TI 2022-035/VII
Tinbergen Institute Discussion Paper

Public transport investments, commuting and gentrification: Evidence from Copenhagen

Ismir Mulalic[1]
Jan Rouwendal[2]

1 Copenhagen Business School

2 Vrije Universiteit Amsterdam, Tinbergen Research Institute

Public transport investments, commuting and gentrification: Evidence from Copenhagen

Ismir Mulalic[1,] **and Jan Rouwendal**[2,3]

Abstract

This paper considers the impact of the introduction of a metro network in the Copenhagen metropolitan area. Using travel surveys from years before and after the opening of the metro network, we observe a significant change in travel times, speeds and mode choice for commutes that can completely or partly be realized by the metro. Interest in the metro among the higher educated is much stronger than among the lower educated. House prices in the vicinity of the metro stations increased significantly. The total additional value of real estate generated by the metro is appr. 40% of the actual construction cost. The government captured a substantial part of the value generated by the metro by concentrating housing construction in some hitherto undeveloped areas close to metro stations. We use a gravity model to explore the implications of the metro for urban structure in an urban equilibrium context and find that all adjustment takes place in the housing market. The lower and medium educated face adjustments in housing attractiveness that counteract the initial impact of the metro. We find no evidence for such adverse effects on the higher educated, which suggest a close connection between the impact of the metro and gentrification in the Copenhagen.

Key words: underground transportation, urban structure, public transport investment, commuting, gentrification.

JEL codes R4, R1, D1

Acknowledgement

Earlier versions of the paper have been presented at the 8th European Meeting of the Urban Economics Association in Dusseldorf (2017), the closing workshop of the project ANR ELITISME: Economics of the Family, Urban Economics, Transportation and Time Use, ENS Paris-Saclay (2017), 3rd Kraks Fond Workshop on Urban Economics in Copenhagen (2018), ITEA conference in Paris (2018), workshop on Neighbourhoods, Externalities and Inequality, Copenhagen Business School (CBS) (2017), workshop on Matters of mobility - Present and Future, Tinbergen Institute Amsterdam (2017) and in seminars at the University of Valencia and Copenhagen Business School (CBS). The authors thank to Statistics Denmark for providing the data. Research support from Kraks Fond is gratefully acknowledged. The usual disclaimer applies.

[1] Department of Economics, Copenhagen Business School, Porcelænshaven 16A, DK-2000 Frederiksberg, Denmark.

[2] Department of Spatial Economics, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam.

[3] Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands.

1. Introduction

Many metropolitan areas in Europe have ancient inner cities in which a large share of their shops, restaurants, cafés, cinemas and concert halls are concentrated. Since these parts of the city were designed and constructed when city sizes and traffic flows were much smaller than nowadays, their road networks are problematic for car traffic. Improving this situation by changing overground transportation networks inevitably requires demolition of many buildings and the transformation of street patterns. This will change the design and atmosphere, which are generally regarded as the core amenities of such ancient inner cities, suggesting the risk that the cure may be worse than the disease. In this situation underground transportation systems, although notoriously expensive, offer a potentially attractive solution since they offer the possibility to leave almost all of the existing urban fabric unchanged, while still being able to provide a substantial increase in the capacity of the transportation network. Moreover, high quality public transport may be as convenient as the car and reduces the demand for parking space in what is usually the most expensive and densely built-up part of the agglomeration. It may therefore be conjectured that metro networks play an important role in the ongoing upsurge of the interest in urban living, especially among the young and higher educated.

To investigate this issue, this paper analyses the impact of the opening of the metro network in Copenhagen, the capital of Denmark, on commuting trips on the basis of travel surveys conducted before and after the opening of the metro network that connected the centre with three suburban locations, including Copenhagen airport.

2. The Greater Copenhagen Area, gentrification and the introduction of metro

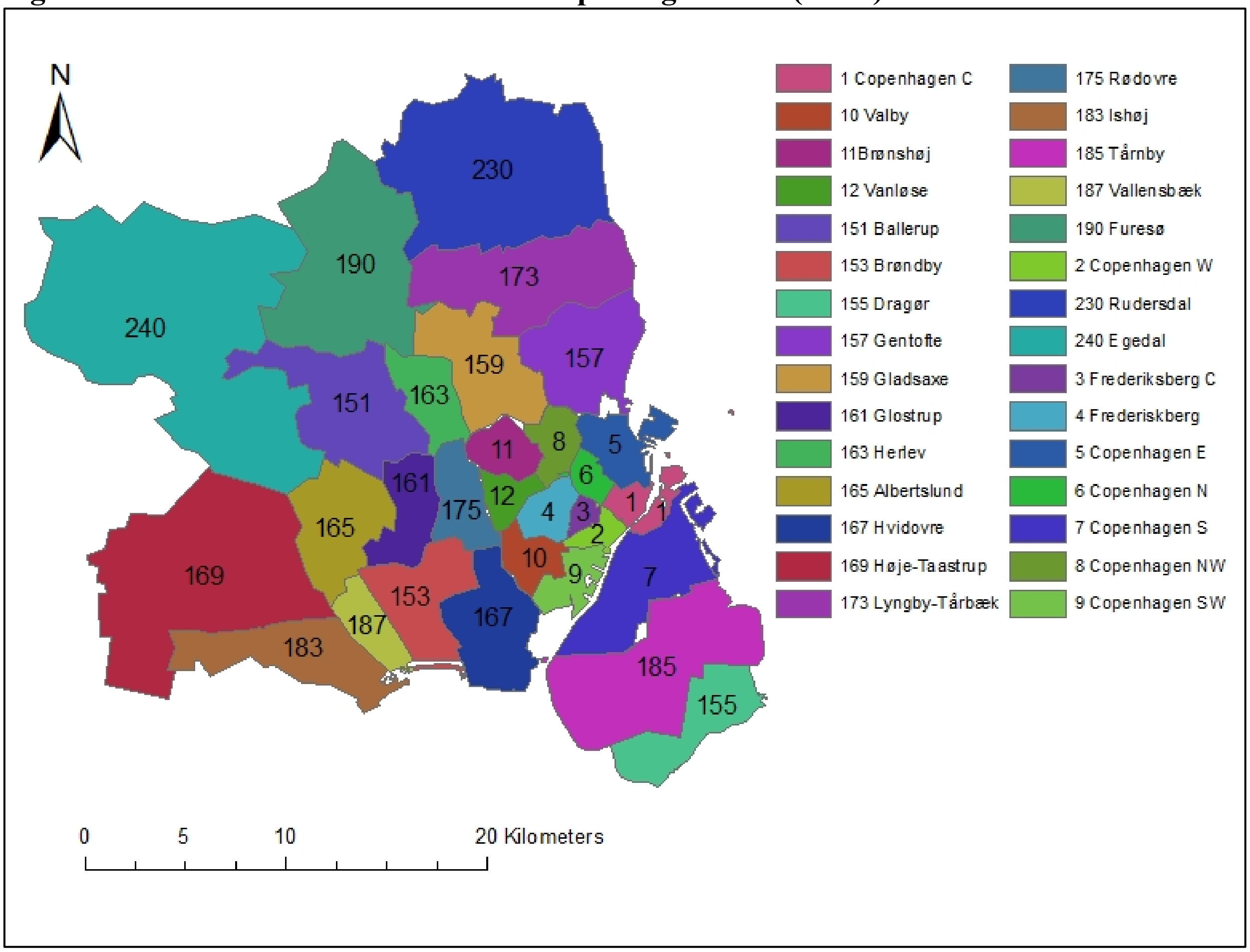
2.1. The Greater Copenhagen Area

The Greater Copenhagen Area (GCA) is part of the Danish island Zealand. The municipality of Copenhagen (the capital city of Denmark) is its centre.[1] The GCA spreads over 20 municipalities with on average 66,288 inhabitants per municipality and an average size of 38.8 km^2. We fragment the two core municipalities, the city of Copenhagen and Frederiksberg, where the metro stations are located into 12 smaller areas, see Figure 2.1. Zones 1-12 on the map refer to the municipalities Copenhagen and Frederiksberg, which together are the core of the metropolitan area. They have

[1] The GCA accounts for more than 40% of Denmark's GDP, app. one third of Danish population, and app. one half of workplaces in Denmark.

been subdivided into smaller administrative units referred to transportation zones. The other zones are the remaining municipalities of the metropolitan area. Information about locations of workers and jobs is available at the level of these areas.

Figure 2.1. Travel zones in the Greater Copenhagen Area (GCA)



Notes: the municipality of Copenhagen incudes: Copenhagen C, Copenhagen W, Copenhagen E, Copenhagen N, Copenhagen S, Copenhagen NW, Copenhagen NS, Valby, Brønshøj and Vanløse. The municipality of Frederiksberg includes Frederiksberg and Frederiksberg C.

Copenhagen is the capital of Denmark since 1443 and the central city partly dates back to the 13th century. The older city is located within a defence belt, which still complicates entrance to the centre of the city. Old maps show that most of the present street pattern in the inner city dates back to the 19th century or earlier. Car infrastructure was created in the course of the 20th century, the most important element being highways that surround the older central part of the metropolitan area. Accessibility of the centre of the metropolitan area decreased, at least in a relative sense, and in the 1970s and 1980s Copenhagen – like many other cities – experienced a period of decline while suburban development flourished.

Urban planning played an important role in the development of the Copenhagen metropolitan area. Since the first urban plan for the Copenhagen metropolitan area – the Finger Plan -- was published in 1948, urban planners have focused on an overall regional structure where urban development is concentrated along city fingers linked to the railway system and radial road networks spread like fingers on a hand from the "palm" of central Copenhagen that are separated by undeveloped green areas (the Danish Nature Agency, 2015).[2] The first version of the Finger Plan did not include the island of Amager between the city center and Copenhagen's international airport (zone 7, Copenhagen S), which at the time did not have much transport infrastructure. Today, Amager is a developed area of Copenhagen and is considered to be the "extra finger."

The counter-development that resulted in the 'triumph of cities' (Glaeser, 2012) also turned the tide for the ancient Copenhagen city centre. Along with the new interest in working and living in older areas with a characteristic atmosphere determined by history and urban amenities, came important investments in the quality of these environments. Increasing prices and rents made refurbishment of dilapidated housing attractive, and this contributed to further improvement in the quality of the central area, and its attraction for high income households and employment.[3]

2.2. Evolution of commuting patterns

This subsection presents some information on the evolution of commuting patterns in the GCA over the period 2002-2010, that is in the period when the metro stations were opened. Over this period the number of lower educated workers in GCA fell by 13% in the period 2002-2010, while that of the medium educated increased by 12% and that of the higher educated by no less than 55%. The latter group is, nevertheless still smaller than the other two.

Table 2.1 shows the commutes within and between the centre and the suburbs for these three education levels in both years. The centre is defined as the municipalities Copenhagen and Frederiksberg, while the remainder of the GCA are the suburbs. The first to columns indicate the

[2] Already in the year 1923, the Danish Engineers' Association had set up a committee that prepared a *transport infrastructure plan for Copenhagen and the surrounding area*" suggesting the road network in the Greater Copenhagen area designed as a system of ring roads and radial roads.
[3] Hybel and Mulalic (2022) shows that households in Denmark are attracted to high amenity areas and that the quality of the public transport system is particularly important for the quality of life.

size of the commuting flows in the Greater Copenhagen Area in 2002 and 2010. The third column presents the relative changes in the flows, while the last one shows the relative changes in these flows, after correcting for the overall change in the size of the groups.

Table 2.1. Commuting flows between the city center and the suburbs for 2002 and 2010 by workers education

	2002	2010	Change	Deviation from the average
Low education				
Residence city - Work city	14,151	12,230	-14%	0%
Residence city - Work suburbs	8,243	6,395	-22%	-9%
Residence suburbs - Work city	27,126	21,352	-21%	-8%
Residence suburbs - Work suburbs	47,695	44,509	-7%	6%
Total	97,215	84,486	-13%	
Medium education				
	2002	2010	Change	Deviation from the average
Residence city - Work city	7,611	8,757	15%	3%
Residence city - Work suburbs	3,752	4,695	25%	13%
Residence suburbs - Work city	11,716	11,131	05%	-17%
Residence suburbs - Work suburbs	19,389	23,107	19%	7%
Total	42,468	47,690	12%	
High education				
	2002	2010	Change	Deviation from the average
Residence city - Work city	7,135	12,700	78%	23%
Residence city - Work suburbs	3,054	5,031	65%	9%
Residence suburbs - Work city	10,477	14,519	39%	-17%
Residence suburbs - Work suburbs	9,467	14,566	54%	-2%
Total	30,133	46,816	55%	

Notes: city includes all zones in the municipalities Copenhagen and Frederiksberg, i.e. zones 1-12, and suburbs all the other municipalities in the GCA; low education obtained includes: basic school, general upper secondary school, vocational upper secondary school and vocational education; medium education obtained includes: short-cycle higher education and medium-cycle higher education; and high education includes: bachelor, long-cycle higher education and PhD-degree.

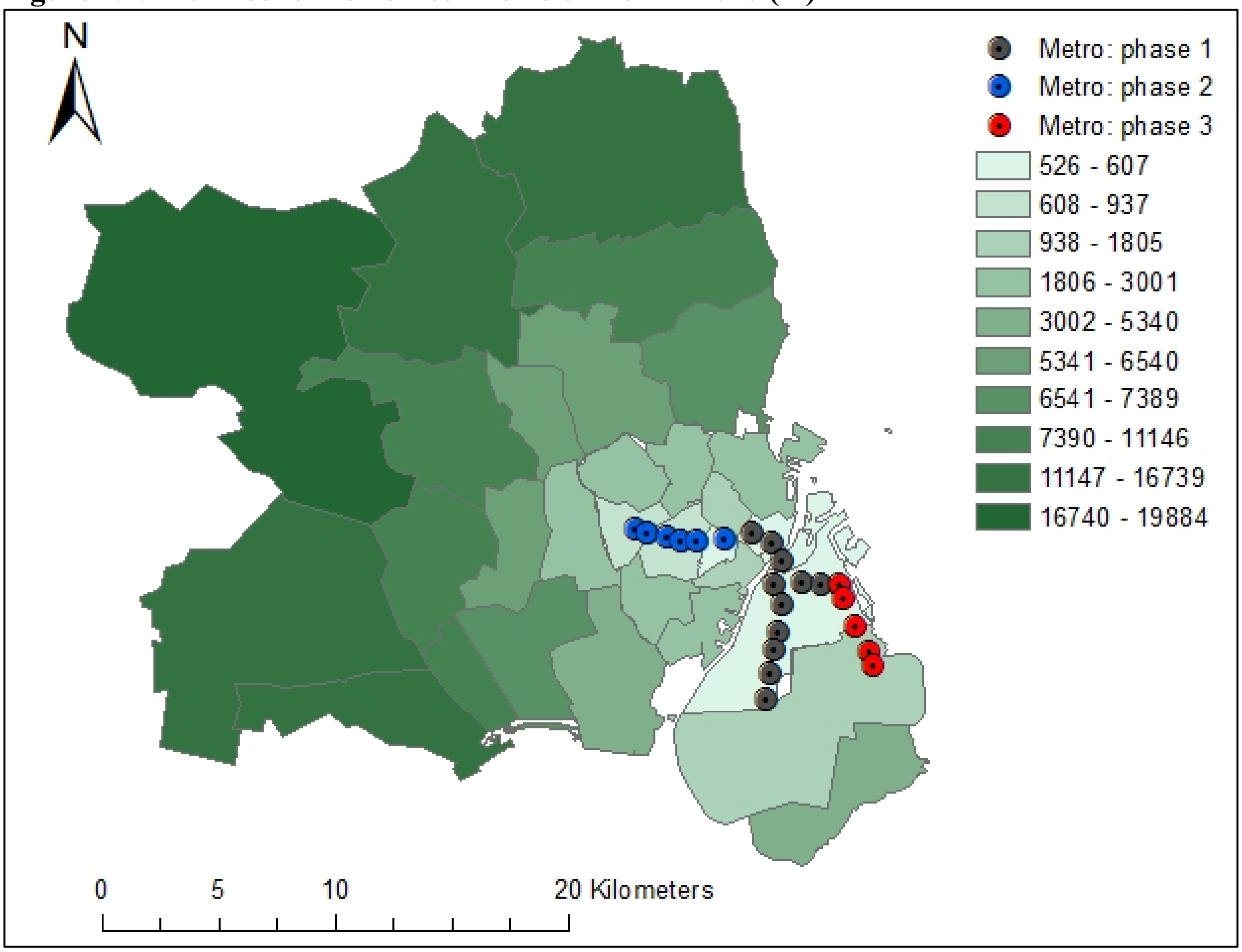
For the lower educated all commuting flows decreased in size, but the one referring to trip with origin and destination in the suburbs decreased much less than the others. The result is that the lower educated live and work more often in the suburbs in 2010. For the higher educated, the pattern is opposite: all flows increased but the one referring to trips with origin and destination in the centre more than the others. The flow referring to higher educated living in the centre and working in the suburbs also increased more than the total number of higher educated. For the medium educated we also see a stronger concentration of residential location in the centre, although for this group jobs in the suburbs appear to be more important than for the higher educated. ‘Reverse commuting’ from the centre to the suburbs thus became more popular for the medium and higher educated, while the conventional pattern of living in the suburbs and working in the centre became less popular for *all* groups.

2.3. The metro

The metro fits into this evolution by offering high quality public transport with frequent services and attractive stations (physical environment). The network that was created before 2010 has a total route length of 20.4 kilometres and 22 stations. It has two lines, M1 and M2, see Figure 2.1. Through the city centre and west to Vanløse, M1 and M2 share a common line. To the southeast, the line M1 runs through the new neighbourhood of Ørestad (site stretching across the island of Amager between the city center and Copenhagen’s international airport (zone 7, Copenhagen S)). The other line, M2, serves the eastern neighbourhoods and Copenhagen Airport. Metro trains run continually (24/7) with the headway varying from two to four minutes, but with longer intervals (up to twenty minutes) during the night only. Metro supplements the larger S-train rapid transit system (commuter rail), and is well integrated with other local trains and buses. Map 2.2 shows the metro system and for each zone the average distance in meters to the nearest metro station in 2010. In contrast to many metro networks elsewhere, that connect distant suburbs with central areas, the Copenhagen metro is especially important for the central part of the agglomeration, including the airport.

The metro network in Copenhagen was introduced in three stages between 2002 and 2007 and currently connects the centre of the metropolitan region with areas to the south. The first stations opened in October 2002, a second set of stations followed in 2003 while the third phase (extending an existing line to the airport) was opened in 2007, see map 2.2. Metro as a new high frequency transportation mode became almost immediately very popular. The metro's ridership reached 20.3 million trips already in 2003 and grew to 44 million in 2008 and 52.5 million in 2010 (Danmarks Statistik, 2014). In the same period, the number of trips in Copenhagen’s S-train increased only by 7%, from 86.8 million passengers in 2002 to 93.0 million passengers in 2010 (Danmarks Statistik, 2014). A distinctive characteristic of the metro is its significant impact on travel times. For instance, with metro it takes less than 15 minutes to travel from Copenhagen Airport to the city centre. The same trip takes more than 40 minutes by bus, about 23 minutes by car, and still more than 20 minutes using the direct train connection from Copenhagen's central station. Although the network is of limited size and covers only a small part of the total metropolitan area, its enormous popularity suggests that it plays an important role in the functioning of Copenhagen's urban fabric. It is the main purpose of this paper to analyse its impact on commuting behaviour.

Figure 2.2. Distance to the nearest metro station in 2010 (m)



3. Trip length, mode choice and the metro

3.1. The travel surveys

We use the Danish National Travel Survey (NTS). The NTS provides information on the travel behaviour of randomly selected individuals who fill out a one-day travel diary. Information is collected continuously throughout the year. We used the 2002 and 2010 versions and selected workers, that is, individuals who report commuting trips. We only use the results referring to these commuting trips. Table 3.1 provides some descriptive statistics.

Table 3.1. Descriptive statistics for Danish national travel survey

	Year=2002		Year=2010	
	Mean	Std. dev.	Mean	Std. dev.
Mode: walk	share = 10.63 %		share = 6.10 %	
Trip length (km)	0.957	0.808	1.293	1.544
Trip time (minutes)	9.422	4.690	11.516	10.518
Trip speed (km/h)	5.861	2.618	6.253	2.151
Mode: bike	share = 31.90 %		share = 31.27 %	
Trip length (km)	4.412	3.258	5.062	3.388
Trip time (minutes)	17.873	10.281	18.786	9.935
Trip speed (km/h)	14.204	3.924	15.541	3.382
Mode: car	share = 37.86 %		share = 44.23 %	
Trip length (km)	12.259	9.903	12.125	8.377
Trip time (minutes)	19.726	10.891	19.598	11.064
Trip speed (km/h)	35.896	14.528	35.928	12.483
Model: public transport	share = 19.60 %		share = 18.39 %	
Trip length (km)	13.539	8.704	12.170	6.579
Trip time (minutes)	40.098	15.708	38.546	14.547
Trip speed (km/h)	19.577	8.503	18.620	6.847
Origin and destination both in center	0.332	0.471	0.323	0.468
Origin outside the center, destination in the center	0.154	0.361	0.146	0.353
Origin in the center, destination outside the center	0.126	0.332	0.148	0.355
Metro station in origin and center			0.111	0.314
Metro station in origin or in center			0.419	0.494
Number of observation	1,561		2,539	

Notes: the monetary values are expressed in real terms (2010 DKK). Center include the City of Copenhagen and Frederiksberg.

We observe in total 4,100 commuting trips, appr. 1600 in 2002 and 2500 in 2010. Between these years the share of walking decreased, that of the car driving increased, while the popularity of biking and public transport remained almost unchanged. It is no surprise that slow transport is mainly used for short trips. For both slow modes, walking and biking, trip lengths and travel speeds have increased over time. In contrast, average trip length for car and public transport have decreased slightly, while the average speed of such trips has also dropped, suggesting that traffic congestion has increased.

The data indicate that public transport has been used, but not whether this was train, bus or (in 2010) the metro. However, since we have detailed information about the origin and destination of the trip, and the location of the metro stations, we can investigate if trips that could have been partly or completely realized by metro in 2010 differ from the others.

In 2002, 23% of the respondents were lower educated, while the shares of medium and higher educated were both 39%. In 2010, the share of lower educated had dropped to 8%, the share of medium educated had decreased slightly to 35%, while the share of higher educated had increased substantially to 57%. Although this may partly reflect peculiarities of the sampling process, we will document later in this paper that it predominantly reflects a substantial change in the composition of the Copenhagen labor force. It is therefore at least potentially important to consider the possibility that preferences for public transport, and in particular the metro, are associated with education.

3.2. Travel time and speed

We start with estimating a series of simple regressions on travel time, see Table 3.2. The regressions have the form of a simple dif-in-dif design in which the variables of interest indicate to the presence of a metro in the origin or the destination of the commuting trip, or in both in the year 2010.

We estimate linear and log-linear equations. In the first models ([1]-[4]) we use untransformed trip length, measured in km, and its square as explanatory variables, and in the second set of models ([5]-[8]) the natural log of trip length. In both sets of equations, we control for the origin and destination of the trip: within the centre, from suburbs to centre and the other way around ('reverse commuting'). These control variables are included to take into account congestion in the centre.

Estimation results confirm the presence of congestion for bike and car. Controlling for distance, we find that trips using these modes that originate or end in the center (or both) take more time. Biking is popular in Copenhagen and the greater density of traffic and (associated with this that) of traffic lights in central parts of the city is probably the main reason for the longer trip durations. In contrast to this, trips by public transport involving the center do not suffer from congestion. To the contrary, trips occurring within the centre (origin and destination both in the center) appear to be *faster* than others, although the effect is only weakly significant for the loglinear equation.

The dummies for metro stations present in origin or destination are expected to have an effect on public transport, via the use of the metro. We note that this effect may be due to the metro itself, but also in changes in the offered services by other types of public transport that are related to the opening of the metro network.[4] We have also included these dummies for the other transport modes to check for the possibility that congestion for cars or bikes has changed because of the opening of the metro network, or because of changes in the networks relevant for these modes. We only find a significant change in the travel time of commute trips made by public transport, which suggests that there were no substantial changes in the travel times of other transport modes for trips potentially served by the metro.[5] The effect of the metro on travel time in public transport is only present for trips that have a metro station present in the origin as well as in the destination. The linear equation suggests that for such trips travel time decreased by almost 6 minutes due to the metro. The log-linear equation suggests a decrease of about 20%, which is of similar magnitude. Our lack of finding a significant effect for public transport trips that originate or end in a zone with a metro station, but not both is likely due to the time it takes to switch from another traffic mode to the metro, if the latter is used for part of the trip only.

We carried out a similar analysis for average speed per trip using similar (linear and loglinear) specifications of the estimating equations. The results are reported in Table 3.3 and confirm the presence of congestion for cars and bikes on trips in which the center is involved as well as the absence of congestion on such trips for public transport. When public transport is used, trips with origin as well as destination in the centre are significantly faster than those in which other parts of the metropolitan area are involved. The arrival of the metro had a substantial impact on travel speed of trips in the centre.

We conclude that the opening of the metro network constituted an important change in the transportation network of GCA. There was a substantial gain in travel times and speed for public transport trips occurring between zones that had a metro station while no such changes for other transport modes on the same trajectories were observed. In the next subsection look at the impact this had on mode choice.

[4] For instance, bus routes largely overlapping with the new metro network may have been abolished in 2010, while some bus routes may have been adjusted to connect better with the new metro network.
[5] Note that such changes would confounding the analysis of the impact of the metro.

Table 3.2. Travel time and the metro

Mode	Walk	Bike	Car	Public transport	Walk	Bike	Car	Public transport
Dependent variable	*Time*	*Time*	*Time*	*Time*	*Ln(time)*	*Ln(time)*	*Ln(time)*	*Ln(time)*
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Length	9.60***	3.60***	1.19***	1.97***				
	(0.59)	(0.11)	(0.03)	(0.18)				
Length^2	-0.61***	-0.05***	-0.004***	-0.02***				
	(0.08)	(0.01)	(0.0004)	(0.004)				
Ln(Length)					0.70***	0.74***	0.65***	0.44***
					(0.02)	(0.01)	(0.01)	(0.02)
Origin and destination in center	0.44	1.79***	4.25***	-4.34***	0.05	0.10***	0.33***	-0.11*
	(0.75)	(0.39)	(0.65)	(1.40)	(0.06)	(0.02)	(0.03)	(0.03)
Origin outside the center, destination	3.71	1.90***	3.76***	-2.49*	0.15	0.10***	0.18***	-0.03
in the center	(2.34)	(0.61)	(0.54)	(1.39)	(0.17)	(0.04)	(0.03)	(0.03)
Origin in the center, destination	1.37	1.14*	3.40***	-1.73	0.01	0.08**	0.16***	-0.02
outside the center	(2.41)	(0.62)	(0.54)	(1.42)	(0.18)	(0.04)	(0.03)	(0.04)
Metro station in origin and destination,	0.18	0.25	2.31	-5.77***	0.10	0.02	0.12	-0.19***
year = 2010	(1.64)	(0.67)	(1.66)	(2.25)	(0.12)	(0.04)	(0.08)	(0.06)
Metro station in origin or destination,	0.44	-0.40	0.63	0.39	0.09	-0.05	0.02	0.01
year=2010	(1.11)	(0.46)	(0.59)	(1.30)	(0.08)	(0.03)	(0.03)	(0.03)
Year 2010	0.11	-0.93**	-0.19	0.41	-0.13**	-0.06	-0.03	0.01
	(0.78)	(0.36)	(0.42)	(1.16)	(0.06)	(0.02)	(0.02)	(0.03)
Constant	1.16	2.54***	4.39***	20.98***	2.30***	1.74***	1.26	2.60***
	(0.74)	(0.44)	(0.47)	(1.73)	(0.05)	(0.02)	(0.03)	(0.05)
Number of obs.	321	1,292	1,714	773	321	1,292	1,714	773
R-squared	0.66	0.79	0.64	0.48	0.74	0.79	0.74	0.55

Notes: ***, **, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and 0.10 levels respectively; standard errors are in parentheses.

Table 3.3. Travel speed and the metro

Mode	Walk	Bike	Car	Public transport	Walk	Bike	Car	Public transport
Dependent variable	*Speed*	*Speed*	*Speed*	*Speed*	*Ln(speed)*	*Ln(speed)*	*Ln(speed)*	*Ln(speed)*
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Length	2.06***	1.10***	1.29***	1.21***				
	(0.30)	(0.10)	(.05)	(0.09)				
Length²	-0.12***	-0.03***	-0.01***	-0.011***				
	(0.04)	(0.01)	(0.001)	(0.002)				
Ln(Length)					0.31***	0.26***	0.34***	0.56***
					(0.02)	(0.01)	(0.01)	(0.02)
Origin and destination in center	-0.35	-1.39***	-8.73***	1.94***	-0.05	-0.10***	-0.33***	0.11***
	(0.37)	(0.34)	(1.11)	(0.65)	(0.06)	(0.02)	(0.03)	(0.03)
Origin outside the center, destination	-1.17	-1.55***	-5.00***	0.79	-0.15	-0.10***	-0.17***	0.03
in the center	(1.17)	(0.53)	(0.91)	(0.65)	(0.17)	(0.04)	(0.03)	(0.03)
Origin in the center, destination	-0.31	-1.17**	-4.62***	0.60	-0.12	-0.08**	-0.16***	0.02
outside the center	(1.12)	(0.54)	(0.92)	(0.66)	(0.18)	(0.04)	(0.03)	(0.04)
Metro station in origin and destination,	-0.26	0.01	-3.17	2.83***	-0.10	-0.02	-0.12	0.19***
year = 2010	(.82)	(0.59)	(2.80)	(1.05)	(0.12)	(0.04)	(0.08)	(0.06)
Metro station in origin or destination,	-0.12	0.79**	-1.53	0.06	-0.09	0.05	-0.02	-0.01
year=2010	(0.55)	(0.40)	(1.00)	(0.61)	(0.08)	(0.03)	(0.03)	(0.03)
Year 2010	0.34	0.62*	0.90	-0.50	0.13**	0.06***	0.03	-0.01
	(0.39)	(0.32)	(0.70)	(0.54)	(0.06)	(0.02)	(0.02)	(0.03)
Constant	4.37	11.44***	25.34***	5.51***	1.80***	2.36***	2.84***	1.49***
	(0.38)	(0.38)	(0.80)	(0.81)	(0.04)	(0.02)	(0.03)	(0.05)
Number of obs.	321	1.292	1.714	773	321	1,292	1,714	773
R-squared	0.27	0.24	0.39	0.57	0.34	0.32	0.46	0.58

Notes: ***, **, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and 0.10 levels respectively; standard errors are in parentheses.

3.3. Modal split

We study modal split of commuting trips in two ways: by estimating linear probability models for each of the four modes and by a simple multinomial logit model that uses the same variables. Results are reported in Table 3.4.

The left part of that Table shows the results of estimating the linear probability models. Although the set of linear probability model equations does not have a choice-theoretic foundation in the sense that there is no known preference relation that leads to this specification, it is worthwhile to note that they are a consistent set of share equations: the equations add up to exactly one.[6] Related to this, we can interpret the estimates as indicating the true 'marginal effects'.

We see a significant negative impact (-11%) of the metro on car use for trips with a metro station at both the origin and the destination, and a somewhat smaller (-6%) – but equally significant – impact on trips with a metro station at the origin or the destination, but not both. The latter result may come as a surprise given the absence of an effect of an impact of the metro on the travel time and speed of such trips. To interpret it, note that the metro does not only have an impact on travel time, but also on the quality and reliability of transport. It is therefore perfectly possible that commuters experience the possibility to make part of their trip by metro as an improvement relative to car use, even though the travel time by public transport on this route did not change.

The positive impact of the metro on the modal share of the public transport is the mirror image of the negative effect on the car. Note that the coefficients estimated for the metro, indicating increases of 12% and 10%, are somewhat larger in absolute value than those estimated for the car, which is consistent with some additional substitution from walking and biking, although the coefficients estimated for the impact of the metro on these two modes are not significant, except for impact on bike trips with a metro station at the origin or the destination. The results thus provide strong evidence of substitution, especially but not exclusively from the car to the metro.

[6] The constants in the equations add up to 1, the coefficients of the explanatory variables to 0. The situation is analogous to that encountered in demand analysis with a set of budget share equations that are linear in the coefficients to be estimated (see Worswick and Champernowne (1954) or Deaton and Muellbauer, (1980)). The error terms of the share equations should add up to zero, which implies a SUR model. Since all equations have the same set of explanatory variables, equation by equation OLS is equivalent to FGLS (see Wooldridge, 2002, Theorem 7.5).

Table 3.4. Modal split and the metro

	Linear probability models				Multinomial logit model (car is the reference alternative)		
Mode	Walk	Bike	Car	Public transport	Walk	Bike	Public transport
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Ln(Length)	-0.140***	-0.080***	0.128***	0.092***	-2.88***	-1.18***	0.52***
	(0.004)	(0.007)	(0.007)	(0.006)	(0.11)	(0.06)	(0.07)
O and D in center[1]	-0.009	0.226***	-0.297***	0.080***	1.88***	1.74***	1.72***
	(0.010)	(0.020)	(0.020)	(0016)	(0.22)	(0.13)	(0.15)
O outside and D in the center	0.050***	0.017	-0.131***	0.064***	1.09**	0.73***	0.64***
	(0.012)	(0.023)	(0.023)	(0.019)	(0.50)	(0.15)	(0.14)
O in and D outside the center	0.049***	0.015	-0.128***	0.063***	0.69	0.72***	0.64***
	(0.012)	(0.023)	(0.024)	(0.020)	(0.57)	(0.15)	(0.14)
Metro st. in O and D	-0.010	-0.004	-0.107***	0.121***	0.72	0.19	1.05***
	(0.019)	(0.039)	(0.039)	(0.033)	(0.44)	(0.27)	(0.30)
Metro st. in O or D	0.015	-0.048**	-0.064***	0.098***	0.33	-0.07	0.57***
	(0.011)	(0.022)	(0.022)	(0.018)	(0.30)	(0.14)	(0.14)
Year 2010	-0.023***	0.029*	0.067***	-0.073***	-0.38*	-0.03	-0.57***
	(0.008)	(0.017)	(0.017)	(0.014)	(0.21)	(0.10)	(0.12)
Constant	0.319***	0.367***	0.309***	0.004	0.50***	1.08***	-2.45***
	(0.009)	(0.019)	(0.019)	(0.016)	(0.19)	(0.13)	(0.20)
Number of obs.	4,100	4,100	4,100	4,100		4,100	
R-squared[2]	0.30	0.11	0.20	0.10		0.25	

Notes: ***, **, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and 0.10 levels respectively; standard errors are in parentheses.
[1] O stands for origin and D for destination
[2] For MNL model this is the pseudo R^2 commonly reported for discrete choice models.

The right part of the Table 3.4 shows the results of estimating a multinomial logit (MNL) model using the same explanatory variables. The car is used as the reference alternative, so all estimates should be interpreted as referring to the attractiveness of a mode relative to the car. Consistent with this interpretation, we do not find significant effects of the metro on walking and biking. For public transport we find strongly significant effects of the metro. The effect on trips with a metro station in the origin as well as in the destination are approximately twice as large as those for trip with a metro station in either the origin or the destination. These results confirm the findings of the linear probability models. Note also that the MNL model has strongly significant and large coefficients for almost all trips in which the centre is involved, which confirms that for such trips the car is in a relatively bad position.[7] A particular advantage of the logit model over the linear probability models is that is has a choice theoretic interpretation (McFadden, 1973).

[7] We have also estimated the linear probability and logit models using trip length and trip length squared instead of the log of trip length as explanatory variables. The results were similar.

3.4. Modal split and education

The models estimated thus far treat workers as essentially homogeneous. Although this is convenient for many purposes, it can also give the potentially misleading impression that workers do not differ in the choices they make with respect to public transport. In this subsection we investigate this issue by distinguishing workers on the basis of their education. The main reason for doing this is the potential link between high quality public transport and the choices of residential and job location of higher educated workers, as discussed in the introduction. To investigate this issue, we interact all coefficients of the mode choice model of the previous section with dummies for lower, medium and higher education. The results are reported in Table 3.5. It indicates the presence of substantial differences in mode choice behaviour that are associated with education. For instance, the linear probability models show important differences in the sensitivity of the choice to walk or bike among the three educational groups. The lower educated use the car less often, but are also less inclined than the two other groups to switch to a different mode for trips in which the centre is involved. Most relevant for the purposes of the present study is that there are clear differences in the impact of the metro on mode choice behaviour. The higher educated appear to be much more inclined to use the metro than the other two groups.

The MNL model confirms the picture that emerges from the linear probability equations. We find a highly significant coefficient for public transport trips by the higher educated that start and end in zones with a metro, and a smaller but also very significant coefficient for trips by the same group that either start or end in a zone with a metro station (but not both). The coefficients for the other two groups are marginally significant at best.

Table 3.5. Education, modal split and the metro

		Linear probability models				Multinomial logit model (car is the reference alternative)		
Mode		Walk	Bike	Car	Public transport	Walk	Bike	Public transport
Variables	Education level	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Ln(Length)	Low	-0.205***	-0.047***	0.105***	0.147***	-3.28***	-1.74***	0.61***
		(0.008)	(0.016)	(0.016)	(0.013)	(0.28)	(0.21)	(0.21)
	Medium	-0.124***	-0.081***	0.110***	0.095***	-2.89**	-1.17***	0.52***
		(0.006)	(0.012)	(0.012)	(0.010)	(0.20)	(0.10)	(0.12)
	High	-0.108***	-0.097***	0.113***	0.091***	-2.75***	-0.99***	0.69***
		(.005)	(0.011)	(0.011)	(0.009)	(0.17)	(0.08)	(0.12)
O and D in center[1]	Low	-0.114***	-0.128***	-0.080**	0.094**	1.66***	0.63	1.10***
		(0.023)	(0.046)	(0.046)	(0.040)	(0.50)	(0.42)	(0.41)
	Medium	-0.011	0.248***	-0.350***	0.113***	1.74***	1.89***	1.80***
		(0.016)	(0.033)	(0.034)	(0.028)	(0.44)	(0.21)	(0.24)
	High	-0.013	0.309***	-0.379***	0.083***	2.45***	2.00***	2.13***
		(0.014)	(0.029)	(0.029)	(0.024)	(0.40)	(0.18)	(0.24)
O outside and D in the center	Low	0.091**	-0.123	0.165**	0.133**	-0.83	0.59	-1.12**
		(0.036)	(0.071)	(0.071)	(0.060)	(1.34)	(0.50)	(0.46)
	Medium	0.029	0.048	-0.176***	0.099***	0.55	0.88***	0.82***
		(0.018)	(0.036)	(0.037)	(0.031)	(1.00)	(0.25)	(0.21)
	High	0.048***	0.046	-0.185***	0.090***	2.27***	0.75***	1.00***
		(0.016)	(0.032)	(0.033)	(0.027)	(0.67)	(0.20)	(0.21)
O in and D outside the center	Low	0.112***	-0.109	-0.097	-0.101	0.22	0.50	0.83*
		(0.036)	(0.073)	(0.073)	(0.061)	(1.21)	(0.52)	(0.48)
	Medium	0.026	0.048	-0.159***	0.086***	0.18	0.81***	0.74***
		(0.019)	(0.038)	(0.038)	(0.032)	(0.99)	(0.26)	(0.22)
	High	0.043***	0.040	-0.175***	0.092***	1.68**	0.69***	1.00***
		(0.016)	(0.033)	(0.033)	(0.028)	(0.85)	(0.20)	(0.21)
Metro station in O and D	Low	-0.143	0.386**	-0.010*	-0.233	-12.07	0.44	-13.25
		(0.093)	(0.185)	(0.187)	(0.158)	(8.43)	(1.04)	(7.46)
	Medium	0.029	-0.048	-0.078	0.097*	1.30*	0.16	0.74
		(0.034)	(0.068)	(0.069)	(0.058)	(0.74)	(0.45)	(0.48)
	High	-0.005	-0.076	-0.070	0.151***	0.20	0.09	1.34***
		(0.025)	(0.048)	(0.050)	(0.042)	(0.60)	(0.36)	(0.40)
Metro station in O or D	Low	0.028	-0.083	-0.141*	0.030	1.27	0.73	1.12*
		(.037)	(0.074)	(0.076)	(0.063)	(0.81)	(0.52)	(0.58)
	Medium	0.024	-0.045	-0.060	0.081**	0.71	0.05	0.40*
		(0.018)	(0.037)	(0.037)	(0.031)	(0.55)	(0.24)	(0.22)
	High	0.006	-0.099***	-0.012	0.105***	-0.40	-0.35	0.50***
		(0.014)	(0.029)	(0.029)	(0.024)	(0.43)	(0.18)	(0.19)
Year 2010	Low	-0.093***	0.066	0.199***	-0.171***	-1.82***	-0.53	-2.03***
		(0.023)	(0.046)	(0.046)	(0.039)	(0.51)	(0.33)	(0.42)
	Medium	-0.015	-0.019	0.058**	-0.024	-0.32	-0.37	-0.26
		(0.014)	(0.027)	(0.028)	(0.023)	(0.40)	(0.18)**	(0.18)
	High	0.016	0.041	-0.017	-0.040*	1.04***	0.33	-0.20
		(0.012)	(0.024)	(0.025)	(0.021)	(0.38)	(0.15)	(0.18)
Constant	Ref	0.319***	0.410***	0.075***	0.100***	2.50***	2.70***	-0.80*
		(0.009)	(0.033)	(0.033)	(0.028)	(0.38)	(0.35)	(0.45)
	Medium	-0.126***	-0.067	0.330***	-0.136***	-2.28***	-1.81***	-1.87***
		(0.023)	(0.046)	(0.047)	(0.040)	(0.55)	(0.42)	(0.56)
	High	0.414***	-0.007	0.352***	-0.157***	-3.63***	-2.14***	-2.75***
		(0.016)	(0.045)	(0.045)	(0.038)	(0.55)	(0.40)	(0.57)

Table 2. (continued) Education, modal split and the metro

Wald test for equality of Metro station in O and D[2]	0.37	0.06	0.35	0.0004	0.0052
Wald test for equality of Metro station in O or D[2]	0.50	0.002	0.098	0.0000	0.0016
Number of obs.	4,100	4,100	4,100	4,100	4,100
R-squared[3]	0.35	0.13	0.22	0.11	0.27

Notes: low education obtained includes: basic school, general upper secondary school, vocational upper secondary school and vocational education; medium education obtained includes: short-cycle higher education and medium-cycle higher education; and high education includes: bachelor, long-cycle higher education and PhD-degree. ***, **, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and 0.10 levels respectively; standard errors are in parentheses.

[1] O stands for origin and D for destination.

[2] The p-value of the test statistic is reported.

[3] For MNL model this is the pseudo R^2 commonly reported in discrete choice models.

4. Implications for commuting flows

To further investigate the implications of the metro for commuting in Copenhagen, we use the register data on commuting flows between considered zones in 2002 and 2010. The workers are classified by education. We combine this information with our modal split model to compute the implied impact of the metro on the commuting flows. We compute two counterfactuals: what would commuting flows have been in 2002 if the metro had already been there, and what would commuting flows have been in 2010 if the metro would not have been present.

The total number of commuters increased from 562 thousand to 578 thousand, a modest increase of 3% over 8 years. However, as Table 4.1 shows, splitting these numbers by education reveals much larger differences. In particular we see a growth in the number of higher educated by 48%, which is associated with an increase in their share of the labour force from 0.14 to 0.20. The number of medium educated workers increased by 12%. The number of lower educated workers dropped by 11%, and their share in the labour force decreased from 65% to 57%.

Table 4.1 Total numbers of commuters by education

	2002	Share in 2002	2010	Share in 2010	Growth
Lower	367,753	0.65	327,616	0.57	-11%
Medium	105,263	0.19	117,632	0.20	12%
Higher	76,934	0.14	114,062	0.20	48%
Not class.	11,594	0.03	18,336	0.03	58%
Total	561,544		577,646		

To compute the impact of the metro on the commutes of these workers, we use the mode choice model discussed in the previous section. According to the model the metro can have an impact on commutes that originate or end in a zone with a metro station. The impact of the metro is therefore limited to 324 of the 900 possible commutes that we can distinguish on the basis of 30 zones. Between 2002 and 2010 the total number of workers realizing these commutes increased by 5%, which is more than the 3% realized by the total number of commuters. However, the higher than average growth was only realized on commutes that could be entirely realized by metro.[8] For these 30 combinations of destination and zones the number of commuters increased by no less than 24%, while for the commutes that could only partly be realized by metro the growth was 1.6%. For the commutes that could partly or completely be realized by metro the share of higher educated increased from 18 to 26%. For the commutes that can be completely realized by metro the corresponding figures are 22 and 21%, respectively.

We assess the impact of the metro through two computations, one using the commuting pattern of 2002, the other using that of 2010. In both cases we apply the mode choice models of the previous section – the one referring to all workers and those referring to specific education groups – with and without the dummies for metro stations. We use only the significant coefficients (p-value 0.05 or better) that have been estimated for these dummies. The results of these exercises are reported in Table 4.2.

Table 4.2. Additional commuting trips using public transport due to the metro

		2002 commutes			2010 commutes		
		Total	O and D	O or D	Total	O and D	O or D
All	Lin. Prob.	24,345	4,560	19,794	25,760	5,654	20,105
	Logit	32,578	6,790	25,788	27,244	6,647	20,596
Lower ed	Lin. Prob.	0	0	0	0	0	0
	Logit	26,556	0	26,556	20,269	0	20,269
Medium ed	Lin. Prob.	3,667	647	3020	4,033	808	3225
	Logit	2,435	-55	2,490	2,453	-70	2,523
Higher ed	Lin. Prob.	5,026	1,274	3,752	7,520	2,161	5,359
	Logit	5,469	2,108	3,361	6,793	2,940	3,814

Notes: 'Total' refers to all commutes that can be partly or completely realized by metro, 'O and D' to commutes with a metro station in zone of origin and in zone of destination and 'O or D' to commutes with a metro station in the zone of origin or in the zone of destination, but not in both.

The table shows that the metro generates between 27 and 32 thousand additional commutes by public transport. Somewhat surprisingly, the large majority of them are realized by lower

[8] The coefficients for the impact of the metro on trips that cold only partly be realized by the new mode were insignificant.

educated workers using the metro on part of the home-work trip. Due to the changing commuting pattern the computed number of additional commutes by public transport is lower in 2010 than in 2002. This difference is entirely due to the lower educated workers. For the higher educated we compute a large number of trips generated by the metro based on the 2000 commuting pattern, while for the medium educated there is hardly any difference.

Table 4.3 reports the results of similar computations for non-realized trips by car due to the metro. Here the linear probability model is not informative when we distinguish between educational groups because no significant coefficients are estimated.[9] We compute a reduction in the commuting trips by car that amounts to approximately two-thirds of the computed increase in public transport trips, presumably realized at least partly by metro. It appears therefore that the opening of the metro network implied a non-negligible reduction of car travel during peak hours. Consistent with our results for public transport, we find that most of this effect is due to lower educated workers.

Table 4.3. Non-realized commuting trips using car due to the metro

		2002 commutes			2010 commutes		
		Total	O and D	O or D	Total	O and D	O or D
All	Lin. Prob.	-16,959	-4,033	-12,927	-18,130	-5,000	-13,130
	Logit	-16,339	-2,797	-13,542	-13,117	-3,229	-9,887
Lower ed	Lin. Prob.	-17,810	-217	-17,592	-15,263	-219	-15,044
	Logit	-14,459	0	-14,459	-13,10	0	-13,190
Medium ed	Lin. Prob.	0	0	0	0	0	0
	Logit	-1,411	-62	-1,350	-1,557	-106	-1,451
Higher ed	Lin. Prob.	0	0	0	0	0	0
	Logit	-2,544	-566	-1,978	-2,338	-561	-2,223

Notes: 'Total' refers to all commutes that can be partly or completely realized by metro, 'O and D' to commutes with a metro station in zone of origin and in zone of destination and 'O or D' to commutes with a metro station in the zone of origin or in the zone of destination, but not in both.

5. Implications for housing prices and land value capture

Our findings in the previous section suggest that the metro had considerable value for commuters in the Copenhagen region. One would therefore expect that consumer's willingness to pay for the metro is reflected in the price of houses in the vicinity of the metro stations. To investigate this, we carry out a hedonic price analysis. Economic theory suggests that the social value generated by the

[9] The logit model produces significant estimates, perhaps because it makes more explicit use of the fact that commutes have to be realized by a specific mode, whereas the linear probability model uses three independently estimated regressions.

metro is reflected in land prices. Unless the land is owned by the government this capitalization effect is a windfall profit for private landlords. In the Copenhagen the government has made an attempt to capture at least some of the increase inland prices generated by the metro by making available hitherto undeveloped land close to some metro stations for housing construction. In the second part of this section we show that this policy was successful, at least in the sense that it had a substantial impact on housing supply and on the prices of new housing in the areas on which the policy concentrated.

5.1. Property prices and the metro

Property prices are an important indicator of the attractiveness of neighbourhoods. To study the impact of the metro on property prices, we use Danish administrative register data on realized real estate transactions, i.e. information on the price paid for specific dwellings, for the period 1993-2016. We focus again on the GCA. The sample includes 366,762 realized real estate transactions. Our data also includes structural dwelling attributes extracted from the Building and Dwelling Register (BBR). These include the size of the housing unit, the number of rooms, and the age of the building. Additionally, variables describing the type of building (residential or business) and dwelling usage and type (rented or occupied by the owner, and multifamily housing or single-family house) are available. Table A.1 in Appendix A shows the descriptive statistics for the sample.

We first estimate simple difference in differences (DD) models in which we regress the log of house price p_{ht} on the indicator for treated areas after opening of a metro station, that are areas very close to new metro stations:

$$p_{ht} = \alpha + \gamma MT_{ht} + \lambda T_{ht} + \beta' X_{ht} + \delta_{t\times Z(h)} + \varepsilon_{ht} \tag{5.1}$$

where p_{ht} is the log of house price of house h in year t, MT is a dichotomous variable that is 1 for treated houses in proximity to a metro station for the period after the metro station opening and 0 otherwise, T is an indicator for treated areas that equals 1 if the house is located close to a metro station and is transacted after that station opened, X denotes structural dwelling attributes, $\delta_{t\times Z(h)}$ are zone-specific year fixed effects, and ε_{ht} is a random error term. Our DD models account for the level differences in the house prices between areas in proximity to a metro station (treatment areas) and areas further away from the metro (control areas) because the metro led to only minor effects on the entire system beyond those that accrued to the treatment area.[10] Other changes in the local economic environment may confound the effects from the metro on property values. We therefore control for local year fixed effects. Moreover, we control for structural dwelling attributes because the changes over time in the average attributes of traded properties might systematically differ between treatment and control areas and consequently impact the estimate of the metro opening. We account for the spatial autocorrelation of errors by clustering standard errors at the zone level.

We need to define the treatment areas in order to compare them to control areas which are less or no affected by the metro. We therefore calculate for each property the shortest distance to a metro station using centroid of hectare-net location for each property in our sample and the exact geographical location of the metro stations. The shortest calculated distance is 18 m and the longest about 31 km. We then estimate model (5.1) for different intensities of treatment, i.e. from 100 m up to 1 km. Table 5.1 shows the results. The key coefficient of interest is that on the treated areas after the metro station opening (γ) that captures the price impact of the metro. This coefficient is as expected positive and strongly significant for houses in the 300 m catchment area around the metro stations. The magnitude of this coefficient for the 100 m catchment area is almost twice as large as for the 200 m catchment area and more than four times as large as for the 400 m catchment area. For longer distances from the metro stations (400 m – 1 km) this coefficient is small and often insignificant. The coefficient for the 300 m catchment area suggests that the metro opening is associated with an economically significant price premium of 3.7% ($exp(0.036) - 1$).[11] The price premium for the 100 m catchment area is 5.4%. These empirical findings suggest that the metro construction was associated with a substantial value creation. The treated areas (λ) coefficient captures the value differential associated with being in treated areas in general. This effect is often significantly negative.

To further explore our definition of the treatment area, we also estimate a DD model similar to eq. (5.1) but now using concentric rings as treatment areas:

[10] We ignore here general equilibrium effects resulting from the metro (Ahlfeldt et al., 2015; Heblich et al., 2020). These effects bias down our estimates. Our estimates of the impact of the metro on housing prices are therefore conservative.
[11] This is in line with Mulalic and Rouwendal (2020) who simulate the impact of an extension of the metro network in Copenhagen and show the housing prices increase in the areas closer to the new metro line and decrease in other areas that become relatively less attractive.

$$p_{ht} = \alpha + \sum_{k=1}^{10} \gamma_k \, CR_{kht} + \sum_{k}^{10} \lambda_k CR_{hk} + \delta_{t\times Z(h)} + \beta' X_{ht} + \varepsilon_{ht} \tag{5.2}$$

where, CR is a dichotomous variable that is 1 for treated houses in concentric ring k with a fix radius to a metro station for the period after the metro station opening and 0 otherwise, and CR is an indicator for concentric rings. We create a series of ten concentric rings with 100-meter increments beginning from a radius of 100 m to 1 km. We find again positive and strongly significant coefficients for the first three concentric rings, so within 300 m catchment area around the metro stations, see Table 5.2. Moreover, the price premium for the first three concentric rings is monotonically decreasing from 5.4% to 2.8%, while it is small or even zero for other considered concentric rings. We therefore define the treatment areas to be all the houses in the 300 m catchment area around the metro stations.

Table 5.1. Treatment effect on house price for different Metro catchment areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Metro catchment area									
	<100 m	<200 m	<300 m	<400 m	<500 m	<600 m	<700 m	<800 m	<900 m	<1000 m
Treated areas × post metro opening (γ)	0.053***	0.034***	0.036***	0.012**	0.013***	0.016***	0.010**	0.010**	0.004	-0.001
	(0.017)	(0.009)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Treated areas (λ)	-0.048***	-0.053***	-0.006	-0.003	-0.006*	-0.005	-0.001	-0.005*	0.001	0.017***
	(0.014)	(0.007)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Zone specific year fixed effect $\left(\delta_{t\times Z(h)}\right)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.799	0.799	0.799	0.799	0.799	0.799	0.799	0.799	0.799	0.799
N	365,179	365,179	365,179	365,179	365,179	365,179	365,179	365,179	365,179	365,179

Notes: Dependent variable is the natural logarithm of house. ***,**, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and at the 0.10 level, respectively. Standard errors clustered at the hectare-net level are in parentheses.

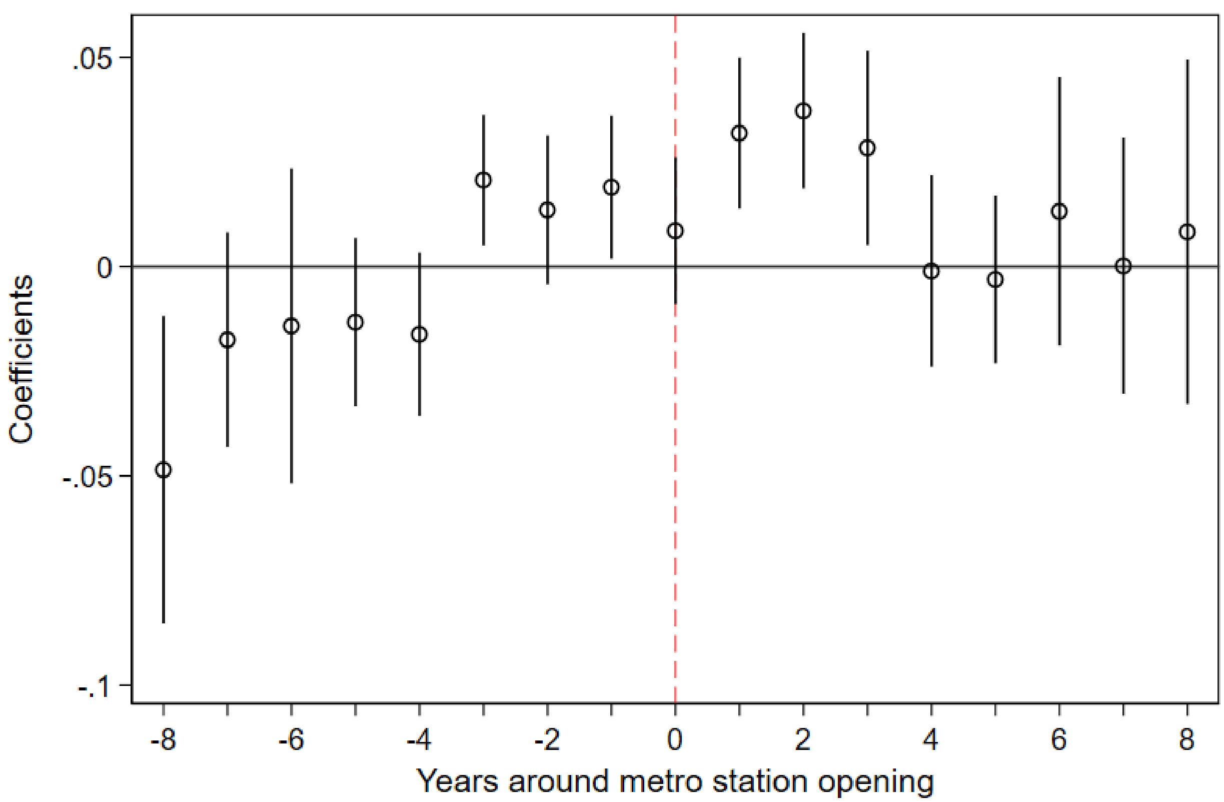
Table 5.2. Treatment effect on house price for concentric rings

	Concentric rings									
	0-100 m	101-200 m	201-300 m	301-400 m	401-500 m	501-600 m	601-700 m	701-800 m	801-900 m	901 m-1 km
Treated × post metro opening	0.053***	0.030***	0.028***	-0.019**	0.010*	0.008	-0.016	0.003**	-0.016**	-0.014
(γ_k)	(0.018)	(0.011)	(0.009)	(0.008)	(0.009)	(0.007)	(0.008)	(0.010)	(0.008)	(0.010)
Treated areas	-0.038***	-0.039***	0.034***	0.013**	0.001	0.012	0.019**	-0.003	0.025***	0.062***
(λ_k)	(0.014)	(0.008)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.007)
Zone specific year fixed effect $\left(\delta_{t\times Z(h)}\right)$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2					0.799					
N					365,179					

Notes: Dependent variable is the natural logarithm of house. ***,**, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and at the 0.10 level, respectively. Standard errors clustered at the hectare-net level are in parentheses.

Finally, we also estimate a dynamic DD model in which the treatment effects are estimated year-by-year. We focus on 8 years before and 8 years after the metro station openings. Figure 5.1 illustrates the coefficient estimates from this specification, in which each year period is allowed to have its own treatment effect. It shows results of a specification similar to column 3 in Table 5.1. The anticipation effect is positive and precisely estimated at about 2% for the three years before the metro station openings, while it is slightly higher (about 2.5%) for the three years after the metro station openings. Almost all the estimated coefficients are not for Figure 5.1 shows also that there are no pre-trends.

Figure 5.1. Dynamic DD specification on the log of house price



Notes: House price event time effects around the opening of the metro. The red vertical line marks the time of the metro station opening. The dark vertical lines indicate 95 percent confidence intervals.

We can use the estimated price premium resulting from the metro opening to calculate the value created on the housing market. We observe for all properties in the GCA the assessed values provided by Danish Property Assessment Agency.[12] In the DD model where the treatment areas are based on a series of ten concentric rings with 100-meter increments beginning from a radius of 100 m to 1 km, the price premiums for the first three concentric rings are 5.4%, 3.0% and 2.8%, respectively. Combining these two pieces of information we find that the metro generated a total value of DKK 5.0 billion. As this would pay for only 41% of the construction costs (see next section), it suggests that the Copenhagen metro project does not pass a cost-benefit test. This estimated gain largely accrues to private landowners, not to the government.

To put this number into its proper perspective, recall that the metro reduced one-way commutes by 5.7 minutes. A typical commuter in the GCA has an hourly wage of about DKK 250/hour and the median treatment property in our sample of DKK 2 million. The average commuter saves thus 2,092 DKK per year, when the one-way commuting is reduced by 5.7 minutes (see Table 3.2).[13] Assuming this value accrues every year into the future and a discount rate of human capital of $r = 2.5\%$ (see e.g. Lustig et al. (2013)) results in a present value of VOT saved of about DKK 84,000. This is similar to the baseline estimated price premium of 3-4% for properties in our treatment area, which amounts to 74,000 DKK (3.9% of DKK 1.9 million, the average house price in central Copenhagen). Notice here that our estimates ignore the general equilibrium effects of transit expansion which might be capitalized in the real estate prices but not yet precisely measured through commuting costs.

5.2. Land value capture and metro network

The city of Copenhagen applied a well-known strategy of using land value capture to raise funds for the metro network.[14] In 1992 the Danish parliament passed the Ørestad act which defined a financing model implying that the metro network will be financed by income generated from developing the Ørestad area, 310-hectare site stretching across the island of Amager between the city center and Copenhagen's international airport (zone 7, Copenhagen S).[15] More precisely, a

[12] For more information about the Danish Property Assessment Agency see https://www.vurdst.dk/english/about-the-danish-property-assessment-agency/.

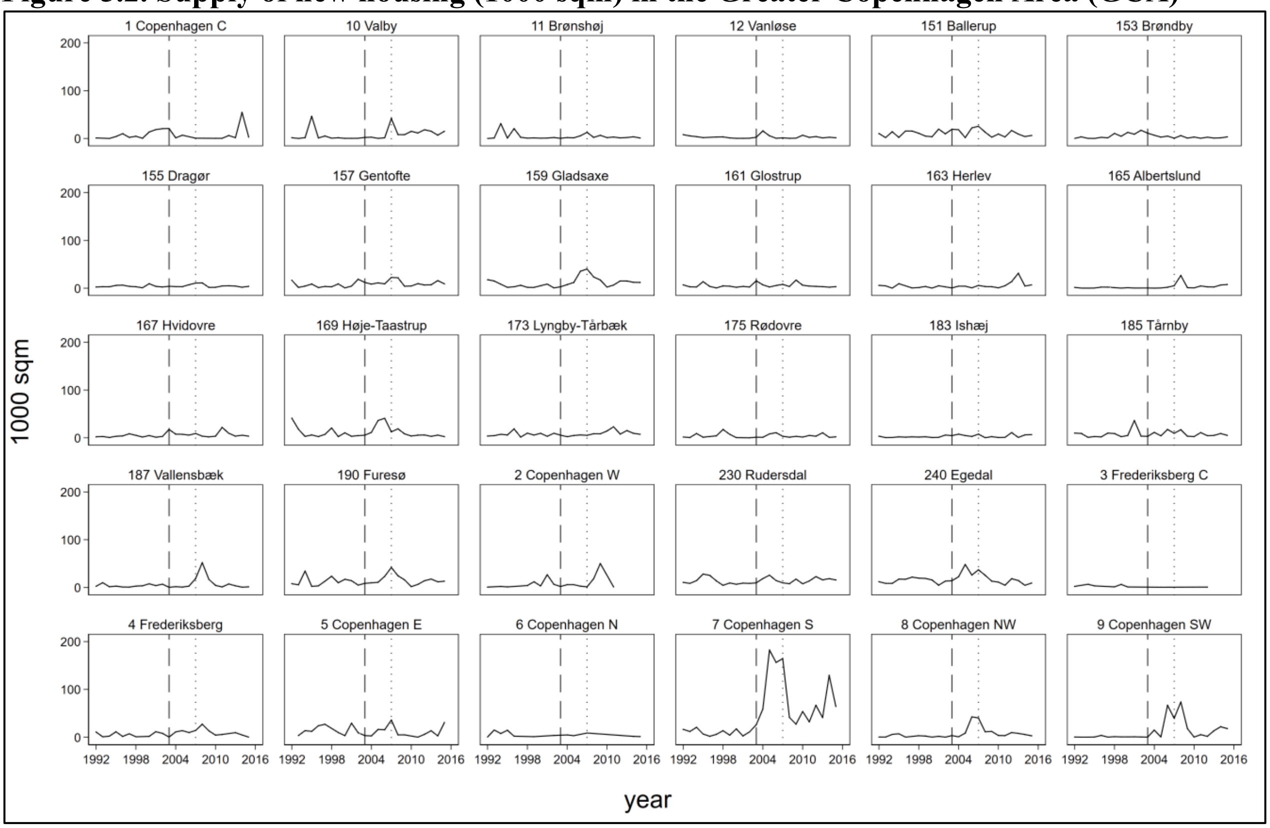
[13] The daily time saving for public transport commutes is 11.4 minutes (2 × 5.7 minutes). For an average commuter the daily travel time saving is 2.3 minutes (0.20 * 11.4 minutes), see Table 3.1. This corresponds to 502 minutes per year (2.3 minutes × 220 days). The value of the annual time saving is then DKK 2,092 (502 minutes × ((250 DKK/t) /60 minutes)).

[14] Other cities have also used similar strategies. For example, private transit companies internalize the spillover benefits of public transit projects by purchasing and developing land prior to station openings in Tokyo. For more examples see Medda (2012).

[15] The full development of Ørestad is expected to take 20–30 years.

state-owned enterprise, Ørestad Development Corporation (ØDC) operated by Copenhagen Municipality (55%) and the Danish Government (45%) was created in 1992 with the explicit goal of using the revenues generated by land development, rezoning and land sales to finance the construction of the metro. This involved several steps (Bruns-Berentelg et al., 2022).[16] First, the ownership of the vacant land was transferred to ØDC. The local government then rezoned the land for residential and commercial use and then borrowed money with loans on favourable terms from the Danish National Bank against the increased value of the land assets to fund the metro construction. Finally, ØDC raised revenues by land sales to service the debt. Although the ØDC was mandated to sell to the highest bidder, the expected profit was partly reduced due to the requirement that new-build developments have to contain 30% affordable and social housing.

Figure 5.2. Supply of new housing (1000 sqm) in the Greater Copenhagen Area (GCA)



Notes: dashed vertical lines indicate the metro line openings in 2003 and 2007.

[16] In 2007, Ørestad Development Corporation was restructured into City & Port. For an exhaustive description of Ørestad Development Corporation and City & Port see Noring (2019).

The initial budget of the metro was DKK 7.9 billion (approximately € 1 billion), while the final costs were DKK 12.2 billion (Vuk, 2005), so about 50% higher than initially expected. According to Enoch et al. (2005) the raised revenues by land sale failed to cover the construction costs.[17] We do not have data on these revenues, but we have information from the Building and Dwelling Register on housing supply in the GCA. Figure 5.2 shows expansion of the housing supply for considered zones. It is easy to see that the supply of new hosing increased significantly in zone 7 (Copenhagen S). The total new supply of housing for the period 2000-2010 in this zone sums to about 740 thousand sqm. Its value amounts to about DKK 15.3 billion, which is 25% above the final costs of the metro construction.[18] We also observe a substantial increase in zone 9 (Copenhagen SW) and a less pronounced but still significant raise in zone 8 (Copenhagen NW). Even if all this new housing supply was developed on land owned by ØDC, a substantial share of the value created is absorbed by construction costs.

5.3. Rezoning, new housing and the heterogeneous treatment

Our estimation results suggest substantial impact of the Metro on housing prices on average in all treated areas. Since anecdotal evidence suggests that the simultaneous arrival of the metro and the development of new housing boosted house prices in zone 7 (Copenhagen S), we now investigate if this zone indeed occupies a special position. We estimate a DD specification with the triple interaction effect $zone\ 7 \times treated\ areas \times post\ metro\ opening$ where $zone\ 7$ is a dichotomous variable that is 1 for properties in zone 7 (Copenhagen S) and 0 otherwise, see column 1 in Table 5.3. We find a 13% larger appreciation for properties in the treatment area in zone 7 after the metro opening and a negative impact for all other treatment areas.

Using the indicator for development of new housing in zone 7 confounds the effect of new housing supply and the effect of older buildings. To focus on the effect of newer buildings we estimate an additional specification in which we interact the treatment variable with a dichotomous variable $new\ properties$. We define new properties to be those not older than 5 years. Not surprisingly, newer properties trade at a substantial premium of 24.9% (exp (0.222)-1) compared with existing properties, see column 2 in Table 5.3. More interestingly, we find a 66.4%

[17] Moreover, the delay in the metro construction also resulted in additional interest payments on the loans (Enoch et al., 2005).

[18] We estimate value of the total new supply of housing in this zone by multiplying annual supply of new housing in this zone with year specific mean sqm price for the zone.

(exp (0.509)-1) larger appreciation for newer treated properties than for older properties.[19] Finally, column 3 in Table 5.3 reports a specification in which the treatment is interacted with property size, measured in log(sqm). The coefficient associated with this variable is small and not significantly different from zero. The much stronger effect in zone 7 is therefore not due to the larger size of the houses created there. It appears that the most important channel through which the metro created increases in real estate values in zone 7 (Copenhagen S) was through the combination of better accessibility and the development of new housing, which was probably associated with the creation of endogenous consumer amenities.

Table 5.3. Heterogeneous treatment

	(1)	(2)	(3)
Treated areas × post metro opening (γ)	-0.016**	0.010*	0.080*
	(0.007)	(0.015)	(0.047)
Zone 7 × treated areas × post metro opening	0.131***		
	(0.012)		
New properties × treated areas × post metro opening		0.509***	
		(0.069)	
log(sqm) × treated areas × post metro opening			-0.010
			(0.011)
Treated areas (λ)	-0.005	-0.007*	-0.006
	(0.004)	(0.004)	(0.004)
New properties		0.222***	
		(0.012)	
log(sqm)			0.872***
			(0.007)
Zone specific year fixed effect $\left(\delta_{t\times Z(h)}\right)$	Yes	Yes	Yes
House attributes	Yes	Yes	Yes
R^2	0.799	0.799	0.799
N	365,179	365,179	365,179

Notes: Dependent variable is the natural logarithm of house. ***,**, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and at the 0.10 level, respectively. Standard errors clustered at the hectare-net level are in parentheses.

[19] We have also estimated a specification with the triple interaction effect "property age × treated areas × post metro opening" as a robustness check. The coefficient associated with this variable is negative, confirming our results with respect to newer properties.

6. Implications for urban structure

6.1. A gravity model for commuting

To explore the implications of the metro network for urban structure, we use the register data on commuting flows between considered zones in 2002 and 2010 and construct a gravity model for commuting behaviour of the type that has become popular in the literature on Quantitative Economic Geography (see Redding and Rossi-Hansberg, 2017 for a review).[20] This commuting model can be regarded as the core of the urban economic models built in the quantitative spatial economics papers as it connects the jobs and the houses. In the reduced form version we use here the attractiveness of zones for the purposes of living and working are both summarized in a single parameter. We refer to these two sets of parameters as the urban structure. These parameters are directly related to the numbers of workers and jobs per zone.

The quantitative spatial economics literature emphasizes the connection between this gravity model and Eaton and Kortum's (2002) model of international trade, but transportation economists will recognize the model as being formally equivalent to a multinomial logit (MNL) model for the simultaneous choice of residence and workplace.[21] As will become clear below, the latter interpretation is convenient for connecting the commuting pattern to mode choice.

The gravity model is interpreted a referring to a general urban equilibrium setting. That is, the choices of the workers with respect to residential and work locations must satisfy constraints imposed on them by the housing and labour market. The number of households residing in a particular residential zone cannot exceed the number of houses available there and the number of workers employed in a particular residential zone cannot exceed the number of jobs that is available there. A change in the transportation infrastructure, such as the opening of a metro network, will have as an immediate, or direct, effect that the demand for the various home-work combinations changes. Some combinations of residential and employment areas become more attractive because of the arrival of the metro. However, the urban fabric is not necessarily able to facilitate these

[20] Gravity equations are also standard tools for the analysis of spatial relationships like trade flows and migration flows (see Anderson, 2011).

[21] Eaton and Kortum (2002) derived the model on the basis of multiplicative random utility, whereas McFadden (1973) worked with additive random utility. A logarithmic transformation of the random part of the utility suffices to transform the Eaton-Kortum formulation into the multinomial logit specification. The MNL is generally regarded as restrictive because of its 'independence of irrelevant alternatives' and the related 'red bus – blue bus' problem and the Eaton-Kortum formulation, which is formally equivalent, has the same properties. It is also well-known that – like the gravity model based on Eaton-Kortum's formulation - the MNL often performs well in empirical work and is easier to handle than generalizations like nested or mixed logit.

changes in commuting behaviour at the prevailing levels of attractiveness of residential and employment zone, and we should therefore in general expect the urban structure to change. It is the purpose of the present section to explore the changes in attractiveness that occurred in response to the opening of the metro network

6.2. The model

It is natural to consider the mode choice decision as conditional on the choice of the residential and work locations. In the discrete choice literature such 'decision trees' have been modelled through nested logit models and this structure is also convenient here. In the previous section we have studied the lower part of the decision tree, the mode choice, and here we connect it to the upper part. To do this, we introduce a summary measure of the mode choice model as an explanatory variable in the commuting choice model.

To see how this works, we postulate that the utility of a commute with residential location $i, i = 1 \ldots N$, and work location $j, j = 1 \ldots N$, equals:

$$u_{ij} = a_i + b_j + c_{ij} + \varepsilon_{ij} \,. \tag{6.1}$$

In this equation a_i is the utility of the residential location, b_j that of the work location, c_{ij} the generalized travel cost which is usually specified as a function of the travel time or distance, e.g. $c_{ij} = \gamma \; dist_{ij}$ with $dist_{ij}$ the distance between i and j, and ε_{ij} the idiosyncratic random part of the utility. The multinomial logit model requires that the ε_{ij}s are identically and independently Extreme Value type I distributed.

This model can be linked to the mode choice model discussed earlier in this paper by adding the inclusive value (or logsum) of that model as part of the travel cost.[22] That is, we further specify (6.1) as:

$$u_{ij} = a_i + b_j + \gamma \; dist_{ij} + \delta \; INCL_{ij} + \varepsilon^*_{ij} \tag{6.2}$$

where $INCL_{ij}$ denotes the inclusive value evaluate for the trip from i to j.

We have estimated models (1) and (2) for 2002 and 2010, allowing the attractiveness variables a_i and b_j to be different in both years. Column [1] in Table 6.1 presents the estimated coefficients for travel cost for the baseline gravity model that does not use the results of the mode choice model. Travel cost is modelled as a linear function of the distance between the residential

[22] See, for instance, Train (2003) for an extensive treatment of discrete choice models.

and work locations. We find a strongly significant negative coefficient for the travel distance. The cross effect with a dummy for the year 2010 is also significantly negative, suggesting that the resistance against long commutes has increased over time. In column [2] the logsum of the mode choice model is added. We now find a slightly smaller coefficient for the travel distance, while the coefficient for the logsum is strongly significant and relatively close to 1. The incorporation of the mode choice improves the performance of the model as is clear from the large increase in the loglikelihood. The negative coefficient for the cross effect with the year 2010 suggests that workers increasingly dislike long commutes in the course of time, which is consistent with a higher value of commuting time, that is possibly related to income growth.[23] This is confirmed by the significantly positive coefficient between the logsum and the year 2010.

Table 6.1 Estimation results of the commuting model

	[1]	[2]
Travel distance (km)	-0.105***	-0.094***
	(0.0004)	(0.0005)
Travel distance (km) * year 2010	-0.024***	-0.017***
	(0.0006)	(0.0007)
Logsum		0.849***
		(0.014)
Logsum * year 2010		0.041**
		(0.019)
Number of obs.	900	900
Log likelihood	-2,158,700	-2,154,812

Notes: ***, **, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and 0.10 levels respectively; standard errors are in parentheses

6.3. The metro and urban structure

As noted earlier the gravity model we estimated can be interpreted as a reduced form representation of a general equilibrium model of the urban economy in which the as and bs indicate the attractiveness of the various zones for working and living. They are determined by the labour and housing and land markets that are not made explicit here.

The direct impact of the metro on commuting behaviour is incorporated in the logsum variable. Some combination of residential and work locations become more attractive. This may have implications for house prices and wages and through them to the attractiveness of the various

[23] Gutiéerrez-i-Puigarnau et al. (2016) show that in Denmark the causal effect of household income on commuting distance is negative and in order of -0.18.

zones for living and working. We should therefore expect the a_is and b_js to change in response to the introduction of the metro network.

We can use our estimated model (6.2) to get a first idea of the impact of the metro on Copenhagen's spatial economy between 2002 and 2010 by carrying out a simple counterfactual analysis. Suppose that the parameters a_i and b_j all kept their initial – 2002 – values, and that only the value of the logsum changes, reflecting the direct impact of the metro. What would then have been the impact of the metro on commuting flows? It turns out that the changes in the commuting shares thus predicted by our model and the actual changes in the shares of Copenhagen commuters in the period 2002-2010 have a correlation coefficient of 0.25. This suggests that a large share of the impact of the metro has been absorbed by changes in housing and labour markets.

To see this, observe that in our counterfactual the attractiveness of all zones for living and working remains constant, while all changes in commuting flows induced by the metro are fully accommodated in each zone. Hence the counterfactual should be interpreted as referring to a situation in which supply of housing and demand for labour are perfectly elastic, while the realized changes in the numbers of houses and jobs do not affect the attractiveness of these zones. This is clearly unrealistic, but it provides, nevertheless, a useful benchmark. That the correlation between the actual changes in the commuting flows and the difference in flow between the counterfactual and the initial situation is lower than 1 confirms the well-known inelasticity of housing supply and labour demand in built-up areas. On the other hand, the fact that the correlation coefficient is positive, statistically significant and the difference in commuting flows non-negligible suggests that the metro had some impact on the commuting pattern in the period considered and thus confirms the analysis of housing supply induced by the opening of the metro network of the previous section.

Simple regressions also confirm that the changes in the shares of commutes between given residential and work locations are significantly related to changes in the logsum of the commute concerned.[24] There must have been changes in the numbers of workers or jobs (or both) in the various zones of Copenhagen to make this possible. In particular combinations of residential and work locations that became better connected were on average more popular in 2010.

[24] There is also a significant relationship between the change in the commute share and a dummy indicating that a metro station is present in both the zone of origin and destination.

6.4. Analysis

To investigate this issue in a more systematic way we consider the changes in the attractiveness parameters of the gravity model (6.2) in detail. In the Appendix B we show that the changes in the attractiveness parameters Δa_i can be decomposed into four parts:

i. the change in natural logarithm of the share of housing in the zone i, $\Delta\, ln\, s_i$,
ii. the change in accessibility of the zone that occurred because of the metro $\Delta ACCRM_i$,
iii. the further change in accessibility that is due to changes in the attractiveness of zones for working, $\Delta ACCRE_i$, and
iv. a term reflecting the size of the housing market that is identical for all zones $\Delta\, ln\, N$.

Hence we can write:

$$\Delta a_i = \Delta\, ln\, s_i - \Delta ACCRM_i - \Delta ACCRE_i + \Delta\, ln\, N \tag{6.3}$$

There is a similar decomposition for the changes in attractiveness parameters for employment zones:

$$\Delta b_j = \Delta \ln e_j - \Delta ACCEM_j - \Delta ACCER_j + \Delta \ln M \tag{6.4}$$

with (i) $\Delta \ln e_j$ the change in the share of jobs in zone j, (ii) $\Delta ACCEM_j$ the change in accessibility of zone j die to the opening of the metro network, (iii) $\Delta ACCER_j$ the further change in accessibility of employment zone j that is due to changes in the attractiveness of residential zones, and (iv) $\Delta \ln M$ the change in the number of workers.

These equations are identities that must always hold. Their significance lies in the light they shed on the validity simple benchmark cases like the simple counterfactual discussed above where urban structure – as incorporated in the values of the a_is and b_js - remains unchanged. If this case were valid we would have $\Delta a_i = 0,\ \Delta b_j = 0$ for all i and all j, respectively, and by implication $\Delta ACCRE_i = 0, \Delta ACCER_j = 0$, also for all i and all j, respectively. (6.3) and (6.4) then imply that the immediate impact of the metro, represented by the second terms $\Delta ACCRM_i$ and $\Delta ACCEM_j$, is exactly compensated by adjustments in the numbers of houses and jobs:

$$\Delta \ln s_i = \Delta ACCRM_i,\ \ i = 1 \dots I \tag{6.5}$$

$$\Delta \ln e_j = \Delta ACCEM_j,\ \ i = 1 \dots I \tag{6.6}$$

A linear regression with the left-hand variable of these equations as the dependent and the right-hand variable as the only independent results in a significant positive coefficient for the first equation (the estimate is 0.37 with a standard deviation 0.13), and no significant coefficient for the second. See the panel I of Table 6.2. This suggests that there has been an impact of the metro on the number of workers living in areas that realized an increase in accessibility due to its arrival, but no relationship with changes in the number of jobs. That is, it seems like housing supply reacted to some extent to the arrival of the metro, whereas labour demand did not.

Table 6.2. The benchmark cases

	Residential		Employment
		I	
	$\Delta \ln s_i$		$\Delta \ln e_j$
$\Delta ACCRM_i$	0.37**	$\Delta ACCEM_i$	0.41
	(0.13)		(0.86)
Constant	0.05	Constant	0.02
	(0.04)		(0.16)
R^2	0.18	R^2	0.08
		II	
	Δa_i		Δb_j
$\Delta ACCRM_i$	-0.60***	$\Delta ACCEM_i$	-0.94
	(0.15)		(1.10)
$\Delta ACCER_i$	-0.75*	$\Delta ACCRE_i$	-2.49
	(0.39)		(2.70)
Constant	-0.05	Constant	0.01
	(0.04)		(0.16)
R^2	0.39	R^2	0.04
		III	
	Δa_i		Δb_j
$\Delta \ln s_i$	0.46**	$\Delta \ln e_j$	0.99***
	(0.19)		(0.03)
Constant	0.23**	Constant	0.14***
	(0.02)		(0.01)
R^2	0.16	R^2	0.97
	$\Delta ACCRM_i$		$\Delta ACCEM_i$
$\Delta ACCER_i$	-0.97**	$\Delta ACCRE_i$	-1.47***
	(0.48)		(0.38)
Constant	-0.25***	Constant	-0.12***
	(0.03)		(0.02)
R^2	0.13	R^2	0.34

Notes: ***, **, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and 0.10 levels respectively; standard errors are in parentheses

A second useful benchmark is a completely inelastic supply of housing and demand for labour. The housing stock and also commercial real estate is given in the short run and requires substantial investment to change in the long run. This second benchmark imposes: $\Delta \ln s_i = \Delta \ln e_j = 0$. Substitution in (6.3) and (6.4) gives:

$$\Delta a_i = -\Delta ACCRE_i - \Delta ACCRM_i + \Delta\, ln\, N \qquad (6.7)$$

$$\Delta b_j = -\Delta ACCER_j - \Delta ACCEM_j + \Delta \ln M. \quad (6.8)$$

When estimating the corresponding regression equations, we find a significant coefficient for $\Delta ACCRM_i$, but not for $\Delta ACCEM_j$. This is consistent with the incomplete adjustment of housing supply suggested by our previous estimation, which now appears to be complemented by adjustment in house prices or amenities. See panel II of Table 6.2.

Our results thus far suggest that all adjustments of urban structure to the opening of the metro network took place on the housing market. If true, this would imply that the change in accessibility of employment zones that was caused by the metro, $\Delta ACCEM_j$, was completely compensated by the change in accessibility of employment zones that was due to the changes in the attractiveness of residential zones, $\Delta ACCER_j$. Hence Δb_j and $\Delta \ln e_j$ would be perfectly correlated, as would $\Delta ACCER_j$ and $\Delta ACCEM_j$. This separation of (6.4) into two independent parts can be regarded as a third benchmark situation. To investigate its relevance, we regress one variable of each pair on the other. We do the same for the corresponding pairs in equation (6.3) for the changes in residential attractiveness. For the labour market we find indeed the supposed perfect correlation between the changes in attractiveness and the changes in the numbers of workers. The standard errors in this equation are very small and the R^2 close to zero. The slope coefficient in the corresponding equation for the two accessibility changes is less precisely estimated but consistent with this finding. For the housing market we find a slope coefficient of roughly 0.5 in the equation for the changes in attractiveness, while the slope coefficient in the other equation is estimated to be close to 1, but with a large standard error. These estimation results are collected in panel III of Table 6.2.

We can summarize our findings as indicating that the labour market didn't adjust at all to the opening of the metro network, while the housing market reacted partly by changes in the number of houses, and partly by changes in the attractiveness of residential zones. These changes in attractiveness neutralized the direct impact of the metro on the attractiveness of residential zones.

6.5. Education and the impact of the metro

In section 3.4 it became apparent that the impact of the metro on commuting is closely associated with education. To analyse the impact of the metro on the location patterns of residences and jobs

for the three educational groups, we have repeated the analysis of the previous section each of them separately. To save space, we do not discuss the estimation results of the gravity model.[25]
For each group of workers we have carried out the same decomposition as we did for the model with homogeneous workers in the previous subsection. Note that urban structure refers to the numbers of workers and jobs of a specific educational groups and can therefore not be identified with the housing stock or total employment per zone. For instance, changes in the residential location pattern of the higher educated may be compensated by opposite changes in the location pattern of the medium and lower educated, while the total housing stock remains unchanged.

However, we find no support for the idea that the residential or employment location patterns of any group have changed in response to the opening of the metro network. See panel I of Table 6.3 below where we report the results of estimating (6.5) and (6.6) for the three educational groups. We do not find a single significant coefficient, which may suggest that the evolution of the number of workers and employment per zone was largely independent of the metro.

The changes in the attractiveness of residential zones are negatively related to those in accessibility. This effect is strongly significant for lower and medium educated workers, but not for the highly educated. See panel II of the Table 6.3. This suggests that for the lower and medium educated, the benefits of the metro were counteracted by increases in house prices and other changes that made living in the areas benefitting from improved accessibility less attractive. For the medium educated the coefficient for the change in accessibility due to the metro is smaller than for the lower educated. For the higher educated the coefficient is insignificant suggesting that for this group the changes in amenities were less detrimental. For the labour market we only find a weakly significant coefficient for the change in accessibility due to the metro for the lower educated. This may signal a negative response of wages or other job characteristics for this group in areas benefitting to improved accessibility from the metro. The general impression is that for the higher educated the benefits brought by the metro were not significantly counteracted by decreasing attractiveness of the residential or work locations concerned, whereas for the medium and especially for the lower educated they were. The opening of the metro network thus seems to have contributed to the gentrification of central Copenhagen.

[25] They are available from the authors upon request.

Table 6.3. Education and the impact of the metro on the attractiveness of residential and work locations

	Low		Medium		High	
	Residential	Employment	Residential	Employment	Residential	Employment
			I			
	$\Delta \ln s_i$	$\Delta \ln e_j$	$\Delta \ln s_i$	$\Delta \ln e_j$	$\Delta \ln s_i$	$\Delta \ln e_j$
$\Delta ACCRM_i$	-0.07		0.13		0.32	
	(0.07)		(0.19)		(.45)	
$\Delta ACCEM_i$		0.11		-0.004		0.45
		(.35)		(0.11)		(.85)
Constant	-0.05	0.057	0.12*	-0.01	0.24***	0.16
	(0.04)	(.15)	(0.06)	(0.17)	(.079)	(.10)
R^2	0.03	0.003	0.02	0.001	0.02	0.01
			II			
	Δa_i	Δb_j	Δa_i	Δb_j	Δa_i	Δb_j
$\Delta ACCRM_i$	-0.98***		-0.69***		-0.57	
	(0.08)		(0.22)		(0.43)	
$\Delta ACCEM_i$		-1.60*		-1.25		-0.27
		(0.80)		(1.86)		(1.50)
$\Delta ACCER_i$	-0.61***		-0.35		2.36	
	(0.19)		(0.50)		(2.20)	
$\Delta ACCRE_i$		-3.66		-1.81		2.17
		(2.70)		(4.9)		(4.82)
Constant	-0.04	0.05	0.14**	-0.03	0.21***	0.20
	(0.04)	(0.15)	(0.06)	(0.18)	(0.08)	(0.12)
R^2	0.87 5	0.16	0.30	0.02	0.11	0.01
			III			
	Δa_i	Δb_j	Δa_i	Δb_j	Δa_i	Δb_j
$\Delta \ln s_i$	0.51***		0.81***		0.92***	
	(0.04)		(0.15)		(0.08)	
$\Delta \ln e_j$		0.99***		1.00***		0.98***
		(0.08)		(0.02)		(0.03)
Constant	1.07**	0.27	0.27***	0.10	0.15***	0.07***
	(0.44)	(0.03)	(0.03)	(0.01)	(.03)	(0.01)
R^2	0.17	0.86	0.50	0.98	0.83	0.97
	$\Delta ACCRM_i$	$\Delta ACCEM_i$	$\Delta ACCRM_i$	$\Delta ACCEM_i$	$\Delta ACCRM_i$	$\Delta ACCEM_i$
$\Delta ACCER_i$	-0.14		-1.29***		-0.88	
	(0.39)		(0.34)		(0.96)	
$\Delta ACCRE_i$		-3.00***		-2.10***		-2.62***
		(0.28)		(0.30)		(0.33)
Constant	-0.46***	-0.08**	-0.24***	-0.08***	-0.13***	-0.05***
	(0.06)	(0.03)	(0.03)	(0.01)	(0.02)	(0.10)
R^2	0.32	0.81	0.35	0.64	0.03	0.68

Notes: ***, **, * indicate that estimates are significantly different from zero at the 0.01, 0.05 and 0.10 levels respectively; standard errors are in parentheses

Finally, panel III of Table 6.3 or all three groups of workers the change in the employment attractiveness is completely explained by that in the shares of jobs, suggesting that for each educational group the impact of the metro on the attractiveness of employment areas was completely neutralized by changes in the attractiveness of residential areas as we found for the population as a whole. For housing we find that changes in the share of housing explains a smaller part of the change in attractiveness, especially for the lower and medium educated. The change in employment accessibility due to the metro is closely correlated to the changes in housing attractiveness. For housing we don't find such a close correlation.

Summarizing, if we consider the three educational groups separately, we find that the earlier conclusion that the labour market did not react to the opening of the metro network is confirmed, whereas we now find that the counteracting changes in attractiveness on the labour market were largely limited to the lower and medium educated. The benefits of the metro thus appear to have been reaped especially by the higher educated.

8. Conclusion

This paper analysed the impact of the metro on the Greater Copenhagen metropolitan Area. We started with an examination of its impact on travel times and pursued the investigation up to a preliminary analysis of the change in urban structure. We combined information of travel surveys with register data on commuting behaviour. Moreover, we were able to distinguish between three educational groups. Our most important findings are:

- The metro implied substantial travel time savings, about 5 minutes, for commuting trips that can be realized completely with this new transport mode.
- We find a significant increase in the use public transport for commuting purposes, not only for trips that can be realized completely by the metro, but also for those that can be realized partly by metro. Since travel times did not change significantly for the later trips, other aspects of metro transport like frequency, reliability and comfort appear to be important characteristics of the metro.
- The metro had a significant impact on house prices close to metro stations. The total increase in real estate values was estimated as appr. 40% of the construction cost.
- For trips that can be partly of completely realized by metro, car use decreased, suggesting additional social benefits in terms of reduced commuting.
- Associated with the arrival of the metro was an increase in housing construction activity, induced primarily by the government in an attempt to capture some of the associated social gains. This attempt seems to have been successful.

- In the period 2002-2010 the labour market did not react to the opening of the metro network. Zones with metro stations in general experienced a decrease in residential attractiveness that counteracted the better accessibility resulting from the opening of the metro.
- Although the lower educated make intense use of the metro, the higher educated seem to prefer this transportation mode most. Moreover, the compensating changes in attractiveness in the urban equilibrium after the opening of the metro line seem to have been most detrimental to the lower educated, while they were negligible for the higher educated.

The paper thus sketches an interesting picture of the impact of a new transport mode in a gentrifying city with a local government actively attempting to realize part of the social gains for public purposes.

References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2016). The economics of density: Evidence from the Berlin wall. *Econometrica*, *83*(6), 1217–2189.

Anderson, James E. (2011). The Gravity Model. *Annual Review of Economics*, vol. 3, 133-160.

Bruns-Berentelg, J., L. Noring and A. Grydehøj. 2022. Developing urban growth and urban quality: Entrepreneurial governance and urban redevelopment projects in Copenhagen and Hamburg. *Urban Studies*, vol. 59(1), pp. 161–177.

Deaton, A. and J. Muellbauer (1980) An almost ideal demand system. *American Economic Review*, **70**, 312-326.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, *70*(5), 1741–1779. http://doi.org/10.1111/1468-0262.00352

Enoch, M. S. Potter and S. Ison. 2005. A Strategic Approach to Financing Public Transport Through Property Values. *Public Money and Management*, 25:3, pp. 147-154.

Glaeser, E. L. (2012). *Triumph of the City*. Penguin.

Gutiérrez-i-Puigarnau, E., I. Mulalic and J.N. van Ommeren (2016). Do rich households live farther away from their workplaces? *The Journal of Economic Geography*, 16, pp. 177-201, 2016.

Heblich, S., Redding, S.J., Sturm, D.M. (2020). The making of the modern Metropolis: Evidence from London*. *The Quarterly Journal of Economics,* 135 (4), 2059–2133.

Hybel, J and I. Mulalic (2022). Transportation and quality of life – Evidence from Denmark. *Transportation Research Part A*, 157, pp. 107-125.

Lustig, H. , Van Nieuwerburgh, S. , Verdelhan, A. (2013). The wealth-consumption ratio. *Review of Asset Pricing Studies*, 3 (1), 38–94.

McFadden, D. (1973) Conditional logit analysis of qualitative choice behaviour. In: P. Zarembka (ed) *Frontiers in econometrics.* Academic Press, 105-142.

Medda, F. (2012). Land value capture finance for transport accessibility: a review. *Journal of Transp Geography*, 25, pp. 154–161.

Mulalic, I. and J. Rouwendal (2020). Does improving public transport decrease car ownership? Evidence from a residential sorting model for the Copenhagen metropolitan area. *Regional Science and Urban Economics*, 83, 103543.

Noring, L. (2019). Public asset corporation: A new vehicle for urban regeneration and infrastructure finance. *Cities*, 88. pp. 125–135.

Redding, S. J., & Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, *9*, 1–47.

Train, K. (2003). Discrete choice models with simulations. Cambridge University Press.

The Danish Nature Agency, 2015. The Finger Plan - A Strategy for the Development of the Greater Copenhagen Area. Ministry of the Environment, ISBN: 978-87-7091-975-3.

Vuk, G. 2005. Transport impacts of the Copenhagen Metro. *Journal of Transport Geography*, vol. 13, issue 3, pp. 223-233.

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Worswick, G.D.N. and D.G. Champernowne (1954-55) A note on the adding-up criterion. *Review of Economic Studies*, **22**, 57-60.

Appendix

Appendix A. Data

Table A.1. Descriptive statistics – dwellings for period 1993-2016

Attributes	Mean	Std. dev.
Price (DKK M.)	2.022	6.684
Space (m^2)	99.00	46.88
Age (year)	61.06	36.96
Number of rooms	4.45	1.58
Multifamily housing, share	0.61	0.49
Residential housing, share	98.8	0.11
Occupied by the owner, share	0.68	0.46
Number of observations	366,762	

Appendix B. Decomposition of the changes in attractiveness in the gravity model

The gravity model is:

$$\pi_{ij} = e^{a_i+b_j+c_{ij}}/N,\ N = \sum_i \sum_j e^{a_i+b_j+c_{ij}} \qquad \text{(B.1)}$$

In this equation c_{ij} summarizes the transportation cost terms. Summation over all employment locations gives:

$$\left(\sum_j \pi_{ij} =\right) s_i = e^{a_i} \sum_j e^{b_j+c_{ij}} /N \qquad \text{(B.2)}$$

where s_i denotes the share of the workers located in i. If workers are homogeneous s_i is he share of the housing stock located in i. After taking logarithms, we can rewrite this as:

$$a_i = \ln s_i \ - \ln\left(\sum_j e^{b_j+c_{ij}}\right) + \ln N \qquad \text{(B.3)}$$

This equation states that the alternative-specific constant for residential location i equals the log of the share of the housing stock located in i minus a measure of employment accessibility from i and a constant.

Comparing two periods, 1 and 0, we can decompose the change in a_i as:

$$\Delta a_i = \Delta\, ln\, s_i - \Delta ACCRE_i - \Delta ACCRM_i + \Delta\, ln\, N \qquad \text{(B.4)}$$

In this equation $\Delta \ln s_i$ is the change in the share of the housing stock located in zone i. The next two terms refer to the change in employment accessibility of zone i which has been split in two parts, one referring to the change in employment attractivities b_j and the other to the change in transportation costs, as explained in (B.5) below, where we use superscripts for the periods: 0 refers to 2002, 1 to 2010. The term $\Delta \ln N$ is common to all changes in the attractiveness of residential zones and therefore of no interest for the present analysis.[26]

$$\begin{aligned}\Delta \ln\left(\sum_j e^{b_j + c_{ij}}\right) &= \ln\left(\sum_j e^{b_j^1 + c_{ij}^1}\right) - \ln\left(\sum_j e^{b_j^0 + c_{ij}^0}\right) \\ &= \left\{\ln\left(\sum_j e^{b_j^1 + c_{ij}^1}\right) - \ln\left(\sum_j e^{b_j^0 + c_{ij}^1}\right)\right\} + \left\{\ln\left(\sum_j e^{b_j^0 + c_{ij}^1}\right) - \ln\left(\sum_j e^{b_j^0 + c_{ij}^0}\right)\right\} \\ &= \qquad \Delta ACCRE_i \qquad + \qquad \Delta ACCRM_i\end{aligned} \tag{B.5}$$

The symbols in the last line of (B.5) refer to the expressions in curly brackets in the previous line. $\Delta ACCRE_i$ is the change in employment accessibility from i evaluated due to changes in the bs between 2002 and 2010. $\Delta ACCRM_i$ is the change in employment accessibility from i due to changes in travel cost between 2002 and 2010.

There is a similar decomposition for the changes in the b_j's, the attractivities of the employment locations:

$$\Delta b_j = \Delta \ln e_j - \Delta ACCER_j - \Delta ACCEM_j + \Delta \ln M \tag{B.6}$$

Were the interpretation of the symbols used is analogous to those in (B.4). Equations (B.4) and (B.6) are identities that always hold exactly in the context of the model.

[26] Note that the model only determines the a_is up to an additive constant.