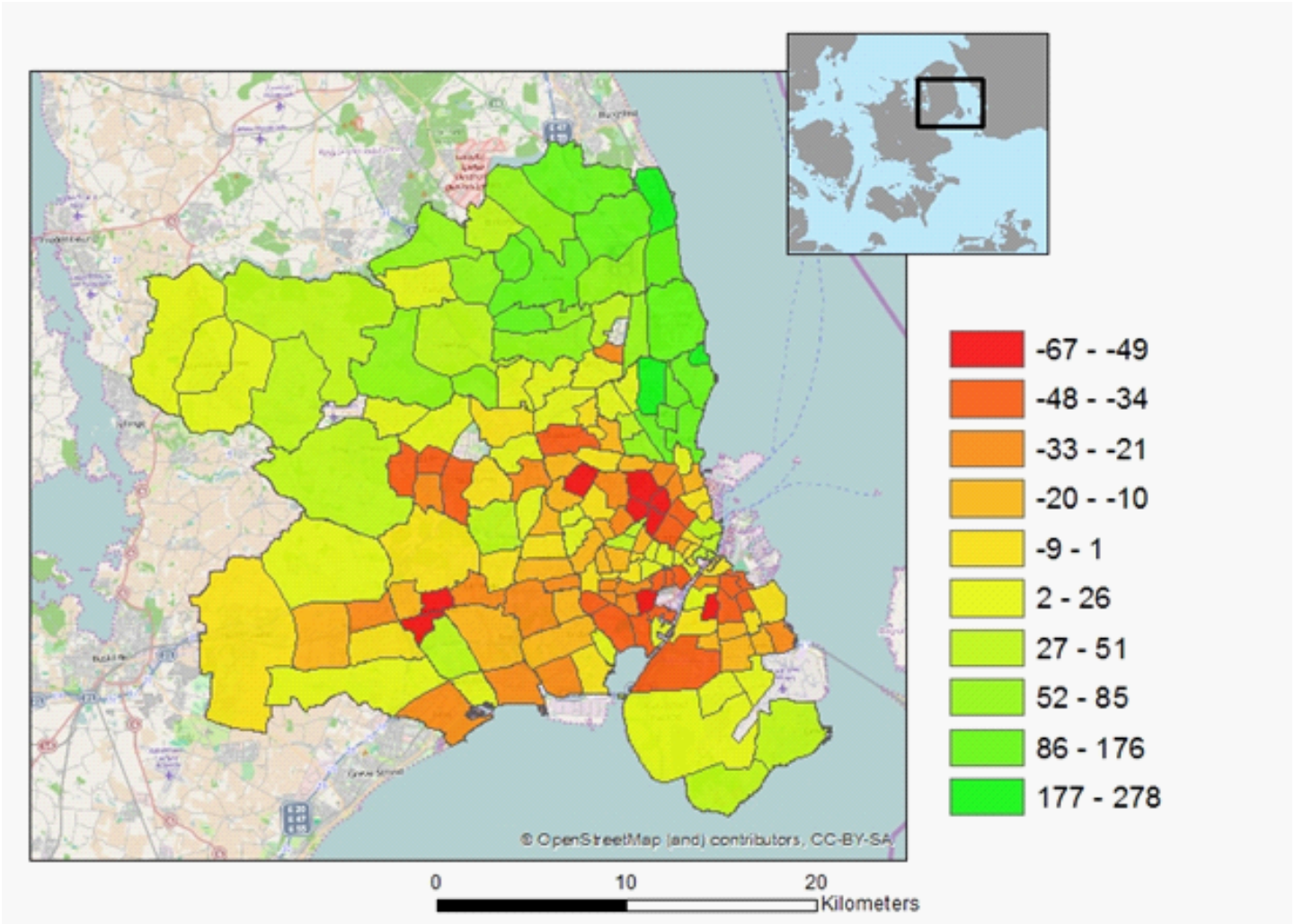
Std. house/apartment price (1000 DKK)



Higher educated (share)



Households income (dev. from the average income)



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

Example

	OLS		IV	
	Estimate	Std.err.	Estimate	Std.err.
Log(standarized house/apartment price)	-2.098	0.324	-3.838	0.357
Share of higher educated	1.800	0.533	1.438	0.666
Number of conserved/protected buildings per sq.km.	0.942	0.167	0.795	0.143
Employment acces reached by public transport * (1-car)	0.007	0.003	0.010	0.003
Proximity to the nearest motorway ramp * (1-car)	0.457	0.208	0.402	0.180
Dummy indicating car owership	0.823	0.228	1.012	0.196

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

Example

	OLS		IV	
	Estimate	Std.err.	Estimate	Std.err.
Log(standarized house/apartment price)	-2.357	0.363	-6.046	1.178
Share of higher educated	2.848	0.599	7.202	2.365
Number of conserved/protected buildings per sq.km.	0.676	0.214	0.794	0.147
Proximity to the nearest metro station * nocar	-0.003	0.003	0.988	0.240
Dummy variable indicating one car	0.154	0.420	1.863	0.291
Dummy variable indicating two cars	-0.254	0.202	0.867	0.344

Decomposition of the mean utilities

- Coefficients associated with accessibility of jobs through public transport and accessibility of the metro network are all positive and significant.
- Coefficients of the housing price index and the share of higher educated are significantly negative and positive, respectively.
- The coefficients associated with indicators for cultural amenities are all as expected significant and positive.
- The average household prefers to live in a house and do not care about the distance to the CBD.
- Car ownership has significant effect on the average households' valuation of the considered amenities.
- *Dealing with the endogeneity makes a huge difference for the estimation results*.

Interaction parameter estimates (single wage-earners)

Example

	Share of higher educated	Proximity to the nearest metro station * nocar	Dummy variable indicating one car
Log(households income)	1.898 (0.162)	-0.003 (0.056)	0.759 (0.073)
Age	-0.076 (0.02)	0.017(0.004)	-0.030 (0.011)
Age squared/1000	0.945 (0.261)	-0.213 (0.0003)	0.297 (0.120)
Number of children in household	0.111 (0.108)	-0.047 (0.041)	0.196 (0.052)
Medium education	3.061 (0.176)	-0.074 (0.057)	0.142 (0.078)
Higher education	5.685 (0.197)	-0.004 (0.058)	-0.076 (0.086)

Interaction parameter estimates (dual wage-earners)

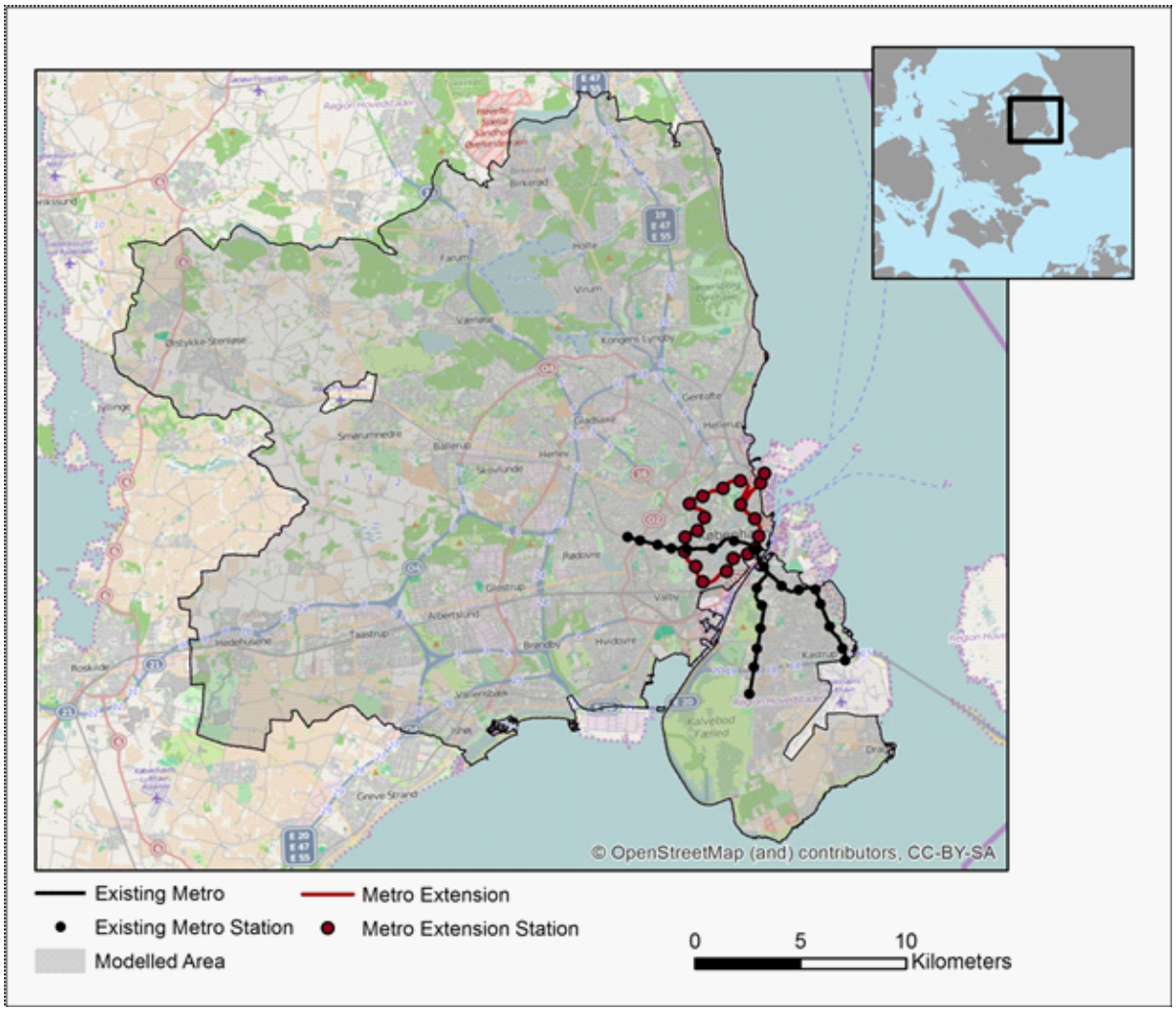
Example

	Share of higher educated	Proximity to the nearest metro station * nocar
Log(households income)	3.854 (0.222)	-0.543 (0.106)
Age, head of the household	-0.148 (0.118)	0.043 (0.049)
Age squared / 1000, head of the household	1.638 (1.193)	-0.448 (0.519)
Medium education, head of the household	2.681 (0.196)	0.131 (0.091)
Higher education, head of the household	5.587 (0.203)	0.297 (0.095)
Age, partner	0.287 (0.123)	0.036 (0.052)
Age squared / 1000, partner	-2.369 (1.343)	-0.608 (0.588)
Medium education, partner	2.579 (0.188)	0.281 (0.088)
Higher education, partner	4.193 (0.213)	0.246 (0.097)
Number of children in household	0.307 (0.090)	-0.165 (0.039)

Interaction parameter estimates

- The estimated coefficients indicate how different households value considered amenities.
- The estimated coefficients reveal a great amount of heterogeneity between observed households:
  - *the proximity to the high quality public transport (metro) is less appreciated by wealthier households and households with more children but more by the higher educated.*
  - households with more children and wealthier households are less sensitive to higher house prices compared to the average household in our sample,
  - wealthier households, households with children and older individuals prefer residence further away from the CDB, while the higher educated prefer residence closer to the CDB,
  - accessibility to the employment access for dual earners households is more appreciated among households with children and higher educated.

The metro system extension in 2019

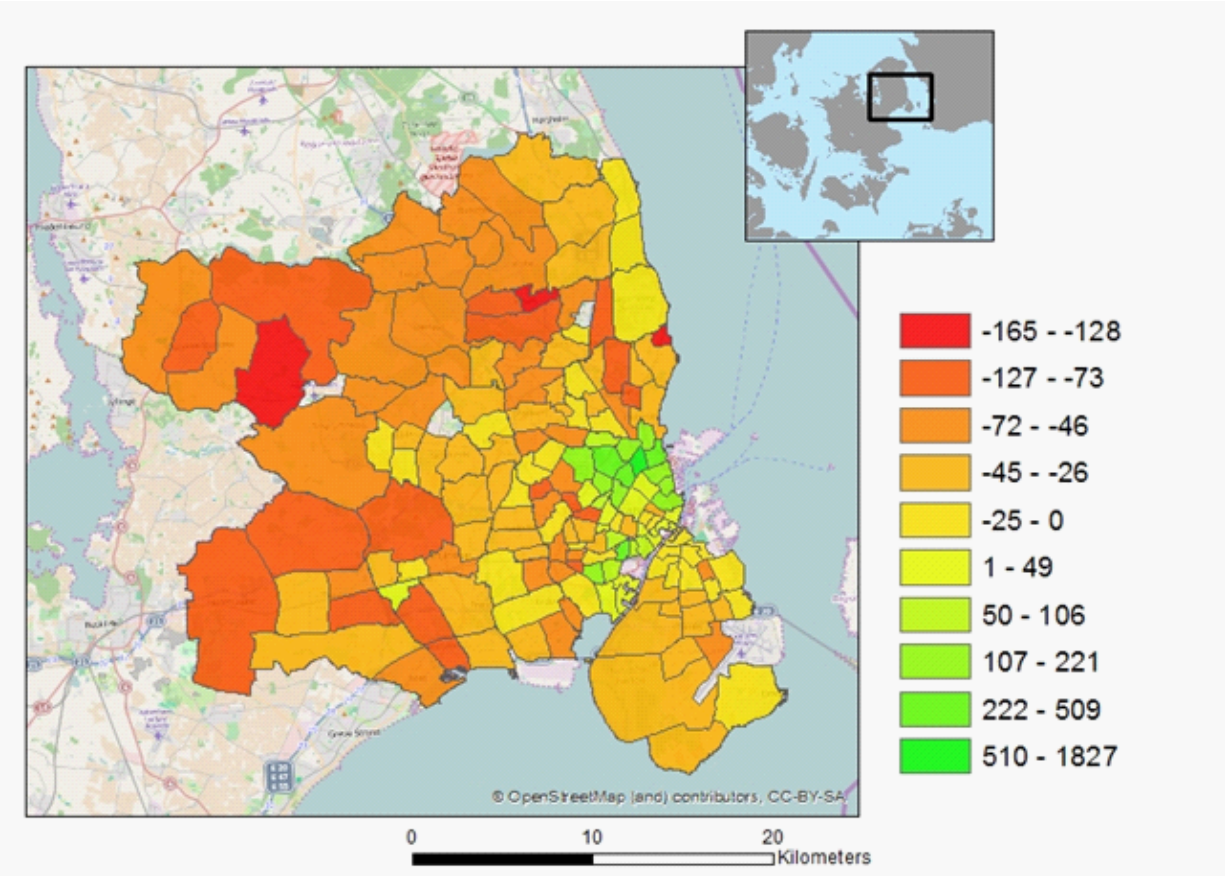


Results: Excess demand

- Extension of the metro system will attract both types of households, i.e. both single earner households and dual earners households will tend to relocate closer to the CBD.

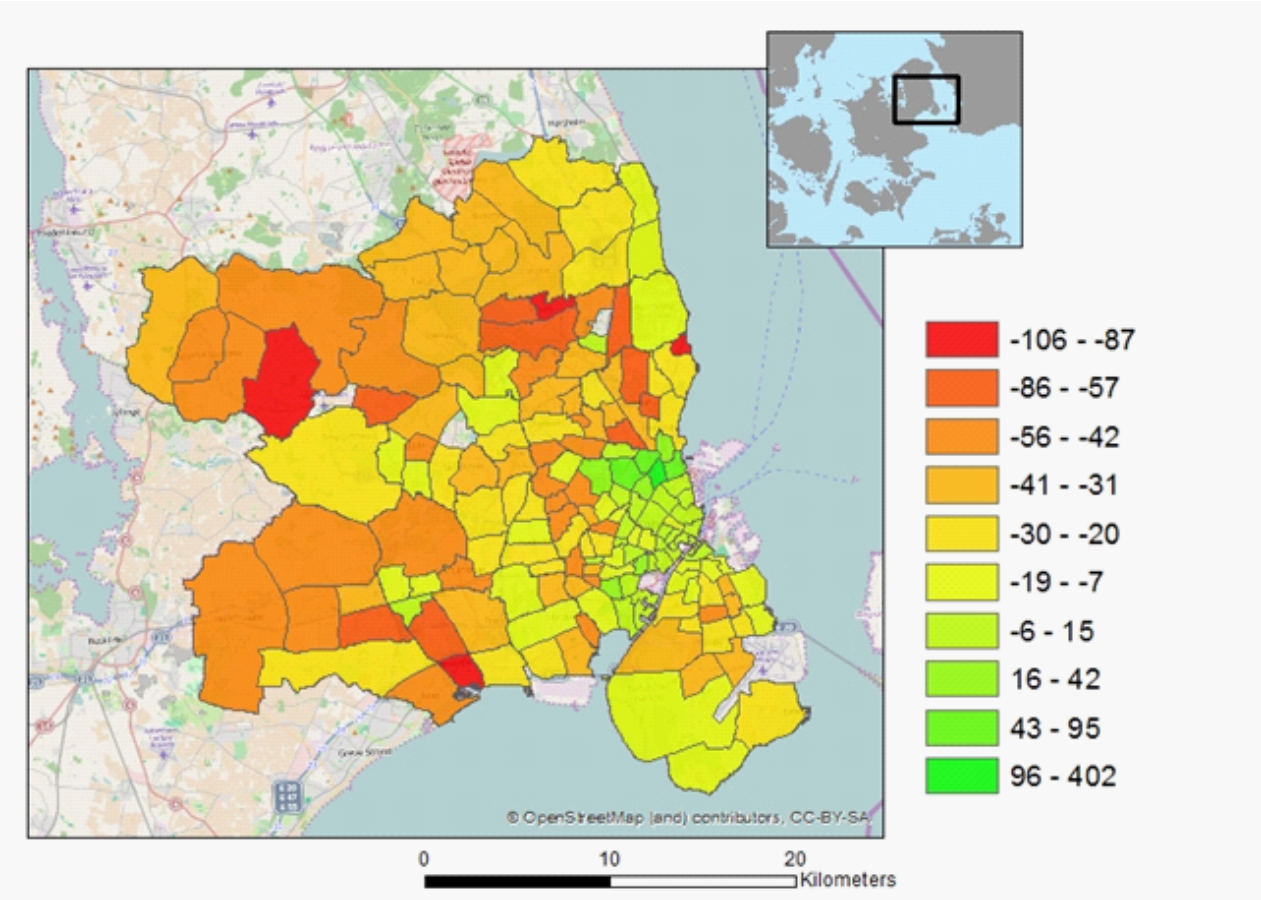
Results: Excess demand

Households will tend to relocate closer to the CBD



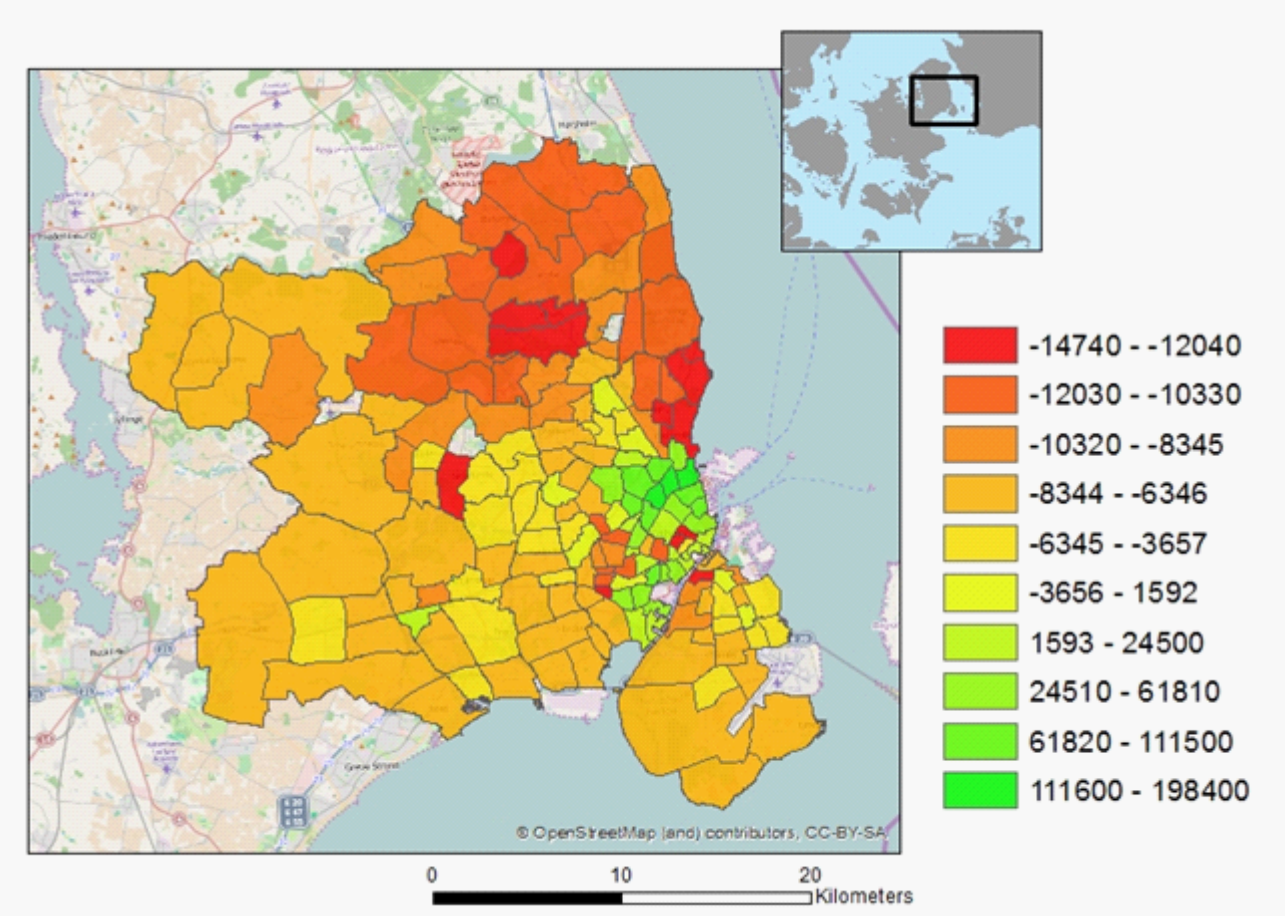
Results: Excess demand

Change in car ownership in the GCA caused by the metro extension



Results: Excess demand

Change in mean household income caused by the metro extension

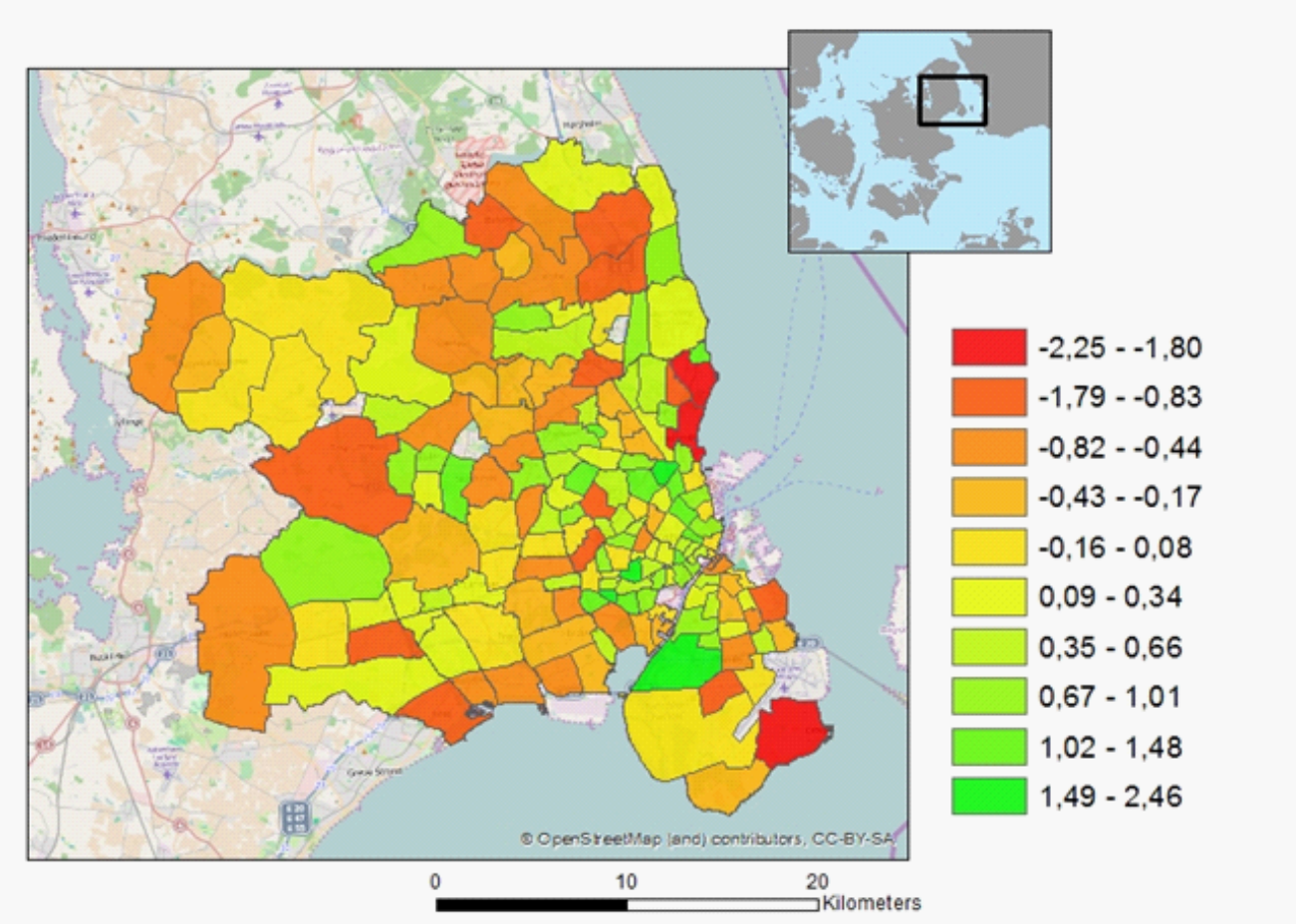


Results: Fixed supply

- Extension of the metro system implies a lot of sorting.
- The number of car owners will most likely increase as a result of the metro extension.

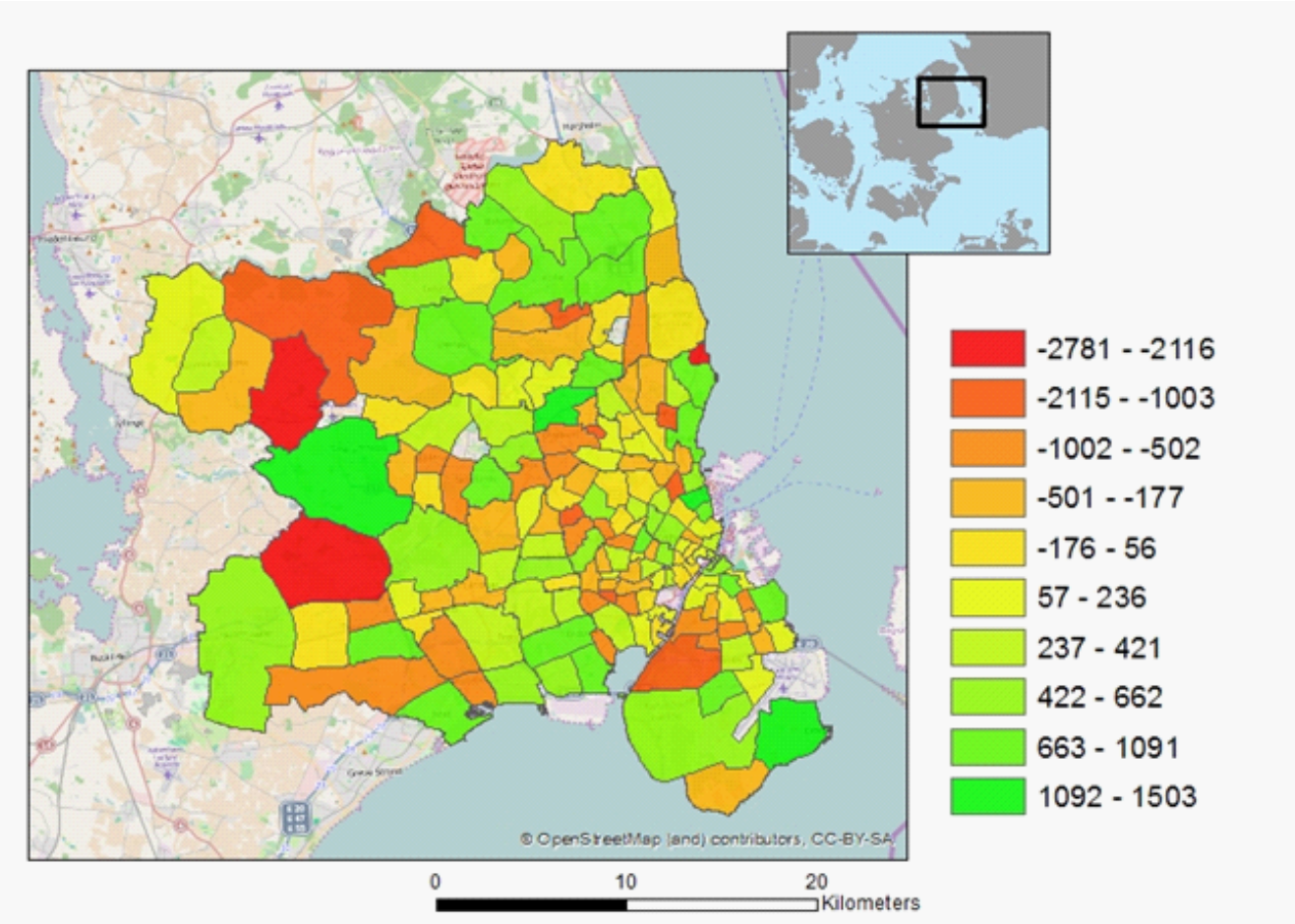
Results: Fixed supply

Change in average housing price



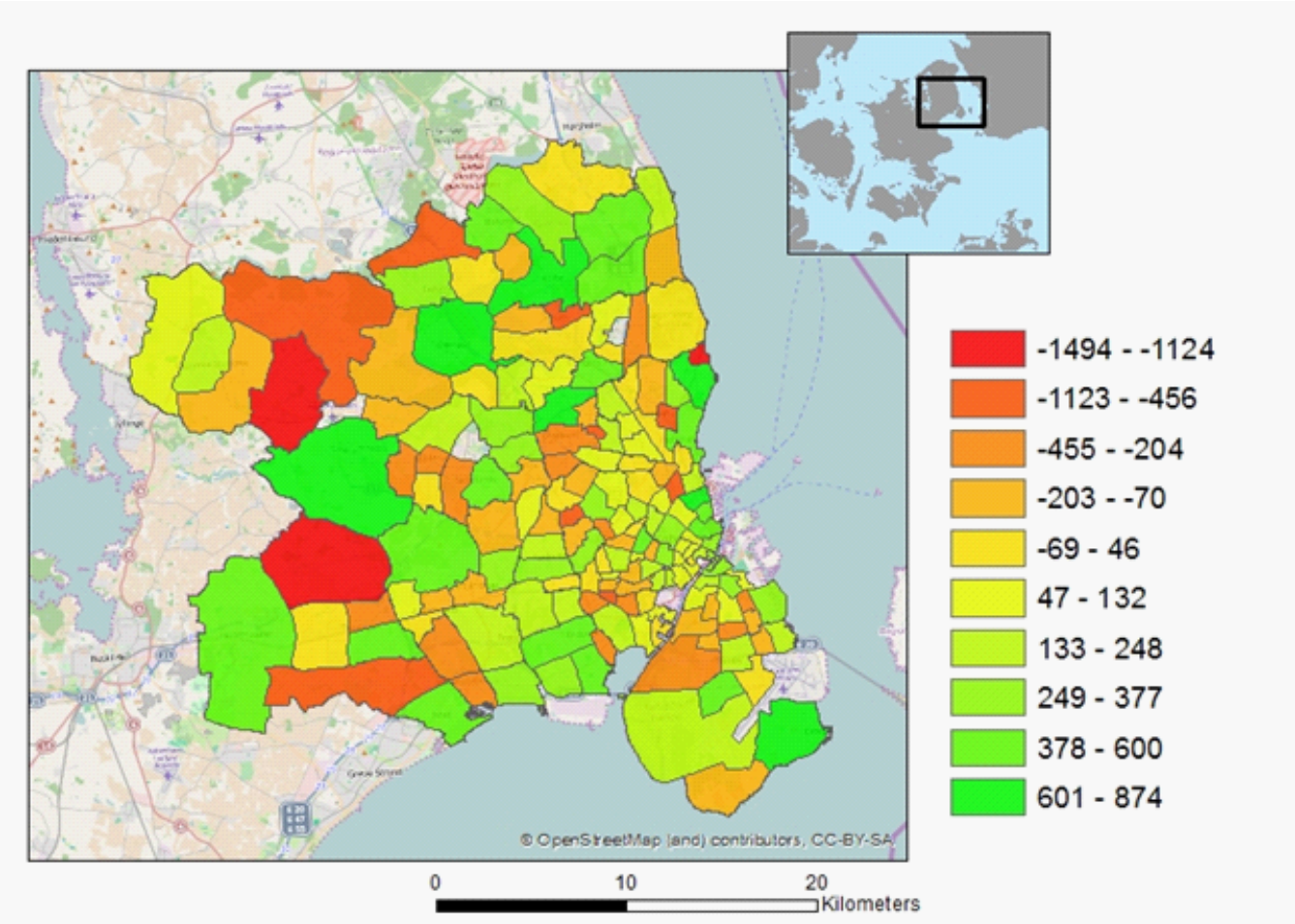
Results: Fixed supply

More sorting!!!



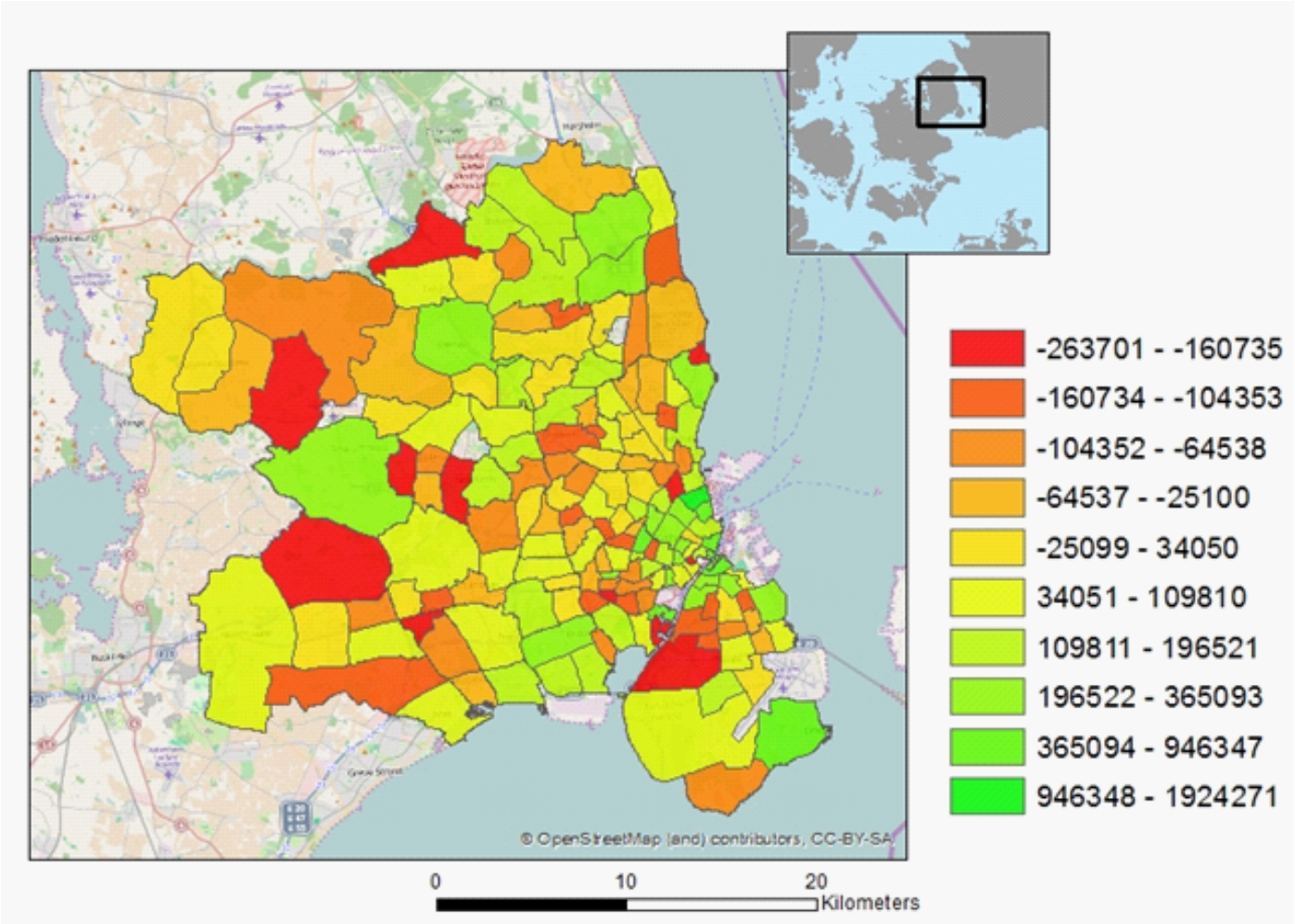
Results: Fixed supply

Change in car ownership in the GCA caused by the metro extension



Results: Fixed supply

Change in mean household income caused by the metro extension



The simulation results

- The model predicts a substantial increase in the interest of single as well as dual earner households for living in the centre of the area, that is, close to the metro network.
- The impact of the extension of the metro-network interacts significantly with the location choices of high income households, who often own two cars.

Conclusion

- The demographic composition of location is as much the result of choice behavior as it is one of its determinants.
- Dealing with the endogeneity makes a huge difference for the estimation results of the MWTP's.
- Car ownership has significant effect on the average households' valuation of the considered amenities.

Future work

- Investigate how workers sort over jobs and houses in an urban area:
  - the residential area and the employment area are connected by the commute and travel time is determined by the quality of the infrastructure and congestion.
  - households have to choose a combination of residential area and an employment area connected by a reasonable commute.
- Firms!
- Dynamics!