IFAU
Institute for Evaluation of Labour
Market and Education Policy

Essays on urban economics

Ina Blind

DISSERTATION SERIES 2015:2

Presented at the Department of Economics, Uppsala University

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P O Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

This doctoral dissertation was defended for the degree of Doctor in Philosophy at the Department of Economics, Uppsala University, October 14, 2015. Essay 1 has recently been published by IFAU as Working paper 2015:25.

ISSN 1651-4149

Dissertation presented at Uppsala University to be publicly examined in Ekonomikum, Hörsal 2, Kyrkogårdsgatan 10 C, Uppsala, Wednesday, 14 October 2015 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Christine Whitehead (The London School of Economics and Political Science (LSE)).

Abstract

Blind, I. 2015. Essays on Urban Economics. *Economic studies* 153. xii+199 pp. Uppsala: Department of Economics. ISBN 978-91-85519-60-6.

This thesis consists of four self-contained essays.

Essay 1 (with Olof Åslund and Matz Dahlberg): In this essay we investigate the impact of commuter train access on individual labor market outcomes. Our study considers the exogenous introduction of a commuter train linking locations in the northern part of Uppsala County (Sweden) to the regional employment center, considerably decreasing commuting times by public transit to the center for those living close to the pre-existing railroad. Using difference-in-differences matching techniques on comprehensive individual panel data spanning over a decade, our intention-to-treat estimates show that the reform had mainly no impact on the earnings and employment development among the affected individuals.

Essay 2: In this essay I look into the role of public transit for residential sorting by studying how the introduction of a commuter train linking locations in the northern part of Uppsala County (Sweden) to the regional employment center affected migration patterns in the areas served. Using a difference-in-difference(-in-difference) approach and comprehensive individual level data, I find that the commuter train had a positive effect on overall in-migration to the areas served and no effect on the average out-migration rate from these areas. With regards to sorting based on labor market status, I find no evidence of sorting based on employment status but some evidence that the train introduction increased the probability of moving out of the areas served for individuals with high labor incomes relative to the probability for individuals with lower income. Considering sorting along other lines than labor market status, the analysis suggests that people born in non-western countries came to be particularly attracted towards the areas served by the commuter train as compared to other similar areas.

Essay 3: In this essay I look into the relation between housing mix and social mix in metropolitan Stockholm (Sweden) over the period 1990-2008. Using entropy measures, I find that although the distribution of tenure types over metropolitan Stockholm became somewhat more even over the studied period, people living in different tenure types still to a large extent tended to live in different parts of the city in 2008. The degree of residential segregation was much lower between different population groups. I further find that the mix of family types, and over time also of birth region groups and income groups, was rather different between different tenure types in the same municipality. The mix of different groups however tended to be similar within different tenure types in the same neighborhood. While the entropy measures provide a purely descriptive picture, the findings thus suggest that tenure type mix could be more useful for creating social mix at the municipal level than for creating social mix at the neighborhood level.

Essay 4 (with Matz Dahlberg): The last decade's immigration to western European countries has resulted in a culturally and religiously more diverse population in these countries. This diversification manifests itself in several ways, where one is through new features in the cityscape. Using a quasi-experimental approach, essay 4 examines how one such new feature, public calls to prayer, affects neighborhood dynamics (house prices and migration). The quasi-experiment is based on an unexpected political process that lead way to the first public call to prayer from a mosque in Sweden combined with rich (daily) information on housing sales. While our results indicate that the public calls to prayer increased house prices closer to the mosque, we find no evidence that the public calls to prayer served as a driver of residential segregation between natives and people born abroad around the mosque in question (no significant effects on migration behavior). Our findings are consistent with a story where some people have a willingness to pay for the possibility to more fully exert their religion which puts an upward pressure on housing in the vicinity of a mosque with public calls to prayer.

Acknowledgements

First and foremost I would like to thank my main supervisor Matz Dahlberg. Matz has been invaluable for the completion of this thesis. He has provided me with just the right mix of pressure and encouragement and taught me much about economic research and academic writing. I would also like to thank my assistant supervisor Olof Åslund for intellectually challenging discussions.

My graduate studies however started under the guidance of Bengt Turner who introduced me to housing and urban economics and who is greatly missed after passing away. I am grateful to Matz Wilhelmsson who temporarily was my supervisor in the emptiness left by Bengt.

I would also like to thank the discussants of my licentiate thesis, Peter Fredriksson, and on my final seminar, Peter Nilsson, who both carefully read my papers and gave useful feedback. A further person who has carefully read several of the papers in this thesis and given useful feedback is Roger Andersson, who I also would like to thank.

While writing this thesis, I have shared my time between Gävle/Uppsala and Toulouse. In Gävle/Uppsala, I have been working at the Institute for Housing and Urban Research (IBF). I would like to thank everyone there for making it such a good working place. Special thanks to Kerstin Larsson and Christina Kjerrman-Meyer who have helped me with administrative issues at many, many, times. In Uppsala, I have also enjoyed time at the Department of Economics and I am grateful for that opportunity. In Toulouse, I spent time at University Toulouse - Jean Jaurès. I am grateful to Christiane Thouzellier for welcoming me there and I would also like to thank the graduate students who shared their office with me and provided precious company.

I would not have joined the Ph. D. program in Economics at Uppsala University, would it not have been for the people working at the Department of Economics at Umeå University when I did my undergraduate studies there. They initiated my interest in Economics and were the ones who first planted the idea of writing a Ph. D. thesis in my head. Umeå University has continued to be important for my thesis work thanks to the Graduate School in Population Dynamics and Public Policy at Umeå University, of which I had the privilege of making part.

Sofi and Lisa, thank you for assuring that my stays in Uppsala and Gävle were not only about work.

Thanks to my sisters Anna and Viktoria who mean so much to me and always make me laugh. Thanks to my parents for advice and always believing in me. Finally, thanks to my sons Kimon and Olof, and to Grégory for everyday love and support.

Toulouse, August 2015
Ina Blind

Contents

Introduction

In 2014, 54 percent of the world's population lived in urban areas (United Nations, 2015). While urban areas potentially can offer many advantages – economic, social and environmental – they also present a lot of challenges. Apart from fundamental factors such as the provision of clean water and preventing the spread of diseases, other issues are for example transport infrastructure, equality and social cohesion within cities. All people in urban areas are not equally well off. Furthermore, while spatial proximity between people lies in the definition of an urban area, people with for example different incomes and different origins still tend to be segregated to some degree. As most of the countries in the world continue to urbanize (United Nations, 2015), a better understanding of the workings of cities thus seems important. This is the main aim of the research conducted within urban economics.

The field of urban economics "emphasizes the spatial arrangements of households, firms, and capital in metropolitan areas, the externalities which arise from the proximity of households and land uses, and the policy issues which arise from the interplay of these economic forces" (Quigley, 2008). Within this broad field, the four essays in this thesis are empirical studies that evolve around the questions of households' location choices and whether the location of households matter for other individual outcomes, e.g. in the labor market.

Households' location choices

A standard theoretic model of the spatial arrangements of households is the one developed by Alonso (1964), Mills (1967), and Muth (1969). In this model, housing and land prices decline with distance from the central business district (CBD) to compensate individuals for longer commutes. In this monocentric urban model, high-income workers consume more land and therefore choose to live where land is cheap, i.e., far from the CBD, while poor workers live close to the CBD.[1]

[1] The key condition for this is that the elasticity of land with respect to income is greater than the elasticity of the value of time with respect to income (see Becker, 1965). The validity of this condition has been questioned, see e.g., Wheaton (1977), LeRoy and Sonstelie (1983) and Glaeser, Kahn and Rappaport (2008).

The standard model just described predicts residential segregation by income, in a pattern corresponding to US cities at the time the model was developed. However, it corresponds less well to the spatial pattern of European cities where people with high income tend to live closer to the CBD and people with low income further away (see e.g., Wheaton, 1977). Also, the pattern does not correspond to all US cities (see e.g., Glaeser et al., 2008). The Alonso (1964), Mills (1967), and Muth (1969) model has therefore been developed to include for example multi-centric employment (see e.g. White, 1976), different transport modes (LeRoy & Sonstelie, 1983), housing stock deterioration and redevelopment (Brueckner & Rosenthal, 2009), and amenities (Brueckner, Thisse, & Zenou, 1999).

Further, while the Alonso (1964), Mills (1967), and Muth (1969) model predicts residential segregation by for example ethnicity to the extent that ethnicity and income are correlated, actual ethnic segregation often seems to be larger than what can be explained by differences in income and demographic characteristics (see e.g., Bayer, McMillan, & Rueben, 2004; Hårsman & Quigley, 1995). To explain this, later research has focused on factors such as discrimination on the housing market (see e.g., Ahmed & Hammarstedt, 2008; Yinger, 1998) and preferences for living with people with the same ethnicity as oneself (see e.g., Card, Mas, & Rothstein, 2008; Cutler, Glaeser, & Vigdor, 1999; Schelling, 1971).

Households' locations and other outcomes

While the above mentioned factors help to explain the location choices of households with different characteristics, a related question is whether the location matters for other individual outcomes, such as labor market outcomes and health.

Focusing on labor market outcomes, the literature discusses mainly two channels through which individual residential location may matter for the individual's outcomes. The first channel is through job access, where it is often thought that being closer to jobs has a positive effect on the probability of being employed and labor income. While the Alonso (1964), Mills (1967), and Muth (1969) model includes a relation between job access, i.e., distance to the central business district, and income, the relation is one way. Within the model, the labor market is fully competitive, productivity and wages are given and there is no unemployment. In the middle of the 1990s efforts therefore started to combine urban economic models with labor economic theories and develop models in which workers' location (land market), as well as wages and unemployment (labor market) are determined in equilibrium. One branch of this literature introduces spatial frictions to efficiency wage models (see e.g., Brueckner & Zenou, 2003; Ross & Zenou,

2008; Zenou, 2002, 2009; Zenou & Smith, 1995).[2] Another branch of the urban labor economics literature introduces spatial frictions to search-matching models (see e.g., Gobillon, Selod, & Zenou, 2007; Smith & Zenou, 2003; Wasmer & Zenou, 2002, 2006).[3]

The second channel through which it is thought that where an individual lives may matter for other individual outcomes is through neighborhood effects. The idea is that the neighborhood environment, in particular the population composition, may influence the individual, for example through exposure to peer norms or access to resources such as information about job openings (see e.g., Jencks & Mayer, 1990, for an early overview of empirical studies, and Edin, Fredriksson, & Åslund, 2003; Åslund & Fredriksson, 2009, and Ludwig et al., 2013, for more recent empirical evidence.)

The essays in this thesis

The first essay in this thesis connects to the urban labor market literature and the issue of whether households' location matters for labor market outcomes. More precisely, it examines if living close to good public transit infrastructure has any effect on individuals' labor income and employment status.

The other three essays aim at providing a better understanding of households' location choices. In Essay 2, I examine which types of people choose to live close to public transit infrastructure. In Essay 3, I look into the relation between the distribution of different tenure types over a city and the social mix within the city. Finally, in Essay 4, the importance of new religious attributes in the cityscape for households' location choices is examined. More specifically, it is examined if the start of public calls to prayer from a Mosque in the Stockholm region affected households willingness to live in the neighborhoods close to the Mosque and whether this changed the migration patterns of natives and foreign-born in the neighborhoods. The essays are described in more detail below.

Public transit infrastructure

The importance of job access through public transit for improving the functioning of the labor market and strengthening the economic position especially for marginal workers is a topic receiving considerable political attention. The infrastructural investments required are substantial and relatively easy to compute. The gains are harder to estimate and knowledge about how

[2] For the initial efficiency wage model see Shapiro and Stiglitz (1984).
[3] For the initial search-matching model see Mortensen and Pissarides (1999) and Pissarides (2000).

improvements in public transit affects firms and workers is limited (see e.g. Gibbons & Machin, 2006, for a literature review on transport and labor market linkages). Recent theoretical work however points to the importance of transport modes for generating differences in economic outcomes across groups (Gautier & Zenou, 2010) and some studies argue that availability of public transport is a key determinant for residential sorting (Glaeser et al., 2008).

To investigate these issues (in Essay 1 and Essay 2), I take advantage of the introduction of a commuter train on a pre-existing railroad in Sweden, considerably decreasing commuting times by public transit from the areas served to the local employment center (Uppsala). The studied case, *Upptåget*, was introduced in the early 1990s and connected locations north of Uppsala city to the local center and further to the greater Stockholm area. The institutional features suggest that the case is well suited for overcoming many of the methodological challenges typically present in this type of research (e.g. endogeneity and reversed causality). The stretch of Upptåget was governed by already existing railroad tracks and the timing was related to a legal change. Further, the train altered commuting opportunities and travel times by public transit in some areas, while leaving conditions unchanged for other similar areas on the same local labor market that can thus be used for comparisons.

In Essay 1 (written with Matz Dahlberg and Olof Åslund), to study the impact of commuter train access on individual labor market outcomes, we compare the development of labor market outcomes for individuals living in treated and non-treated areas just before the introduction of Upptåget. That is, we conduct an intention-to-treat (ITT) analysis. To control for observed and unobserved differences between the group of treatment and control individuals unrelated to the commuter train, the analysis is conducted using a difference-in-differences matching estimator. Of importance here is that we can follow each individual over a long time period, both before and after the introduction of the commuter train, which enables us to match on past outcomes. For the study we rely on population-wide longitudinal register data, compiled for research purposes by Statistics Sweden, and held by the Institute for Evaluation of Labour Market and Education Policy (IFAU). Among other things, the data contain rich and detailed information on demographic characteristics, income, employment, and education as well as detailed geographic information on the workplace and residential location of each individual. We find that the introduction of the commuter train essentially had no significant effects on the employment probability or labor earnings for those individuals that lived in the treated area before the commuter train was introduced.

In Essay 2, to look into the role of public transit infrastructure for residential sorting, I use a difference-in-differences(-in-differences) type of analysis to compare migration patterns in the areas treated with commuter train ac-

cess to the migration patterns in non-treated areas before and after the introduction of Upptåget. The primary focus in the study is whether there is sorting between people with different employment status and labor income, but I also consider sorting based on education level, age, sex, and birth region. For the study I rely on the same type of comprehensive individual data as in Essay 1, which allows me to separately identify in-migrants, out-migrants and stayers. I find that the commuter train Upptåget had a positive effect on overall in-migration to the areas served and no effect on the average out-migration rate from these areas. With regards to sorting based on labor market status, I find no evidence of sorting based on employment status but some evidence that the train introduction increased the probability of moving out of the areas served for individuals with high labor incomes relative to the probability for individuals with lower income. Considering sorting along other lines than labor market status, the analysis suggests that people born in non-western countries came to be particularly attracted towards the areas served by the commuter train as compared to other similar areas.

Housing mix and social mix

Another factor that might influence households' location choices is the spatial distribution of different types of housing. In Sweden like in many other countries, policies to create neighborhoods with mixed housing have been advocated as a means to obtain socially mixed neighborhoods. There is however little empirical evidence on the relation between housing mix and social mix. The aim of Essay 3 is to study this issue.

I focus on metropolitan Stockholm over the period 1990-2008 and the mix of housing with different tenure types. Stockholm is interesting in this respect for several reasons. First, in Sweden, socially mixed neighborhoods was stated a national housing policy goal in the mid-1970s, and housing mix a primary mean advocated to achieve it (see e.g. Holmqvist, 2009). Second, the goal has a general feature that it is partly motivated by the wish to counter overall residential segregation and to obtain social equality (see e.g. Bergsten & Holmqvist, 2007; Holmqvist, 2009). In their interviews with municipal planning departments and housing companies, Bergsten and Holmqvist (2007) find that the understanding and practice of the social mix policy in Sweden have been rather consistent since it was introduced as a national housing policy goal in the middle of the 1970's, with social mix policy remaining a general policy for counteracting socioeconomic segregation rather than ethnic segregation and with age groups and family types as other categories frequently cited as desirable to mix. Third, the tenure types of buildings are rather fixed in Sweden, where some buildings almost exclusively contain apartments inhabited by tenant-owners, other buildings exclusively contain apartments inhabited by renters, and private houses to a large

extent are owner-occupied. To the extent that different population groups tend to be over-represented in different tenure types, it thus seems reasonable to assume that the distribution of tenure types could affect the distribution of different population groups.

For the study, I rely on population-wide register data, compiled for research purposes by Statistics Sweden, and held by the Institute for Housing and Urban Research (IBF) at Uppsala University. Among other things, the data contain detailed information on demographic characteristics and income of each individual. Important for this study, the data also contain detailed geographic information on the residential location of each individual as well as information on the real estate the individual lives in. From information on legal form of ownership and housing type, it is possible to classify people into tenure types.

Using entropy measures on the data described above, I first calculate the degree of residential segregation between tenure types and between population groups –birth region groups, income groups, age groups and family types – and then whether on average the population mix is different within different tenures in the same area.

I find that although the distribution of tenure types over metropolitan Stockholm became more even over the studied period, people living in different tenure types still to a large extent tended to live in different parts of the city in 2008. The degree of residential segregation was much lower between different population groups. I further find that the mix of family types, and over time also of birth region groups and income groups, was rather different between different tenure types in the same municipality. The mix of different groups however tended to be similar within different tenure types in the same neighborhood. While the entropy measures provide a purely descriptive picture, the findings thus suggest that tenure type mix could be more useful for creating social mix at the municipal level than for creating social mix at the neighborhood level.

New features in the cityscape

As mentioned above, yet another factor that can affect households' location choices is different types of amenities or features in the cityscape. Essay 4 (written with Matz Dahlberg) examines how a new religious feature, public calls to prayer from a mosque in a Western country (Sweden), affects neighborhood dynamics in terms of house prices and migration behavior. We take advantage of an unexpected political decision that lead way to the first public calls to prayer from a mosque in Sweden (the Fittja Mosque in Botkyrka municipality in the Stockholm region). This allows us to examine the issue at hand by combining the hedonic price theory of house price capitalization with a quasi-experimental approach, yielding a hedonic difference-in-

differences estimator. By using data on housing sales with precise information on the date when an object was sold and where the object was geographically located, we are in a good position to estimate the effect of the public call to prayer events on house prices. Likewise, by using data on monthly in- and out-migration from each neighborhood, we are able to estimate the effects of the public call to prayer events on migration and sorting patterns close to the mosque.

It should be stressed that we do not think that it is the sound of the public calls to prayer in itself that is important for house prices or sorting patterns. There are few public calls to prayer (one every Friday at 1 pm), the loudspeakers are directed away from residential housing, and there is only a limited number of houses in the direct vicinity of the mosque. Instead, we think of the public calls to prayer as an expression of Islam that can be important for some Muslims, whereas some non-Muslims/natives might want to avoid expressions of Islam or Muslims/immigrants.

Regarding house prices, estimates from a distance-motivated (i.e., distance from the mosque) difference-in-differences specification indicate that, within Botkyrka municipality, the public call to prayer events made housing closer to the mosque relatively more expensive. Also, estimates obtained through the synthetic control method (Abadie, Diamond, & Hainmueller, 2010; Abadie & Gardeazabal, 2003) indicate that the public call to prayer events had a positive effect on house prices in Botkyrka municipality as a whole. Regarding sorting, we find no indications of either native flight/native avoidance or a relative increase of people born abroad in the neighborhoods close to the Mosque. Given the original character of Botkyrka municipality with 38 % immigrants (the highest share among the municipalities in the Stockholm region) our findings are consistent with a local revitalization story in neighborhoods where native-immigrant sorting has already taken place.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association, 105*(490), 493-505.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review, 93*(1), 113-132.

Ahmed, A. M., & Hammarstedt, M. (2008). Discrimination in the rental housing market: A field experiment on the Internet. *Journal of Urban Economics, 64*(2), 362-372.

Alonso, W. (1964). *Location and land use. Toward a general theory of land rent*: Harvard University Press Cambridge, MA.

Åslund, O., & Fredriksson, P. (2009). Peer effects in welfare dependence quasi-experimental evidence. *Journal of Human Resources, 44*(3), 798-825.

Bayer, P., McMillan, R., & Rueben, K. S. (2004). What drives racial segregation? New evidence using Census microdata. *Journal of Urban Economics, 56*(3), 514-535.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal, 75*(299), 493-517.

Bergsten, Z., & Holmqvist, E. (2007). *Att blanda?: En undersökning av planerares och allmännyttiga bostadsbolags syn på planering för en allsidig hushållssammansättning* IBF Researach report, 2007:1. Gävle: Uppsala university, Institute for Housing and Urban Research

Brueckner, J. K., & Rosenthal, S. S. (2009). Gentrification and neighborhood housing cycles: will America's future downtowns be rich? *Review of Economics and Statistics, 91*(4), 725-743.

Brueckner, J. K., Thisse, J.-F., & Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review, 43*(1), 91-107.

Brueckner, J. K., & Zenou, Y. (2003). Space and unemployment: The labor market effects of spatial mismatch. *Journal of Labor Economics, 21*(1), 242-262.

Card, D., Mas, A., & Rothstein, J. (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics, 123*(1), 177-218.

Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (1999). The rise and decline of the American ghetto. *Journal of Political Economy, 107*(3), 455-506.

Edin, P.-A., Fredriksson, P., & Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants - Evidence from a natural experiment. *Quarterly Journal of Economics, 118*(1), 329-357.

Gautier, P. A., & Zenou, Y. (2010). Car ownership and the labor market of ethnic minorities. *Journal of Urban Economics, 67*(3), 392-403.

Gibbons, S., & Machin, S. (2006). *Transport and labour market linkages: empirical evidence, implications for policy and scope for further UK research*: Commissioned for the Eddington Study.

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics, 63*(1), 1-24.

Gobillon, L., Selod, H., & Zenou, Y. (2007). The mechanisms of spatial mismatch. *Urban Studies, 44*(12), 2401-2427.

Hårsman, B., & Quigley, J. M. (1995). The spatial segregation of ethnic and demographic groups: Comparative evidence from Stockholm and San Francisco. *Journal of Urban Economics, 37*(1), 1-16.

Holmqvist, E. (2009). *Politik och planering för ett blandat boende och minskad boendesegregation. - ett mål utan medel? [Policy and planning for social and housing mix and decreased housing segregation. - A goal without means?]*. PhD Thesis, Uppsala university, Uppsala.

Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In L. E. Lynn & M. G. H. McGeary (Eds.), *Inner city poverty in the United States* (pp. 111-186). Washington, DC: National Academy Press.

LeRoy, S. F., & Sonstelie, J. (1983). Paradise lost and regained: Transportation innovation, income and residential location. *Journal of Urban Economics, 13*(1), 67-89.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R., & Sanbonmatsu, L. (2013). Long-term neighborhood effects on Low-income families: Evidence from Moving to Opportunity. *American Economic Review, 103*(3), 226-231.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review, 57*(2), 197-210.

Mortensen, D. T., & Pissarides, C. A. (1999). New developments in models of search in the labor market. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (Vol. 3, pp. 2567-2627). Amsterdam: Elsevier Science.

Muth, R. F. (1969). *Cities and housing*: University of Chicago press.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. Cambridge, MA: MIT press.

Quigley, J. M. (2008). "urban economics". In S. N. Durlauf and L. E. Blume (Eds), *The New Palgrave Dictionary of Economics*. Second Edition. Palgrave Macmillan. The New Palgrave Dictionary of Economics Online. 13 August 2015. <http://www.dictionaryofeconomics.com/article?id=pde2008_U000035> doi:10.1057/9780230226203.1770

Ross, S. L., & Zenou, Y. (2008). Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory. *Regional Science and Urban Economics, 38*(5), 498-517.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*, 143-186.

Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review, 74*(3), 433-444.

Smith, T. E., & Zenou, Y. (2003). Spatial mismatch, search effort, and urban spatial structure. *Journal of Urban Economics, 54*(1), 129-156.

United Nations. (2015). *World Urbanization Prospects. 2014 Revision*. New York.

Wasmer, E., & Zenou, Y. (2002). Does city structure affect job search and welfare? *Journal of Urban Economics, 51*(3), 515-541.

Wasmer, E., & Zenou, Y. (2006). Equilibrium search unemployment with explicit spatial frictions. *Labour Economics, 13*(2), 143-165.

Wheaton, W. C. (1977). Income and urban residence: An analysis of consumer demand for location. *American Economic Review, 67*(4), 620-631.

White, M. J. (1976). Firm suburbanization and urban subcenters. *Journal of Urban Economics, 3*(4), 323-343.

Yinger, J. (1998). Evidence on discrimination in consumer markets. *Journal of Economic Perspectives, 12*(2), 23-40.

Zenou, Y. (2002). How do firms redline workers? *Journal of Urban Economics, 52*(3), 391-408.

Zenou, Y. (2009). *Urban Labor Economics*. Cambridge: Cambridge University Press.

Zenou, Y., & Smith, T. E. (1995). Efficiency wages, involuntary unemployment and urban spatial structure. *Regional Science and Urban Economics, 25*(4), 547-573.

Essay 1: All aboard? Commuter Train Access and Labor Market Outcomes*

Co-authored with Olof Åslund and Matz Dahlberg

* We thank John Östh for providing a distance table for locations (SAMS) in the local labor market of Uppsala, Eva Jirner for creating the present map, and Åsa Bråmå for creating maps that were used in the study design. We are grateful for comments from Peter Fredriksson, Elisabet Viladecans Marsal and seminar participants at the 2011 European Meeting of the Urban Economics Association in Barcelona and at the 2011 National Conference in Economics in Uppsala. Ina Blind also thanks the Graduate School in Population Dynamics and Public Policy at Umeå University for financial support.

1 Introduction

The importance of job access via public transit in improving the functioning of the labor market and strengthening the economic position of workers, is a topic receiving considerable political attention. The infrastructural investments required are substantial and relatively easy to compute. The gains are harder to estimate, and knowledge, for example about the labor market impact of changes in job access and commuting times, is limited. However, recent theoretical work points to the importance of transport modes for generating differences in economic outcomes across groups (Gautier & Zenou, 2010) and some studies argue that the availability of public transit is a key determinant for cross-group differences in geographical distribution (Glaeser, Kahn, & Rappaport, 2008).

We investigate these issues, studying the individual labor market effects of a commuter train considerably decreasing commuting times by public transit to the employment center for those living close to the pre-existing railroad on which the commuter train was introduced. *Upptåget* (the case we study), was inaugurated in the early 1990s, connecting locations north of the city of Uppsala, Sweden, to the local center and further to the greater Stockholm area. We argue below that the institutional features suggest that the case is well suited for overcoming many of the methodological challenges typically present in this type of research.[1] While the location of the train was governed by pre-existing railroad, the timing was related to a legal change. The train altered commuting opportunities and travel times for some areas, while leaving conditions unchanged for other areas included in the same local labor market. It can also be argued that—at least in a European context—we address the more policy-relevant margin: the effects of improving public transport rather than introducing it in a context where only private transport has been available previously.

Theory suggests a number of reasons why commuting opportunities may affect the employment and wages of individual workers. The literature is described in more detail in the next section, but let us here only point out a few potential mechanisms. First, shorter travel times or less expensive commutes may increase the optimal job search area (Gobillon, Selod, & Zenou, 2007) and may also decrease the reservation wage (Brueckner & Zenou, 2003; Coulson, Laing, & Wang, 2001), leading to decreased unemployment. In other words, the effective labor market is increased, which should im-

[1] Our approach is in line with the suggestions made by Gibbons and Machin (2006), in their literature study on transport and labor market linkages, on how to deal with the problems of endogeneity and to identify a causal impact of transport or transportation policy on labor market outcomes.

prove matching. Long commutes may also affect the productivity of workers, either because the commute itself requires a lot of effort (Zenou, 2002), or because the commute affects the flexibility between work and other commitments (Ross & Zenou, 2008). This may make employers reluctant to hire people living too far away, or induce workers to shirk, which increases the risk of unemployment.

Although there is a large number of empirical studies on the impact of job access on labor market outcomes following Kain (1968), it is only relatively recently that transport modes and transport infrastructure have been taken into account in this line of research.[2] There are thus some studies on the importance of car ownership or car access (e.g., Gurley & Bruce, 2005; Ong, 2002; Raphael & Rice, 2002; Shen & Sanchez, 2005) and a few studies focusing on job access by public transit. Some US studies suggest no or little relation between job access by public transit and employment (e.g., Cervero, Sandoval, & Landis, 2002; Sanchez, Shen, & Peng, 2004); whereas others find a positive association (e.g., Kawabata, 2003; Ong & Houston, 2002; Sanchez, 1999).

Evidence regarding the importance of job access by public transit is thus mixed and pertains mainly to the US. It is possible that the effects of new public transit infrastructure are different in Europe where the public transit network in and around cities is generally more extended than in the US. Matas, Raymond and Roig (2010) study the importance of job access by public transit in Barcelona and Madrid, Spain, and find a positive effect on employment probability among women, primarily among the low-educated.

We use comprehensive, individual panel data for the years 1985–1996, including detailed geographical information on residential and workplace location, as well as on labor market outcomes. We combine a difference-in-differences approach with matching methods to compare the development of labor market outcomes for individuals living in treated and non-treated areas before the introduction of the train.

We find that the introduction of the commuter train essentially had no significant effects on employment probability or labor earnings for those individuals living in the treated areas before the new commuter train was introduced. A large set of robustness checks and supplementary analyses confirm the impression that getting access to the commuter train did not significantly alter the labor market development of the treated individuals.

[2] Kain (1968) suggested that the high unemployment rate of African-Americans in US metropolitan areas was aggravated by the movement of low-skilled jobs from the central cities to the suburbs, worsening job access for African-American workers constrained to central cities by housing market discrimination (the spatial mismatch hypothesis). Since the study by Kain (1968) a large number of empirical studies have been carried out which attempt to test the relation between job access and labor market outcomes in general and the spatial mismatch hypothesis in particular. The collected evidence suggests that poor access to jobs does indeed lead to worse labor market outcomes (for literature surveys see Gobillon et al., 2007; Ihlanfeldt, 2006; Ihlanfeldt & Sjoquist, 1998; Zenou, 2009).

The rest of the paper is organized as follows: in Section 2 we briefly present some theories as to why job access can matter for individual labor market outcomes. Section 3 describes the development of the commuter train Upptåget and defines the treatment group and the potential control group. Section 4 explains the empirical strategy and presents the data used as well as some descriptive statistics. The results from the empirical analysis are presented in Sections 5 and 6, and Section 7 concludes.

2 Theories

The introduction of the commuter train Upptåget considerably decreased commuting times by public transit from the stations towards the employment center of Uppsala city and further south towards Stockholm, and thus led to improved job access close to the stations. What does theory lead us to expect about the effect of the commuter train on individuals' labor market outcomes?

In the standard urban economic model developed by Alonso (1964), Mills (1967), and Muth (1969), housing and land prices decline with distance from the central business district (CBD) to compensate individuals for longer commutes. In this monocentric urban model, high-income workers consume more land and therefore choose to live where land is cheap, i.e., far from the CBD, while poor workers live close to the CBD.[3] In the model and versions thereof, for example including different transport modes (LeRoy & Sonstelie, 1983) and decentralized or multi-centric employment (e.g., White, 1976), the labor market is fully competitive, productivity and wages are given and there is no unemployment. Thus, although the models include a relation between job access and income, length of commute cannot affect individuals' labor market outcomes.

In the middle of the 1990s efforts began to combine urban economic models with labor economic theories and develop models in which workers' location (land market), as well as wages and unemployment (labor market) are determined in equilibrium (for a synthesis, see Zenou, 2009). While most of these models do not take transport modes into account, they may still be relevant at least to the extent that people rely on public transit.

One branch of this literature introduces spatial frictions to efficiency wage models (see e.g. Brueckner & Zenou, 2003; Ross & Zenou, 2008; Zenou, 2002, 2009; Zenou & Smith, 1995).[4] In some of these models work effort and thus productivity is allowed to vary with the length of commute, either

[3] The key condition for this is that the elasticity of land with respect to income is greater than the elasticity of the value of time with respect to income (see Becker, 1965). The validity of this condition has been questioned, see e.g. LeRoy and Sonstelie (1983) and Glaeser, Kahn and Rappaport (2008).

[4] For the initial efficiency wage model see Shapiro and Stiglitz (1984).

because the commute itself requires a lot of effort (Zenou, 2002), or because the commute affects the flexibility between work and other commitments (Ross & Zenou, 2008). The implications of this for employment and wages depend on the ability of employers to observe workers' commuting costs and to anticipate workers' behavior. Within these models, the new and faster commuting opportunity could reduce work related fatigue and increase workers' leisure time, thus making people with longer commutes more productive at work. This could in turn make employers more willing to hire workers living far away/reduce the workers' risk of being caught shirking and with that the number of unemployment spells. Increased leisure time and a reduction of commuting related fatigue could have some effect even on people who did not rely on public transit but on car before the introduction of Upptåget. Switching from driving a car to riding a train would allow them to rest, work or do some errands (e.g. send mails, pay bills, make phone calls) during the commute.

Another branch of the urban labor economics literature introduces spatial frictions to search-matching models (see e.g., Gobillon et al., 2007; Smith & Zenou, 2003; Wasmer & Zenou, 2002, 2006).[5] Studies in this vein suggest that the introduction of the commuter train could help people to higher search efficiency and search intensity, increasing their employment probabilities and probabilities of finding better paying jobs.

A spatial search-matching model that includes mode-choice is provided by Gautier and Zenou (2010). In the model, because of initial wealth differences, whites can buy cars while ethnic minorities have to rely on public transit. Since the set of jobs that can be reached by car is larger than the set that can be reached by public transport, whites find jobs more quickly and experience shorter unemployment spells. Furthermore, a worker's bargaining position depends on what employers know or suppose about car ownership among white and ethnic minorities (statistical discrimination), resulting in higher wages for whites. In this model, better public transport such as the commuter train Upptåget should reduce differences in labor-market outcomes between whites and ethnic minorities.

To the extent that workers' residential locations are fixed, there are also models where workers may refuse jobs involving commutes that are too long because commuting to that job would be too costly in view of the proposed wage (Brueckner & Zenou, 2003; Coulson et al., 2001). This can depress both wages and employment rates in areas where the number of jobs are few relative to the labor pool. The new and faster commuting opportunity could here allow people to accept jobs that they would previously not have accepted, positively affecting their employment and earnings. These models could be most relevant for groups with stronger residential constraints, e.g., people

[5] For the initial search-matching model see Mortensen and Pissarides (1999) and Pissarides (2000).

with low incomes for whom credit constraints can limit residential choices and immigrants for whom different types of discrimination can be limiting (see e.g. Ahmed & Hammarstedt, 2008, for evidence on discrimination on the Swedish rental market).

In the long run it is possible that better labor market outcomes brought about by Upptåget are amplified if the train helps people's careers take off. However, it is also possible that people trade off the improved commuting opportunities for other things, e.g., larger housing further away. It is also possible that with time there is increased job competition from people moving into the station localities. Competition for jobs close to the stations could also come from people who now reverse commute from Uppsala city to the other station localities. Theoretically, a large in-migration could also raise housing costs and force some people to move and give up the improved job access. In practice, however, there is little direct pressure on housing costs for those who own their housing and since rents are regulated in Sweden and location and job access are not very important in the rent setting, the changes in housing costs were probably also small for people renting their housing. On the other hand, the improved commuting opportunities could retain or attract firms to the station localities, thus further increasing job access in these places.

There is thus some uncertainty regarding the direction of the effects, even though most mechanisms would indicate that better access to jobs would theoretically mean better labor market outcomes. What is more unclear, however, is the magnitude (or even presence) of the empirical impact. We essentially do not know how people value or are able to take advantage of a given decrease in expected commuting time by public transit. The empirical study performed below aims to provide some information on this topic, which is of core relevance for policy in the area.

3 Upptåget and the research design

We employ a quasi-experimental research design that builds on the introduction of a commuter train, Upptåget, to the Swedish city of Uppsala. Our definition of treatment and control groups is based on residential location in 1989; before public discussion on the commuter train began, and two years before the trains started running. We compare those who then lived in the part of the local labor market of Uppsala where the commuter train was introduced, to individuals who lived in two other parts of the same local labor market, but which were not subject to changes in transport infrastructure. That is, we will conduct an intention-to-treat (ITT) analysis. To control for observed and unobserved differences between the group of treatment and control individuals unrelated to the commuter train, the analysis is conducted using a difference-in-differences matching estimator. Below, we first give

some institutional detail and describe the decision process and the implementation of the commuter train. We then proceed to a more detailed description of the research design.

3.1 Upptåget

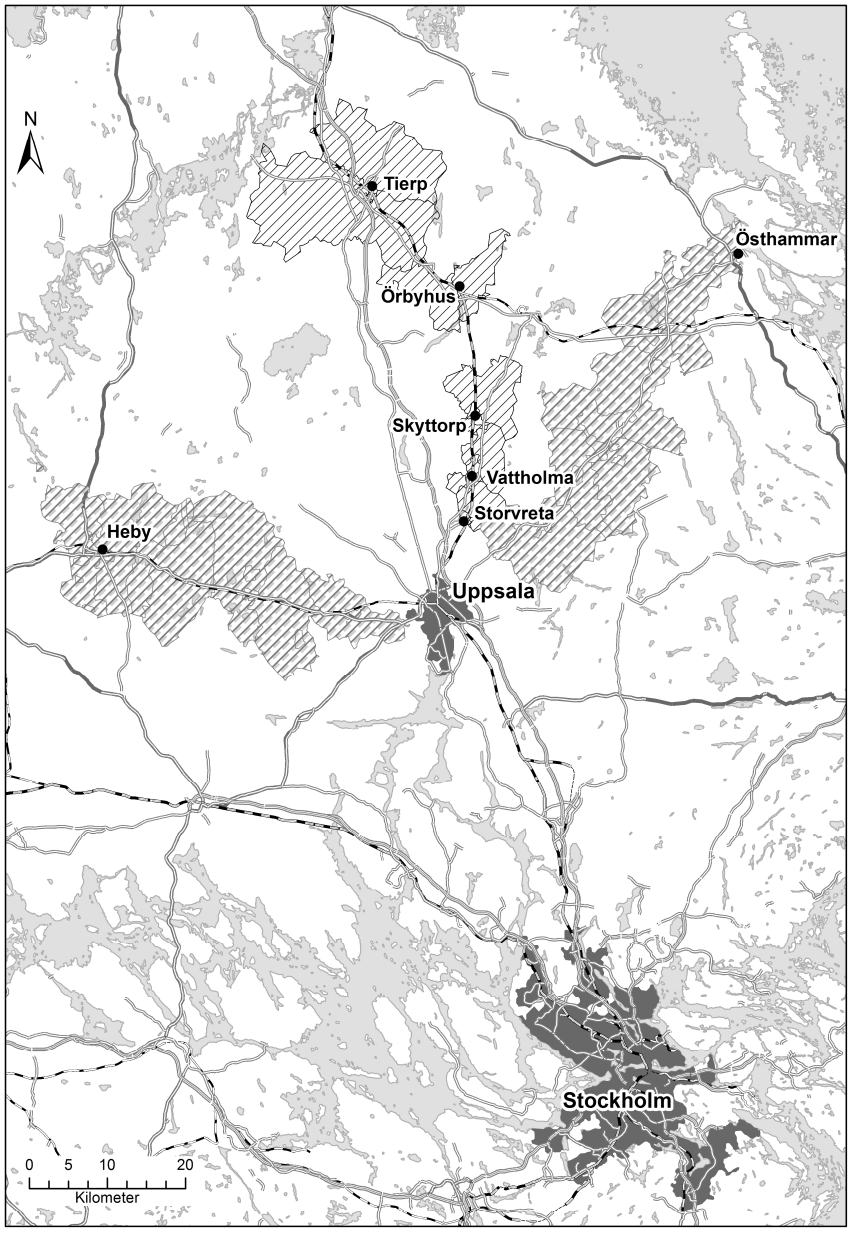
The central node of the local labor market of Uppsala is Uppsala city, which is the main destination for labor commuting from the surrounding municipalities of Enköping, Heby, Tierp, and Östhammar.[6] Upptåget runs between Uppsala city and the principal localities in Tierp municipality, Tierp town, 54 km north, see *Figure 1*. Before the introduction of Upptåget, all public transit within the municipalities of Heby, Tierp, and Östhammar, as well as between these municipalities and Uppsala municipality was by bus.[7] The only exception was a long-distance train with stops in Tierp town and in Uppsala city. Public transit by road was coordinated and purchased by Upplands Lokaltrafik (UL), a firm jointly owned by Uppsala County Council and the municipalities in the county.[8]

[6] The local labor markets are defined by Statistics Sweden based on commuting patterns. From 1996, Uppsala municipality and with it the municipalities of Enköping, Heby, Tierp, and Östhammar came to belong to the local labor market of Stockholm. However, Uppsala city continued to be the main destination for labor commuting from the municipalities of Enköping, Heby, Tierp, and Östhammar.

[7] The information on the public transit network in this and the following paragraphs is from timetables and annual reports from UL unless otherwise indicated. We thank Mats O Karlsson, member of Uppsala County Council 1988-2006 for checking the accuracy of the information.

[8] Before 2007 the municipality of Heby belonged to Västmanland County and not to Uppsala County, so public transit within Heby municipality was not organized by UL.

Figure 1. The Upptåget, Heby, and Östhammar-corridors



Note: The shaded (north-bound) area between Uppsala and Tierp constitutes the "treatment corridor". The shaded areas between Uppsala and Heby (west-bound) and between Uppsala and Östhammar (northeast-bound) constitute the "control-corridors". When inaugurated in 1991, Upptåget stopped in Storvreta, Vattholma, Skyttorp, and Örbyhus between the end stations of Uppsala and Tierp. The map was created by Eva Jirner.

The first concrete plans for Upptåget seem to have been outlined in 1988 and were accompanied by a trial trip in the same year. In a study preceding the train (Upplands Lokaltrafik (UL), 1988) it can be read that two elements brought to the fore the interest for a commuter train: in an analysis on express buses from the northern part of Uppsala County towards Uppsala city it was found that a commuter train on already existing railroad tracks would not be more expensive while it would radically shorten commuting times; and government bill 1987/88:50 opening for counties to take own initiatives regarding train services by renting railroad from the administering State organization (Banverket).

It was first suggested that the train should start running in the autumn of 1990 (Arbetsgruppen Projekt Upptåget, 1988b). However, this proved impossible since another project was also planned on the same link (a high-speed train between Stockholm and Sundsvall) making double tracks and some other track work necessary (see Arbetsgruppen Projekt Upptåget, 1988b and UL annual report 1989/1990). In the summer/autumn 1990, the necessary decisions regarding the division of operation and capital costs for the train between Uppsala County Council, Tierp municipality and Uppsala municipality were reached (Documents Uppsala County Council meeting, 1990 & UL annual report 1989/1990). Around the same time, UL announced that they intended the train to start running in August 1991, but that this depended on the progress of the track work. We have not been able to track down at what time the inauguration date was finally fixed, but in accordance with the aim of UL, Upptåget was inaugurated in August 1991. At this first stage there were 15 trips per weekday between Tierp and Uppsala, with stops in the localities of Örbyhus, Skyttorp, Vattholma, and Storvreta (see Figure 1 for the locations of these stations). From January 1994 Upptåget also stopped in the locality of Tobo.[9]

Upptåget did not directly replace a particular bus line but stopped at places previously served by three other bus lines. These bus lines also served places not crossed by Upptåget and were to a large extent maintained after the introduction of Upptåget. After the inauguration of Upptåget, the track work continued between Tierp and Uppsala until 1997, and the number of trips and speed were steadily increased over the 1990s, despite occasional delays due to the track work.

[9] The next step in the development of Upptåget was official discussions, starting in 1994, about a continuation of Upptåget northwards towards the city of Gävle. In 1999, however, it became clear that the track work needed for the northward continuation could be conducted no earlier than 2005/2006. In 2002 official discussions also began about a southwards continuation of Upptåget towards Arlanda airport and the northern parts of Stockholm. In 2004, a decision was finally reached that the northward and southward continuations should be inaugurated in August 2006, which was also achieved. Given that these further developments of Upptåget were at most at the discussion stage during the time period we are studying (1985-1996), we do not think that they affected the studied outcomes.

An examination of timetables for public transit suggests a substantial reduction in travel times. In 1990, i.e., before Upptåget started running, the fastest bus transit between Tierp town and Uppsala city took 65 minutes, while Upptåget took 47 minutes in 1992, and 40 minutes in 1996. For Skyttorp (located approximately in the middle of the route), travel times were more than cut in half; the corresponding figures were 42, 20, and 18 minutes respectively.[10] Here it can also be noted that the fare for a ride on Upptåget between any of the stations and Uppsala city was the same as for a bus ride, and no more expensive than corresponding bus rides to Uppsala city from other parts of the Uppsala local labor market.

From the investigations preceding the inauguration of Upptåget (Arbetsgruppen Projekt Upptåget, 1988a, 1988b) it appears that it was thought that Upptåget would help handle regional imbalances, counter the county's dependence on Stockholm, and create an integrated county with its own identity. The regional imbalances mentioned are the strong growth of job opportunities and population in Uppsala municipality, with ensuing pressure on housing provision and municipal services, and the stagnating number of job opportunities and population decrease in the municipalities of Tierp and Älvkarleby (the municipality just north of Tierp municipality), with ensuing under-use of existing municipal infrastructure. It was hoped that Upptåget would give the inhabitants in the northern parts of the county access to a larger labor market and that the population pressure on Uppsala could be distributed over the other stations.[11] However, Upptåget was also seen as one of several means to alleviate the problems of traffic congestion in the inner city of Uppsala. It was estimated that Upptåget could replace 10 percent of the buses entering the inner city, which would also help postpone an expansion of the bus terminal. It was regarded as more reasonable to make better use of the land already reserved for public transit, i.e., the railroad tracks crossing Uppsala city, than to convert new land within Uppsala city to accommodate public transit.

In conclusion, the introduction of Upptåget seems to be well suited to use in a quasi-experimental approach to study the effect of improved commuting opportunities and thereby job access on individuals' labor market outcomes. The introduction was not primarily motivated by the labor market outcomes in the areas served, although there were hopes that the train would help workers in stagnating areas. Instead, the stretch covered by Upptåget was governed by existing railroad tracks, and the timing was related to a legal

[10] Both before and after the introduction of Upptåget the fastest public transit between Tierp town and Uppsala city was the long distance train that took about 40 minutes. However, the number of connections with the long distance train were few.

[11] Other benefits hoped for from Upptåget mentioned in the investigations were decreased commuting by car, larger choice of housing especially for two-earner households, and a higher share of people with higher education in the northern part of the county thanks to better access to the higher education institutions in the city of Uppsala.

change. Furthermore, population pressure and congestion problems in Uppsala city seem to have been as important for the introduction of Upptåget as the labor market outcomes in the areas served. Some other elements add to the suitability of using Upptåget in a quasi-experimental setting. The time period between the idea and the realization of Upptåget was relatively short (3.5 years) and characterized by uncertainty about the launch date, making anticipatory behavior less likely. Also, changes in commuting times between the stations and Uppsala city were large, and there are potential control groups who lived in areas in the same local labor market but that were not subject to changes in transport infrastructure (see Section 3.2).

3.2 Treatment and control groups

In this study we consider the treatment group to be individuals who at the end of 1989 lived in a SAMS (Small Area for Market Statistics) with a population center within 4,500 meters of one of the stations served by Upptåget, and more than 10,000 meters from the central parts of Uppsala city.[12] The SAMS classification was created by Statistics Sweden to satisfy demand for small area statistics from users other than municipalities. The objective was to create fairly homogeneous residential areas of about 1000 inhabitants each, implying that the classification divides Sweden into about 9,000 units.[13] The SAMS have been used frequently in Swedish studies as the formal division closest to neighbourhoods. The choice of 4,500 meters is somewhat arbitrary but captures the areas where Upptåget came to be the main public transport mode. The (end of) year 1989 is before major decisions about the commuter train were taken (which was in the summer of 1990) and before one could tell for sure if/when the train would come into being. We therefore think it is unlikely that the people who in 1989 lived close to an Upptåget station had chosen to do so because of the train. At the same time, 1989 is close enough to the decisions for most people not to have moved before the decisions were reached, meaning that most of the people who lived close to an Upptåget station in 1989 actually received the offer of improved commuting opportunities.

As potential control group we have chosen individuals from two other parts of the local labor market of Uppsala. These areas were not subject to

[12] However, we have excluded the station locality of Tobo, despite the fact that it lies 4400 m from one of the other stations, since the Upptåget commuter train did not stop in Tobo until 1994.

[13] In larger municipalities, the SAMS classification is based on municipal subdivisions used for intra-municipal and sometimes regional planning and administration and in smaller municipalities it is based on election districts. The SAMS classification came into use in 1994 and has remained unchanged since then apart from minor adjustments, for example to adapt the SAMS borders to municipal borders. Information from before 1994 can be located to a SAMS by use of the more precise coordinates that real estates have in Sweden. For more information, see Statistics Sweden (2005).

changes in transport infrastructure, but also exhibit frequent commutes to Uppsala city. The first is the corridor between the principal locality in Heby municipality, Heby town, and Uppsala, where a commuter train on existing railroad tracks was discussed at the same time as Upptåget but not put into practice.[14] The second is the corridor between the principal locality in Östhammar municipality, Östhammar town, and Uppsala.[15] We thus have a "treatment corridor" north of Uppsala, and two "control corridors" to the west and the northeast respectively. Analogous to the proximity to stations for the treatment group, we define the potential control group to be individuals who at the end of 1989 lived in a SAMS with a population centers within 4,500 meters of the main road between Uppsala and Heby on one hand and between Uppsala and Östhammar on the other, and more than 10,000 meters from the central parts of Uppsala city.[16] The "treatment" and "control" corridors are shown by the shaded areas in *Figure 1*.

The corridor between Uppsala and Tierp, where Upptåget was introduced, and the corridor between Uppsala and Heby and Uppsala and Östhammar respectively are countryside with some smaller localities. In 1990, Tierp had around 5,000 inhabitants, Heby around 2,500 inhabitants, and Östhammar around 6,000 inhabitants, while Uppsala city had around 110,000 inhabitants (Statistics Sweden, 1992). During the period we use in our analysis, 1985–1996, regional buses were the only type of public transit in the corridor between Uppsala and Heby and between Uppsala and Östhammar. In 1997 some long-distance/regional trains between Linköping and Uppsala, the "UVEN trains", began stopping in Heby and another locality in the corridor between Heby and Uppsala (Morgongåva). The number of connections were few, and the frequency and maintenance not in the hands of Heby municipality or Uppsala County Council. Nevertheless, to avoid the risk that the UVEN trains could influence our estimates, we have chosen 1996 as the last year in which we investigate the effects of the introduction of Upptåget. We think that a period of 5.5 years from the offer of improved commuting opportunities, whereof Upptåget was up and running 4.5 years, should be enough to detect effects from the train on the labor market outcomes of the treatment individuals.

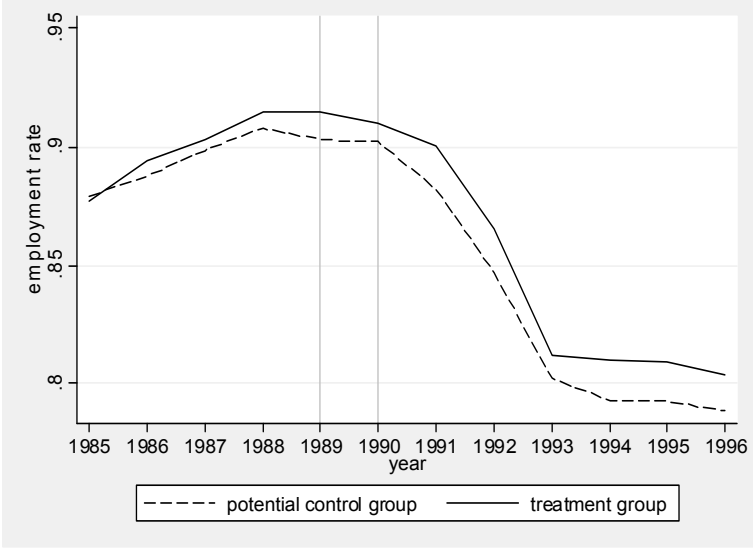
[14] The railroad tracks between Heby town and Uppsala city were used for long distance trains that did not stop in Heby town or anywhere between Heby town and Uppsala city. In the investigation preceding Upptåget it was found that half-hour traffic Heby-Uppsala would be very uneconomic and require four trains, while half-hour traffic Tierp-Uppsala would only require three trains.

[15] Henceforth, Tierp, Östhammar, and Heby will be shorthand for the towns with the same name.

[16] The largest part, 95 percent (88 percent), of the individuals in the potential control group in the end of 1989 lived in a SAMS with a population center no further than 4500 meters (2000 meters) from one of the localities along the main roads.

Here it can also be mentioned that to our knowledge the road network in and around the Upptåget corridor, the Heby corridor, and the Östhammar corridor remained largely unchanged from 1985 until well after 2000.[17]

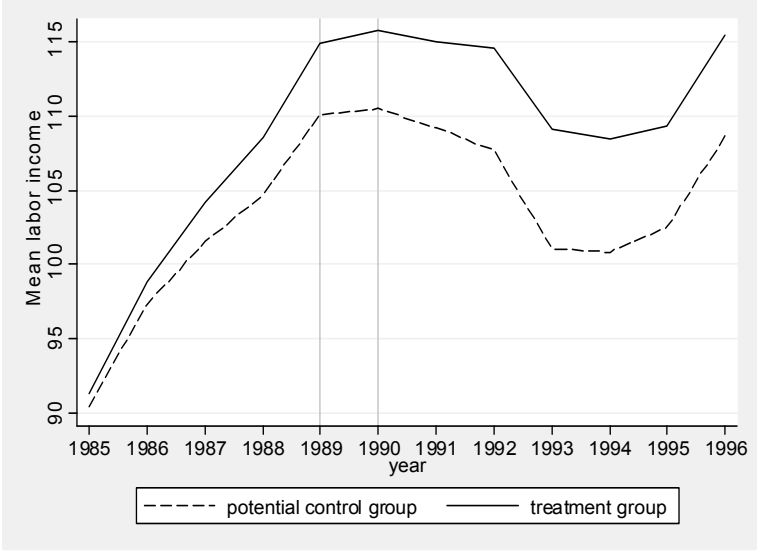
Figure 2. Employment rate for the treatment group and the potential control group



Notes: The measures refer to individuals aged 22-57 in 1989 and for whom we have data for at least the years 1987-1989. The vertical line for 1989 shows when the individuals were selected. The vertical line for 1990 shows the last year before Upptåget started running.

[17] From the late 1990s, plans for a new stretch of the European E4 route northwards from Uppsala were outlined, but construction did not start until 2002 and it was only in 2007 that the route was entirely ready for use.

Figure 3. Mean annual labor earnings for the treatment group and the potential control group (1000 SEK in 1989 prices)



Notes: The measures refer to individuals aged 22-57 in 1989 and for whom we have data for at least the years 1987-1989. The vertical line for 1989 shows when the individuals were selected. The vertical line for 1990 shows the last year before Upptåget started running.

Figure 2 and *Figure 3* show the raw employment rate and the mean annual earnings from labor for the treatment group and the potential control group for the period 1985–1996 (see Appendix A for variable definitions). From the figures it can be seen that both in the treatment group and the potential control group, the employment rate and the mean labor income decreased markedly at the beginning of the 1990s—Sweden went from a booming economy in the late 1980s to a deep recession in the early 1990s. Furthermore, it can be seen that although the employment rate and the mean labor income were similar in the treatment group and the potential control group in 1985, the trends in the two groups before the introduction of the commuter train Upptåget were somewhat different. To handle the different trends, we combine a difference-in-differences approach with a matching strategy in order to obtain treatment and control groups that are balanced in terms of labor market history as well as in terms of age, sex, education, birth region, and industry. The empirical strategy is further described in the next section.

4 Empirical strategy and data

4.1 Intention to treat

There are several well-known methodological problems associated with empirical investigations of the effect of commuting opportunities and job access on labor market outcomes. First, endogeneity can be a problem. Economic theory suggests that individuals simultaneously choose their job, residential location, and commuting behavior. Self-selection and unobserved heterogeneity, e.g., unobserved productivity such as motivation or perseverance may be the reason that individuals living in places with better commuting opportunities/job access have better labor market outcomes. As noted by Åslund, Östh, and Zenou (2010), residential sorting can also lead researchers to understate the impact of job access, e.g., if residential amenities are better in locations with worse job access, or if the low-skilled are forced to live close to jobs due to transportation restrictions. Similarly, workers with jobs or higher earnings may choose residential locations with poor job access in order to consume larger amounts of housing at a lower price, as hypothesized by the standard urban economic model.

Furthermore, as noted by Ihlanfeldt (2006), there may also be reversed causality running from labor market outcomes to job access, so that the better labor market outcomes of workers in some areas attract firms to these locations, implying better job access there. With regards to transport links and commuting opportunities there could be reversed causality running from labor market outcomes to the introduction of transport links or other changes in commuting opportunities. Policy makers could improve commuting opportunities from areas where workers have bad labor market outcomes to help these areas. Alternatively, policy makers could improve commuting opportunities from areas where workers have good labor market outcomes to further strengthen these areas or, by improving public transit, to decrease commuting by car. Also, profit-maximizing transport companies could choose to provide more services to areas with high employment rates and high incomes, implying more trips and the possibility to set higher fares.

To obtain a better estimate of the effect of public transit infrastructure on individual labor market outcomes, as explained in Section 3.2, we compare the outcomes of individuals who in 1989 lived in proximity to the pre-existing railroad upon which Upptåget began running with the outcomes of individuals in the same local labor market who in 1989 lived in places that were not subject to changes in transport infrastructure. Given the institutional setting described in Section 3.1, we think it is a reasonable assumption that these individuals did not choose their 1989 location based on the commuter train. Conditional on a set of control variables (see Section 4.2), the introduction of the commuter train thus provides variation in (offered) commuting opportunities exogenous to individuals' labor market outcomes. Fur-

thermore, the labor market outcomes of the 1989 individuals do not seem to have been very important for the location and timing of the train. Studying the 1989 individuals thus alleviates both omitted variable bias and the problem of reversed causality.

In our analysis, the individuals in the treatment and potential control groups are traced forward to 1996 regardless of where they lived in other periods. We thus perform a reduced-form analysis, allowing the introduction of the commuter train to influence labor market outcomes through any channel. In other words, we estimate the intention-to-treat (ITT) effect, which here is the effect of being offered improved commuting opportunities by public transit, regardless of whether an individual actually came to live in proximity to an Upptåget station once the train was up and running. From a policy perspective, this parameter should be of direct interest.

4.2 Methodology

As explained above, the individuals in the treatment and potential control groups did not choose whether or not they would be offered improved commuting opportunities, which should alleviate omitted variable bias. Nevertheless, it turns out that the two groups are rather different with respect to observed characteristics in 1989 (see Table 1, the rows "unmatched"), and potentially still with respect to unobserved characteristics. Given that, for example, people with different ages and education levels can be expected to have different developments of employment and labor income, this could help explain the diverging trends between the treatment and potential control groups before treatment assignment (as shown in *Figure 2* and *Figure 3*). For the exogeneity assumption to hold, it thus seems important to control for these differences when estimating the ITT-effect (the effect of being offered improved commuting opportunities by public transit). To estimate the ITT we therefore use a difference-in-differences matching estimator (DIDM) (Heckman, Ichimura and Todd, 1997).[18] This type of estimator is analogous to the standard difference-in-differences (DID) regression estimator, but does not impose functional form restrictions in estimating the conditional expectations of the outcome variable, and reweights the observations according to the weighting function implied by the matching estimator.[19] The matching thus ensures that the treatment and control group are balanced in

[18] The description of difference-in-difference matching estimators in this and following paragraphs relies heavily on Smith and Todd (2005).

[19] Matching techniques are traditionally used to overcome selection bias in non-experimental settings. Here, however, we mainly use matching to balance the treatment and potential control groups in terms of observed characteristics.

terms of observed characteristics, while the DID approach controls for unobserved but temporally invariant characteristics remaining after matching.[20]

To be precise, we use a DID propensity score matching estimator. The estimator requires that:

$$E(Y_{0t} - Y_{0t'} | P, Z = 1) = E\left(Y_{0t} - Y_{0t'} | P, Z = 0\right) \quad (1)$$

where Y_0 is outcome conditional on non-assignment to treatment, t and t' are time periods before and after the treatment assignment respectively, P is the propensity score i.e., the probability of treatment assignment, Z is treatment assignment status, with $Z=1$ for the treatment group, and $Z=0$ for the control group. The estimator also requires that a match can be found for each individual in the treatment group:

$$\Pr\left(Z = 1 | X\right) < 1 \quad (2)$$

where X is a set of observable conditioning variables. Equation 2 must hold in both period t and period t'.

The difference-in-differences propensity score matching estimator we use is constructed in the following way: first, the propensity score is estimated using a logistic model. Second, nearest neighbor matching, with replacement and ties, on the propensity score is used to match each treatment group individual to an individual in the potential control group. Third, the difference in outcome between the treatment and matched control groups after treatment assignment is compared to the mean difference in outcome between the treatment and matched control groups for 1985-1989. The estimator can be written as:

$$\hat{\alpha}_{DIDM} = \frac{1}{n_1}\left\{ \sum_{i \in I_1 \cap S_P} \left(Y_{1ti} - \overline{Y_{0t'i}}\right) - \sum_{j \in I_0 \cap S_P} W(i,j)\left(Y_{0tj} - \overline{Y_{0t'j}}\right) \right\} \quad (3)$$

where i and j denote individuals, I_1 denotes the set of individuals in the treatment group, I_0 the set of individuals in the potential control group, S_P the region of common support, n_1 the number of persons in the set $I_1 \cap S_p$, $W(i,j)$ are weights given by the nearest neighbor propensity score matching, and $Y_{0t'i}$-bar and $Y_{0t'j}$-bar are the average outcome 1985-1989 for individual i and j respectively.[21,22]

[20] For comparison, in the results section 5.1 we also show some results using only a matching estimator and only a difference-in-differences estimator.

[21] We think that taking the average over 1985-1989 gives a better estimate of unobserved, temporally invariant characteristics than using a single year.

How to specify the propensity score is not obvious, and the specification of the propensity score could be important. For example, Smith and Todd (2005) and Heckman et al. (1997) find that, in their data, which variables are included in the estimation of the propensity score can make a substantial difference to the performance of an estimator. To choose the specification of the propensity score we use a version of an algorithm for stepwise regression proposed in Imbens and Rubin (2014). Briefly described, we start by estimating a logistic model with only a constant and then iteratively try adding variables to the model, first linear variables and then interaction variables, from a set of *X*. Variables are selected for inclusion in the propensity score depending on the likelihood ratio test statistic. The set of *X* consists of pre-treatment assignment variables that we think may affect the development of labor market outcomes even in the absence of the introduction of the commuter train: sex, birth region, dummy variables for age group in 1989 and education level in 1990.[23] Given that the deep recession that Sweden experienced in the early 1990s, mirrored in *Figure 2* and *Figure 3*, affected some industries more than others, the set of *X* also contains dummy variables for the broadly defined industry in which an individual worked in 1989. Finally, the set of *X* also includes pre-treatment assignment, i.e., 1985-1989, labor income and employment status, to handle differences in trends not accounted for by the other variables in *X*. See Table 1 for a list of the *X* variables.[24]

[22] In the estimations, we use the robust standard errors derived by Abadie and Imbens (Abadie & Imbens, 2006, 2011, 2012) with two matches. It should be noted that these standard errors do not correct for potential correlations in unobserved shocks across individuals. Taking account of clustering in matching models is not a straightforward task (see Hanson & Sunderam, 2012, for a discussion about this). Hanson & Sunderam (2012) propose a version of the Abadie and Imbens standard errors in the presence of clustering. Like the Abadie and Imbens standard errors, the proposition of Hanson and Sunderam builds on matching within the treatment group, with the difference that the matching is done between different clusters, e.g., SAMS, within the treatment group. So far, the Hanson and Sunderam way of estimating the standard errors has not been much used in applied work. One reason for this might be that there is a trade off in their way of estimating the standard errors; matching across clusters decreases the potential clustering problems but might, at the same time, also decrease the match quality and thereby the overall quality of the standard errors. For the estimations, we have used the Stata 13 command "teffects psmatch".

[23] The reason for using education level in 1990, which is not strictly a pre-treatment assignment variable, is that the 1990 data on education can be thought to be of higher quality than the 1989 data on education. The reason for the difference between the 1989 and 1990 data on education is that the 1989 data only rely on administrative registers while the 1990 data are supplemented with information from the 1990 census. Information on education in 1989 is missing for about 3% (4%) of the individuals in the treatment group (the potential control group), while education in 1990 is missing for less than 1% of the individuals in both the treatment and potential control groups. Given the choice between using incomplete data, not conditioning on education at all, and using education in 1990 which could to some extent be influenced by treatment assignment (but not by treatment), we have chosen the latter option.

[24] The variables labor income in 1985, 1986, 1987, 1988, and 1989 are tried separately as are the variables employment in 1985, 1986, 1987, 1988, and 1989. The dummy variables for 1990 education level, age group in 1989, birth region, and industry in 1989 respectively, are tried as a group, e.g., either dummy variables for all education levels are included or no dum-

We estimate Equation 3 for the whole treatment group as well as for a range of sub-groups, each time re-estimating the propensity scores for the sub-group under study.[25]

4.3 Data

The study uses population-wide longitudinal register data, compiled for research purposes by Statistics Sweden, and held by IFAU. Among other things, the data contain rich and detailed information on demographic characteristics, income, employment, and education. For this study, it is particularly important that we also have access to geographic information on the workplace and residential location of each individual. This information is available at the SAMS level (see discussion on SAMS in Section 3.2). The individuals' locations are measured at the end of each year.

The individuals who lived in the Upptåget corridor in 1989, i.e., the treatment group, and the individuals who lived in the Heby corridor and the Östhammar corridor in 1989, i.e., the potential control group, are traced backwards to 1985 and forward to 1996 regardless of where they lived in years other than 1989. We only consider the individuals who were of employable age (18-64) over the whole period 1985-1996, i.e., who were at least 22 and no older than 57 at the end of 1989. Furthermore, in order to have some pre-treatment information for the matching and difference-in-differences analysis, we limit our sample to individuals for whom we have data for at least the years 1987-1989, i.e., who lived in Sweden during that period.[26] The number of individuals aged 22-57 living in the Upptåget corridor in late 1989 was 7,989, and of these 7,934 lived in Sweden in 1987-1989. The corresponding number for the Heby and Östhammar corridors together are 11,493 and 11,341. Concerning attrition, people should only disappear from the original data set if they die or leave Sweden. We have coded individuals missing in a given year as not being employed and as re-

---

my variables for education level are included. A linear variable/group of linear variables is included if its likelihood ratio test statistic is larger than the likelihood ratio test statistic for the other tried linear variables/groups of linear variables and larger than 1 ("Clin" in Imbens, 2014). Concerning the interaction variables, only interactions between the linear variables/groups of linear variables selected for inclusion in the propensity score are tried. Interaction variables/groups of interaction variables are included if its likelihood ratio test statistic is larger than for the other tried interaction variables/groups of interaction variables and larger than 3 ("Cqua" in Imbens, 2014). The choices of Clin and Cqua are somewhat arbitrary; Clin is the same as in Imbens 2014 while Cqua is set slightly higher than in Imbens 2014 to limit the number of interaction variables included. The labor income variables are tried in levels and not in natural logarithms in the algorithm.

[25] The densities of the estimated propensity score for the groups analyzed are given in Appendix B.

[26] To be able to examine how the very youngest (those aged 19-21 in 1989) are affected by the introduction of Upptåget, we also carry out another selection of individuals in the analysis of sub-populations (see Section 5.2; last panel in Figure 8).

ceiving no labor income that year. Of the individuals in the data for 1989, less than 1 percent were missing from the data for 1985 and just above 2 percent from the data for 1996.

Table 1. Summary statistics

		mean			% reduction	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.50	0.48	4.2		2.85	0.004
	Matched	0.50	0.50	-0.1	97	-0.08	0.937
age group 1989							
20-29 years	Unmatched	0.17	0.20	-7.5		-5.06	0
	Matched	0.17	0.18	-2.3	68.8	-1.5	0.135
30-39 years	Unmatched	0.32	0.30	4.7		3.24	0.001
	Matched	0.32	0.32	0.2	96.5	0.1	0.919
40-49 years	Unmatched	0.36	0.34	5.5		3.73	0
	Matched	0.36	0.36	1.1	79.6	0.7	0.487
50-57 years	Unmatched	0.15	0.17	-5.2		-3.51	0
	Matched	0.15	0.14	0.8	83.8	0.54	0.588
education 1990							
10 years or	Unmatched	0.29	0.34	-10.3		-7	0
less	Matched	0.29	0.28	2.4	76.7	1.55	0.121
secondary	Unmatched	0.47	0.48	-2.6		-1.77	0.077
	Matched	0.47	0.47	-0.7	73.7	-0.43	0.667
tertiary	Unmatched	0.24	0.18	15.1		10.41	0
	Matched	0.24	0.25	-1.6	89.5	-0.94	0.345
missing	Unmatched	0.01	0.01	-0.8		-0.56	0.578
	Matched	0.01	0.01	-1.6	-93.7	-0.97	0.33
birth region							
Sweden	Unmatched	0.94	0.92	10		6.7	0
	Matched	0.94	0.94	1.8	82.1	1.21	0.228
western	Unmatched	0.04	0.07	-11.3		-7.55	0
country	Matched	0.04	0.05	-1.9	82.7	-1.35	0.178
non-western	Unmatched	0.01	0.01	0.8		0.55	0.583
country	Matched	0.01	0.01	0	100	0	1
industry 1989							
manufacturing	Unmatched	0.16	0.18	-5.9		-4.02	0
	Matched	0.16	0.17	-1.3	77.9	-0.84	0.402
construction	Unmatched	0.07	0.07	1.4		0.93	0.35
	Matched	0.07	0.07	0.7	52.3	0.41	0.684
trade	Unmatched	0.10	0.09	2.4		1.65	0.098
	Matched	0.10	0.10	-0.7	69.2	-0.46	0.647
education	Unmatched	0.06	0.05	3.8		2.6	0.009
	Matched	0.06	0.07	-2	48.2	-1.17	0.241
health care	Unmatched	0.12	0.09	8.2		5.68	0
	Matched	0.12	0.12	0.1	99	0.05	0.96
other types of	Unmatched	0.07	0.08	-4.9		-3.29	0.001
care	Matched	0.07	0.07	0	99	0.03	0.974
public admin-	Unmatched	0.06	0.04	6.7		4.64	0
istration	Matched	0.06	0.05	1.8	73.5	1.07	0.287
other	Unmatched	0.21	0.25	-9.2		-6.23	0
	Matched	0.21	0.20	0.8	91.8	0.49	0.623

unknown/did	Unmatched	0.17	0.15	3.9		2.66	0.008
not work	Matched	0.17	0.16	0.7	82.2	0.43	0.669
employed	Unmatched	0.88	0.88	-0.8		-0.56	0.579
1985	Matched	0.88	0.87	2.8	-246.2	1.74	0.083
employed	Unmatched	0.89	0.89	2.1		1.42	0.156
1986	Matched	0.89	0.89	1.7	20.5	1.05	0.296
employed	Unmatched	0.90	0.90	1.5		1.03	0.304
1986	Matched	0.90	0.90	1.5	-1.8	0.97	0.334
employed	Unmatched	0.91	0.91	2.4		1.64	0.101
1988	Matched	0.91	0.91	1.2	48.4	0.79	0.43
employed	Unmatched	0.91	0.90	3.9		2.67	0.008
1989	Matched	0.91	0.91	2.9	26.5	1.83	0.067
ln labor in-	Unmatched	3.62	3.62	0		-0.03	0.976
come 1985	Matched	3.62	3.53	2.9	-6599.2	1.8	0.072
ln labor in-	Unmatched	3.75	3.73	0.7		0.48	0.631
come 1986	Matched	3.75	3.70	1.7	-141.7	1.06	0.291
ln labor in-	Unmatched	3.83	3.78	1.8		1.26	0.209
come 1987	Matched	3.83	3.81	0.9	50.5	0.57	0.566
ln labor in-	Unmatched	3.92	3.84	3.1		2.13	0.033
come 1988	Matched	3.92	3.89	1.3	58.2	0.82	0.41
ln labor in-	Unmatched	3.99	3.84	5.4		3.7	0
come 1989	Matched	3.99	3.96	1.3	76.8	0.82	0.414
Individuals in sample	Unmatched	7933	11341				

Notes: a) Western countries include: Finland, Norway, Denmark, Iceland, Ireland, Great Britain, the Netherlands, Belgium, France, Monaco, Luxembourg, Lichtenstein, Spain, Portugal, Andorra, Germany, Switzerland, Italy, San Marino, Vatican City State, Austria, Greece, Canada, the USA, Australia, Papua New Guinea, New Zealand and other Oceanian countries.
b) To be able to take logarithms, individuals with zero labor income were ascribed 1 SEK.
c) The full sample includes one more individual with a covariate, working in education in 1989 combined with education for 1990 missing, which predicts treatment perfectly. This individual is therefore dropped before the analysis.

Table 1 presents some summary statistics for the treatment group, the potential control group and the matched control group.[27] Table 1 shows for each variable the mean in the treatment group, the potential (unmatched) control group, and the matched control group; the t-test for equality of means in the unmatched and matched samples, and the percentage standardized bias in the unmatched and matched samples along with the percentage reduction in absolute bias between the samples.[28] The standardized bias is a scale and sample size free way of assessing overlap (Imbens 2014).[29]

[27] Corresponding tables for the sub-groups can be found in Appendix C.

[28] The standardized percentage bias, suggested by Rosenbaum & Rubin (1985) is the difference of the sample means in the treatment group and the control group (unmatched or matched) as a percentage of the square root of the average of the sample variances in the treatment and control groups, $((X_1-X_0)/(0.5*(V_1(X)+ V_0(X))^{1/2})*100$, where X_1 (V_1) is the mean (variance) in the treatment group and X_0 (V_0) the analogue for the control group. This measure seems to go under different names. Imbens (2014) calls it normalized differences and Smith & Todd (2005) standardized differences.

[29] As Imbens (2014) explains, the t-statistic for testing the null hypothesis that a difference is zero may be large in absolute value simply because the sample is large and, as a result, small differences between the two samples' means are statistically significant even if they are sub-

From Table 1, it can be seen that the standardized biases indicate that there are some substantial differences in the average covariate values between the treatment and the unmatched control group. This is especially the case for education: the percentage standardized bias for tertiary education (at most 10 years of education) is 15.1 (-10.3). It is also the case for birth region where the percentage standardized bias for being born in Sweden (a western country) is 10 (-11.3). There are also some differences with respect to the industry in which the treatment and unmatched control groups worked in 1989: the percentage standardized bias for health care (public administration) (manufacturing) is 8.2 (6.7) (-5.9). Furthermore, there are some differences with respect to age groups where the absolute values of the percentage standardized bias are between 4.7 and 7.5, and for gender where the percentage standardized bias is 4.2. With respect to employment in 1985-1989, the differences in standardized bias between the treatment and control groups were relatively small in the unmatched sample; the absolute values of the standardized bias were between 0.8 and 3.9. The same holds for ln labor income 1985-1989.

In the matched sample the absolute values of the percentage standardized bias are always under 3 percent and for most covariates smaller than in the unmatched sample. One exception is "education 1990 missing" where the absolute value of the percentage standardized bias has increased from 0.8 to 1.6. The bias can, however, be considered small also in the matched sample. Another exception is employment, where the absolute value of the percentage standardized bias is higher in the matched sample than in the unmatched sample in 1985, although still under 3 percent. More important, however, is the fact that the trends in the treatment and matched control groups are similar: in both the treatment and matched control group the employment rate increased by 0.04 from 1985 to 1989 while the increase in the unmatched control group was only 0.02. For labor income, the absolute values of the percentage standardized bias for labor income are higher in the matched than in the unmatched sample in 1985 and 1986, although once again under 3 percent. Furthermore, the trends in the treatment and matched control groups are similar: mean ln labor income increased by 0.38 in the treatment group and 0.42 in the matched control group while the increase in the unmatched control group was 0.23. In sum, the matching strategy we employ seems to do a good job in creating a sample that is well balanced in terms of age, education, sex, birth region, and 1989 industry, and in ensuring that the devel-

stantively small. Large values for the standardized bias, in contrast, indicate that the average covariate values in the two groups are substantially different. There are no clear indications of what is a "large" difference, but in his applications Imbens (2014) seems to consider a standardized bias of above 10 percent as substantial, whereas Caliendo & Kopeinig (2008) write that in most empirical studies a standardized bias below 3 or 5 percent is seen as sufficiently small.

opments in employment and labor income between the treatment and the matched control groups were similar prior to treatment assignment.

5 Baseline results

In this section we present the "intention to treat" (ITT) estimates, i.e., the effects of the introduction of Upptåget on the employment status and earnings of the individuals who lived close to the railroad tracks just before the commuter train was instigated. The first section presents the mean effects in the whole population, the second section presents the mean effects in different subpopulations, and the final section provides some sensitivity analyses of the baseline results.

5.1 ITT estimations for the whole population

Figure 4 presents the ITT estimates on the employment probability and on the natural logarithm of labor income for the whole studied population living in the treated area in the year before it was revealed that the commuter train was going to be instigated (1989—the year indicated by the first vertical line in the figure; the second vertical line indicates the last year before the commuter train started operating (1990)).[30] The year-specific estimates, linked by the solid line, are obtained from the matched difference-in-differences estimator given in equation (3). That is, each year specific-estimate shows the difference in outcome between the treatment group and the matched control group in that year, over and above the mean difference in outcome 1985-1989. The dashed lines show the 95 percent confidence interval.

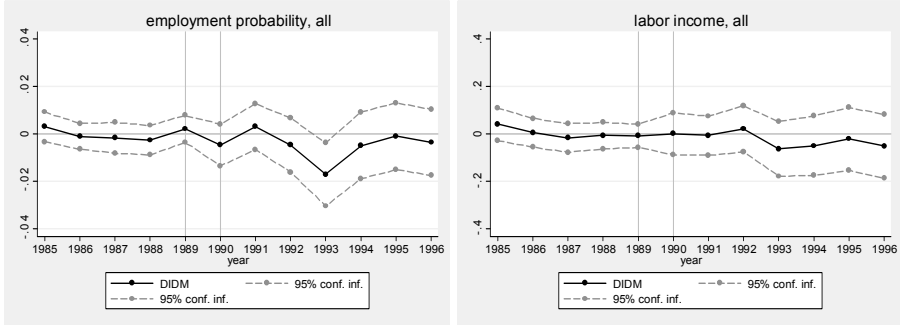
From the figure, it can first be noted from the pre-trends that the matching procedure does a good job in balancing the observations in the treatment and control groups**.** The point estimates are insignificant and close to zero in the whole pre-period, indicating that the development of employment probability and annual labor earnings among the individuals in the treated area are very similar to their matched "twins" in the control areas before the news about the future commuter train was released. This was not unexpected given the summary statistics in Table 1.

Second, it seems like the introduction of the commuter train had, on average, no effects on employment probability and labor income for those individuals that were intended to be treated. For earnings, the point estimates are insignificant at the five percent significance level for all years. For the employment probability, the point estimates are insignificant for all years but one (it is barely significant at the five percent significance level in 1993). The only significant ITT estimate is negative, and indicates a 1.5 percentage

[30] To be able to take logarithms, individuals with zero labor income were ascribed 1 SEK.

point decrease in the employment probability in 1993.[31] We consider the general message from Figure 4 to be that for the overall population, the introduction of the commuter train had very little impact on the employment probabilities and earnings among treated workers.

Figure 4. Effects on employment probability and labor earnings for the whole population.



Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see Equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 7933 treatment individuals and 11341 potential control individuals.

It can be instructive to compare the matched difference-in-differences estimates with the estimates obtained when using only a matching estimator (Figure 5) or only a difference-in-differences estimator (Figure 6). From Figure 5 it can be seen that, as should be the case, the point estimates from the matching estimator follows the same patterns as the point estimates from the matched difference-in-differences estimator (cf. Figure 4). However, in the pre-treatment period the point estimates from the matching estimator lies further from zero than the point-estimates from the matched difference-in-differences estimator, indicating that there could be unobserved differences between treated and untreated observations important to take into account. Furthermore, the 95 % confidence interval is somewhat tighter from the matched difference-in-differences estimator than from the matching estimator.

Likewise, the pre-treatment estimates for the difference-in-differences estimator (Figure 6) show – as expected given the differences observed in Table 1 and Figures 2 and 3 – that the observations in the treatment and control groups are not well balanced before treatment. In particular, there seems to be a positive pre-treatment trend, with significant differences in the last

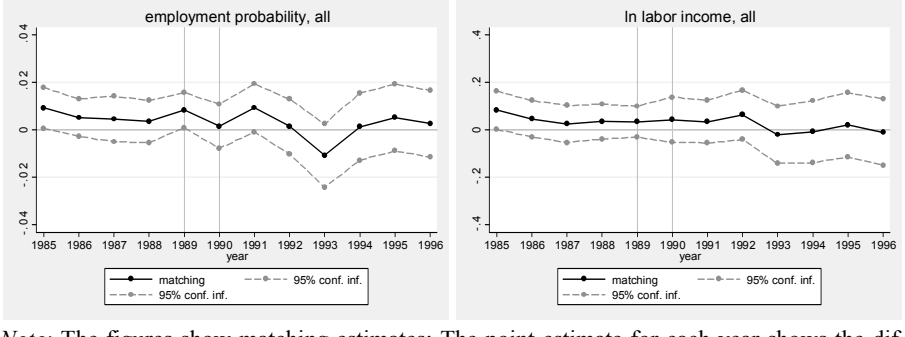
[31] Of course, we cannot rule out that the significant estimate is obtained by chance; since we estimate many point-estimates, some of them will, by chance, turn out to be falsely significant.

years of the pre-treatment period for labor earnings, which cast serious doubt on the positive and significant point estimates observed in the post-treatment period.

Hence, the matched difference-in-differences estimator yields more reliable pre-treatment trends than the matching and the difference-in-differences estimators and should thus provide the most reliable estimates of the effects of the introduction of the commuter train on the affected individuals' labor market outcomes.

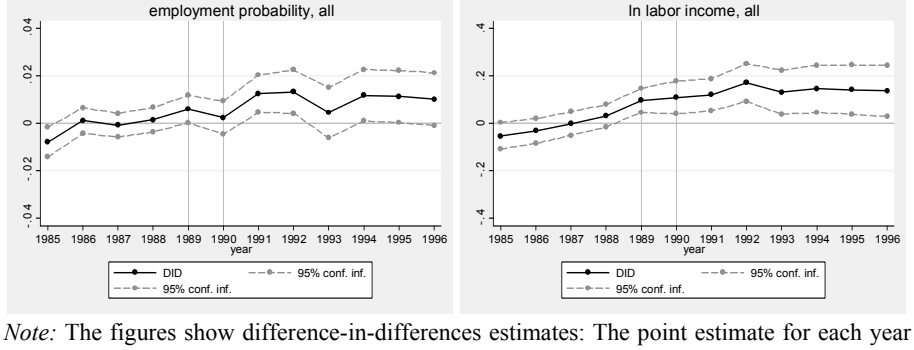
Next we will examine whether some sub-populations are more affected than others.

Figure 5. Effects on employment probability and labor earnings for the whole population when using only a matching estimator.



Note: The figures show matching estimates: The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year. The sample used for the estimations contains 7933 treatment individuals and 11341 potential control individuals.

Figure 6. Effects on employment probability and labor earnings for the whole population when using only a difference-in-differences estimator.



Note: The figures show difference-in-differences estimates: The point estimate for each year shows the difference in mean outcome between the treatment group and the *un*matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 7934 treatment individuals and 11341 control individuals.

5.2 ITT estimates for subpopulations

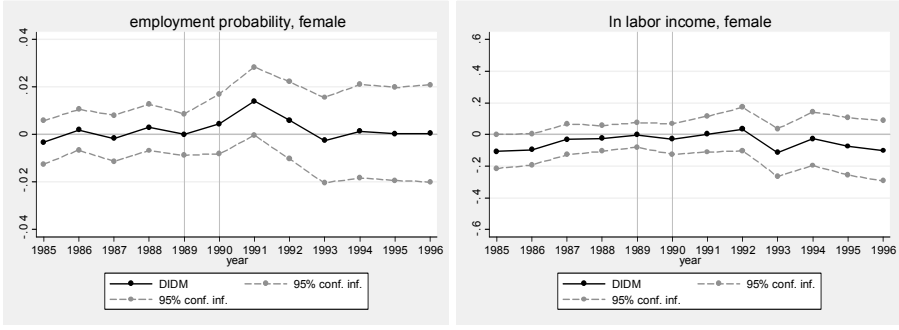
As discussed in the theoretical section we would expect the introduction of a commuter train to have larger effects on labor market outcomes among the groups that rely more on public transit and/or have stronger residential constraints, e.g., ethnic minorities and people with low incomes. In this section we will therefore report results from estimations on separate sub-populations to examine whether there are any heterogeneous effects. The sub-groups we consider are women (Figure 7), different age-groups (Figure 8), individuals who had low labor income before the introduction of the train (Figure 9), and different birth-regions (Figure 10). Generally speaking, two common themes emerge from the estimates in these figures. First, the pre-trends also look reasonable for the sub-groups, providing further support for the assumption that we have a good comparison group when using the matched difference-in-differences estimator. Second, there is very little support for the hypotheses that the commuter train increases employment probability and earnings from labor among these sub-groups.

For individuals age 22–29 in 1989 in the treated areas, there is a tendency to a negative development relative to their counterparts in the control group (cf. the panels in Figure 8).[32] However, only a few of the estimates are statistically significant. When we focus on those very young at the time of treatment, standard errors become large, and there is not as clear a trend in the point estimates. As for the scattered positive estimates found in other groups (women, age 40–49), we do not interpret this as considerable evidence on any impact.

For those born in a non-western country (cf. the last panel in Figure 10), the point estimates indicate a fairly stable, positive and large effect of the commuter train on employment as well as on earnings, but there is too much uncertainty in the estimates to be able to draw any clear conclusions (an uncertainty that probably stems from the fact that the group is fairly small; there are 94 individuals in the treatment group). One reason for highlighting the estimated pattern for the non-western group is that this is perhaps the group where we would *a priori* be most likely to find an impact: employment outcomes are poor, meaning that only a few people entering employment may make a difference, and access to alternative transport (car ownership, co-driving, etc.) may be lower.

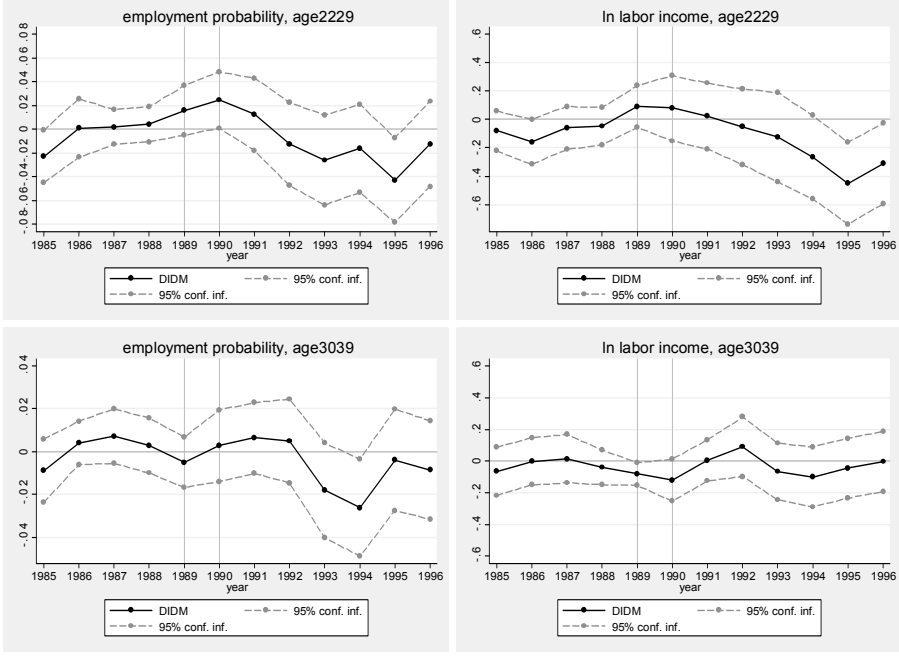
[32] Note that in the last panel in Figure 8, we carry out a selection of individuals other than in the baseline analysis to be able to examine how the very youngest (those aged 19-21 in 1989) are affected by the introduction of Upptåget.

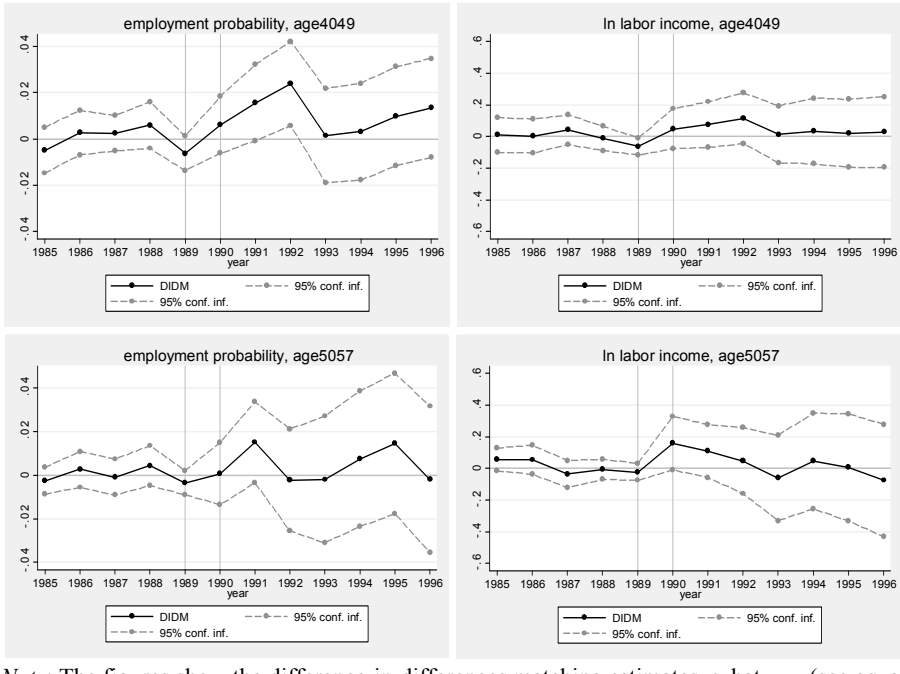
Figure 7. ITT estimates on employment probability and labor earnings: women.



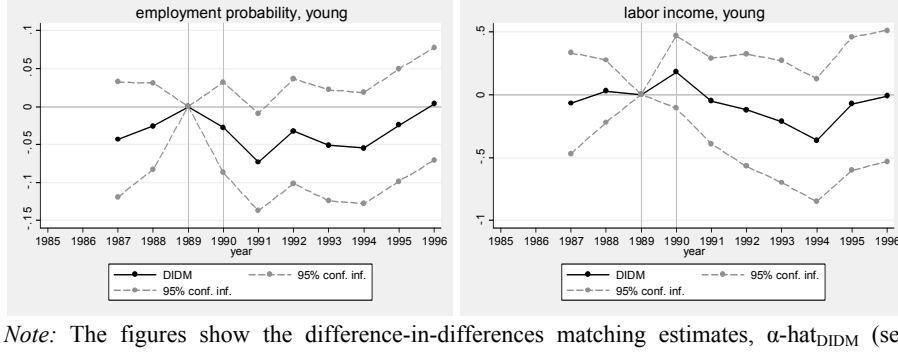
Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 3954 treatment individuals and 5419 potential control individuals.

Figure 8. ITT estimates on employment probability and labor earnings: by age.



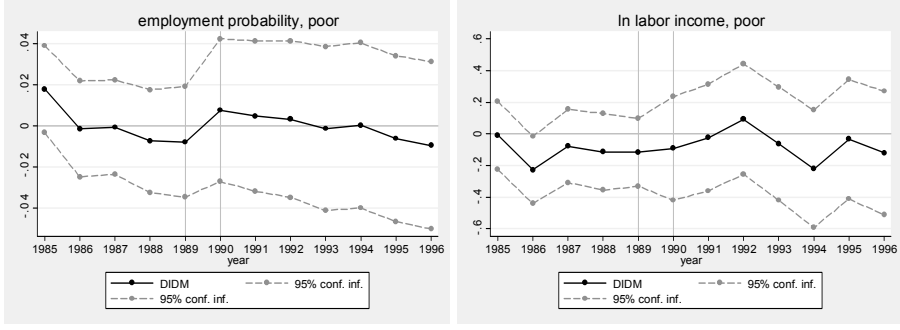


Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations for age 2229 contains 1369 treatment individuals and 2282 potential control individuals. The corresponding figures for age 3039 (4049) [5057] are 2526 (2859) [1857] and 3362 (3800) [1152] respectively.



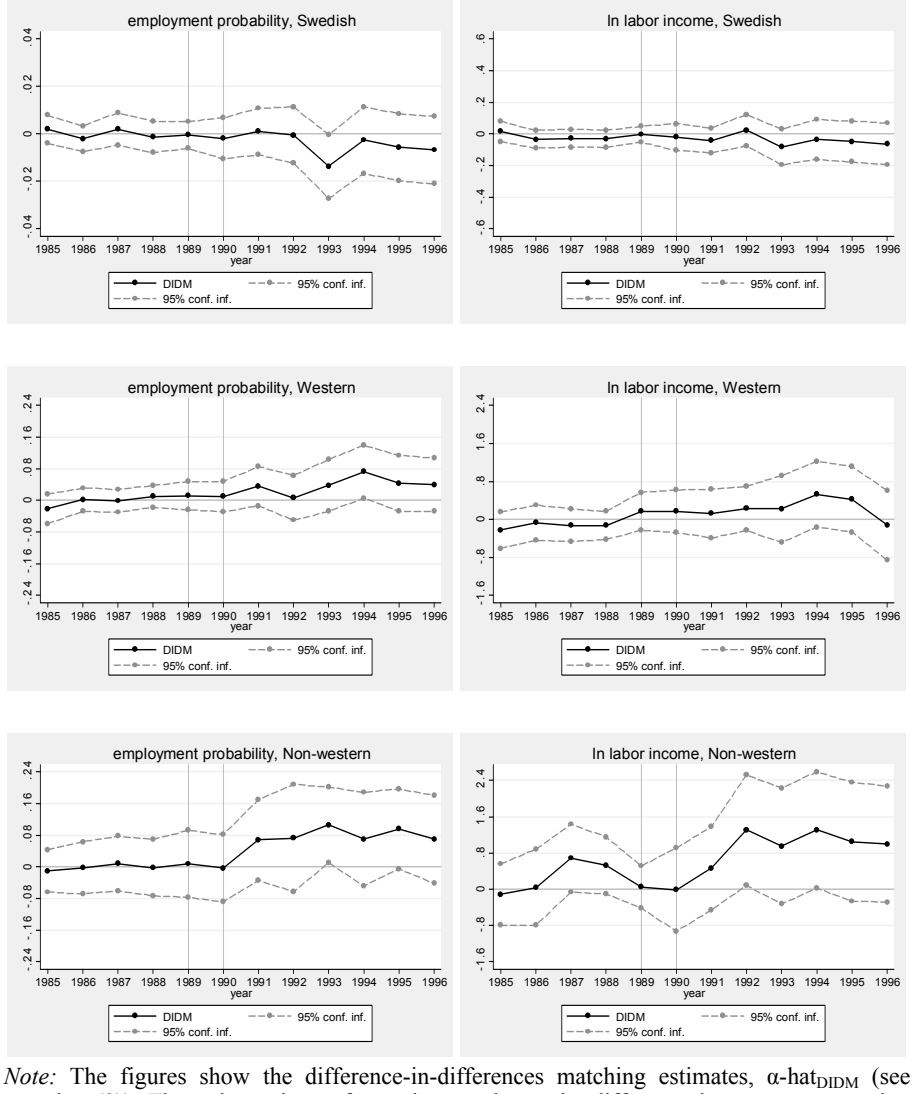
Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1989. The sample used for the estimations contains 647 treatment individuals and 897 potential control individuals. Young is defined as being 19-21 years old in 1989. For young people, the labor income and employment history variables in the set of *X* used in the algorithm to choose the propensity score only include values for 1988 and 1989.

Figure 9. ITT estimates on employment probability and labor earnings: for those with low incomes in 1989.



Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 1484 treatment individuals and 2472 potential control individuals. An individual is defined as having low income if his or her labor income in 1989 was less than 50 percent of the median taxable income in 1989 among those for whom the Swedish Tax Agency had information (82% of the population) (see Statistical Yearbook 1992, Table 221 "Income-earners by total net income and age").

Figure 10. ITT estimates on employment probability and labor earnings: by birth-region



Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations in the Swedish (Western) [Non-Western] subsample contains 7477 (351) [94] treatment individuals and 10401 (782) [125] potential control individuals.

6 Robustness checks and further analysis

In this section, we will conduct two further analyses. First, in Section 6.1, we perform some sensitivity checks by (i) dropping the observations from an

area (Tierp) where commuting time was less affected by Upptåget than in other treated areas and (ii) dropping the observations from an area (Storvreta) that was already more integrated with Uppsala city before the introduction of Upptåget. Then, in Section 6.2, we examine which effects the introduction of the commuter train had on commuting behavior. Finally, in Section 6.3, we check whether there are any indications of heterogeneous labor market shocks (i.e., shocks that hit the treatment area more strongly or differently than the control areas).

6.1 Robustness checks: treatment intensity

All individuals in the treatment area were offered proximity to the commuter train Upptåget and thereby improved commuting opportunities by public transit. While the commuter train was substantially faster than bus connections, the decrease in travel time by public transit was not the same everywhere in the Upptåget corridor. For the train stations between Tierp and Uppsala, travel times by public transit to Uppsala were approximately cut in half, which in 1996 represented an absolute gain of between 17 and 36 minutes depending on station. From Tierp, on the other hand, Upptåget decreased travel time to Uppsala by less than 30 percent compared to bus in 1992 and by less than 40 percent in 1996. Furthermore, the fastest public transit between Tierp and Uppsala was, both before and after the introduction of Upptåget, a long distance train that took about 40 minutes. It can therefore be argued that Upptåget had less of an effect on travel time by public transit between Tierp and Uppsala, even though it indeed increased the number of fast connections.[33]

Likewise, it can be argued that Storvreta, which is the Upptåget station closest to Uppsala, might have been less intensively treated than the other localities since it was already more integrated with Uppsala before the introduction of Upptåget.

In this section we will therefore examine whether the results are sensitive to excluding either Tierp or Storvreta from the analyses.[34]

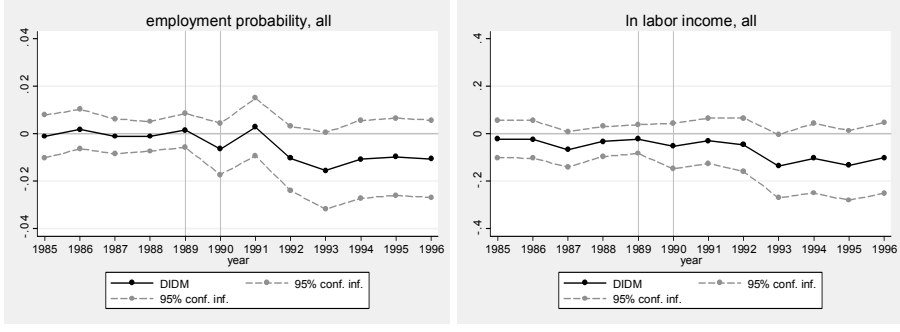
The results when using a sample of all individuals but excluding those who lived in Tierp (Storvreta) and corresponding control areas in 1989 are presented in Figure 11 (Figure 12). When comparing with the baseline results (cf. Figure 4), it is clear that we reach very similar conclusions; the pretrends look reasonable, and there are essentially no significant effects (either

[33] In 1989 long distance trains did 8 trips per weekday from Tierp to Uppsala. In 1992, Upptåget alone accounted for 15 trips, which had increased to 32 in 1996.

[34] We have also estimated models in which we make other exclusions, such as using only the individuals living in Uppsala municipality but not in Storvreta (and hence also excluding Tierp) and using only the individuals living in the municipalities Tierp, Heby, and Östhammar; none of these alterations changes the conclusions. These results are available upon request.

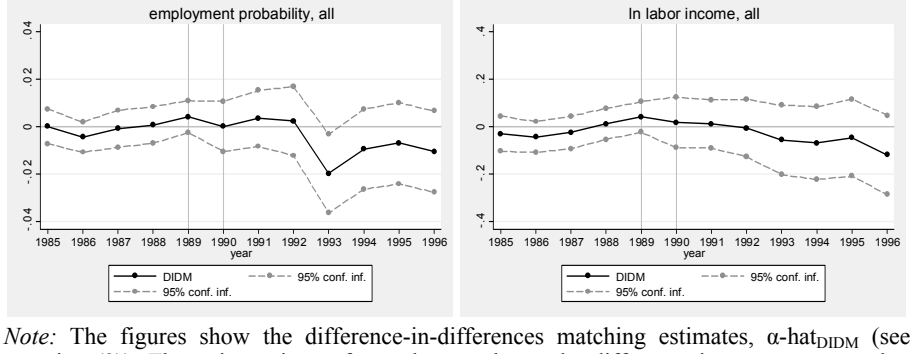
statistically or economically) from the commuter train on the individuals' employment probability or labor earnings.

Figure 11. ITT estimates on employment probability and labor earnings: whole sample but excluding Tierp.



Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 5085 treatment individuals and 8729 potential control individuals.

Figure 12. ITT estimates on employment probability and labor earnings: whole sample but excluding Storvreta.



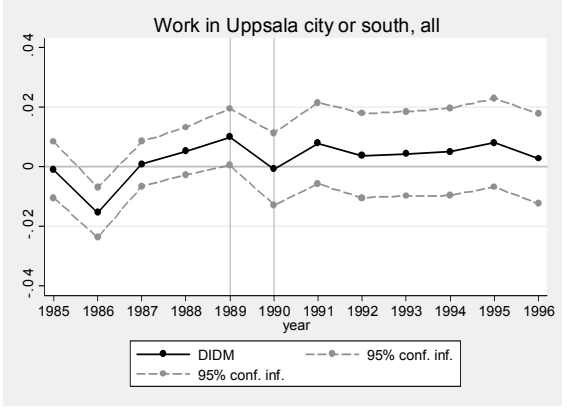
Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 5127 treatment individuals and 10462 potential control individuals.

6.2 Commuting patterns

Even though there appears to be little effect on employment and earnings among those who got access to the commuter train, it is possible that people altered their commuting behavior. To investigate this, we performed an analysis similar to those above, but with the probability of working in Upp-

sala city or further south as outcome variable (it should be recalled that the treatment and control areas are all somewhat north of the city, whereas the commuter-receiving Stockholm region is to the south). We found very little impact on this probability, again suggesting small effects of the reform.

Figure 13. ITT estimates on employment probability and labor earnings: part of sample who worked elsewhere in 1989.



Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the for groups 1985-1989. The sample used for the estimations contains 7933 treatment individuals and 11341 potential control individuals.

6.3 Heterogeneous labor market shocks

The identifying assumption in the analysis presented above is that without the introduction of the commuter train, the development of employment and earnings in the treatment group would have been the same as in the matched control group. As discussed above, several facts support this assumption: the institutional details and our study design suggest that we should not worry about self-selection into locations; the treatment and control areas are all part of a local labor market sharing the same employment center; the pre-reform comparisons indicate that the matched sample contains individuals with comparable development and responses to economic fluctuations (where the matching is also conducted on the pre-reform industry that the individuals worked in).

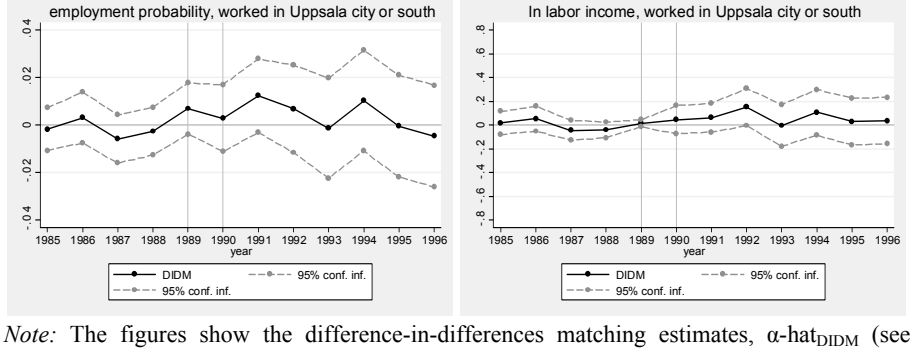
However, it should be noted that if there are geographically heterogeneous economic shocks over treatment and control areas that are not fully captured by the research design, we risk confusing the impact of the train with changes that would have happened anyway. The zero effect could then, for example, be the sum of a negative local labor market shock and a positive effect of the train. This is in principle an untestable assumption; we can nev-

er fully rule out the possibility that the treated areas are affected by different shocks than the control areas, even though the factors mentioned above point in another direction.

However, one way to check whether there are any indications of unaccounted for, heterogeneous, labor market shocks is to examine if the estimated effects for the treated individuals who were employed in the treatment area are different from those for the treated individuals who were employed outside the treatment area. The presumption is that, in the presence of heterogeneous negative labor market shocks in the treatment area, the treated individuals working in the treatment area would be more negatively affected than the treated individuals working outside the treatment area.

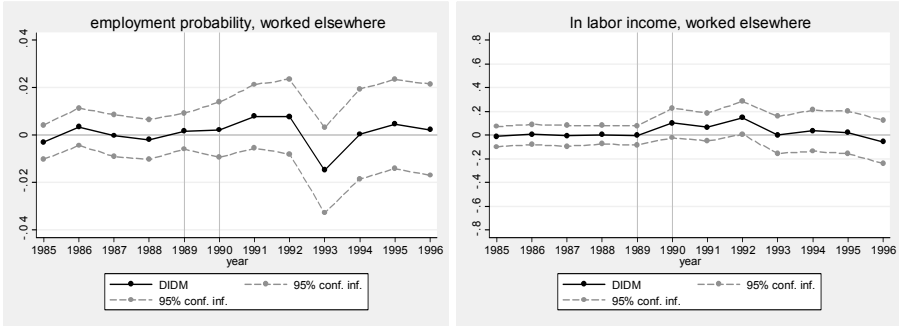
To examine this, we divide the treated individuals into those working in Uppsala or further south, towards the Stockholm region (these are the individuals working outside the treatment area) and those not working in Uppsala or further south (these are the individuals that mainly work close to the home, in the treatment area). As is clear from Figures 14 and Figure 15, the estimated coefficients for these two groups are similar and not significantly different from each other in any of the time periods; for none of the groups can we reject the null hypothesis that Upptåget had no effect on the treated individuals' labor market outcomes. This strengthens the assumption that, given the research design, there were no labor market shocks that affected the treatment area more strongly or differently than the control areas.

Figure 14. ITT estimates on employment probability and labor earnings: part of sample who worked in Uppsala city or south in 1989 (i.e., outside the treatment area).



Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 3441 treatment individuals and 3288 potential control individuals.

Figure 15. ITT estimates on employment probability and labor earnings: part of sample who did *not* work in Uppsala city or south in 1989 (i.e., mainly within the treatment area).



Note: The figures show the difference-in-differences matching estimates, $\alpha\text{-hat}_{DIDM}$ (see equation (3)): The point estimate for each year shows the difference in mean outcome between the treatment group and the matched control group that year, above the difference in mean outcome between the groups for 1985-1989. The sample used for the estimations contains 4485 treatment individuals and 8051 potential control individuals.

7 Concluding discussion

In this paper, we have investigated the importance of improved public transit for individual labor market outcomes. The introduction of a commuter train, Upptåget, between Uppsala and Tierp (54 km north of Uppsala) in Sweden in the early 1990s meant that individuals living in some areas were offered considerably decreased commuting times by public transit and increased job access to the regional labor center in Uppsala, whereas other individuals competing for jobs in the same local labor market did not experience a similar change.

We argue that institutional features suggest that the setting is suitable for evaluating the labor market impact of transport opportunities: timing was affected by a change of national law, the stretch was determined by a pre-existing railroad, and the time between the first discussion and implementation was relatively short. Our empirical analysis uses detailed longitudinal individual data to compare the development for individuals who lived in treated and non-treated areas the year before the information about the new commuter train was released. The intention-to-treat (ITT) estimates are obtained through a matched difference-in-differences estimator.

Our results suggest that the introduction of the commuter train essentially had no significant effects (either statistically or economically) on the employment probability or labor earnings for those individuals who lived in the treated area before the new commuter train was announced. The only potential exception to this result is for the group of individuals who were born in a non-western country. For this group, the patterns of the ITT point estimates

tend to a positive and large effect on both their employment probability and their labor earnings. However, since the non-westerners constitute a fairly small group in the studied area, there is a large uncertainty in the point estimates and the results should be interpreted cautiously. Nevertheless, the results are interesting in that the group is also one where we could expect greater effects according to theory: the economic position is on average poor, and access to alternative transport (car, co-driving etc.) is likely to be lower than for the average worker. This is a group that warrants more analysis in future studies.

We can only speculate on the reasons for the absence of empirical effects, despite rather clear theoretical effects. For the average worker, one could perhaps argue that it is reasonable to find limited effects in a context where public transportation is also available prior to the introduction of the train, and considering that many individuals in the treated areas use private transport to get to work. We do of course not know whether effects would be more present in a context where public transportation was provided to a market with initially no or very limited public transportation. On the other hand, the type of case we study is a very common example facing policy makers.

When analyzing the costs and benefits of major infrastructural investments such as railroads and commuter trains, there are of course aspects other than increased employment and higher earnings to take into account. Less time spent on commutes is arguably also a welfare gain for those whose job and wage prospects are not affected at all. But our analysis provides a piece of the puzzle that is to a large extent missing in previous research and which is essential to any cost-benefit analysis. However, it is also important to acknowledge that we restrict our attention to the individual consequences for workers directly exposed to the reform. From a societal perspective, the effects on in- and out-migration and the regional economic impact are probably at least as relevant. This is a topic of another paper.

References

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica, 74*(1), 235-267.

Abadie, A., & Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics, 29*(1), 1-11.

Abadie, A., & Imbens, G. W. (2012). *Matching on the estimated propensity score*. http://www.hks.harvard.edu/fs/aabadie/pscore.pdf. Retrieved 2015-06-15.

Ahmed, A. M., & Hammarstedt, M. (2008). Discrimination in the rental housing market: A field experiment on the Internet. *Journal of Urban Economics, 64*(2), 362-372.

Alonso, W. (1964). *Location and land use: toward a general theory of land rent*, Harvard University Press. Cambridge, MA.

Arbetsgruppen Projekt Upptåget. (1988a). *Upptåget - En förstudie*.
Arbetsgruppen Projekt Upptåget. (1988b). *Upptåget - förslag till länspendel Uppsala-Tierp*.
Åslund, O., Östh, J., & Zenou, Y. (2010). How important is access to jobs? Old question—improved answer. *Journal of Economic Geography, 10*(3), 389-422.
Becker, G. S. (1965). A theory of the allocation of time. E*conomic Journal, 75*(299), 493-517.
Brueckner, J. K., & Zenou, Y. (2003). Space and unemployment: The Labor market effects of spatial mismatch. *Journal of Labor Economics, 21*(1), 242-262.
Caliendo, M. & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31-72.
Cervero, R., Sandoval, O., & Landis, J. (2002). Transportation as a stimulus of welfare-to-work. *Journal of Planning Education and Research, 22*(1), 50-63.
Coulson, N. E., Laing, D., & Wang, P. (2001). Spatial mismatch in search equilibrium. *Journal of Labor Economics, 19*(4), 949-972.
Gautier, P. A., & Zenou, Y. (2010). Car ownership and the labor market of ethnic minorities. *Journal of Urban Economics, 67*(3), 392-403.
Gibbons, S., & Machin, S. (2006). *Transport and labour market linkages: empirical evidence, implications for policy and scope for further UK research*: Commissioned for the Eddington Study.
Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics, 63*(1), 1-24.
Gobillon, L., Selod, H., & Zenou, Y. (2007). The mechanisms of spatial mismatch. *Urban Studies, 44*(12), 2401-2427.
Gurley, T., & Bruce, D. (2005). The effects of car access on employment outcomes for welfare recipients. *Journal of Urban Economics, 58*(2), 250-272.
Heckman, J., Ichimura, H., & Todd, P.E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training program. *Review of Economic Studies*, *64*(4), 605-654.
Ihlanfeldt, K. R. (2006). A primer on spatial mismatch within urban labor markets. In R. J. Arnott & D. P. McMillen (Eds.), *A companion to urban economics* (pp 404-417). Boston: Blackwell Publishing.
Ihlanfeldt, K. R., & Sjoquist, D. L. (1998). The spatial mismatch hypothesis: a review of recent studies and their implications for welfare reform. *Housing Policy Debate, 9*(4), 849-892.
Imbens, G. (2014). Matching methods in practice: three examples. NBER Working Paper no. 19959.
Imbens, G., & Rubin, D. (2014). Causal Inference, Cambridge University Press.
Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *Quarterly Journal of Economics, 82*(2), 175-197.
Kawabata, M. (2003). Job access and employment among low-skilled autoless workers in US metropolitan areas. *Environment and Planning A, 35*(9), 1651-1668.
LeRoy, S. F., & Sonstelie, J. (1983). Paradise lost and regained: Transportation innovation, income and residential location. *Journal of Urban Economics, 13*(1), 67-89.
Matas, A., Raymond, J.-L., & Roig, J.-L. (2010). Job Accessibility and Female Employment Probability: The Cases of Barcelona and Madrid. *Urban Studies, 47*(4), 769-787.
Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review, 57*(2), 197-210.

Mortensen, D. T., & Pissarides, C. A. (1999). New developments in models of search in the labor market. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (Vol. 3, pp. 2567-2627). Amsterdam: Elsevier Science.
Muth, R. F. (1969). *Cities and housing*: University of Chicago press.
Ong, P. M. (2002). Car ownership and welfare to work. *Journal of Policy Analysis and Management, 21*(2), 239-252.
Ong, P. M., & Houston, D. (2002). Transit, employment and women on welfare. *Urban Geography, 23*(4), 344-364.
Pissarides, C. A. (2000). *Equilibrium unemployment theory*. Cambridge, MA: MIT press.
Raphael, S., & Rice, L. (2002). Car ownership, employment, and earnings. *Journal of Urban Economics, 52*(1), 109-130.
Rosenbaum, P. & Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*(1), 33-38.
Ross, S. L., & Zenou, Y. (2007). *Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory*. CEPR Discussion Paper No. 6128.
Ross, S. L., & Zenou, Y. (2008). Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory. *Regional Science and Urban Economics, 38*(5), 498-517.
Sanchez, T. W. (1999). The connection between public transit and employment. *Journal of the American Planning Association, 65*(3), 284-296.
Sanchez, T. W., Shen, Q., & Peng, Z.-R. (2004). Transit mobility, jobs access and low-income labour participation in US metropolitan areas. *Urban Studies, 41*(7), 1313-1331.
Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review, 74*(3), 433-444.
Shen, Q., & Sanchez, T. W. (2005). Residential location, transportation, and welfare to work in the United States: A case study of Milwaukee. *Housing Policy Debate, 16*(3-4), 393-431.
Smith, J. A., & Todd, E.P. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125, 305-353.
Smith, T. E., & Zenou, Y. (2003). Spatial mismatch, search effort, and urban spatial structure. *Journal of Urban Economics, 54*(1), 129-156.
SOU. (2003:67). *Kollektivtrafik med människan i centrum* Stockholm: Näringsdepartementet.
Statistics Sweden. (1992). *Tätorter 1990 - Befolkning och areal i tätorter och glesbygd - Reviderade uppgifter*. Statistiska meddelanden Na 38 SM 9201 [*Localities 1990 - Population and area in urban and rural areas - Revised figures.* Statistical reports Na 38 SM 9201]. Örebro.
Statistics Sweden. (2005). *Geografin i statistiken - regionala indelningar i Sverige.* Meddelanden i samordningsfrågor för Sveriges officiella statistik 2005:2 [*Geography in statistics - regional divisions in Sweden.* Reports on Statistical Coordination for the Official Statistics of Sweden 2005:2]. Örebro.
Upplands Lokaltrafik (UL). Time-tables 1986-2008.
Upplands Lokaltrafik (UL). (1988). Minutes board meeting February 11, 1988.
Uppsala County Council. (1990). Documents County Council meeting September 24, 1990.
Wasmer, E., & Zenou, Y. (2002). Does city structure affect job search and welfare? *Journal of Urban Economics, 51*(3), 515-541.

Wasmer, E., & Zenou, Y. (2006). Equilibrium search unemployment with explicit spatial frictions. *Labour Economics, 13*(2), 143-165.
White, M. J. (1976). Firm suburbanization and urban subcenters. *Journal of Urban Economics, 3*(4), 323-343.
Zenou, Y. (2002). How do firms redline workers? *Journal of Urban Economics, 52*(3), 391-408.
Zenou, Y. (2009). *Urban Labor Economics*. Cambridge: Cambridge University Press.
Zenou, Y., & Smith, T. E. (1995). Efficiency wages, involuntary unemployment and urban spatial structure. *Regional Science and Urban Economics, 25*(4), 547-573.

Appendix A: Definition of some variables

Definition of outcome variables

Employment: Employment is based on the official annual employment statistics and refers to status during November each year. A person is classified as employed if he or she did paid work for at least one hour per week. If data for an individual is missing in a given year, we consider the individual as not employed in that year.

Labor income, labor earnings, earnings (the words are used interchangeably): Annual earnings from work, including self-employment and employer's income, in 1000 SEK in 1989 prices. If data for an individual is missing in a given year, we consider the individual to have no labor income in that year.

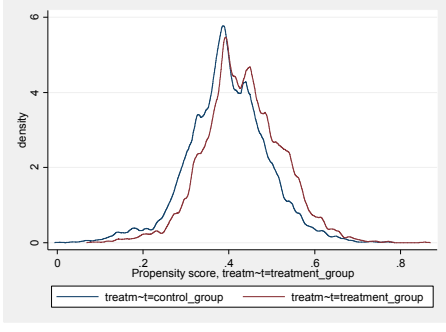
Definition of "western countries":
Finland, Norway, Denmark, Iceland, Ireland, Great Britain, the Netherlands, Belgium, France, Monaco, Luxembourg, Lichtenstein, Spain, Portugal, Andorra, Germany, Switzerland, Italy, San Marino, Vatican City State, Austria, Greece, Canada, the USA, Australia, Papua New Guinea, New Zealand and the other Oceanian countries.

Definition of low income in 1989: An individual is defined as having low income in 1989 if his or her labor income in 1989 was less than 50 percent of the median taxable income in 1989 among those for whom the Swedish Tax Agency had information (82% of the population) (see Statistical Yearbook 1992, Table 221 "Income-earners by total net income and age").

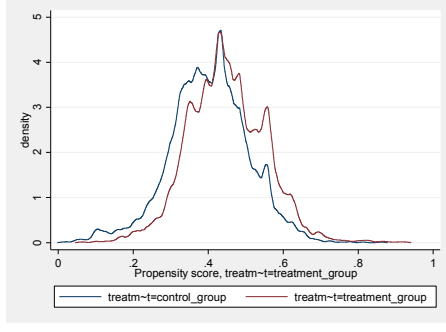
Definition of young: Being 19-21 years old in 1989.

Appendix B: propensity score densities of estimated propensity scores

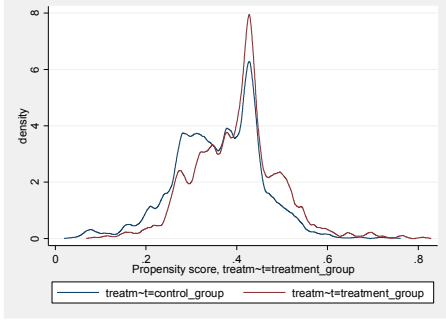
The whole sample



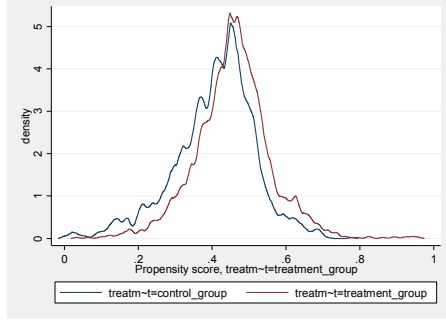
Women



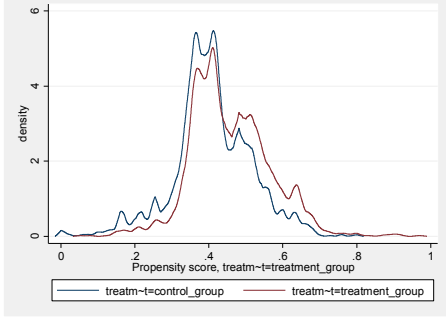
Age 22-29 years in 1989



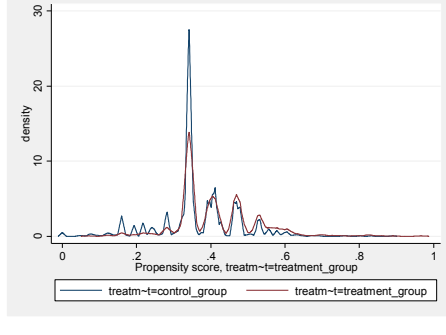
Age 30-39 years in 1989



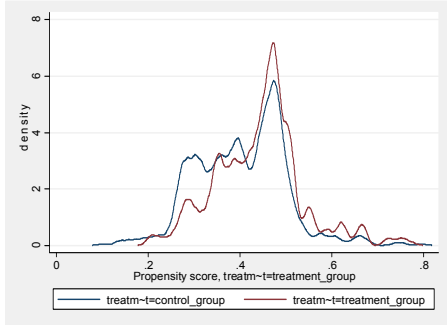
Age 40-49 years in 1989



Age 50-57 years in 1989

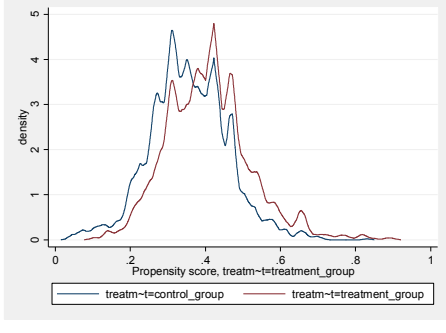


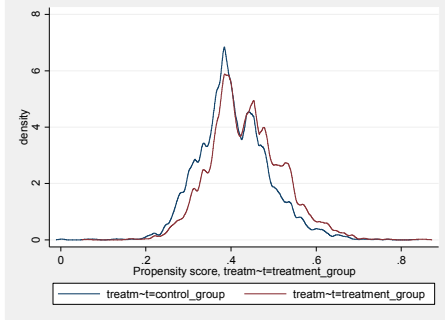
young



Low income in 1989

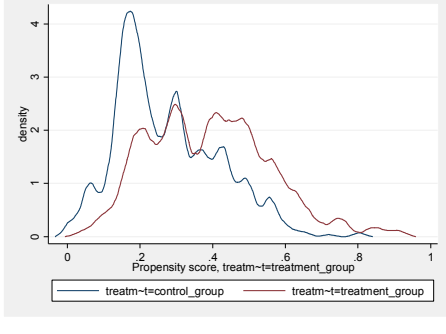
Born in Sweden

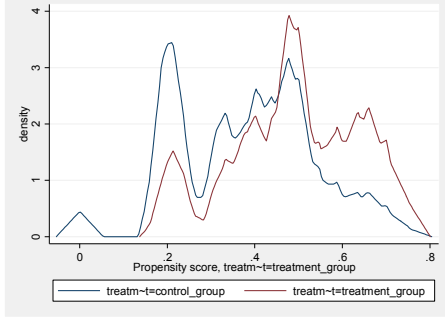




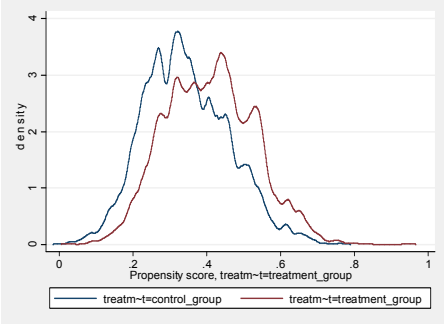
Born in a western country

Born in a non-western country

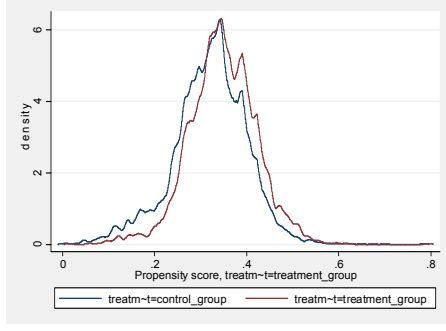




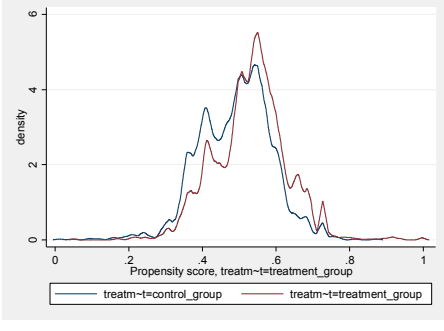
Without Tierp



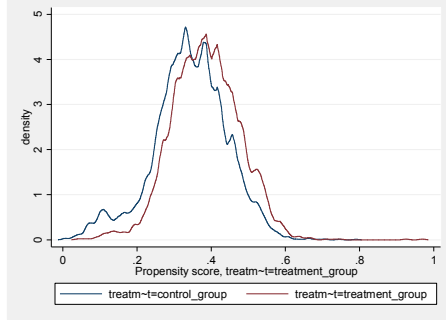
Without Storvreta



Worked in Uppsala city or south



Worked elsewhere



Appendix C: Summary statistics for sub-populations

Women

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
age group 1989							
20-29 years	Unmatched	0.17	0.20	-6.8		-3.26	0.001
	Matched	0.17	0.17	0.5	92.3	0.24	0.81
30-39 years	Unmatched	0.33	0.31	4.2		2.03	0.042
	Matched	0.33	0.32	2.2	47.8	0.98	0.326
40-49 years	Unmatched	0.35	0.33	5.4		2.57	0.01
	Matched	0.35	0.37	-2.2	59.2	-0.96	0.337
50-57 years	Unmatched	0.14	0.16	-5.3		-2.51	0.012
	Matched	0.14	0.14	-0.6	89.3	-0.26	0.797
education 1990							
10 years or	Unmatched	0.27	0.32	-10.8		-5.12	0
less	Matched	0.27	0.26	1.1	89.6	0.51	0.611
secondary	Unmatched	0.47	0.48	-1.4		-0.67	0.502

	Matched	0.47	0.48	-1.5	-4.5	-0.65	0.514
tertiary	Unmatched	0.25	0.20	13.5		6.51	0
	Matched	0.25	0.25	1	92.4	0.44	0.66
missing	Unmatched	0.00	0.01	-0.8		-0.37	0.714
	Matched	0.00	0.01	-2.8	-270.4	-1.18	0.237
birth region							
Sweden	Unmatched	0.94	0.91	8.2		3.9	0
	Matched	0.94	0.94	-0.7	91.8	-0.32	0.746
western	Unmatched	0.05	0.07	-9		-4.24	0
country	Matched	0.05	0.05	3	66.4	1.52	0.13
non-western	Unmatched	0.01	0.01	0.2		0.09	0.931
country	Matched	0.01	0.02	-5.2	-2789.4	-2.08	0.038
industry 1989							
manufacturing	Unmatched	0.08	0.12	-13.3		-6.27	0
	Matched	0.08	0.08	-0.8	93.7	-0.41	0.685
construction	Unmatched	0.01	0.01	1.9		0.91	0.362
	Matched	0.01	0.01	0.3	86.7	0.11	0.914
trade	Unmatched	0.08	0.08	-1.3		-0.63	0.53
	Matched	0.08	0.07	3.5	-166.4	1.62	0.106
education	Unmatched	0.08	0.08	2.3		1.08	0.279
	Matched	0.08	0.09	-1.2	46.3	-0.53	0.599
health care	Unmatched	0.20	0.17	8.7		4.16	0
	Matched	0.20	0.21	-3.5	60	-1.47	0.141
other types of	Unmatched	0.12	0.15	-8.2		-3.87	0
care	Matched	0.12	0.12	0.5	93.6	0.24	0.807
public admin-	Unmatched	0.06	0.05	4.8		2.32	0.021
istration	Matched	0.06	0.06	-0.6	88.1	-0.24	0.81
other	Unmatched	0.16	0.18	-7		-3.32	0.001
	Matched	0.16	0.15	0.5	92.3	0.25	0.804
unknown/did	Unmatched	0.21	0.17	11.7		5.61	0
not work	Matched	0.21	0.20	1.7	85	0.75	0.455
employed	Unmatched	0.84	0.85	-3.1		-1.48	0.139
1985	Matched	0.84	0.85	-2.4	23.6	-1.05	0.296
employed	Unmatched	0.86	0.85	0		0.02	0.986
1986	Matched	0.86	0.86	-0.9	-2455.2	-0.42	0.677
employed	Unmatched	0.87	0.87	0.3		0.15	0.884
1986	Matched	0.87	0.87	-2	-569.7	-0.92	0.359
employed	Unmatched	0.89	0.88	1.6		0.78	0.437
1988	Matched	0.89	0.89	-0.7	56.4	-0.32	0.749
employed	Unmatched	0.89	0.87	3.2		1.55	0.122
1989	Matched	0.89	0.89	-1.6	49.6	-0.75	0.454
ln labor in-	Unmatched	2.98	3.07	-2.9		-1.41	0.159
come 1985	Matched	2.98	3.08	-3.1	-6.7	-1.4	0.163
ln labor in-	Unmatched	3.14	3.19	-1.4		-0.68	0.495
come 1986	Matched	3.14	3.23	-2.8	-94.7	-1.24	0.214
ln labor in-	Unmatched	3.27	3.26	0.4		0.2	0.841
come 1987	Matched	3.27	3.30	-0.8	-89.6	-0.35	0.724
ln labor in-	Unmatched	3.40	3.33	2.3		1.11	0.267
come 1988	Matched	3.40	3.42	-0.6	74.5	-0.27	0.791
ln labor in-	Unmatched	3.47	3.33	4.8		2.3	0.022
come 1989	Matched	3.47	3.47	0.2	96.6	0.08	0.94
Individuals in sample	Unmatched	3954	5419				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Age 22-29 years in 1989

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.49	0.46	5.1		1.5	0.133
	Matched	0.49	0.47	2.8	45.9	0.73	0.467
education 1990							
10 years or	Unmatched	0.19	0.20	-2.7		-0.79	0.428
less	Matched	0.19	0.20	-2.6	5.1	-0.68	0.5
secondary	Unmatched	0.67	0.67	-1.7		-0.5	0.615
	Matched	0.67	0.67	-0.6	63.9	-0.16	0.871
tertiary	Unmatched	0.14	0.12	5.2		1.52	0.129
	Matched	0.14	0.13	3.9	25.1	1	0.316
missing	Unmatched	0.00	0.00	3.4		1.04	0.298
	Matched	0.00	0.00	0	100	0	1
birth region							
Sweden	Unmatched	0.97	0.94	13.1		3.71	0
	Matched	0.97	0.97	1.5	88.9	0.44	0.658
western	Unmatched	0.02	0.05	-15.8		-4.39	0
country	Matched	0.02	0.02	-0.4	97.4	-0.14	0.892
non-western	Unmatched	0.01	0.01	2.2		0.66	0.507
country	Matched	0.01	0.01	-2.4	-7.1	-0.56	0.576
industry 1989							
manufacturing	Unmatched	0.22	0.20	4.6		1.36	0.174
	Matched	0.22	0.22	0	100	0	1
construction	Unmatched	0.09	0.08	4.5		1.32	0.185
	Matched	0.09	0.09	-2.7	40.5	-0.66	0.507
trade	Unmatched	0.11	0.11	2.3		0.68	0.494
	Matched	0.11	0.11	0.9	60.1	0.24	0.809
education	Unmatched	0.02	0.01	2.6		0.78	0.436
	Matched	0.02	0.01	4	-52.1	1.08	0.282
health care	Unmatched	0.11	0.10	3.8		1.13	0.259
	Matched	0.11	0.09	4.4	-13.8	1.15	0.252
other types of	Unmatched	0.06	0.08	-5.8		-1.68	0.092
care	Matched	0.06	0.05	5.1	12.6	1.48	0.14
public admin-	Unmatched	0.03	0.03	1.8		0.52	0.604
istration	Matched	0.03	0.03	0	100	0	1
other	Unmatched	0.19	0.25	-13.6		-3.93	0
	Matched	0.19	0.21	-3.5	73.9	-0.96	0.338
unknown/did	Unmatched	0.17	0.15	4		1.18	0.239
not work	Matched	0.17	0.18	-3.4	16	-0.85	0.393
employed	Unmatched	0.78	0.80	-5		-1.46	0.144
1985	Matched	0.78	0.80	-4.8	3	-1.26	0.208
employed	Unmatched	0.83	0.83	-1.1		-0.33	0.74
1986	Matched	0.83	0.82	1.1	2.9	0.29	0.775
employed	Unmatched	0.87	0.86	4.1		1.19	0.236
1986	Matched	0.87	0.87	1.5	63.1	0.4	0.69
employed	Unmatched	0.88	0.89	-3.6		-1.07	0.285
1988	Matched	0.88	0.87	2.3	36.8	0.58	0.561
employed	Unmatched	0.88	0.89	-0.6		-0.19	0.85
1989	Matched	0.88	0.86	6	-824	1.5	0.134
ln labor in-	Unmatched	3.16	3.33	-6.8		-2.02	0.044
come 1985	Matched	3.16	3.15	0.1	98.9	0.02	0.985
ln labor in-	Unmatched	3.44	3.60	-6.5		-1.93	0.053

come 1986	Matched	3.44	3.52	-3	54.2	-0.77	0.442
ln labor in-	Unmatched	3.59	3.62	-1.2		-0.35	0.725
come 1987	Matched	3.59	3.57	0.9	23.3	0.24	0.811
ln labor in-	Unmatched	3.62	3.76	-5.6		-1.65	0.099
come 1988	Matched	3.62	3.59	1.4	75.4	0.34	0.734
ln labor in-	Unmatched	3.70	3.74	-1.7		-0.5	0.614
come 1989	Matched	3.70	3.52	6.8	-292.7	1.67	0.096
Individuals in sample	Unmatched	1369	2282				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Age 30-39 years in 1989

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.52	0.51	3.3		1.24	0.215
	Matched	0.52	0.51	1.4	56.9	0.5	0.617
education 1990							
10 years or	Unmatched	0.21	0.27	-12.5		-4.7	0
less	Matched	0.21	0.23	-4	67.9	-1.46	0.145
secondary	Unmatched	0.51	0.52	-2.3		-0.86	0.39
	Matched	0.51	0.51	0.1	96.5	0.03	0.978
tertiary	Unmatched	0.28	0.21	15.9		6.08	0
	Matched	0.28	0.26	4.2	73.7	1.43	0.152
missing	Unmatched	0.00	0.01	-4.6		-1.72	0.086
	Matched	0.00	0.01	-1.6	65.8	-0.63	0.531
birth region							
Sweden	Unmatched	0.94	0.92	7.5		2.82	0.005
	Matched	0.94	0.93	3.1	58.6	1.15	0.252
western	Unmatched	0.04	0.06	-8.8		-3.32	0.001
country	Matched	0.04	0.05	-1.6	82.3	-0.6	0.548
non-western	Unmatched	0.02	0.01	0.8		0.3	0.766
country	Matched	0.02	0.02	-3.5	-352.1	-1.16	0.245
industry 1989							
manufacturing	Unmatched	0.14	0.17	-8.1		-3.06	0.002
	Matched	0.14	0.14	0.5	93.3	0.2	0.84
construction	Unmatched	0.07	0.07	-1.8		-0.7	0.486
	Matched	0.07	0.07	0	100	0	1
trade	Unmatched	0.09	0.09	0.8		0.31	0.758
	Matched	0.09	0.09	0.3	65.8	0.1	0.922
education	Unmatched	0.05	0.05	0.6		0.23	0.816
	Matched	0.05	0.04	2.2	-264.9	0.81	0.42
health care	Unmatched	0.13	0.10	8.6		3.3	0.001
	Matched	0.13	0.14	-1.6	81.5	-0.54	0.592
other types of	Unmatched	0.08	0.09	-2.4		-0.92	0.356
care	Matched	0.08	0.08	0.1	94.2	0.05	0.959
public admin-	Unmatched	0.06	0.05	6.5		2.49	0.013
istration	Matched	0.06	0.06	1.5	76.2	0.53	0.6
other	Unmatched	0.21	0.24	-6.4		-2.41	0.016
	Matched	0.21	0.23	-2.8	55.4	-1.02	0.308
unknown/did	Unmatched	0.16	0.14	6		2.27	0.023
not work	Matched	0.16	0.15	1.6	73.9	0.54	0.588
employed	Unmatched	0.88	0.88	0.7		0.25	0.799
1985	Matched	0.88	0.89	-4.1	-511.9	-1.5	0.134

employed	Unmatched	0.90	0.89	2.6		1	0.319
1986	Matched	0.90	0.90	-0.1	95.1	-0.05	0.963
employed	Unmatched	0.90	0.90	1.3		0.51	0.61
1986	Matched	0.90	0.90	0.9	31.5	0.33	0.743
employed	Unmatched	0.92	0.91	2.4		0.93	0.354
1988	Matched	0.92	0.92	-0.6	76.7	-0.21	0.835
employed	Unmatched	0.93	0.92	2.9		1.11	0.269
1989	Matched	0.93	0.93	-3.5	-20.8	-1.32	0.186
ln labor in-	Unmatched	3.50	3.52	-0.9		-0.35	0.724
come 1985	Matched	3.50	3.57	-2.7	-189.6	-0.98	0.325
ln labor in-	Unmatched	3.65	3.63	0.6		0.22	0.826
come 1986	Matched	3.65	3.66	-0.5	8.2	-0.19	0.849
ln labor in-	Unmatched	3.72	3.71	0.2		0.06	0.95
come 1987	Matched	3.72	3.72	0	85.9	0.01	0.993
ln labor in-	Unmatched	3.89	3.82	2.7		1.02	0.306
come 1988	Matched	3.89	3.94	-2	24.8	-0.74	0.457
ln labor in-	Unmatched	4.03	3.96	2.8		1.06	0.288
come 1989	Matched	4.03	4.13	-3.9	-39.5	-1.46	0.144
Individuals in sample	Unmatched	2526	3362				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Age 40-49 years in 1989

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.49	0.47	4.5		1.8	0.071
	Matched	0.49	0.48	2.6	42	0.98	0.328
education 1990							
10 years or	Unmatched	0.31	0.37	-11.3		-4.56	0
less	Matched	0.31	0.30	2.7	75.9	1.06	0.289
secondary	Unmatched	0.40	0.42	-4		-1.6	0.111
	Matched	0.40	0.39	1.5	62.2	0.57	0.57
tertiary	Unmatched	0.29	0.21	17.1		6.97	0
	Matched	0.29	0.30	-4.5	73.9	-1.6	0.111
missing	Unmatched	0.00	0.00	-1.2		-0.46	0.644
	Matched	0.00	0.00	-1.7	-47.2	-0.63	0.531
birth region							
Sweden	Unmatched	0.94	0.91	9.3		3.73	0
	Matched	0.94	0.94	-1.7	81.4	-0.73	0.468
western	Unmatched	0.05	0.08	-10.2		-4.08	0
country	Matched	0.05	0.05	0.3	97.2	0.12	0.905
non-western	Unmatched	0.01	0.01	0.7		0.28	0.783
country	Matched	0.01	0.01	3.9	-474.5	1.61	0.107
industry 1989							
manufacturing	Unmatched	0.14	0.18	-9.5		-3.81	0
	Matched	0.14	0.13	4.3	54.8	1.74	0.082
construction	Unmatched	0.07	0.06	1.9		0.76	0.45
	Matched	0.07	0.06	2.3	-21.4	0.86	0.39
trade	Unmatched	0.09	0.08	4.2		1.7	0.09
	Matched	0.09	0.08	4.3	-2.9	1.63	0.104
education	Unmatched	0.08	0.07	6.9		2.79	0.005
	Matched	0.08	0.08	2	71.1	0.72	0.471
health care	Unmatched	0.12	0.08	10.6		4.33	0

	Matched	0.12	0.13	-5.4	49.3	-1.85	0.064
other types of	Unmatched	0.05	0.08	-8.8		-3.53	0
care	Matched	0.05	0.06	-1.6	82.5	-0.63	0.53
public admin-	Unmatched	0.07	0.05	8.8		3.58	0
istration	Matched	0.07	0.08	-2.7	69.6	-0.91	0.361
other	Unmatched	0.22	0.26	-9.4		-3.78	0
	Matched	0.22	0.23	-2.1	78.1	-0.8	0.426
unknown/did	Unmatched	0.16	0.15	2.7		1.11	0.268
not work	Matched	0.16	0.16	-1	64.5	-0.36	0.718
employed	Unmatched	0.92	0.92	-0.7		-0.29	0.775
1985	Matched	0.92	0.92	-2.2	-207.5	-0.83	0.405
employed	Unmatched	0.92	0.92	2		0.8	0.422
1986	Matched	0.92	0.92	0.6	69	0.24	0.813
employed	Unmatched	0.93	0.93	-0.7		-0.3	0.762
1986	Matched	0.93	0.92	0.5	28.5	0.2	0.841
employed	Unmatched	0.94	0.93	3.4		1.37	0.169
1988	Matched	0.94	0.93	1.9	43.2	0.74	0.457
employed	Unmatched	0.93	0.92	5.3		2.12	0.034
1989	Matched	0.93	0.94	-2.8	46.1	-1.16	0.247
ln labor in-	Unmatched	3.95	3.86	3.4		1.36	0.173
come 1985	Matched	3.95	3.93	0.6	81.7	0.23	0.814
ln labor in-	Unmatched	4.03	3.95	3		1.19	0.233
come 1986	Matched	4.03	4.03	0.3	90.3	0.11	0.914
ln labor in-	Unmatched	4.14	4.04	3.9		1.57	0.116
come 1987	Matched	4.14	4.09	1.9	51.1	0.71	0.476
ln labor in-	Unmatched	4.21	4.07	5.6		2.24	0.025
come 1988	Matched	4.21	4.22	-0.2	95.6	-0.1	0.924
ln labor in-	Unmatched	4.28	4.07	8.3		3.31	0.001
come 1989	Matched	4.28	4.34	-2.2	73.1	-0.91	0.364
Individuals in sample	Unmatched	2859	3800				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Age 50-57 years in 1989

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.48	0.47	3		0.79	0.43
	Matched	0.48	0.47	3	-2.4	0.73	0.467
education 1990							
10 years or	Unmatched	0.51	0.57	-11.9		-3.19	0.001
less	Matched	0.51	0.50	2.8	76.6	0.67	0.505
secondary	Unmatched	0.31	0.30	2.9		0.77	0.444
	Matched	0.31	0.31	-0.9	67.1	-0.22	0.822
tertiary	Unmatched	0.17	0.12	12.8		3.46	0.001
	Matched	0.17	0.17	-1.7	86.5	-0.39	0.699
missing	Unmatched	0.01	0.01	1.3		0.34	0.732
	Matched	0.01	0.01	-3.9	-204.2	-0.82	0.412
birth region							
Sweden	Unmatched	0.93	0.90	12.4		3.24	0.001
	Matched	0.93	0.94	-2.8	77.3	-0.77	0.44
western	Unmatched	0.06	0.09	-14		-3.64	0
country	Matched	0.06	0.05	1	93	0.27	0.786
non-western	Unmatched	0.01	0.01	2.8		0.76	0.445

country	Matched	0.01	0.00	5.7	-104.4	1.51	0.132
industry 1989							
manufacturing	Unmatched	0.18	0.20	-4.5		-1.19	0.233
	Matched	0.18	0.19	-2	55.6	-0.48	0.629
construction	Unmatched	0.05	0.05	2.6		0.69	0.492
	Matched	0.05	0.06	-0.8	69.5	-0.18	0.856
trade	Unmatched	0.09	0.08	3.5		0.95	0.344
	Matched	0.09	0.08	3.6	-3.4	0.88	0.381
education	Unmatched	0.08	0.08	2.1		0.56	0.573
	Matched	0.08	0.08	1	54.1	0.23	0.818
health care	Unmatched	0.09	0.07	7.1		1.93	0.054
	Matched	0.09	0.09	-0.6	91.1	-0.14	0.885
other types of	Unmatched	0.06	0.06	-0.6		-0.17	0.868
care	Matched	0.06	0.06	-0.4	41.9	-0.09	0.93
public admin-	Unmatched	0.05	0.04	3		0.82	0.414
istration	Matched	0.05	0.05	-1.7	44.7	-0.38	0.701
other	Unmatched	0.19	0.23	-10.3		-2.72	0.007
	Matched	0.19	0.19	0	100	0	1
unknown/did	Unmatched	0.20	0.19	3.5		0.94	0.347
not work	Matched	0.20	0.20	0.7	81.2	0.16	0.876
employed	Unmatched	0.89	0.91	-4		-1.07	0.285
1985	Matched	0.89	0.90	-2	49.3	-0.48	0.633
employed	Unmatched	0.90	0.90	0.6		0.16	0.872
1986	Matched	0.90	0.90	-0.3	52.6	-0.07	0.945
employed	Unmatched	0.89	0.90	-1.9		-0.52	0.604
1986	Matched	0.89	0.89	-1.4	27.4	-0.34	0.736
employed	Unmatched	0.90	0.88	4.4		1.16	0.246
1988	Matched	0.90	0.89	0.3	93.7	0.07	0.946
employed	Unmatched	0.88	0.86	5.8		1.53	0.127
1989	Matched	0.88	0.89	-2.1	63.9	-0.52	0.601
ln labor in-	Unmatched	3.64	3.67	-1		-0.26	0.795
come 1985	Matched	3.64	3.59	1.7	-74.2	0.39	0.694
ln labor in-	Unmatched	3.69	3.65	1.4		0.36	0.719
come 1986	Matched	3.69	3.64	1.6	-19	0.38	0.704
ln labor in-	Unmatched	3.66	3.63	0.8		0.21	0.835
come 1987	Matched	3.66	3.70	-1.3	-61.3	-0.3	0.762
ln labor in-	Unmatched	3.64	3.51	3.9		1.04	0.298
come 1988	Matched	3.64	3.65	-0.3	92.4	-0.07	0.942
ln labor in-	Unmatched	3.56	3.30	7.2		1.91	0.056
come 1989	Matched	3.56	3.59	-0.8	89.5	-0.19	0.849
Individuals in sample	Unmatched	1152	1857				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

young

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.47	0.47	0.8		0.15	0.881
	Matched	0.47	0.47	-0.9	-18.1	-0.16	0.87
education 1990							
10 years or	Unmatched	0.12	0.16	-11.7		-2.24	0.025
less	Matched	0.12	0.09	7.2	38.8	1.44	0.15
secondary	Unmatched	0.79	0.74	12.3		2.38	0.018

	Matched	0.79	0.82	-6.6	46.5	-1.27	0.204
tertiary	Unmatched	0.09	0.10	-4		-0.77	0.442
	Matched	0.09	0.08	1.1	73.2	0.2	0.842
missing	Unmatched	-	-				
	Matched	-	-				
birth region							
Sweden	Unmatched	0.99	0.98	8.2		1.55	0.122
	Matched	0.99	0.99	-2.6	68.1	-0.63	0.526
western	Unmatched	0.01	0.02	-8.2		-1.55	0.122
country	Matched	0.01	0.01	2.6	68.1	0.63	0.526
non-western	Unmatched	-	-				
country	Matched	-	-				
industry 1989							
manufacturing	Unmatched	0.22	0.23	-3.1		-0.6	0.552
	Matched	0.22	0.21	3	3.7	0.54	0.587
construction	Unmatched	0.08	0.08	1.4		0.28	0.781
	Matched	0.08	0.07	4	-178.9	0.73	0.465
trade	Unmatched	0.16	0.13	10.4		2.04	0.042
	Matched	0.16	0.17	-2.2	79	-0.37	0.71
education	Unmatched	0.02	0.01	7.1		1.42	0.155
	Matched	0.02	0.00	10.9	-53	2.15	0.032
health care	Unmatched	0.09	0.10	-1.6		-0.31	0.756
	Matched	0.09	0.10	-4.2	-163.9	-0.75	0.454
other types of	Unmatched	0.09	0.08	4.3		0.84	0.401
care	Matched	0.09	0.08	2.8	35.8	0.49	0.623
public admin-	Unmatched	0.01	0.02	-4.8		-0.91	0.363
istration	Matched	0.01	0.03	-9.6	-100.8	-1.59	0.113
other	Unmatched	0.18	0.23	-10.5		-2.02	0.043
	Matched	0.18	0.21	-6.1	41.7	-1.12	0.263
unknown/did	Unmatched	0.14	0.13	1.6		0.3	0.764
not work	Matched	0.14	0.13	4	-161.3	0.74	0.46
employed	Unmatched	0.10	0.14	-10.3		-1.98	0.048
1985	Matched	0.10	0.13	-8.2	20.2	-1.5	0.134
employed	Unmatched	0.29	0.34	-10.7		-2.06	0.04
1986	Matched	0.29	0.31	-5.1	52.1	-0.93	0.354
employed	Unmatched	0.48	0.52	-8.5		-1.65	0.099
1986	Matched	0.48	0.50	-4.4	48	-0.8	0.426
employed	Unmatched	0.73	0.69	7.8		1.5	0.134
1988	Matched	0.73	0.73	-1	86.8	-0.19	0.851
employed	Unmatched	0.79	0.78	1		0.19	0.847
1989	Matched	0.79	0.77	5.2	-421.6	0.92	0.357
ln labor in-	Unmatched	-3.30	-2.39	-20.7		-4.01	0
come 1985	Matched	-3.30	-2.51	-17.9	13.6	-3.24	0.001
ln labor in-	Unmatched	0.18	1.12	-25.8		-5.07	0
come 1986	Matched	0.18	1.12	-25.8	0	-4.71	0
ln labor in-	Unmatched	2.05	2.41	-12.8		-2.49	0.013
come 1987	Matched	2.05	2.06	-0.4	97.2	-0.06	0.951
ln labor in-	Unmatched	3.18	3.31	-6.3		-1.23	0.218
come 1988	Matched	3.18	3.10	4.2	33.9	0.7	0.486
ln labor in-	Unmatched	3.64	3.59	2.2		0.42	0.673
come 1989	Matched	3.64	3.58	2.9	-35.3	0.52	0.6
Individuals in sample	Unmatched	647	897				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Low income in 1989

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.74	0.69	10.3		3.12	0.002
	Matched	0.74	0.72	5.5	46.3	1.53	0.127
age group 1989							
20-29 years	Unmatched	0.27	0.27	-0.2		-0.05	0.963
	Matched	0.27	0.27	-1.7	-984.2	-0.45	0.65
30-39 years	Unmatched	0.33	0.30	5.2		1.59	0.111
	Matched	0.33	0.33	0	100	0	1
40-49 years	Unmatched	0.23	0.25	-5.1		-1.53	0.125
	Matched	0.23	0.23	0	100	0	1
50-57 years	Unmatched	0.17	0.17	-0.5		-0.15	0.882
	Matched	0.17	0.17	2	-301.5	0.54	0.59
education 1990							
10 years or	Unmatched	0.36	0.38	-4.2		-1.27	0.205
less	Matched	0.36	0.35	2.9	29.8	0.81	0.42
secondary	Unmatched	0.48	0.48	-1.1		-0.34	0.731
	Matched	0.48	0.48	-0.5	52.3	-0.15	0.883
tertiary	Unmatched	0.15	0.12	8.5		2.62	0.009
	Matched	0.15	0.15	-2.2	74.2	-0.56	0.573
missing	Unmatched	0.02	0.02	-2.2		-0.67	0.504
	Matched	0.02	0.02	-3	-35.9	-0.81	0.418
birth region							
Sweden	Unmatched	0.91	0.89	5.2		1.56	0.12
	Matched	0.91	0.91	0.2	95.7	0.06	0.95
western	Unmatched	0.07	0.08	-6.4		-1.91	0.056
country	Matched	0.07	0.07	-0.3	96	-0.07	0.942
non-western	Unmatched	0.03	0.02	0.9		0.26	0.794
country	Matched	0.03	0.03	0	100	0	1
industry 1989							
manufacturing	Unmatched	0.09	0.09	-1.1		-0.35	0.728
	Matched	0.09	0.09	-1.2	-2.6	-0.32	0.749
construction	Unmatched	0.02	0.02	-2.5		-0.77	0.444
	Matched	0.02	0.02	-2.3	11	-0.62	0.537
trade	Unmatched	0.07	0.07	0.4		0.11	0.915
	Matched	0.07	0.09	-6.8	-1835.9	-1.75	0.08
education	Unmatched	0.04	0.04	0.1		0.02	0.985
	Matched	0.04	0.04	-0.4	-488.6	-0.1	0.924
health care	Unmatched	0.11	0.09	6.5		1.99	0.047
	Matched	0.11	0.11	-2	68.5	-0.53	0.599
other types of	Unmatched	0.08	0.09	-3.5		-1.06	0.287
care	Matched	0.08	0.06	7.3	-108.4	2.18	0.03
public admin-	Unmatched	0.02	0.02	3.5		1.07	0.286
istration	Matched	0.02	0.03	-2.7	20.9	-0.68	0.497
other	Unmatched	0.15	0.18	-9.4		-2.83	0.005
	Matched	0.15	0.14	0.9	90.3	0.26	0.795
unknown/did	Unmatched	0.42	0.39	5.4		1.65	0.098
not work	Matched	0.42	0.40	2.2	59.5	0.6	0.551
employed	Unmatched	0.64	0.70	-11.5		-3.52	0
1985	Matched	0.64	0.62	4.4	61.3	1.18	0.238
employed	Unmatched	0.65	0.69	-7.1		-2.18	0.029

1986	Matched	0.65	0.65	0.4	95	0.1	0.923
employed	Unmatched	0.64	0.68	-8.9		-2.71	0.007
1986	Matched	0.64	0.64	0.5	94.2	0.14	0.891
employed	Unmatched	0.62	0.65	-4.7		-1.45	0.148
1988	Matched	0.62	0.63	-0.9	81.2	-0.24	0.809
employed	Unmatched	0.58	0.59	-2.3		-0.7	0.482
1989	Matched	0.58	0.59	-1	58.5	-0.26	0.794
ln labor in-	Unmatched	1.18	1.67	-11.3		-3.48	0.001
come 1985	Matched	1.18	1.13	1.1	90.3	0.29	0.771
ln labor in-	Unmatched	1.12	1.55	-9.5		-2.91	0.004
come 1986	Matched	1.12	1.29	-3.8	59.5	-1.04	0.298
ln labor in-	Unmatched	1.05	1.34	-6.4		-1.94	0.052
come 1987	Matched	1.05	1.07	-0.4	93.1	-0.12	0.906
ln labor in-	Unmatched	0.80	0.98	-4		-1.23	0.219
come 1988	Matched	0.80	0.85	-1.3	68.6	-0.34	0.731
ln labor in-	Unmatched	0.30	0.30	0		-0.01	0.991
come 1989	Matched	0.30	0.36	-1.3	-3391.4	-0.36	0.718
Individuals in sample	Unmatched	1484	2472				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Born in Sweden

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.49	0.48	3.7		2.45	0.014
	Matched	0.49	0.49	1.9	48.2	1.18	0.239
age group 1989							
20-29 years	Unmatched	0.18	0.21	-7.5		-4.9	0
	Matched	0.18	0.18	0.7	90.5	0.45	0.653
30-39 years	Unmatched	0.32	0.30	4.3		2.87	0.004
	Matched	0.32	0.32	-1.1	74	-0.68	0.494
40-49 years	Unmatched	0.36	0.33	5.3		3.52	0
	Matched	0.36	0.36	0.8	84.2	0.51	0.609
50-57 years	Unmatched	0.14	0.16	-4.4		-2.91	0.004
	Matched	0.14	0.15	-0.4	89.9	-0.28	0.781
education 1990							
10 years or	Unmatched	0.29	0.33	-9.1		-5.98	0
less	Matched	0.29	0.29	-0.6	93.3	-0.38	0.705
secondary	Unmatched	0.47	0.48	-3.3		-2.2	0.028
	Matched	0.47	0.47	-0.4	88.7	-0.23	0.819
tertiary	Unmatched	0.24	0.18	14.3		9.54	0
	Matched	0.24	0.23	1.3	91	0.75	0.452
missing	Unmatched	0.00	0.00	0.8		0.51	0.61
	Matched	0.00	0.01	-0.8	-1.7	-0.45	0.65
industry 1989							
manufacturing	Unmatched	0.16	0.17	-2.6		-1.72	0.086
	Matched	0.16	0.16	0.8	68.1	0.52	0.606
construction	Unmatched	0.07	0.07	2		1.3	0.195
	Matched	0.07	0.07	0.2	92	0.09	0.925
trade	Unmatched	0.10	0.09	1.4		0.9	0.366
	Matched	0.10	0.10	-1.6	-20.5	-0.99	0.324
education	Unmatched	0.06	0.05	3.4		2.24	0.025

	Matched	0.06	0.06	0.2	94.9	0.1	0.919
health care	Unmatched	0.12	0.09	7.9		5.26	0
	Matched	0.12	0.12	-2.3	70.5	-1.34	0.18
other types of	Unmatched	0.06	0.08	-5.6		-3.68	0
care	Matched	0.06	0.06	0.9	84.3	0.57	0.568
public admin-	Unmatched	0.06	0.05	6.2		4.13	0
istration	Matched	0.06	0.06	0.9	85.5	0.52	0.602
other	Unmatched	0.21	0.25	-10.9		-7.15	0
	Matched	0.21	0.21	-0.6	94.2	-0.4	0.687
unknown/did	Unmatched	0.16	0.15	4.2		2.76	0.006
not work	Matched	0.16	0.16	1.8	57.6	1.07	0.285
employed	Unmatched	0.88	0.89	-1.1		-0.73	0.464
1985	Matched	0.88	0.88	1.3	-21	0.81	0.419
employed	Unmatched	0.90	0.89	2.2		1.42	0.155
1986	Matched	0.90	0.90	0.1	95.9	0.05	0.957
employed	Unmatched	0.91	0.91	1.1		0.72	0.471
1986	Matched	0.91	0.90	1.5	-36.4	0.91	0.362
employed	Unmatched	0.92	0.91	2.2		1.45	0.147
1988	Matched	0.92	0.92	0.4	83.8	0.22	0.825
employed	Unmatched	0.92	0.91	3.5		2.31	0.021
1989	Matched	0.92	0.92	0.7	81.1	0.42	0.678
ln labor in-	Unmatched	3.69	3.69	-0.3		-0.18	0.855
come 1985	Matched	3.69	3.64	1.6	-480.9	0.96	0.338
ln labor in-	Unmatched	3.82	3.81	0.7		0.48	0.632
come 1986	Matched	3.82	3.83	-0.3	65.3	-0.15	0.878
ln labor in-	Unmatched	3.89	3.86	1.3		0.85	0.397
come 1987	Matched	3.89	3.89	0	99.7	0	0.998
ln labor in-	Unmatched	3.97	3.90	2.8		1.85	0.064
come 1988	Matched	3.97	3.97	-0.1	95	-0.09	0.931
ln labor in-	Unmatched	4.03	3.90	5		3.28	0.001
come 1989	Matched	4.03	4.01	1	80.5	0.61	0.544
Individuals in sample	Unmatched	7477	10401				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Born in a western country

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.58	0.51	15.9		2.47	0.014
	Matched	0.58	0.62	-8	49.4	-1.08	0.28
age group 1989							
20-29 years	Unmatched	0.08	0.15	-22		-3.26	0.001
	Matched	0.08	0.07	0.9	95.9	0.14	0.887
30-39 years	Unmatched	0.32	0.27	11.1		1.75	0.08
	Matched	0.32	0.30	5.6	49.6	0.73	0.463
40-49 years	Unmatched	0.42	0.36	11.9		1.87	0.062
	Matched	0.42	0.42	0	100	0	1
50-57 years	Unmatched	0.18	0.22	-10		-1.54	0.125
	Matched	0.18	0.21	-7.1	29.1	-0.95	0.343
education 1990							
10 years or	Unmatched	0.30	0.42	-24.9		-3.82	0
less	Matched	0.30	0.28	5.4	78.4	0.75	0.454

secondary	Unmatched	0.44	0.42	3.4		0.54	0.592
	Matched	0.44	0.44	0	100	0	1
tertiary	Unmatched	0.25	0.15	24.1		3.88	0
	Matched	0.25	0.27	-6.4	73.2	-0.77	0.44
missing	Unmatched	0.01	0.00	8.1		1.41	0.16
	Matched	0.01	0.01	0	100	0	1
industry 1989							
manufacturing	Unmatched	0.20	0.37	-39.1		-5.87	0
	Matched	0.20	0.16	9	76.9	1.39	0.166
construction	Unmatched	0.02	0.05	-15.8		-2.29	0.022
	Matched	0.02	0.03	-3.1	80.1	-0.51	0.614
trade	Unmatched	0.07	0.04	11.3		1.84	0.066
	Matched	0.07	0.05	9.8	13.9	1.27	0.204
education	Unmatched	0.05	0.03	10.9		1.79	0.074
	Matched	0.05	0.08	-12.6	-15.7	-1.36	0.175
health care	Unmatched	0.15	0.08	19.6		3.2	0.001
	Matched	0.15	0.14	0.9	95.4	0.11	0.915
other types of	Unmatched	0.07	0.07	1.4		0.23	0.821
care	Matched	0.07	0.08	-3.3	-128.4	-0.42	0.674
public admin-	Unmatched	0.03	0.02	9.3		1.53	0.126
istration	Matched	0.03	0.07	-23	-146.8	-2.2	0.028
other	Unmatched	0.20	0.16	11		1.75	0.081
	Matched	0.20	0.19	2.2	79.9	0.28	0.777
unknown/did	Unmatched	0.20	0.17	7.6		1.2	0.232
not work	Matched	0.20	0.20	0.7	90.4	0.09	0.925
employed	Unmatched	0.81	0.85	-11		-1.74	0.082
1985	Matched	0.81	0.84	-9.2	16.2	-1.21	0.227
employed	Unmatched	0.83	0.87	-11.7		-1.86	0.063
1986	Matched	0.83	0.84	-3.2	72.7	-0.41	0.685
employed	Unmatched	0.84	0.87	-6.4		-1	0.317
1986	Matched	0.84	0.86	-4	36.5	-0.53	0.597
employed	Unmatched	0.86	0.88	-4.9		-0.78	0.437
1988	Matched	0.86	0.87	-1	79.9	-0.13	0.898
employed	Unmatched	0.86	0.87	-1.9		-0.29	0.772
1989	Matched	0.86	0.86	-0.4	77.5	-0.05	0.956
ln labor in-	Unmatched	2.85	3.30	-12.8		-2.03	0.043
come 1985	Matched	2.85	3.06	-6.1	52.5	-0.78	0.436
ln labor in-	Unmatched	2.97	3.50	-15.4		-2.47	0.014
come 1986	Matched	2.97	3.02	-1.6	89.4	-0.2	0.839
ln labor in-	Unmatched	3.11	3.45	-9.9		-1.56	0.118
come 1987	Matched	3.11	3.22	-3.3	66.7	-0.43	0.67
ln labor in-	Unmatched	3.20	3.52	-9.4		-1.49	0.136
come 1988	Matched	3.20	3.32	-3.4	63.3	-0.44	0.658
ln labor in-	Unmatched	3.35	3.40	-1.6		-0.24	0.809
come 1989	Matched	3.35	3.17	5.4	-248.2	0.68	0.494
Individuals in sample	Unmatched	351	782				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Born in a non-western country

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.50	0.46	8.8		0.64	0.521
	Matched	0.50	0.44	12.1	-37.9	0.83	0.407
age group 1989							
20-29 years	Unmatched	0.14	0.14	0.7		0.05	0.961
	Matched	0.14	0.10	11.6	-1640.5	0.85	0.399
30-39 years	Unmatched	0.41	0.40	3		0.22	0.825
	Matched	0.41	0.41	0.8	73.6	0.05	0.957
40-49 years	Unmatched	0.32	0.30	3.3		0.24	0.812
	Matched	0.32	0.34	-5.3	-62.2	-0.36	0.722
50-57 years	Unmatched	0.13	0.16	-9.2		-0.67	0.505
	Matched	0.13	0.15	-5.5	40.2	-0.38	0.702
education 1990							
10 years or	Unmatched	0.22	0.31	-20		-1.46	0.147
less	Matched	0.22	0.22	0	100	0	1
secondary	Unmatched	0.38	0.40	-3.5		-0.25	0.8
	Matched	0.38	0.38	0	100	0	1
tertiary	Unmatched	0.35	0.21	32.1		2.38	0.018
	Matched	0.35	0.35	0	100	0	1
missing	Unmatched	0.04	0.08	-15.6		-1.12	0.264
	Matched	0.04	0.04	0	100	0	1
industry 1989							
manufacturing	Unmatched	0.18	0.26	-20		-1.45	0.148
	Matched	0.18	0.18	0	100	0	1
construction	Unmatched	0.01	0.01	2.7		0.2	0.84
	Matched	0.01	0.01	0	100	0	1
trade	Unmatched	0.04	0.02	10.3		0.77	0.442
	Matched	0.04	0.04	0	100	0	1
education	Unmatched	0.03	0.06	-11.7		-0.84	0.4
	Matched	0.03	0.03	0	100	0	1
health care	Unmatched	0.03	0.04	-4.3		-0.31	0.754
	Matched	0.03	0.03	0	100	0	1
other types of	Unmatched	0.15	0.09	18.8		1.4	0.162
care	Matched	0.15	0.15	0	100	0	1
public admin-	Unmatched	0.01	0.01	2.7		0.2	0.84
istration	Matched	0.01	0.01	0	100	0	1
other	Unmatched	0.28	0.23	10.2		0.75	0.454
	Matched	0.28	0.28	0	100	0	1
unknown/did	Unmatched	0.27	0.28	-3.1		-0.23	0.819
not work	Matched	0.27	0.27	0	100	0	1
employed	Unmatched	0.63	0.50	25		1.83	0.069
1985	Matched	0.63	0.64	-2.9	88.5	-0.2	0.841
employed	Unmatched	0.68	0.57	23.4		1.7	0.09
1986	Matched	0.68	0.69	-1.3	94.3	-0.1	0.924
employed	Unmatched	0.72	0.64	17.9		1.3	0.194
1986	Matched	0.72	0.72	1	94.3	0.07	0.943
employed	Unmatched	0.80	0.76	9.1		0.66	0.508
1988	Matched	0.80	0.80	-1.4	84.3	-0.1	0.919
employed	Unmatched	0.79	0.75	8.3		0.61	0.544
1989	Matched	0.79	0.78	0.9	89.7	0.06	0.952

ln labor income 1985	Unmatched	0.95	-0.20	22.2		1.62	0.107
	Matched	0.95	1.07	-2.5	89	-0.18	0.861
ln labor income 1986	Unmatched	1.15	-0.09	24		1.74	0.083
	Matched	1.15	1.12	0.6	97.7	0.04	0.968
ln labor income 1987	Unmatched	2.09	0.56	31.8		2.3	0.022
	Matched	2.09	1.41	14	56	0.98	0.326
ln labor income 1988	Unmatched	2.84	1.78	25.6		1.85	0.066
	Matched	2.84	2.33	12.4	51.5	0.88	0.38
ln labor income 1989	Unmatched	2.92	2.47	11.7		0.85	0.394
	Matched	2.92	2.88	1	91.4	0.07	0.944
Individuals in sample	Unmatched	94	125				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Without Tierp

		mean			% reduction	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.50	0.47	5.3		2.98	0.003
	Matched	0.50	0.49	1.3	76	0.63	0.526
age group 1989							
20-29 years	Unmatched	0.15	0.20	-14.6		-8.16	0
	Matched	0.15	0.14	2.4	83.3	1.33	0.182
30-39 years	Unmatched	0.36	0.30	12.1		6.9	0
	Matched	0.36	0.36	0.3	97.2	0.17	0.869
40-49 years	Unmatched	0.38	0.33	9.6		5.44	0
	Matched	0.38	0.39	-2.7	71.6	-1.35	0.178
50-57 years	Unmatched	0.12	0.17	-13.5		-7.52	0
	Matched	0.12	0.12	0.6	95.4	0.34	0.735
education 1990							
10 years or less	Unmatched	0.26	0.35	-20.6		-11.54	0
	Matched	0.26	0.26	0.8	96.2	0.41	0.683
secondary	Unmatched	0.44	0.47	-5.1		-2.88	0.004
	Matched	0.44	0.45	-0.8	83.7	-0.42	0.675
tertiary	Unmatched	0.29	0.17	29.2		16.95	0
	Matched	0.29	0.29	0	100	0	1
missing	Unmatched	0.00	0.01	-2.8		-1.58	0.115
	Matched	0.00	0.00	0.8	72.2	0.46	0.647
birth region							
Sweden	Unmatched	0.93	0.91	7.5		4.2	0
	Matched	0.93	0.93	1.4	81.2	0.75	0.453
western country	Unmatched	0.05	0.07	-9.7		-5.39	0
	Matched	0.05	0.05	-0.9	90.7	-0.5	0.62
non-western country	Unmatched	0.02	0.01	2.9		1.68	0.094
	Matched	0.02	0.02	-1.3	54.3	-0.62	0.536
industry 1989							
manufacturing	Unmatched	0.13	0.19	-16.4		-9.1	0
	Matched	0.13	0.13	0.6	96	0.36	0.721
construction	Unmatched	0.06	0.07	-2.6		-1.45	0.148
	Matched	0.06	0.07	-1.1	57.5	-0.56	0.577
trade	Unmatched	0.10	0.10	0.4		0.25	0.804
	Matched	0.10	0.09	2.5	-461.6	1.25	0.21
education	Unmatched	0.06	0.05	5		2.84	0.004

	Matched	0.06	0.06	-1.4	72.3	-0.65	0.514
health care	Unmatched	0.12	0.08	12.2		7.08	0
	Matched	0.12	0.12	1.6	86.7	0.77	0.444
other types of	Unmatched	0.07	0.08	-2.3		-1.3	0.192
care	Matched	0.07	0.08	-2.3	2.3	-1.14	0.254
public admin-	Unmatched	0.06	0.04	10.9		6.38	0
istration	Matched	0.06	0.06	1.9	82.7	0.87	0.382
other	Unmatched	0.22	0.23	-1.8		-1.01	0.31
	Matched	0.22	0.23	-1.2	31.7	-0.62	0.536
unknown/did	Unmatched	0.17	0.17	1.6		0.89	0.374
not work	Matched	0.17	0.17	-0.5	69.8	-0.24	0.813
employed	Unmatched	0.89	0.88	2.9		1.63	0.103
1985	Matched	0.89	0.89	-1.9	34.3	-0.99	0.323
employed	Unmatched	0.90	0.89	3.7		2.11	0.035
1986	Matched	0.90	0.90	-1	73.2	-0.52	0.601
employed	Unmatched	0.91	0.90	3.3		1.85	0.065
1986	Matched	0.91	0.92	-2.1	37.1	-1.08	0.28
employed	Unmatched	0.92	0.91	4		2.26	0.024
1988	Matched	0.92	0.92	-2.2	46.2	-1.14	0.254
employed	Unmatched	0.92	0.90	7.4		4.12	0
1989	Matched	0.92	0.92	-1.2	83.3	-0.66	0.507
ln labor in-	Unmatched	3.72	3.59	4.5		2.55	0.011
come 1985	Matched	3.72	3.75	-1.2	73.4	-0.62	0.533
ln labor in-	Unmatched	3.84	3.71	4.5		2.57	0.01
come 1986	Matched	3.84	3.87	-1.2	72.9	-0.64	0.524
ln labor in-	Unmatched	3.92	3.75	6.2		3.47	0.001
come 1987	Matched	3.92	3.99	-2.8	53.8	-1.51	0.13
ln labor in-	Unmatched	4.00	3.81	7.4		4.18	0
come 1988	Matched	4.00	4.04	-1.6	78	-0.87	0.386
ln labor in-	Unmatched	4.08	3.79	10.6		5.92	0
come 1989	Matched	4.08	4.11	-1.2	88.4	-0.67	0.505
Individuals in sample	Unmatched	5085	8729				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Without Storvreta

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.49	0.48	2.4		1.41	0.158
	Matched	0.49	0.49	0.3	87	0.16	0.874
age group 1989							
20-29 years	Unmatched	0.20	0.21	-2		-1.15	0.251
	Matched	0.20	0.20	-0.3	85.2	-0.15	0.882
30-39 years	Unmatched	0.29	0.30	-0.8		-0.47	0.64
	Matched	0.29	0.29	-0.1	83.9	-0.07	0.948
40-49 years	Unmatched	0.34	0.33	2.1		1.23	0.22
	Matched	0.34	0.34	1.1	46.8	0.56	0.574
50-57 years	Unmatched	0.16	0.16	0.4		0.26	0.794
	Matched	0.16	0.17	-0.9	-113.5	-0.48	0.633
education 1990							
10 years or	Unmatched	0.33	0.35	-2.4		-1.38	0.169
less	Matched	0.33	0.33	1	57.9	0.5	0.615

secondary	Unmatched	0.49	0.48	0.4		0.24	0.813
	Matched	0.49	0.49	-1.6	-296.5	-0.81	0.418
tertiary	Unmatched	0.17	0.16	2.5		1.46	0.145
	Matched	0.17	0.17	1.3	49.5	0.63	0.529
missing	Unmatched	0.01	0.01	-0.2		-0.12	0.908
	Matched	0.01	0.01	-1.7	-759.9	-0.82	0.411
birth region							
Sweden	Unmatched	0.95	0.91	15		8.41	0
	Matched	0.95	0.95	-0.1	99.5	-0.05	0.963
western	Unmatched	0.04	0.07	-13.7		-7.69	0
country	Matched	0.04	0.04	-0.6	95.7	-0.34	0.733
non-western	Unmatched	0.01	0.01	-5.7		-3.2	0.001
country	Matched	0.01	0.00	1.7	70.7	1.09	0.275
industry 1989							
manufacturing	Unmatched	0.20	0.19	2		1.18	0.237
	Matched	0.20	0.20	-0.1	95.1	-0.05	0.961
construction	Unmatched	0.08	0.07	5.1		3.02	0.003
	Matched	0.08	0.08	1.8	64.7	0.88	0.376
trade	Unmatched	0.08	0.09	-0.7		-0.41	0.682
	Matched	0.08	0.09	-1.6	-129.7	-0.81	0.42
education	Unmatched	0.06	0.05	3.5		2.09	0.037
	Matched	0.06	0.06	-2.3	33.6	-1.12	0.262
health care	Unmatched	0.10	0.09	3.4		2.01	0.045
	Matched	0.10	0.10	0.5	86.2	0.23	0.816
other types of	Unmatched	0.06	0.08	-5.4		-3.1	0.002
care	Matched	0.06	0.07	-1.1	80.2	-0.56	0.578
public admin-	Unmatched	0.05	0.04	2.6		1.54	0.123
istration	Matched	0.05	0.05	-2.6	0	-1.25	0.211
other	Unmatched	0.20	0.25	-11.1		-6.44	0
	Matched	0.20	0.19	1.7	84.8	0.9	0.371
unknown/did	Unmatched	0.17	0.15	4.8		2.83	0.005
not work	Matched	0.17	0.16	1.4	69.9	0.72	0.472
employed	Unmatched	0.86	0.88	-3.8		-2.26	0.024
1985	Matched	0.86	0.86	0.6	84.8	0.29	0.774
employed	Unmatched	0.89	0.89	0.6		0.33	0.743
1986	Matched	0.89	0.89	-0.8	-43	-0.41	0.683
employed	Unmatched	0.90	0.90	-0.4		-0.25	0.802
1986	Matched	0.90	0.90	0.3	25.1	0.16	0.872
employed	Unmatched	0.91	0.91	1.1		0.62	0.535
1988	Matched	0.91	0.91	0.9	17.5	0.44	0.657
employed	Unmatched	0.91	0.90	2		1.19	0.235
1989	Matched	0.91	0.90	2.1	-1	1.04	0.299
ln labor in-	Unmatched	3.49	3.59	-3.6		-2.14	0.033
come 1985	Matched	3.49	3.50	-0.6	84.6	-0.28	0.782
ln labor in-	Unmatched	3.64	3.70	-2.1		-1.26	0.208
come 1986	Matched	3.64	3.67	-1	51.2	-0.53	0.599
ln labor in-	Unmatched	3.72	3.76	-1.2		-0.72	0.469
come 1987	Matched	3.72	3.73	-0.4	68.5	-0.19	0.845
ln labor in-	Unmatched	3.81	3.80	0.1		0.03	0.973
come 1988	Matched	3.81	3.78	0.9	-1488.9	0.45	0.652
ln labor in-	Unmatched	3.86	3.81	1.8		1.07	0.283
come 1989	Matched	3.86	3.81	1.9	-4.8	0.97	0.332
Ind. in sample	Unmatched	5127	10462				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Worked in Uppsala city or further south

		mean			% reduc-tion	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.47	0.44	5.5		2.25	0.024
	Matched	0.47	0.47	1	81.9	0.41	0.681
age group 1989							
20-29 years	Unmatched	0.16	0.23	-19.6		-8.07	0
	Matched	0.16	0.16	-1.3	93.6	-0.56	0.575
30-39 years	Unmatched	0.36	0.32	7.6		3.11	0.002
	Matched	0.36	0.35	1.4	81.4	0.58	0.562
40-49 years	Unmatched	0.38	0.32	12.5		5.12	0
	Matched	0.38	0.40	-3.1	75.1	-1.26	0.207
50-57 years	Unmatched	0.10	0.12	-5.7		-2.32	0.02
	Matched	0.10	0.09	4.1	27.1	1.82	0.068
education 1990							
10 years or	Unmatched	0.22	0.26	-8.4		-3.45	0.001
less	Matched	0.22	0.20	4.6	45.7	1.98	0.048
secondary	Unmatched	0.45	0.50	-9.8		-4.01	0
	Matched	0.45	0.48	-4.7	51.7	-1.96	0.05
tertiary	Unmatched	0.32	0.24	18.9		7.74	0
	Matched	0.32	0.32	0.3	98.3	0.13	0.897
missing	Unmatched	0.00	0.00	0.3		0.13	0.896
	Matched	0.00	0.00	4.9	-1448.5	2.72	0.007
birth region							
Sweden	Unmatched	0.93	0.93	0.4		0.17	0.865
	Matched	0.93	0.94	-2.5	-509.5	-1.07	0.282
western	Unmatched	0.05	0.06	-1.6		-0.66	0.51
country	Matched	0.05	0.05	2.6	-61	1.12	0.262
non-western	Unmatched	0.02	0.01	2.1		0.85	0.397
country	Matched	0.02	0.02	0.5	77.3	0.19	0.85
industry 1989							
manufacturing	Unmatched	0.13	0.13	-1.2		-0.51	0.613
	Matched	0.13	0.13	0.7	44	0.29	0.772
construction	Unmatched	0.07	0.08	-5.2		-2.14	0.032
	Matched	0.07	0.07	1.4	72.9	0.62	0.535
trade	Unmatched	0.12	0.13	-2		-0.81	0.415
	Matched	0.12	0.12	1.6	20.3	0.67	0.503
education	Unmatched	0.04	0.03	4.7		1.91	0.056
	Matched	0.04	0.03	3.7	21.6	1.5	0.135
health care	Unmatched	0.16	0.15	3.6		1.47	0.141
	Matched	0.16	0.16	-0.6	82	-0.26	0.793
other types of	Unmatched	0.06	0.06	2.7		1.09	0.275
care	Matched	0.06	0.07	-2.8	-6.8	-1.12	0.261
public admin-	Unmatched	0.07	0.06	6.7		2.74	0.006
istration	Matched	0.07	0.08	-2.4	64.6	-0.91	0.361
other	Unmatched	0.26	0.28	-3.8		-1.57	0.116
	Matched	0.26	0.27	-2.6	31.8	-1.09	0.276
unknown/did	Unmatched	0.09	0.09	-0.9		-0.36	0.72
not work	Matched	0.09	0.08	3	-243.3	1.28	0.201
employed	Unmatched	0.91	0.89	6.3		2.6	0.009
1985	Matched	0.91	0.91	0.5	92.2	0.21	0.831

employed 1986	Unmatched	0.94	0.92	5.2		2.13	0.033
	Matched	0.94	0.93	2.5	52.1	1.06	0.291
employed 1986	Unmatched	0.94	0.94	1.9		0.78	0.438
	Matched	0.94	0.95	-1.1	41	-0.48	0.634
employed 1988	Unmatched	0.96	0.95	2.3		0.93	0.35
	Matched	0.96	0.96	0.3	87.7	0.12	0.905
employed 1989	Unmatched	0.96	0.95	3.3		1.35	0.177
	Matched	0.96	0.95	5	-50.5	2.02	0.044
ln labor income 1985	Unmatched	4.08	3.98	4.8		1.97	0.049
	Matched	4.08	4.07	0.5	88.6	0.23	0.819
ln labor income 1986	Unmatched	4.22	4.13	4.4		1.8	0.071
	Matched	4.22	4.17	2.3	46.6	0.96	0.336
ln labor income 1987	Unmatched	4.33	4.27	3.1		1.26	0.206
	Matched	4.33	4.38	-2.8	10.8	-1.18	0.237
ln labor income 1988	Unmatched	4.43	4.39	2.8		1.16	0.246
	Matched	4.43	4.48	-2.9	-1.9	-1.23	0.219
ln labor income 1989	Unmatched	4.64	4.56	9.1		3.73	0
	Matched	4.64	4.63	1	89	0.43	0.665
Individuals in sample	Unmatched	3441	3288				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Worked elsewhere

		mean			% reduction	t-test	
Variable	Sample	Treatment	Control	% bias	|bias|	t	p>|t|
female	Unmatched	0.52	0.49	5.6		2.99	0.003
	Matched	0.52	0.50	4.4	20.2	2.1	0.035
age group 1989							
20-29 years	Unmatched	0.19	0.19	-0.9		-0.47	0.637
	Matched	0.19	0.18	1.2	-36.2	0.57	0.567
30-39 years	Unmatched	0.29	0.29	0.4		0.19	0.847
	Matched	0.29	0.28	0.9	-161.2	0.44	0.657
40-49 years	Unmatched	0.35	0.34	1.2		0.63	0.53
	Matched	0.35	0.35	-1.4	-16.3	-0.64	0.521
50-57 years	Unmatched	0.18	0.18	-1		-0.52	0.602
	Matched	0.18	0.18	-0.6	34.5	-0.3	0.762
education 1990							
10 years or less	Unmatched	0.34	0.37	-5.9		-3.13	0.002
	Matched	0.34	0.34	0.4	92.8	0.2	0.841
secondary	Unmatched	0.48	0.47	1.3		0.67	0.501
	Matched	0.48	0.47	0.3	75	0.15	0.882
tertiary	Unmatched	0.18	0.15	6		3.27	0.001
	Matched	0.18	0.18	-0.6	90.1	-0.28	0.782
missing	Unmatched	0.01	0.01	-0.8		-0.44	0.663
	Matched	0.01	0.01	-1.5	-86.9	-0.71	0.478
birth region							
Sweden	Unmatched	0.95	0.91	15.9		8.23	0
	Matched	0.95	0.96	-1.8	88.9	-1.01	0.314
western country	Unmatched	0.04	0.08	-16.4		-8.42	0
	Matched	0.04	0.04	1.9	88.4	1.12	0.265
non-western country	Unmatched	0.01	0.01	-1.8		-0.95	0.343
	Matched	0.01	0.01	0	100	0	1

industry 1989							
manufacturing	Unmatched	0.19	0.20	-4.5		-2.41	0.016
	Matched	0.19	0.18	0.6	87.5	0.27	0.786
construction	Unmatched	0.07	0.06	4.4		2.37	0.018
	Matched	0.07	0.07	-0.1	97.9	-0.04	0.967
trade	Unmatched	0.07	0.07	0.9		0.46	0.644
	Matched	0.07	0.07	-0.3	70	-0.12	0.904
education	Unmatched	0.08	0.06	6.6		3.62	0
	Matched	0.08	0.08	0.2	97.4	0.08	0.937
health care	Unmatched	0.08	0.07	5.4		2.93	0.003
	Matched	0.08	0.08	-0.2	96.9	-0.08	0.939
other types of	Unmatched	0.07	0.09	-7		-3.68	0
care	Matched	0.07	0.07	-0.7	89.3	-0.37	0.709
public admin-	Unmatched	0.05	0.04	4.5		2.46	0.014
istration	Matched	0.05	0.05	1.3	70.9	0.6	0.548
other	Unmatched	0.17	0.23	-16.7		-8.8	0
	Matched	0.17	0.16	0.5	97	0.26	0.798
unknown/did	Unmatched	0.23	0.18	12.4		6.72	0
not work	Matched	0.23	0.23	-1	91.9	-0.45	0.652
employed	Unmatched	0.85	0.87	-6.9		-3.75	0
1985	Matched	0.85	0.85	0.3	96.3	0.12	0.906
employed	Unmatched	0.86	0.87	-3.1		-1.68	0.092
1986	Matched	0.86	0.86	2.2	30.3	1	0.317
employed	Unmatched	0.87	0.88	-3.1		-1.66	0.097
1986	Matched	0.87	0.87	1.1	63.3	0.52	0.601
employed	Unmatched	0.88	0.89	-2.4		-1.28	0.202
1988	Matched	0.88	0.88	0.6	73.3	0.29	0.769
employed	Unmatched	0.88	0.88	-0.9		-0.51	0.612
1989	Matched	0.88	0.87	1.7	-82.8	0.8	0.422
ln labor in-	Unmatched	3.26	3.47	-6.5		-3.54	0
come 1985	Matched	3.26	3.25	0.3	94.9	0.15	0.879
ln labor in-	Unmatched	3.40	3.57	-5.4		-2.93	0.003
come 1986	Matched	3.40	3.37	0.9	83.2	0.42	0.677
ln labor in-	Unmatched	3.46	3.58	-4		-2.19	0.029
come 1987	Matched	3.46	3.44	0.5	86.9	0.24	0.807
ln labor in-	Unmatched	3.53	3.61	-2.7		-1.45	0.148
come 1988	Matched	3.53	3.50	0.8	68.7	0.39	0.699
ln labor in-	Unmatched	3.50	3.55	-1.7		-0.9	0.367
come 1989	Matched	3.50	3.48	0.6	61.9	0.3	0.767
Individuals in sample	Unmatched	4485	8051				

Note: To be able to take logarithms individuals with labor income 0 were ascribed 1 SEK.

Essay 2: Commuter Train Access and Residential Sorting*

* I would like to thank Lina Hedman, Matz Dahlberg, Olof Åslund, Peter Nilsson, and seminar participants at the Institute for Housing and Urban Research at Uppsala University, the 2012 Ph.D. Conference in Economic Demography in Mölle, the workshop Taxing cross-border commuters: Öresund and Beyond in Uppsala, the 2012 ERSA Summer School, the 2012 IIPF conference and the 2012 ERSA conference for helpful comments and discussions. I am grateful to John Östh, Åsa Bråmå, and Eva Jirner who at various stages helped me with geo-coded information and maps. Financial support from the Graduate School in Population Dynamics and Public Policy, Umeå University, is gratefully acknowledged.

1 Introduction

Good public transit in general, and by rail in particular, is often mentioned as important for the functioning of a metropolitan area. However, the construction, expansion, and maintenance of rail transit infrastructure require substantial investments. At the same time knowledge about how improvements in public transit affect firms and workers is limited. One area where evidence is particularly scarce is concerning the role of public transit, as well as other forms of transportation, in generating residential sorting.[1]

From a policy perspective, knowledge about how public transit affects residential sorting should be important for several reasons. For example, if it is thought that access to good public transit is especially important in strengthening the economic position of marginal workers, it would be of interest to know whether marginal workers tend to locate in places with good public transit or if they are outbid from such places. Knowledge about how public transit affects residential sorting could also guide decisions about the location and stretch of public transit infrastructure in order to avoid aspects such as wasteful migration and windfall gains to some residents. Further, knowledge about how public transit affects residential sorting could be of help in making decisions about what kind of housing and services should accompany public transit infrastructure.

This paper looks into the role of public transit in generating residential sorting by studying how the introduction of a commuter train in one part of Uppsala local labor market (Sweden) affected migration patterns in the areas served. The studied case, *Upptåget*, was introduced in the early 1990s and connected locations north of Uppsala city to the local center and further to the greater Stockholm area. To study the effects on in-migration patterns, I employ a difference-in-differences strategy comparing the characteristics of individuals moving into areas treated with commuter train access to the characteristics of individuals moving into non-treated areas on the same local labor market before and after the introduction of Upptåget. To study the effects on out-migration patterns, I employ a difference-in-differences-in-differences strategy comparing the probability of moving out of treated and non-treated areas before and after the introduction of Upptåget for individuals with different characteristics. I focus on whether there is sorting between individuals with different employment status

[1] As noted by Gibbons and Machin (2006, p. 10) in their literature study on transport and labor market linkages, there is "widespread evidence that land and housing prices are higher in places with good (rail) transport access and respond to transport improvements". However, this only says that people in general are attracted towards places with good (rail) transit access, not whether some types of people are more attracted than others.

and labor income, but also consider sorting based on education level, age, sex, and birth region.

The only previous empirical studies that I am aware of on how public transit affects residential sorting are Glaeser, Kahn, and Rappaport (2008) and Kahn (2007). They use the rail transit expansions in a number of major US cities between 1970 and 2000 to compare how the poverty rate (Glaeser et al., 2008), the share of college graduates, and the average household income (Kahn, 2007) changed in census tracts that were treated with rail transit access and in census tracts that were not.[2,3] Glaeser et al. (2008) find that pooling across cities, census tracts that came to be located within a mile of a rail transit line experienced an increase in poverty relative to non-treated census tracts. Kahn (2007) allows the treatment effect to be different for tracts that got access to a "Park and Ride" station and to a "Walk and Ride" station. Kahn (2007) finds that pooling across cities, census tracts that got access to a "Park and Ride" station saw a reduction in the share of college graduates and in average household income while census tracts that got access to a "Walk and Ride" station saw an increase in the share of college graduates and in the average household income. Conducting separate estimations for each city, however, Kahn (2007) finds rather different effects in different cities although the pattern that census tracts which received access to a new "Walk and Ride" station experienced greater gentrification than census tracts that received access to a new "Park and Ride" station holds in most cities.

By their difference-in-differences approach, the studies of Glaeser et al. (2008) and Kahn (2007) control for time effects as well as area specific fixed effects that could influence both public transit access and population composition. However, if individuals' income, employment status, and education level are affected by rail transit access in line with the spatial mismatch hypothesis (Kain, 1968), the estimates in Glaeser et al. (2008) and Kahn

[2] Glaeser et al. (2008) also use a sample of census tracts in New York City's outer boroughs to look at whether at one point in time, (the year 2000), there is a relation between distance to a subway line and the census tracts' median household income. This is done on the basis of the argument that the location of the subway stops are at least to some degree exogenous to the population composition in 2000 since no subway stops had been added in these boroughs since 1942 – any endogeneity of stop locations must thus stem from poverty levels before 1942 and many neighborhoods have changed radically since then. Glaeser et al. (2008) find that subway access appeared to strongly predict poverty, and also to explain a substantial amount of the connection between distance to the central business district and poverty. Further, Glaeser et al. (2008) make cross-city comparisons, finding that the poor are more likely to live in cities with more public transit and that they are less centralized when the suburb-central city gap in public transit is smaller (the gap can be smaller either because the transit mode is almost exclusively car or because subway provides good public transit also at some distance from the city center).

[3] The city samples and the time periods are not identical in the two studies. Glaeser, Kahn and Rappaport (2008) use a 16-city sample and study the time period 1980-2000, while Kahn (2007) uses a 14-city sample and studies the time period 1970-2000.

(2007) – although being unbiased measures of the effect of rail transit access on population composition – could be biased measures of the effect of rail transit access on residential sorting.[4]

One contribution of the present paper is therefore that it focuses directly on the migration patterns in the areas that obtained commuter train access through Upptåget. The focus on migration is possible thanks to the data available: the study uses longitudinal data containing all individuals aged 16-64 residing in Sweden in 1985 or later, with detailed geographical information on residential location as well as on demographic characteristics, education, and labor market outcomes. Another contribution is the European perspective: It is possible that the effects of rail transit access are different in Europe where the public transit network in and around cities is generally more extended than in the US, and where residential mobility is generally lower. A third contribution of the paper is that it considers sorting by employment status, labor income, and education level as well as by age, sex, and birth region. It can also be noted that although the paper only studies one case of commuter train introduction, Upptåget, the institutional features suggest that the case is well suited to overcome problems with reversed causality and omitted variables. The introduction of Upptåget was not primarily motivated by migration patterns in the treatment areas; instead the stretch was governed by already existing railroad tracks and the timing was related to a legal change. Further, the train altered commuting opportunities and travel times by public transit in some areas, while leaving conditions unchanged for other areas in the same local labor market that can thus be used for comparison.

I find that the commuter train Upptåget had a positive effect on overall in-migration to the areas served and no effect on the average out-migration rate from these areas. With regards to sorting based on labor market status, I find no evidence of sorting based on employment status but some evidence that the train introduction increased the probability of moving out of the areas served for individuals with high labor incomes relative to the probability for individuals with lower income. Considering sorting along other lines than labor market status, the analysis suggests that people born in non-western

[4] Kain (1968) suggested that the unemployment rate of African-Americans in US metropolitan areas was aggravated by the movement of low-skilled jobs from the central cities to the suburbs, worsening job access for African-American workers constrained to central cities by housing market discrimination. Since the study of Kain (1968) a large number of empirical studies have been carried out trying to test the relation between job access and labor market outcomes in general and the spatial mismatch hypothesis in particular. The collected evidence suggests that bad job access does worsen labor market outcomes (for literature surveys see Gobillon, Selod, & Zenou, 2007; Ihlanfeldt, 2006; Ihlanfeldt & Sjoquist, 1998; Zenou, 2009). It has been suggested that it is not only the physical distance from jobs that matters but also whether people are disconnected from jobs in the sense that they do not have access to a car, given that a public transit ride between two given points tends to take more time than the car ride (see e.g., Taylor & Ong, 1995).

countries came to be particularly attracted towards the areas served by the commuter train as compared to other similar areas.

The rest of the paper is organized as follows: Section 2 briefly presents theories on the relation between public transit and residential sorting. Section 3 describes the development of the commuter train Upptåget and presents the treatment and control areas. Section 4 presents the data and provides some descriptive statistics. Section 5 explains the estimation strategy. The results are reported in Section 6 and Section 7 concludes.

2 Theories

In this section I briefly review theories on the relation between public transit and residential sorting between people with different employment statuses and labor incomes.

In the standard urban economic model developed by Alonso (1964), Mills (1967), and Muth (1969), land prices decline with distance from the central business district (CBD) to compensate individuals for longer commutes. In this monocentric urban model, high-income people consume more land and therefore choose to live where land is cheap, i.e., far from the CBD, while low-income people live close to the CBD. The key condition for this is that the income elasticity of demand for land is greater than the income elasticity of marginal commuting cost (see Becker, 1965). The validity of this condition has been questioned, see e.g., Glaeser et al. (2008) and LeRoy and Sonstelie (1983). It has also been noted that for the theory to explain the spatial income gradient in European cities where high-income people tend to live closer to the CBD and low-income people further out, the preference pattern in Europe must be the opposite from that in the US (see e.g., Wheaton, 1977). LeRoy and Sonstelie (1983) not only question the empirical validity of the above mentioned key condition but also the model assumption that everyone commutes by the same mode.

LeRoy and Sonstelie (1983) therefore extend the standard urban economic model to incorporate two competing modes of commuting, one fast (e.g., car) and one slow with a lower material cost (e.g., public transit).[5,6] In their model, equilibrium residential locations are determined in a static framework

[5] LeRoy and Sonstelie (1983) describe their model using car as the fast commuting mode with high material cost and bus, i.e., one kind of public transit, as the slow commuting mode with lower material cost. However, as LeRoy and Sonstelie (1983) demonstrate in their interpretation of the historical changes in residential patterns of US cities, before the introduction of the car the fast commuting mode with high material cost was consecutive types of public transit – omnibus and commuter rail road from about 1830 and streetcar from about 1850 – while the slow mode was walking.

[6] Other extensions of the standard urban model include differences in amenities (Brueckner, Thisse, & Zenou, 1999) and housing quality (Brueckner & Rosenthal, 2009) at different distances from the CBD.

where individuals choose residential location and commuting mode simultaneously and where the costs of the competing commuting modes are given. Within the model, LeRoy and Sonstelie (1983) show that if the income elasticity of demand for land is less than that of the marginal commuting cost by either mode, then high-income people will live on the edge of the city only if the faster commuting mode is cheap enough that high-income people opt to use it, but costly enough that low-income people do not. When the material cost of the faster and more expensive commuting mode (car) decreases so that low-income people can afford it, the comparative advantage of high-income people in bearing high material transportation costs and living further from the CBD diminishes while their comparative advantage to live in the innermost urban area where both high- and low-income people commute by the slow and cheap mode (public transit) increases, and the tendency of high-income people to live further out will be reversed. Under the presence of positive fixed material costs, LeRoy and Sonstelie (1983) show that several different residential equilibria can prevail, where people are distributed around the CBD in up to four different rings defined by income-group and commuting mode.

For the case studied in the present paper, the introduction of a commuter train connecting locations north of Uppsala city to the local center and further to the greater Stockholm area, it can be noted that, like many other Swedish cities, Uppsala city and its surroundings seems largely to correspond to the four ring pattern. In the inner-city, high-income people walk or use public transit; in the next ring, low-income people use public transit; in the third ring, high-income people drive cars; and still further away from the city center, low-income people drive cars. The station localities seem to have been in the car rings. The commuter train made commuting by public transit from the station localities to Uppsala city faster and thus made the station localities more likely to be interesting alternatives for low-income people using public transit. The commuter train could thus have the effect of stretching the ring where low-income people use public transit further out along the commuter train line. The LeRoy and Sonstelie (1983) model therefore leads to the hypothesis that, if anything, we should see sorting of low-income people to the areas served following the introduction of the commuter train.

In the standard urban economic model as well as in the version by LeRoy and Sonstelie (1983), the labor market is fully competitive, productivity and wages are given, and there is no unemployment. In the middle of the 1990s, efforts started to combine urban economic models with labor economic theories and develop models in which workers' location (land market), as well as wages and unemployment (labor market) are determined in equilibrium, for a review see Zenou (2009). One branch of this urban labor economics introduces spatial frictions to efficiency wage models (see e.g., Brueckner & Ze-

nou, 2003; Ross & Zenou 2008; Zenou 2002; Zenou & Smith, 1995).[7] Another branch introduces spatial frictions to search-matching models (see e.g., Gobillon et al., 2007; Smith & Zenou, 2003; Wasmer & Zenou, 2002, 2006).[8] None of the above mentioned urban efficiency wage or urban search-matching models include different transport modes. However, these models provide a relation between employment status and access to jobs. To the extent that improved public transit from a place can be seen as increasing general access to jobs from that place, these theories can thus be of interest when studying the effect of commuter train access on residential sorting.

In the standard urban economic model as well as in the version by LeRoy and Sonstelie (1983), the residential equilibrium depends on the trade-off between land consumption and commuting costs for high- and low-income people. In the urban efficiency wage and urban search-matching models, the residential sorting equilibrium of the employed and unemployed depends on other forces instead or as well. Thus, in urban efficiency wage models with costless relocation, Zenou & Smith, 1995 and Zenou 2002, all workers are assumed to consume the same amount of land. In Zenou & Smith (1995) unemployed workers are simply assumed to commute less often to the CBD than the employed and therefore to be less willing to pay for land close to the CBD and consequently live further out. In Zenou (2002) unemployed workers are again assumed to commute less often to the CBD than the employed, but it is now also assumed that workers with longer commutes are more tired so that work effort is a negative function of the length of commute. Living far away from the CBD therefore has offsetting effects on employed workers. They pay a higher commuting cost but lower effort is exerted on the job. Apart from monetary and time costs, the commuting cost for employed workers includes non-work-related fatigue so even though people benefit from working less hard on the job the further away they live they are worse off overall. If the net effect for employed workers is stronger than the commuting cost effect for unemployed workers, they outbid the unemployed close to the CBD; in the opposite case the unemployed will live closer to the CBD.

In the urban search-matching model of Wasmer and Zenou (2002) employed workers travel more frequently to the CBD than unemployed workers, making residence closer to the CBD relatively more attractive for them, while the increase in job search efficiency associated with proximity to jobs attracts unemployed workers towards the job center. The relative strength of these forces decides whether the employed or the unemployed live closer to the CBD in equilibrium. In Wasmer and Zenou (2006) relocation costs are introduced in the model of Wasmer and Zenou (2002) and when these be-

[7] For the initial efficiency wage model see Shapiro and Stiglitz (1984).
[8] For the initial search-matching model see Mortensen and Pissarides (1999) and Pissarides (2000).

come positive a new equilibrium appears with an additional area between the two other areas, in which both employed and unemployed workers co-exist. In Smith and Zenou (2003), endogenous land consumption is introduced in a model similar to that of Wasmer and Zenou (2002, 2006) allowing unemployed workers to compete for locations close to the CBD by consuming small amounts of land if necessary. In the model the unemployed workers face a trade-off when choosing their residential locations: because of high (low) rents, locations close to (further away from) the CBD are costly (cheap) in the short run but allow higher (lower) search intensities which in turn increase the long-run prospects of re-employment (reduces the long run prospects of re-employment). Under certain assumptions, Smith and Zenou (2003) show that in this model, three different equilibria can prevail: the unemployed close to the CBD, the unemployed further away from the CBD and an equilibrium in which short run unemployed workers reside close to the CBD, long run unemployed live far away from the CBD and the employed live in between.

With the exception of Zenou and Smith (1995), the urban efficiency wage models and urban search matching models do not lead to any clear hypothesis on the effect of the introduction of the commuter train Upptåget on the sorting of employed and unemployed to the areas served. Instead they show that theoretically, both are possible. Empirical analyses are thus important.

The theories presented above focus on sorting by income and employment status. However, public transit could also affect sorting along other lines. At least in Sweden, women, young people, and immigrants, in particular immigrants from non-western countries, use public transit to a higher extent than other groups (see e.g., SOU, 2003:67). It is therefore interesting to study whether the introduction of the commuter train Upptåget led to sorting of these groups to the areas served, beyond what can be explained by employment status and labor income.

One further comment can be made about what to expect about the effect of the commuter train Upptåget on migration patterns in the areas served. As noted by Gibbons and Machin (2006), the effect of improved public transit on residential sorting will depend on whether transport costs are capitalized in housing and whether workers actually have to pay for any housing rent increases that occur. There is no adverse income effect of housing price increases for owner-occupiers, and before the introduction of the commuter train Upptåget most inhabitants in the areas subsequently served were owner-occupiers. Further, since rents are regulated in Sweden and location and job access not very important in the rent setting, renters in the areas served should not have suffered any adverse income effects either. At least in the short run, it could thus be expected that the introduction of the commuter train Upptåget would have a larger effect on in-migration patterns than on out-migration patterns.

3 Upptåget, treatment and control areas

To empirically study the role of public transit in generating residential sorting, I employ a quasi-experimental research design that builds on the introduction of a commuter train, Upptåget, to the Swedish city of Uppsala. I compare migration patterns in areas treated with commuter train access to migration patterns in non-treated areas in the same local labor market before and after the introduction of the train.

3.1 Upptåget

Uppsala city is the node of the local labor market of Uppsala, i.e., the main destination for labor commuting from the surrounding municipalities of Enköping, Heby, Tierp, and Östhammar.[9] Before the introduction of Upptåget, all public transit within the municipalities of Heby, Tierp, and Östhammar, as well as between these municipalities and Uppsala municipality was by bus. The only exception was a long-distance train with stops in Tierp town and Uppsala city. Public transit by road was coordinated and purchased by Upplands Lokaltrafik (UL), a firm jointly owned by Uppsala County Council and the municipalities in the county.[10,11]

Upptåget started running in 1991 between Uppsala city and Tierp town, 54 km north (see Figure 1 in Essay 1), with stops in four localities in between (Örbyhus, Skyttorp, Vattholma, Storvreta). In 1994 a fifth stop was added (Tobo). The train improved public transit from the areas served towards Uppsala city but also further towards the greater Stockholm area. Upptåget did not directly replace a particular bus line but stopped in places served by several other bus lines that also served places not crossed by Upptåget. These bus lines were to a large extent maintained after the introduction of Upptåget.

For several reasons, the commuter train Upptåget seems to be well suited to use as a quasi-experiment to study the effect of improved public transit on migration patterns. First, the introduction of Upptåget was not primarily motivated by the population composition or migration flows of different types of people in the areas subsequently served. The stretch of Upptåget was gov-

[9] The local labor markets are defined by Statistics Sweden based on commuting patterns. From 1996, Uppsala municipality and with it the municipalities of Enköping, Heby, Tierp, and Östhammar came to belong to the local labor market of Stockholm. However, Uppsala city continued to be the main destination for labor commuting from the municipalities of Enköping, Heby, Tierp, and Östhammar.

[10] Before 2007, the municipality of Heby belonged to Västmanland County and not to Uppsala County, so public transit within Heby municipality was not organized by UL.

[11] The information on the public transit network and on the development of *Upptåget* in this and the following paragraphs is from time-tables, annual reports and minutes from board meetings from UL, the investigations preceding the introduction of Upptåget and documents from County Council meetings. For more information on the development of Upptåget, see Section 3 in Essay 1.

erned by already existing railroad tracks, and the timing was related to a legal change – government bill 1987/88:50 making it possible for counties to take their own initiatives regarding train services by renting railroad from the administering State organization (Banverket).[12] It is true that the investigations preceding Upptåget mention regional imbalance with population pressure in Uppsala municipality and a stagnating population in the municipalities north of Uppsala municipality as one of the motivations for the commuter train. However, the population growth in the areas subsequently served does not seem to have been much different than the population growth in other similar areas in the same local labor market (cf. *Figure 1* below). Further, it transpires from the investigations preceding Upptåget that it was hoped that the commuter train would provide a wider housing choice especially for two earner households. However, this appears to have been seen as general feature of a commuter train, regardless of the actual stretch, and not related to any particular features of the areas subsequently served.[13] In sum, the institutional features suggest that the introduction of the commuter train Upptåget between Tierp town and Uppsala city was exogenous to subsequent migration patterns in the areas that came to be served.

Another reason that Upptåget seems to be well suited to use as a quasi-experiment to study the effect of improved public transit on migration patterns is that the time period between the first sketches and the inauguration of the train was at the same time short and characterized by uncertainty about the launch date, which diminishes estimation problems linked to anticipatory migration. The first concrete plans for Upptåget seem to have been outlined in 1988. It was first proposed that Upptåget should start running in the autumn of 1990. However, this proved impossible since another project was also planned on the same link (a high-speed long-distance train from Stockholm towards more northern parts of Sweden) making double tracks and some other track work necessary. In summer/autumn 1990 it was still unclear how fast the track work would proceed but a starting date for Upptåget was set for August 1991 and this was also achieved.

A third reason that Upptåget seems to be well suited to use as a quasi-experiment to study the effect of improved public transit on migration patterns is that the changes in travel times by public transit for (some of) the areas served were substantial, which should make it easier to detect any ef-

[12] When studying the effect of rail transit access on a census tract's share of college graduates and average household income, Kahn (2007) argues that communities are more likely to be treated with rail transit access if they anticipate that they will benefit from such treatment which could bias OLS estimates upwards. As a complement to OLS estimates, Kahn (2007) therefore conducts instrumental variables estimation, using a census tract's distance to the closest railroad line as an instrument for commuter train access. The motivation for the instrument is that proximity to a railroad line acts as a cost shifter since it is cheaper to convert old railroad lines into rail transit lines than to build entirely new rail infrastructure.

[13] For information on other motivations for Upptåget, see Section 3 in Essay 1.

fects. In 1990, the 41 kilometers between Örbyhus and Uppsala city took 65 minutes by bus, while the train ride took no more than 40 minutes in 1992, and about 30 minutes in 1996. For Skyttorp, 25 kilometers from Uppsala city, the corresponding figures were 42, 25, and 20 minutes. For Storvreta, 12 km from Uppsala city, the bus ride took 25 minutes, while the travel time with Upptåget was less than 10 minutes. Unlike the train, however, buses between Storvreta and Uppsala city stopped at several places both within Storvreta and Uppsala city. Considering walking time to a public transit station, the actual time gain for a public transit ride between Storvreta and Uppsala city brought about by Upptåget can therefore be thought to have been less important. For Tierp town, Upptåget did not really decrease the travel time by public transit since it was already possible to take a relatively fast long-distance train. However, the number of fast connections was greatly improved.[14] Here it can also be noted that the improvement did not imply more expensive tickets. The fare for a ride on Upptåget between any of the stations and Uppsala city was the same as for a bus ride, and no more expensive than corresponding bus rides to Uppsala city from other parts of the Uppsala local labor market.

A fourth reason that Upptåget seems to be well suited to use as a quasi-experiment to study the effect of improved public transit on migration patterns is that there are suitable control areas where public transit continued to be by bus for a long time, see Section 3.2.

3.2 Treatment and control areas

In this study I consider the treatment area to consist of the SAMS (Small Area for Market Statistics) with a population center within 4,500 meters of one of the stations served by Upptåget, and more than 10,000 meters from the central parts of Uppsala city.[15] The SAMS classification was created by Statistics Sweden to satisfy demand for small area statistics from users other than municipalities. The objective was to create fairly homogeneous residential areas of about 1000 inhabitants each, implying that the classification divides Sweden into about 9,000 units.[16] The SAMS have been used fre-

[14] In 1989 long-distance trains did 8 trips/weekday from Tierp to Uppsala. In 1992, Upptåget alone accounted for 15 trips, which had increased to 32 in 1996.

[15] We have, however, excluded the station settlement Tobo, despite the fact that it lies 4400 m from one of the other stations, since the commuter train Upptåget did not stop in Tobo until 1994.

[16] In larger municipalities, the SAMS classification is based on municipal subdivisions used for intra-municipal and sometimes regional planning and administration and in smaller municipalities it is based on election districts. The SAMS classification came into use in 1994 and has remained unchanged since then apart from minor adjustments, for example to adapt the SAMS-borders to municipal borders. Information from before 1994 can be located to a SAMS by use of the more precise coordinates that real estates have in Sweden. For more information see Statistics Sweden (2005).

quently in Swedish studies as the formal division closest to neighborhoods. The choice of 4,500 meters is somewhat arbitrary but should capture the areas where Upptåget came to be the main public transit mode.

As control area I have chosen two other parts of the Uppsala local labor market that prior to the introduction of the commuter train were similar to the treatment area in terms of size, distance to Uppsala city, population growth, public service, and housing, but where public transit continued to be by bus over the studied period, 1985-1996. The first part is the corridor between the principal locality in Heby municipality, Heby town, and Uppsala, where a commuter train on already existing railroad tracks was discussed at the same time as Upptåget but not realized.[17] The second part is the corridor between the principal locality in Östhammar municipality, Östhammar town, and Uppsala city. There is thus a treatment corridor north of Uppsala city, and two control corridors to the west and to northeast respectively. Analogous to the proximity to the stations for the SAMS that constitute the treatment area, I define the control corridors to consist of the SAMS with a population center within 4,500 meters of the main road between Uppsala city and Heby town on the one hand and between Uppsala city and Östhammar town on the other, and more than 10,000 meters from the central parts of Uppsala city. The treatment and control corridors are given by the shaded areas in *Figure 1* in Essay 1.

Figure 1 below shows the number and growth of the population 18-64 years old in the treatment corridor and the control corridors for the period 1985-1996. As can be seen from the figure, the population growth was similar in the treatment and control corridors 1985-1988. Thereafter the population growth was somewhat higher in the control corridors for three years, 1989-1991. From 1992, the year after Upptåget started running, the population growth was however much higher in the treatment corridor than in the control corridors.

[17] The railroad tracks between Heby town and Uppsala city were used for long distance trains that did not stop in Heby town or anywhere between Heby town and Uppsala city. In the investigations preceding Upptåget it was found that half-hour traffic Heby-Uppsala would be very uneconomic and require four trains, while half-hour traffic Tierp-Uppsala would only require three trains. In 1997, some long-distance/regional trains between Linköping and Uppsala ("the UVEN trains") started to stop in Heby and another locality in the stretch between Heby and Uppsala (Morgongåva). However, the connections were few.

Figure 1. Population aged 18-64 years in the treatment and control corridors 1985-1996

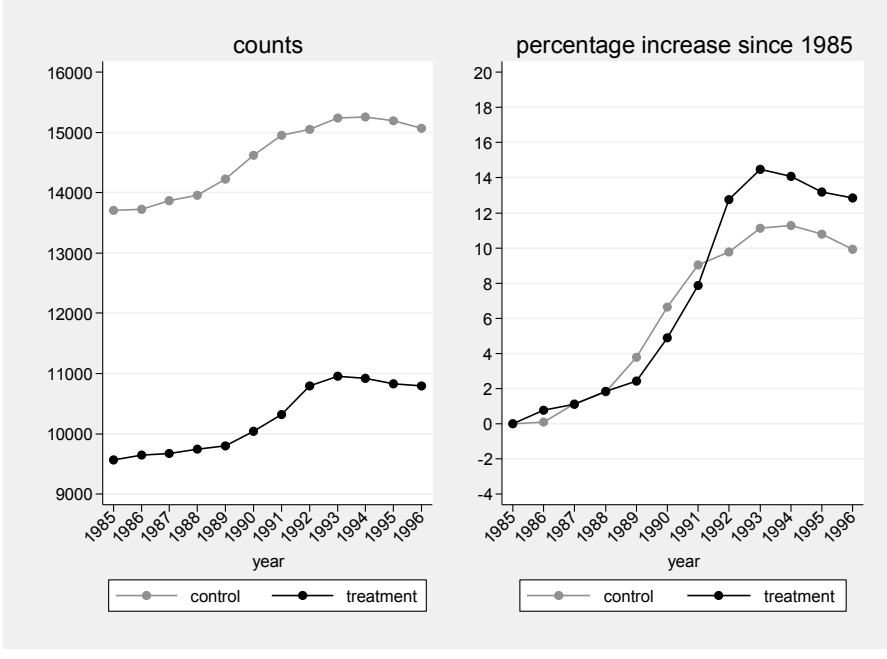
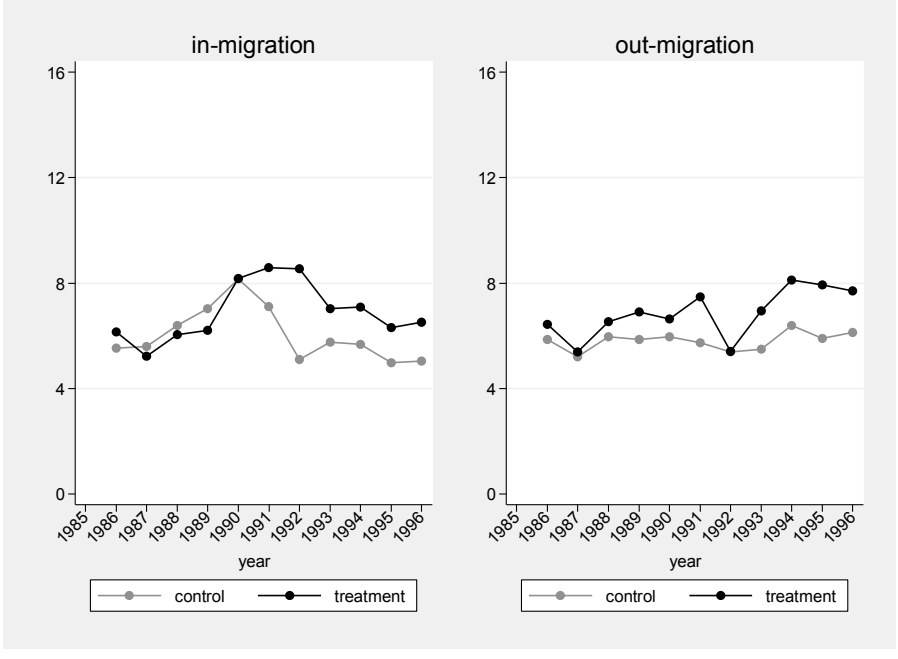


Figure 2 shows the in- and out-migration rate from the treatment and control corridors for the period 1986-1996. From *Figure 2* it seems that the development of the in- and out-migration rate in the treatment and control corridors was similar 1986-1990, although the in-migration (out-migration) rate was somewhat lower (higher) in the treatment corridor than in the control corridors most of the years. After 1990 however, the in-migration rate was higher in the treatment corridor than in the control corridors, while the difference in out-migration rate between the treatment corridor and the control corridors came to be somewhat larger than before, with the exception of 1992.

Figure 2. Migration rates in percent for the population aged 18-64 years in in the treatment and control corridors 1986-1996, percent



A final thing to note in this section is that the difference-in-difference(-in-differences) approach in this study includes differences before and after the introduction of the commuter train. However, it is not evident what is to be regarded as the pre-period and what is to be regarded as the post-period. For example, the introduction could be considered to start at the moment the discussions about the train started, i.e., at the beginning of 1988; at the moment it became clear that the train would in fact become reality although some uncertainty remained about the date, i.e., around summer 1990; or when the train actually started running, i.e., autumn 1991, or in 1992, the first year in which the train ran for the whole year. In the baseline estimations I have chosen 1986-1989 as the pre-period and 1990-1996 as the post-period on the basis that it was in 1990 that information about the certain and imminent running of the train was officially spread. As a sensitivity analysis, the time dynamics have been more closely studied.

4 Data

This paper looks into the role of public transit in generating residential sorting by studying how the introduction of the commuter train Upptåget affected the migration patterns of people with different characteristics in the areas

served. The focus is on whether there is sorting between individuals with different employment status and labor income, but I also consider sorting based on education level, age, sex, and birth region. To be able to analyze the characteristics of individuals moving into, moving out of and staying in the areas that got commuter train access, longitudinal individual-level data with information on residential location is crucial for at least three reasons. First, such data is necessary for the identification of in-migrants, out-migrants and stayers. Second, when studying the characteristics of individuals moving into the neighborhoods that got commuter train access, characteristics such as employment and labor income should refer to the period just before in-migration since an individual's employment status etc. after moving in could be affected by the commuter train access in line with the spatial mismatch hypothesis (Kain, 1968). Third, individual-level data makes it possible to take several characteristics into account at the same time which can be important. For example, if young people are more attracted towards the treatment corridor for reasons unrelated to Upptåget and young people tend to be unemployed to a higher extent than those who are older, it would be important to control for age composition among in-migrants when estimating the effect of Upptåget on the in-migration of employed and unemployed people to the treatment corridor. In the same way, if the population in the treatment corridor is slightly younger than in the control corridors, and the young tend to be both more mobile and more unemployed than those who are older, it would be important to control for age when estimating the effect of Upptåget on the probability of moving out of the treatment corridor for employed and unemployed people. Likewise, to study the effect of the commuter train on the migration patterns of people with different sex, age, education level and birth region, it would be important to control for any effect working through employment status or labor income.

For this study I have access to longitudinal (yearly) register data on all individuals 16-64 years living in Sweden. The data is compiled for research purposes by Statistics Sweden, and held by the Institute for Evaluation of Labour Market and Education Policy (IFAU). Among other things, the data contain rich and detailed information on demographic characteristics, income, employment, and education. For this study, it is particularly important that the data also contain geographic information on the residential location of each individual each year since this allows the identification of in-migrants, stayers, and out-migrants in the treatment and control corridors. In the register data, the residential location is given at SAMS level (see Section 3.2).[18]

From this database I extract two main samples. The first, sample, "in-migrants", is pooled cross-sections of the individuals aged 19-64 years[19] who

[18] The individuals' residential location refers to the location at the end of each year.
[19] The age refers to the age at the end of each year.

moved into the treatment or the control corridors 1986-1996.[20] To each individual in the "in-migrant" sample I tie information on employment status, labor income, and education level in the year before the individual moved into the treatment or the control corridor. Here it should be noted that from 1985 to 1994 there was a refugee placement policy in Sweden where recently arrived refugees were assigned to an initial place of residence.[21] Since I am mainly interested in the effect of the commuter train on voluntary location choices, I would like to exclude placed individuals from the sample of in-migrants. This is operationalized by excluding individuals missing from the register data the year before they appear in the control or treatment corridors. A likely reason that an individual is missing from the register data in a given year is that the individual did not live in Sweden in that year and thus potentially was a placed refugee.[22] The original in-migrant sample contains 17,039 individuals. Of these, 1,759 were missing from the register data the year before appearing in the treatment or control corridors and of those missing 70.3 % were born in a non-western country and thus potentially placed refugees.

The second sample, "stayers and out-migrants" is pooled cross-sections of all individuals aged 18-63 years living in the treatment and control corridors in 1985-1995. For these individuals I study who stayed and who moved out in the following year, i.e., 1986-1996, at the age of 19-64 years.[23]

4.1 The sample of in-migrants

Table 1 provides summary statistics for the sample of in-migrants stratified by whether an individual moved into the treatment corridor or the control corridors and whether this was before (1986-1989) or after (1990-1996) the introduction of the commuter train Upptåget.

[20] Individuals who move between SAMS *within* the treatment corridor or *within* the control corridors are not included in the in-migrant sample.

[21] For information on the refugee reception in Sweden during this period, see Edin, Fredriksson, and Åslund (2004).

[22] Even if some individuals missing from the register data the year before appearing in the treatment and control corridors are not placed refugees, it is reasonable to exclude them from the in-migrant sample since it is uncertain whether they are really in-migrants or were missing from the data for some other reason.

[23] Individuals who move between SAMS *within* the treatment corridor or *within* the control corridors are considered as stayers.

Table 1. Summary statistics sample of in-migrants.

	in-migrants to treatment corridor 1986-1989	in-migrants to control corridors 1986-1989	in-migrants to treatment corridor 1990-1996	in-migrants to control corridors 1990-1996
Employment status[a] *year-1*				
% employed	81.4	82.9	69.1	71.8
% unemployed	18.6	17.1	30.9	28.2
Labor income[b] *year-1*				
% middle	68.2	67.6	65.2	66.3
% low	21.4	19.7	24.9	21.6
% high	10.4	12.7	9.9	12.2
Education *year-1*				
% secondary	47.1	46.3	50.8	52.3
% =< 10 years	26.5	30.6	24.3	25.3
% tertiary	19.2	17.2	20.0	19.2
% information missing	7.2	5.9	4.9	3.3
Sex				
% man	49.1	49.8	50.1	49.5
% woman	50.9	50.2	49.9	50.5
Age group				
% aged 19-29	45.9	46.8	44.4	44.3
% aged 30-39	31.2	27.1	29.7	27.7
% aged 40-49	13.8	15.9	16.2	16.4
% aged 50-64	9.1	10.2	9.7	11.6
Birth region[c]				
% Sweden	90.7	91.2	85.1	89.7
% Western	6.9	6.4	5.1	6.0
% Non-western	2.4	2.4	9.9	4.2
N	2136	2954	4717	5473

Notes: a) Employment is based on the official annual employment statistics and refers to status during November each year. A person is classified as employed if he or she did paid work for at least one hour/week.

b) Labor income refers to annual earnings from work, including self-employment and employer's income. A person is classified as having low (high) labor income if he or she had a labor income lower (higher) than the 20th (80th) percentile of the population 18-64 years in Uppsala city (i.e., SAMS with SAMS-number 3800000-3800167) *year-1*.

c) Western countries include: Finland, Norway, Denmark, Iceland, Ireland, Great Britain, the Netherlands, Belgium, France, Monaco, Luxembourg, Lichtenstein, Spain, Portugal, Andorra, Germany, Switzerland, Italy, San Marino, Vatican City State, Austria, Greece, Canada, the USA, and the countries in Oceania.

From Table 1 it can first be noted that the share of employed was much lower both among people moving into the treatment corridor and among people moving into the control corridors in the period 1990-1996 than in the period 1986-1989. This decrease is probably due to a more general downturn. Sweden went from a booming economy in the late 1980s to a deep recession in the early 1990s. This probably also had an effect on the average labor income of the in-migrants. However, the income groups in Table 1 are defined relative to labor income among people in Uppsala city, who were also sub-

ject to the economic downturn. The economic downturn thus had little effect on the share of in-migrants with low, high, and middle income compared to the labor incomes in Uppsala city.

Comparing the in-migrants to the treatment corridor with the in-migrants to the control corridors, it can be seen from Table 1 that both before and after the introduction of Upptåget, the in-migrants to the treatment corridor were more likely to be unemployed than the in-migrants to the control corridors (18.6% vs. 17.1% and 30.9% vs. 28.2%). They were also more likely to have low labor income (21.4% vs. 19.7% and 24.9% vs. 21.6%) and less likely to have high labor income (10.4% vs. 12.7% and 9.9% vs. 12.2%). With regards to middle labor income, before the introduction of Upptåget the in-migrants to the treatment corridor were more likely to have middle income than the in-migrants to the control corridors (68.2% vs. 67.6%) while after the introduction the opposite was the case (65.2% vs. 66.3%).

From Table 1 it can also be seen that before the introduction of Upptåget the in-migrants to the treatment corridor were more likely to have secondary (47.1% vs. 46.3%) or tertiary education (19.2% vs. 17.2%) or information on education missing (7.2% vs 5.9%) than the in-migrants to the control corridors, and thus less likely to have at most 10 years of education (26.5% vs 30.6%). After the introduction of Upptåget, the in-migrants to the treatment corridor were still more likely to have tertiary education (20.0% vs 19.2%) or information on education missing (4.9% vs 3.3%) than in-migrants to the control corridors, but were now less likely to have secondary education (50.8% vs 52.3%).

Further, while the in-migrants to the treatment corridor were more likely to be women than the in-migrants to the control corridors before the introduction of Upptåget, the opposite was the case after the introduction. The difference between the two groups of in-migrants with respect to sex was however small in both periods.

With respect to age, Table 1 shows that both before and after the introduction of Upptåget, the in-migrants to the treatment corridor were more likely to belong to the age group 30-39 years than in-migrants to the control corridors (31.2% vs. 27.1% and 29.7% vs. 27.7%) and less likely to belong to any of the other age groups.

Finally, Table 1 shows that before the introduction of Upptåget there was little difference between the in-migrants to the treatment corridor and the in-migrants to the control corridors with respect to birth region. However, after the introduction, in-migrants to the treatment corridor were less likely to be born in Sweden (85.1% vs. 89.7%) and more likely to be born in a non-western country (9.9% vs 4.2%) than the in-migrants to the control corridors.

4.2 The sample of stayers and out-migrants

Table 2 provides summary statistics for the sample of stayers and out-migrants stratified by whether an individual stayed or moved out of the treatment corridor or the control corridors before (1986-1989) or after (1990-1996) the introduction of the commuter train Upptåget.

Table 2. Summary statistics sample of stayers and out-migrants.

	1986-1989				1990-1996			
	treatment		control		treatment		control	
	stayers	out-movers	stayers	out-movers	stayers	out-movers	stayers	out-movers
Employment status[a] *year-1*								
% employed	86.6	81.9	86.0	75.6	80.3	63.4	80.2	62.7
% unemployed	13.4	18.1	14.0	24.4	19.7	36.6	19.8	37.3
Labor income[b] *year-1*								
% middle	64.3	67.1	65.5	63.6	64.7	63.3	65.8	63.4
% low	15.8	20.6	16.7	25.1	15.2	25.7	15.9	27.2
% high	19.9	12.3	17.8	11.3	20.1	11.0	18.4	9.4
Education *year-1*								
% secondary	41.7	51.8	40.7	47.1	47.9	51.8	49.0	51.9
% =< 10 years	36.8	26.1	42.1	29.7	29.2	22.2	33.4	24.7
% tertiary	18.4	17.8	12.9	14.7	21.4	19.6	16.3	17.3
% information missing	3.1	4.3	4.2	8.4	1.5	6.4	1.3	6.1
Sex								
% man	50.5	49.0	52.4	49.3	50.8	49.9	52.0	51.2
% woman	49.5	51.0	47.6	50.7	49.2	50.1	48.0	48.8
Age group								
% aged 19-29	19.9	49.0	22.0	51.4	21.0	52.4	21.5	52.6
% aged 30-39	28.8	26.4	26.0	23.7	23.3	20.1	23.6	21.9
% aged 40-49	29.0	16.9	26.8	15.5	31.0	18.2	27.7	15.3
% aged 50-64	22.2	7.8	25.2	9.4	24.7	9.3	27.2	10.1
Birth region[c]								
% Sweden	94.4	92.5	91.9	87.5	92.2	83.7	91.4	85.0
% Western	4.5	5.4	7.1	6.8	4.3	4.5	6.6	4.9
% Non-western	1.1	2.0	1.0	5.7	3.5	11.8	2.0	10.1
N	35731	2349	51324	3041	67611	5196	97134	5992

Notes: a) Employment is based on the official annual employment statistics and refers to status during November each year. A person is classified as employed if he or she did paid work for at least one hour/week.

b) Labor income refers to annual earnings from work, including self-employment and employer's income. A person is classified as having low (high) labor income if he or she had a labor income lower (higher) than the 20th (80th) percentile of the population 18-64 years in Uppsala city (i.e., SAMS with SAMS-number 3800000-3800167) *year-1*.

c) Western countries include: Finland, Norway, Denmark, Iceland, Ireland, Great Britain, the Netherlands, Belgium, France, Monaco, Luxembourg, Lichtenstein, Spain, Portugal, Andorra, Germany, Switzerland, Italy, San Marino, Vatican City State, Austria, Greece, Canada, the USA, and the countries in Oceania.

From Table 2 it can first be noted that the share of employed was much lower among both stayers and out-migrants, in the treatment corridor and in the control corridors, in the period 1990-1996 than in the period 1986-1989, probably due to the economic downturn mentioned in Section 4.1. It can further be noted that in both the treatment corridor and in the control corridors, both before and after the introduction of the commuter train, out-migrants were more likely than stayers to be unemployed and to have low labor income and less likely to have high labor income. Out-migrants were also more likely than stayers to be young, women and to have at most secondary education. Finally, out-migrants were more likely than stayers to be born in a non-western country.

Comparing the stayers and out-migrants in the treatment corridor and in the control corridors, it can be seen from Table 2 that while the share of unemployed among stayers in the treatment and control corridors was rather similar both before and after the introduction of Upptåget (13.4% vs. 14% and (19.7% vs. 19.8%), the share of unemployed was much lower among the out-migrants from the treatment corridor than among the out-migrants from the control corridors before the introduction of the commuter train (18.1% vs 24.4%). After the introduction, the share of unemployed among the out-migrants from the treatment and control corridors became more similar (36.6% vs. 37.3%).

From Table 2 it can also be seen that before the introduction of Upptåget, the share of people with middle income was lower among stayers than among out-migrants in the treatment corridor (64.3% vs. 67.1%) while the opposite was the case for the control corridors both before (65.5% vs. 63.6%) and after (65.8% vs. 63.4%) the train introduction and for the treatment corridor after (64.7% vs. 63.3%) the train introduction.

One final thing that can be noted from Table 2: Before the introduction of the commuter train Upptåget the share of people born in a non-western country was about the double among out-migrants than among stayers in the treatment corridor. However, in the control corridors, the share of people born in a non-western country was almost six times as much among out-migrants than among stayers. After the introduction of Upptåget the relation between the share of people born in a non-western country among out-migrants and stayers in the treatment and in the control corridors became more similar: in the treatment corridor the share of people born in a non-western country among out-migrants was now more than three times as much as the share of people born in a non-western country among stayers, while the corresponding figure for the control corridors was about five times as much.

5. Estimation strategy

To look into the role of public transit for residential sorting I study how the introduction of the commuter train Upptåget affected the migration patterns of people with different characteristics in the areas served. To control for time effects as well as area specific fixed effects unrelated to the commuter train, the migration patterns in the treatment corridor are compared to the migration patterns in the control corridors before and after the introduction of Upptåget. To study the effects on in-migration patterns, I employ a difference-in-differences strategy comparing the characteristics of individuals moving into the treatment corridor to the characteristics of individuals moving into the control corridors before and after the introduction of Upptåget. To study the effects on out-migration patterns, I employ a difference-in-differences-in-differences strategy comparing the probability of moving out of the treatment and control corridors before and after the introduction of Upptåget for individuals with different characteristics. The focus is on whether there is sorting between individuals with different employment statuses and labor incomes, but I also consider sorting based on education level, age, sex, and birth region. The econometric specification used to study the effect of the commuter train on in-migration patterns is presented in Section 5.1 and the econometric specification used to study the effect on out-migration patterns is presented in Section 5.2.

5.1 Econometric specification for in-migration patterns

To study the effect of the commuter train Upptåget on the characteristics of individuals moving into the areas treated with commuter train access, I use the sample of in-migrants described in Section 4.1. I employ a difference-in-differences strategy to compare the characteristics of those moving into the treatment corridor to the characteristics of those moving into the control corridors before and after the introduction of Upptåget.[24] That is, using a linear probability model and the sample of in-migrants described in Section 4.1, I estimate an equation that in its richest specification takes the following form:

[24] The approach thus differs somewhat from the most common form of difference-in-differences analyses where the first difference is over time and the second difference is between treatment and control units, and the outcome variables is some unit characteristics, e.g., labor income. Here, while the first difference is over time, the second difference is between individuals with different characteristic, and the outcome variable is whether an individual moved into the treatment or control corridors.

$$IN_{it} = \alpha + \beta_0 X_{it-1} + \gamma_0' Z_{it-1} + \delta D + \beta_1 X_{it-1} \times D + \gamma_1' Z_{it-1} \times D + \varepsilon_{it-1}. \qquad (1)$$

In equation 1, IN_{it} is a dummy variable that equals one if individual *i* moved into the treatment corridor in year *t* and zero if individual *i* moved into one of the control corridors. In the case of employment status, X_{it-1} is a dummy variable that equals one if individual *i* was unemployed in year *t-1* and zero otherwise (see Appendix for the definition of employed and unemployed). Z_{it-1} is a vector of individual characteristics other than employment status that might affect the probability of moving into the treatment corridor. Z_{it-1} includes dummy variables corresponding to the age groups, education levels, and birth regions that are listed in Table 1, as well as a dummy variable for sex. The omitted, baseline, characteristics are the first group within each category in Table 1, i.e., being aged 18-29 years, having at most secondary education, being male and being born in Sweden. *D* is a dummy variable that switches from zero to one at the introduction of Upptåget. In equation 1, α hence measures the baseline probability that an in-migrant located in the treatment corridor before the introduction of the commuter train. β_0 measures whether this probability was different for unemployed in-migrants, and the coefficients in γ_0 whether this probability was different for in-migrants with the corresponding characteristics. The first three terms in equation (1) thus control for differences in the probability of locating in the treatment corridor between in-migrants with different characteristics that already existed before the train introduction. The coefficients of primary interest are δ, β_1 and the coefficients in γ_1. δ measures the change in the baseline probability that an in-migrant located in the treatment corridor following the train introduction. β_1 measures whether the change was different for unemployed in-migrants and the coefficients in γ_1 whether the change was different for in-migrants with the corresponding characteristics. A positive (negative) β_1 value thus indicates that the introduction of the commuter train Upptåget increased (decreased) the probability of locating in the treatment corridor for unemployed in-migrants as compared to the probability for employed in-migrants, *ceteris paribus*. A positive (negative) γ_1 value, for example corresponding to being born in a non-western country, indicates that the introduction of the commuter train Upptåget increased (decreased) the probability of locating in the treatment corridor for in-migrants born in a non-western country as compared to the probability for in-migrants born in Sweden, *ceteris paribus*.[25] In the case of labor income, X_{it} includes two dummy variables, one for high and one for low labor income instead of a dummy for being unemployed (see

[25] The estimation strategy is inspired by McKinnish, Walsh, and White (2010) who use a similar equation to study the influx of people with different characteristics to gentrifying neighborhoods in the US. However, McKinnish et al. (2010) study only one time period.

Appendix for the definition of low, high and middle labor income), and the corresponding coefficients are interpreted accordingly.[26]

Four things can be noted about the estimation of equation (1). First, the interpretation of β_1 and γ_1 as the effect of the commuter train on the relative probability of locating in the treatment corridor for in-migrants with different characteristics hinges on the assumption that without the train the relative probability of locating in the treatment corridor for in-migrants with different characteristics would have been unchanged.

Second, the absolute probability of locating in the treatment corridor for in-migrants with a certain characteristic can very well increase (decrease) at the same time that the probability of locating in the treatment corridor for in-migrants with that characteristic relative to the probability for in-migrants with the baseline characteristic decreases (increases).

Third, the estimates from equation (1) do not indicate which population groups are most likely to move into the treatment corridor. They tell us which types of people are more likely to move into the treatment corridor *conditional* on the fact that they have chosen to move into a locality of the same type as the treatment corridor (i.e., the treatment or control corridors). For example, it is found that after the train introduction, among those choosing to move into the treatment or control corridors, people born in a non-western country were more likely than natives to actually move into the treatment corridor, *ceteris paribus*. However, this does not mean that individuals moving into the treatment corridor were likely to be born in a non-western country. As can be seen from Table 1, most individuals moving into the treatment corridor both before and after the train introduction were in fact born in Sweden. This is because it was overall more likely that people born in Sweden moved into localities of the treatment-control corridor type than that people born in a non-western country moved into this type of localities. In contrast, if an individual born in a non-western country moved into a locality of the treatment-control corridor type after the train introduction, he or she was more likely to have chosen the treatment corridor than a comparable native. Put otherwise, the inflow of people born in a non-western country is a feature associated with areas treated with commuter train access as compared to non-treated areas.

[26] Here it can be noted that the effect of Upptåget on the in-migration patterns of employed and unemployed people cannot be studied through the area in-migration rates of employed and unemployed. Imagine that exactly the same number of employed and unemployed individuals move into the treatment corridor and the control corridors before and after the introduction of Upptåget. Imagine also that commuter train access increases individuals' employment probability so that some people in the treatment corridor who were previously unemployed become employed following the train introduction. In this case, the in-migration rate of employed would decrease in the areas treated with commuter train access while the in-migration rate of unemployed would increase, although the in-migration patterns are unchanged. Similar reasoning applies to the in-migration patterns of people with different labor incomes and education levels.

A fourth thing to note is that although I try to measure the effects on sorting, the estimations of equation (1) are not immune to selection bias. If the composition of the in-migrant sample changes over time with respect to some unobservable characteristic(s) that is of importance for location choice, this will bias the estimates related to the other characteristics.

5.2 Econometric specification for out-migration patterns

To study the effect of the commuter train Upptåget on out-migration patterns, I employ a difference-in-differences-in-differences strategy comparing the probability of moving out of the treatment and control corridors before and after the introduction of Upptåget for individuals with different characteristics. Using a linear probability model and the sample of stayers and out-migrants described in Section 4.2, I estimate an equation that in its richest specification takes the following form:

$$\begin{aligned} OUT_{it} = {} & \alpha + \beta_0 X_{it-1} + \gamma_0' Z_{it-1} + \xi R_{it-1} + \varphi T + \delta D + \beta_1 (X_{it-1} \times T) + \\ & \gamma_1' (Z_{it-1} \times T) + \beta_2 (X_{it-1} \times D) + \gamma_2' (Z_{it-1} \times D) + \varphi (T \times D) + \\ & \beta_3 (X_{it-1} \times T \times D) + \gamma_3' (Z_{it-1} \times T \times D) + \varepsilon_{it-1} \end{aligned} \quad (2)$$

In equation 2, the dependent variable, OUT_{it} is a dummy variable that equals one if individual i moved out of the corridor where he or she lived in year t and zero otherwise. In the case of employment status, X_{it-1} is a dummy variable that equals one if individual i is unemployed in year $t-1$ and zero otherwise (see Appendix for the definition of employed and unemployed). $Zit-1$ is a vector of individual characteristics other than employment status that might affect the probability of moving out. $Zit-1$ includes dummy variables corresponding to the age groups, education levels, and birth regions that are listed Table 2, as well as a dummy variable for sex. The omitted, baseline, characteristics are the first group within each category in Table 2, i.e., being aged 18-29 years, having at most secondary education, being male, and being born in Sweden. T is a dummy variable that equals one if individual i lived in the treatment corridor in year $t-1$ and zero otherwise. D is a dummy variable that switches from zero to one at the introduction of Upptåget. Finally, since it has been found that subsequent migration is common among recently arrived refugees, whether or not placed by the placement program mentioned in Section 5.1 (see e.g., Åslund, 2000), R_{it-1} is a dummy variable that takes the value one if individual i appeared in the data in year $t-1$, i.e., was missing from the register data in year $t-2$, and was born in a non-western country, i.e., potentially a refugee, and zero otherwise.[27]

[27] The register data I have access to starts with the year 1985. For individuals living in the treatment and control corridors in 1985, i.e., staying or moving out in 1986, I do not know

In equation 2, α hence measures the baseline probability that an individual moved out of the corridor where he or she lived in year *t-1*. β_0 measures whether this probability was different for the unemployed, the coefficients in γ_0 whether this probability was different for inhabitants with the corresponding characteristics, and ξ whether this probability was different for newly arrived refugees. The first four terms in equation (2) thus control for differences in the probability of moving out between individuals with different characteristics that already existed before the train introduction. *T* controls for differences in probability of moving out of the treatment and control corridors that already existed before the train introduction, and *D* controls for change in the probability of moving out between the period before and after the train introduction unrelated to the train. The interaction terms $(X_{it-1} \times T)$ and $(Z_{it-1} \times T)$ control for differences in differences in the probability of moving out between individuals with different characteristics in the treatment and control corridors that already existed before the train introduction. The interaction terms $(X_{it-1} \times D)$ and $(Z_{it-1} \times D)$ control for changes in differences in the probability of moving out between individuals with different characteristics between the period before and the period after the train introduction unrelated to the train.

The coefficients of primary interest are φ, β_3 and the coefficients in γ_3. Given the control variables explained above, φ should measure the effect of the commuter train on the baseline probability of moving out of the treatment corridor. β_3 should measure the effect of the commuter train on the probability of moving out of the treatment corridor for unemployed individuals, beyond the effect for employed individuals. The coefficients in γ_3 should measure the effect of the commuter train on the probability of moving out of the treatment corridor for individuals with the corresponding characteristics, beyond the effect for individuals with the baseline characteristics. A positive (negative) β_3 thus indicates that the introduction of the commuter train Upptåget increased (decreased) the probability of moving out of the treatment corridor for unemployed inhabitants as compared to the probability for employed inhabitants, *ceteris paribus*. A positive (negative) γ_3 value, corresponding, for example to being born in a non-western country, indicates that the introduction of the commuter train Upptåget increased (decreased) the probability of moving out of the treatment corridor for inhabitants born in a non-western country as compared to the probability for inhabitants born in Sweden, *ceteris paribus*.[28] In the case of labor income, X_{it} includes two

whether they are in the register data 1984 or not. In the baseline analysis, I have ascribed a R_i value of zero to the few persons born in a non-western country that were in data 1985, i.e., they are not considered as placed refugees. When looking closer at time dynamics, section 6.2, I do not study who stayed and who moved out in year 1986, but start with year 1987.

[28] The estimation strategy is inspired by Edmark (2009) who uses a similar strategy to estimate the migration response to the implementation of stricter rules for receiving welfare benefits in Stockholm's town districts. Edmark (2009), however, compares a welfare-prone

dummy variables, one for high and one for low labor income instead of a dummy for being unemployed (see Appendix for the definition of low, high, and middle labor income), and the corresponding coefficients are interpreted accordingly.

Three things can be noted about the estimation of equation 2, similar to those that were noted about the estimation of equation 1. First, the interpretation of φ, β_3, and the coefficients in γ_3 as the effect of the commuter train on the difference in probability of moving out from the treatment corridor for people with different characteristics hinges on the assumption that without the train, the difference in probability of moving out of the treatment corridor for people with different characteristics would have followed the development in the control corridors.

Second, the absolute probability that people with a certain characteristic move out of the treatment corridor can very well increase (decrease) at the same time as their relative probability of moving out as compared to the probability for people with the baseline characteristics decrease (increases).

A third thing to note is that estimations of equation 2 are not immune to selection bias. If the population composition in the treatment and/or control corridors changes over time with respect to some unobservable characteristic(s) that is of importance for the decision whether to stay or move out, this will bias the estimates related to the other characteristics. A change in population composition with respect to some unobservable characteristic(s) can result from in-migration, but also come about from people staying. Since the samples "stayers and out-migrants" are pooled cross-sections of the individuals in the treatment and control areas 1985-1995, the people who stay from one year to the next will account for more than one observation.[29]

group and a non-welfare-prone group under the assumption that the non-welfare-prone group will not be affected by the stricter rules and will thus provide a valid counterfactual for the welfare-prone group. The coefficient corresponding to β_3 in her analysis is thus interpreted as the effect of the implementation of the stricter rule on the welfare-prone group. In the context of the present paper it is possible that both the employed and unemployed are affected by the introduction of the commuter train. β_3 should therefore not be interpreted as the effect of the commuter train on the probability of the unemployed to move out, but as the effect on the difference in probability of moving out between the employed and unemployed.

[29] One way to handle the composition bias would be to include individual fixed effects in equation 2. However, if individual fixed effects are included, only the time variation within the observations for each individual is used in the identification. Estimation results will therefore be driven by individuals who are present in the sample both before and after the introduction of the commuter train and who lived in both the treatment and the control area in at least one of the periods. These individuals are likely to be very few and not very representative.

6. Results

In this section I present the results from the analysis of the effect of the introduction of Upptåget on the migration patterns of people with different characteristics in the areas treated with commuter train access. The baseline results are presented in Section 6.1. Section 6.2 contains sensitivity analysis with respect to time dynamics and Section 6.3 sensitivity analysis with respect to treatment intensity/closeness to Uppsala city.

6.1 Baseline results

6.1.1 In-migration

Table 3 and Table 4 present the results from the analysis of how the introduction of the commuter train Upptåget affected the in-migration of people with different characteristics to the areas treated with commuter train access, i.e., from the estimation of equation 1 on the sample of in-migrants. Table 3 shows the results with respect to employment status and Table 4 the results with respect to labor income. The first column in each table presents the overall effect of the train introduction on in-migration, i.e., from the estimation of equation 1 without any individual characteristics (Column 1 is thus the same in the two tables). Column 2 presents the results from the estimation of equation 1 taking only employment status/labor income into account. Column 3 presents the results obtained when controls for sex, age, education, and birth region are added. Column 4, finally, presents the results when the dummies for sex, age, education level, and birth region are also interacted with the dummy for the commuter train introduction (the full equation 1).

Starting with Table 3, the results presented in column 1 indicate that the introduction of the commuter train had a positive effect on overall in-migration to the treatment corridor. Before the introduction of the commuter train, the probability of locating in the treatment corridor for an in-migrant was 42 %. Put otherwise, about 28 % fewer people moved into the treatment corridor than into the control corridors [((0.42-(1-0.42))/(1-0.42))*100]. After the train introduction, the probability of locating in the treatment corridor for an in-migrant increased by 4.3 percentage points, statistically significant at the 1 % significance level. That is, about 15 % fewer people now moved into the treatment corridor than into the control corridors [(((0.42+0.043)-(1-(0.42+0.043))/(1-(0.42+0.043))*100].

Further, the results presented in column 2 indicate that the effect of the train was similar for employed and unemployed in-migrants. Before the train introduction, the probability of locating in the treatment corridor for employed in-migrants was 41.5 % and the probability was not significantly different for unemployed in-migrants. After the train introduction, the probability of locating in the treatment corridor increased by 3.8 percentage points for employed in-migrants, statistically significant at the 1 % signifi-

cance level. The increase was, however, not significantly different for unemployed in-migrants. Controlling for other individual characteristics does not change this conclusion (column 3), and neither does interacting the other individual characteristics with treatment (column 4).

Moreover, the results presented in column 4 indicate that the introduction of the commuter train had a larger effect on the probability of locating in the treatment corridor for in-migrants born in a non-western country than for in-migrants born in Sweden. Before the introduction of the commuter train, the probability of locating in the treatment corridor for in-migrants born in Sweden with the baseline characteristics was 41 %, and the probability was not significantly different for similar in-migrants born in a non-western country. After the introduction of Upptåget, the probability of locating in the treatment corridor increased by 3.7 percentage points for in-migrants born in Sweden, statistically significant at the 1 % significance level. The increase was however much larger – 24.8 percentage points larger, statistically significant at the 1 % significance level – for similar in-migrants born in a non-western country;. That is, considering people with the baseline characteristics, before the introduction of Upptåget about 31 % fewer people born in Sweden moved into the treatment corridor than into the control corridors [((0.41-(1-0.41))/(1-0.41))*100] and also about 31 % fewer people born in a non-western country moved into to the treatment corridor than into the control corridors. After the introduction of the commuter train, about 19 % fewer people born in Sweden moved into the treatment corridor than into the control corridors [(((0.41+0.037)-(1-(0.41+0.037)))/(1-(0.41+0.037)))*100] while about 127 % *more* people born in a non-western country moved into the treatment corridor than into the control corridors. [(((0.41+0.037+0.248)-(1-(0.41+0.037+0.248)))/(1-(0.41+0.037+0.248)))*100]

Finally, the results presented in column 4 suggest that the introduction of the commuter train had a larger effect on the probability of locating in the treatment corridor for in-migrants with at most 10 years of education than for in-migrants with secondary education. The probability of locating in the treatment corridor increased by 3.5 percentage points more for in-migrants with at most 10 years of education than for in-migrants with secondary education, statistically significant at the 10 % significance level.

Table 3. Sample of in-migrants. Probability of locating in the treatment corridor. Employment status.

	(1)	(2)	(3)	(4)
unemployed		0.025	0.023	0.030
		(0.018)	(0.018)	(0.019)
female			0.001	0.008
			(0.008)	(0.014)
western			-0.013	0.005
			(0.017)	(0.029)
non-western			0.169***	-0.040
			(0.019)	(0.047)
aged 30-39			0.017*	0.036**
			(0.010)	(0.017)
aged 40-49			-0.016	-0.030
			(0.012)	(0.021)
aged 50-64			-0.036***	-0.018
			(0.014)	(0.025)
education =< 10 years			-0.012	-0.035**
			(0.010)	(0.017)
tertiary education			0.018	0.022
			(0.011)	(0.020)
education info missing			0.019	0.046
			(0.021)	(0.032)
post	0.043***	0.038***	0.034***	0.037**
	(0.009)	(0.010)	(0.010)	(0.018)
unemployed#post		0.007	-0.002	-0.013
		(0.021)	(0.021)	(0.022)
female#post				-0.010
				(0.017)
western#post				-0.032
				(0.036)
non-western#post				0.248***
				(0.052)
aged 30-39#post				-0.027
				(0.021)
aged 40-49#post				0.020
				(0.025)
aged 50-64#post				-0.026
				(0.030)
education =< 10 years#post				0.035*
				(0.021)
tertiary education#post				-0.007
				(0.024)
education info missing#post				-0.039
				(0.042)
Constant	0.420***	0.415***	0.412***	0.410***
	(0.007)	(0.008)	(0.010)	(0.014)
Observations	15,280	15,280	15,280	15,280
R-squared	0.002	0.002	0.011	0.013

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Looking next at Table 4, the results presented in column 2 indicate that the effect of the commuter train was similar for in-migrants with different labor incomes. Before the introduction of the commuter train, the probability of locating in the treatment corridor was not significantly different for in-migrants with low labor income than for in-migrants with middle labor income. However, the probability of locating in the treatment corridor for in-migrants with high labor income was 5.9 percentage points lower than for in-migrants with middle labor income, statistically significant at the 5 % significance level. After the introduction of Upptåget, the probability of locating in the treatment corridor increased by 3.7 % points for in-migrants with middle labor income, statistically significant at the 1 % significance level. However, this increase was not significantly different for in-migrants with low or high labor income. Controlling for other individual characteristics does not change this conclusion (column 3), and neither does interacting the other individual characteristics with treatment (column 4).

The results presented in column 4 further show that also when controlling for any effect working through labor income, the commuter train seems to have had a much larger effect on the probability of locating in the treatment corridor for in-migrants born in a non-western country than for in-migrants born in Sweden. The size of the estimate, 24.3 percentage points, is very similar to the estimate when controlling for employment status and also statistically significant at the 1 % significance level.

Finally, with respect to in-migrants with at most 10 years of education, the point estimate is similar to the estimate when controlling for employment status, but no longer significant at least at the 10 % significance level.

Table 4. Sample of in-migrants. Probability of locating in the treatment corridor. Labor income.

	(1)	(2)	(3)	(4)
low income		0.019	0.020	0.024
		(0.018)	(0.018)	(0.018)
high income		-0.050**	-0.059***	-0.061***
		(0.021)	(0.022)	(0.023)
female			-0.009	-0.005
			(0.008)	(0.015)
western			-0.015	0.004
			(0.017)	(0.029)
non-western			0.163***	-0.040
			(0.019)	(0.047)
aged 30-39			0.020**	0.040**
			(0.010)	(0.017)
aged 40-49			-0.009	-0.018
			(0.012)	(0.021)
aged 50-64			-0.033**	-0.012
			(0.014)	(0.025)
education =< 10 years			-0.015	-0.037**
			(0.010)	(0.017)
tertiary education			0.026**	0.030
			(0.011)	(0.020)
education info missing			0.014	0.041
			(0.021)	(0.032)
post	0.043***	0.037***	0.032***	0.035*
	(0.009)	(0.010)	(0.010)	(0.018)
low income#post		0.022	0.009	-0.000
		(0.021)	(0.021)	(0.022)
high income#post		0.004	0.006	0.010
		(0.027)	(0.027)	(0.029)
female#post				-0.007
				(0.018)
western#post				-0.033
				(0.036)
non-western#post				0.243***
				(0.052)
aged 30-39#post				-0.029
				(0.021)
aged 40-49#post				0.015
				(0.026)
aged 50-64#post				-0.030
				(0.030)
education =< 10 years#post				0.034
				(0.021)
tertiary education#post				-0.008
				(0.024)
education info missing#post				-0.040
				(0.042)
Constant	0.420***	0.422***	0.422***	0.420***
	(0.007)	(0.008)	(0.011)	(0.015)

Observations	15,280	15,280	15,280	15,280
R-squared	0.002	0.004	0.012	0.014

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In sum, the results presented in Table 3 and Table 4 suggest that the introduction of the commuter train Upptåget had a positive effect on overall in-migration to the areas treated with commuter train access. Further, the train introduction does not seem to have affected the relative in-migration of people with different employment statuses or labor incomes to the treated areas. However, whether controlling for any effect working through employment status or labor income, the results in Table 3 and Table 4 suggest that the introduction of the commuter train Upptåget caused people born in non-western countries to be particularly attracted towards the areas served as compared to other similar areas.

6.1.2 Out-migration

Table 5 and Table 6 present the results from the analysis of whether the introduction of the commuter train Upptåget had the same effect on the probability of moving out of the areas treated with commuter train access for people with different characteristics, i.e., from the estimation of equation 2 on the sample of stayers and out-migrants. Table 5 shows the results with respect to employment status and Table 6 the results with respect to labor income. The first column in each table presents the overall effect of the train introduction on out-migration from the treatment corridor, i.e., from the estimation of equation 2 without any individual characteristics (column 1 is thus the same in the two tables). Column 2 presents the results from estimation of equation 2 taking only employment status/labor income into account. In column 3, controls for sex, age, education level, and birth region are added as well as a dymmy for being born in a non-western country and just arrived (i.e. potentially a refugee). Finally, in column 4, the dummies for sex, age, education level, and birth region are interacted with the dummy for living in the treatment corridor, the dummy for the train introduction (the full equation 4) and both dummies at the same time. To save on space, only the coefficients of primary interest are presented, i.e., the coefficient that measures the effect of Upptåget on the probability of moving out of the treatment corridor for people with the baseline characteristics (φ) and the coefficients that measure whether this effect was different for people with other characteristics (β_3 and the coefficients in γ_3).

Starting with Table 5, the results presented in column 1 indicate that the introduction of the commuter train had no significant effect on the probability of moving out of the treatment corridor. The results presented in column 2 indicate that the introduction of Upptåget increased the probability of moving out of the treatment corridor for unemployed people relative to the probability for employed people. The probability of moving out increased

by 3.1 percentage points more for unemployed than for employed, significant at the 5 % level. However, controlling for other characteristics than employment status (column 3 and column 4), this point estimate becomes much smaller and no longer significant at least at the 10 % level.

The results presented in column 4 further indicate that controlling for any effect working through employment status, the introduction of the commuter train Upptåget did not have a significantly different effect on the probability of moving out of the areas treated with commuter train access for people with different sex, age, education level, or birth region.

Table 5. Sample of stayers and outmigrants. Probabilty of moving out. Employment status.

	(1)	(2)	(3)	(4)
treatment_corridor#post	0.008	0.002	0.001	0.003
	(0.005)	(0.003)	(0.003)	(0.009)
unemployed#treatment_corridor#post		0.031**	0.014	0.012
		(0.013)	(0.010)	(0.008)
female# treatment_corridor#post				0.004
				(0.003)
western# treatment_corridor#post				0.001
				(0.009)
non-western# treatment_corridor#post				0.030
				(0.030)
aged 30-39# treatment_corridor#post				-0.008
				(0.009)
aged 40-49# treatment_corridor#post				-0.002
				(0.008)
aged 50-64# treatment_corridor#post				-0.004
				(0.008)
education =< 10 years # treatment_corridor#post				-0.001
				(0.004)
tertiary education# treatment_corridor#post				0.005
				(0.007)
education info missing # treatment_corridor#post				0.003
				(0.036)
Dummy for treatment_corridor	yes	yes	yes	yes
Dummy for post	yes	yes	yes	yes
Dummy for being unemployed	no	yes	yes	yes
Dummies for being unemployed #treatment_corridor	no	yes	yes	yes
Dummies for being unemployed #post	no	yes	yes	yes
Dummies for sex, age group, education level, and birth region	no	no	yes	yes
Dummy for just arrived non-western	no	no	yes	yes
(Dummies for sex, age group, education level, and birth region)#treatment_corridor	no	no	no	yes
(Dummies for sex, age group, education level, and birth region)#post	no	no	no	yes
Constant	0.056***	0.050***	0.120***	0.110***
	(0.005)	(0.004)	(0.006)	(0.008)
Observations	268,378	268,378	268,378	268,378
R-squared	0.001	0.009	0.052	0.053

Standard errors clustered on SAMS in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Looking next at Table 6, the results presented in column 2 indicate that the introduction of Upptåget did not have a significantly different effect on the probability of moving out of the areas treated with commuter train access for people with different labor incomes. Controlling for characteristics other than labor income (column 3), does not change this conclusion. However, the results from the estimation with full interactions, presented in column 4, indicate that the introduction of the commuter train Upptåget increased the probability of moving out of the treatment corridor for people with high labor income relative to the probability for people with middle income. The probability of moving out increased by 0.9 percentage points more for people with high labor income than for people with middle labor income, statistically significant at the 5 % significance level. The results presented in column 4 further indicate that the introduction of Upptåget increased the probability of moving out of the areas treated with commuter train access by 0.7 percentage points more for women than for men, statistically significant at the 5 % significance level.

Table 6. Sample of stayers and outmigrants. Probabilty of moving out. Labor income.

	(1)	(2)	(3)	(4)
treatment_corridor#post	0.008	0.004	0.001	0.003
	(0.005)	(0.004)	(0.003)	(0.009)
low income#treatment_corridor#post		0.019	0.004	0.003
		(0.014)	(0.009)	(0.008)
high income #treatment_corridor#post		0.003	0.006	0.009**
		(0.005)	(0.004)	(0.004)
female# treatment_corridor#post				0.007**
				(0.003)
western# treatment_corridor#post				0.002
				(0.009)
non-western# treatment_corridor#post				0.034
				(0.031)
aged 30-39# treatment_corridor#post				-0.011
				(0.009)
aged 40-49# treatment_corridor#post				-0.005
				(0.009)
aged 50-64# treatment_corridor#post				-0.006
				(0.009)
education =< 10 years # treatment_corridor#post				0.000
				(0.003)
tertiary education# treatment_corridor#post				0.003
				(0.007)
education info missing # treatment_corridor#post				0.006
				(0.036)
Dummy for treatment_corridor	yes	yes	yes	yes
Dummy for post	yes	yes	yes	yes
Dummies for labor income	no	yes	yes	yes
Dummies for labor income #treatment_corridor	no	yes	yes	yes
Dummies for labor income #post	no	yes	yes	yes
Dummies for sex, age group, education level, and birth region	no	no	yes	yes
Dummy for just arrived non-western	no	no	yes	yes
(Dummies for sex, age group, education level, and birth region)#treatment_corridor	no	no	no	yes
(Dummies for sex, age group, education level, and birth region)#post	no	no	no	yes
Constant	0.056***	0.054***	0.123***	0.113***
	(0.005)	(0.005)	(0.006)	(0.008)
Observations	268,378	268,378	268,378	268,378
R-squared	0.001	0.006	0.051	0.052

Standard errors clustered on SAMS in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

In sum, the results presented in Table 5 and Table 6 suggest that the introduction of the commuter train Upptåget had no effect on the overall out-migration from the areas treated with commuter train access. Based on the richest specifications, the train seems to have had the same (zero) effect on the probability of moving out of the areas treated with commuter train access for employed and unemployed people but to have increased the probability of moving out of the areas served for individuals with high labor incomes relative to the probability for individuals with lower income. Controlling for any effect working through labor income, the commuter train also seems to have increased the relative probability of moving out for women as compared to men.

6.2 Time dynamics

This study covers a rather long time period and a dynamic phenomenon; migration patterns. It is therefore unlikely that no change in migration patterns would have taken place in the absence of the introduction of the commuter train Upptåget. To see if the effects found in the previous sections fall within normal variations/whether there are effects preceding the train introduction, placebo tests where it is pretended that the treatment arrived at an earlier point in time would have been useful. However, since the discussions about Upptåget started in 1988 and I only have information about in- and out-migration from 1986 onwards, there is little room to conduct such a placebo test. To get an idea of whether there are effects preceding the train introduction, and about development over time after the train introduction, I instead conduct separate estimations for each year 1986-1996.

For in-migration, for each year, *t*, I thus estimate

$$IN_{it} = \alpha + \beta_0 X_{it-1} + \gamma_0' Z_{it-1} + \varepsilon_{it-1} \tag{3}$$

using the sample of in-migrants and a linear probability model. I have then graphed β_0 and the coefficients in γ_0 for each year, i.e., the differences in probability of locating in the treatment area between in-migrants with the corresponding characteristics and in-migrants with the baseline characteristics, along with the 95 % confidence interval.

For out-migration, for each year, *t*, I estimate

$$\begin{aligned} OUT_{it} = {} & \alpha + \beta_0 X_{it-1} + \gamma_0' Z_{it-1} + \xi R_{it-1} + \varphi T + \\ & \beta_1 (X_{it-1} \times T) + \gamma_1' (Z_{it-1} \times T) + \varepsilon_{it-1} \end{aligned} \tag{4}$$

using the sample of stayers and out-migrants and a linear probability model. I have then graphed β_1 and the coefficients in γ_1 for each year, i.e., the differ-

ence in probability of moving out of the treatment corridor between inhabitants with the corresponding characteristics and inhabitants with the baseline characteristics, beyond any differences in probability to move out between people with different characteristics that also existed in the control corridors, along with the 95 % confidence interval. Equation 3 and equation 4 are estimated with X_{it-1} being in turn dummies for employment status and labor income.

Below I show the graphs with respect to the characteristics in focus – employment status and labor income – and the cases with respect to the other characteristics where there seem to be systematic, statistically significant, effects over time.[30]

6.2.1 In-migration

The results presented in *Figure 3* and *Figure 4* confirm the findings from the baseline analysis, that the effect of Upptåget on the probability of locating in the treatment corridor was similar for employed and unemployed in-migrants and for in-migrants with different labor income: In *Figure 3*, the difference in probability of locating in the treatment corridor between unemployed and employed in-migrants evolves around zero for the whole period 1986-1996. In *Figure 4*, the left panel, the difference in probability of locating in the treatment corridor between in-migrants with low and middle labor income also evolves around zero. In *Figure 4*, the right panel, the point estimates for the difference in probability of locating in the treatment corridor between in-migrants with high and middle labor income are below zero for most years although only statistically so in 1992 and 1996 and there is no systematic time variation that can be related to the timing of Upptåget.

Further, from *Figure 5* it can be seen that the difference in probability of locating in the treatment corridor between in-migrants born in a non-western country and in Sweden was not significantly different from zero up to and including 1989. For all years 1990-1995, i.e., after the introduction of Upptåget, the difference is, however, well above zero, statistically significant at least at the 5% significance level.

[30] The other graphs are available on request.

Figure 3. Sample of in-migrants. Difference in probability of locating in the treatment corridor between unemployed and employed in-migrants.

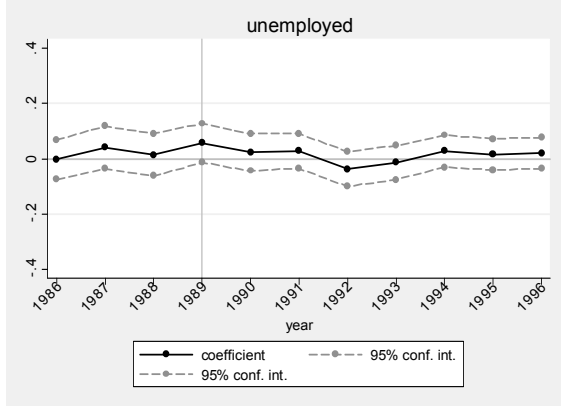


Figure 4. Sample of in-migrants. Left panel: Difference in probability of locating in the treatment corridor between in-migrants with low and middle labor income. Right panel: Difference in probability of locating in the treatment corridor between in-migrants with high and middle labor income.

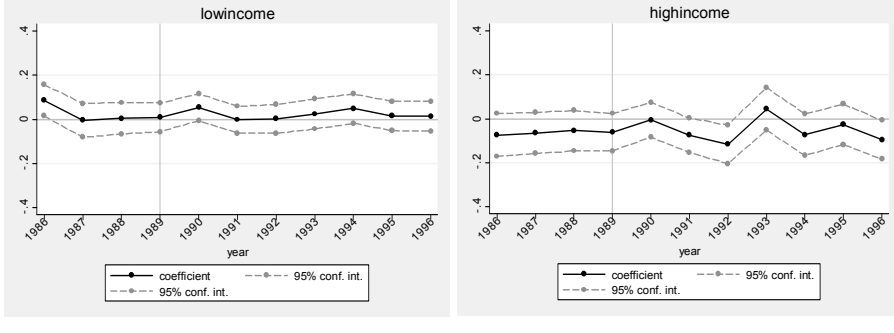
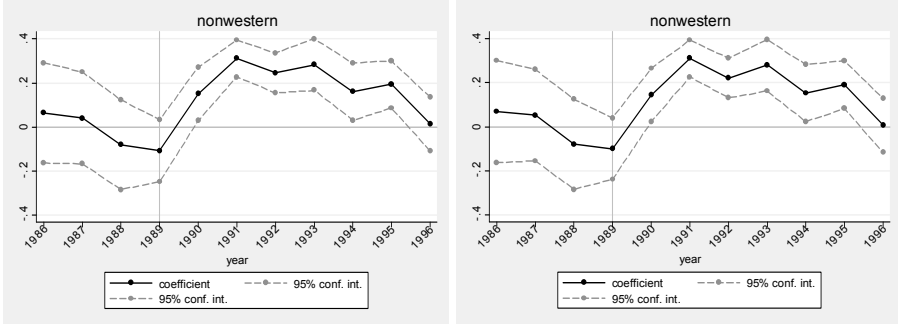


Figure 5. Sample of in-migrants. Difference in probability of locating in the treatment corridor between in-migrants born in a non-western country and in-migrants born in Sweden. The left part of the figure is from estimation of equation 3 with X_{it-1} being a dummy for being unemployed and the right part of the figure is from estimations of equation 3 with X_{it-1} being dummies for low and high labor income.



6.2.2 Out-migration

The results presented in *Figure 6* confirm the finding from the baseline analysis that Upptåget had the same (zero) effect on the probability of moving out of the areas treated with commuter train access for employed and unemployed people: there is no systematic variation over time that seems to be related to the timing of Upptåget or preceding it. On the other hand, the results presented in *Figure 7* cast some doubts on the finding from the baseline analysis that the introduction of Upptåget increased the relative probability of moving out of the treatment corridor for people with high labor income as compared to people with middle income. The difference in probability of moving out of the treatment corridor between inhabitants with high and middle labor income, beyond any differences that also existed in the control corridors, is never significantly different from zero, at least at the 5 % significance level. The same is the case when comparing men and women (graphs not shown).

Figure 6. Sample of stayers and out-migrants. Difference in probability of moving out of the treatment corridor between unemployed and employed inhabitants, beyond any difference that also existed in the control corridors.

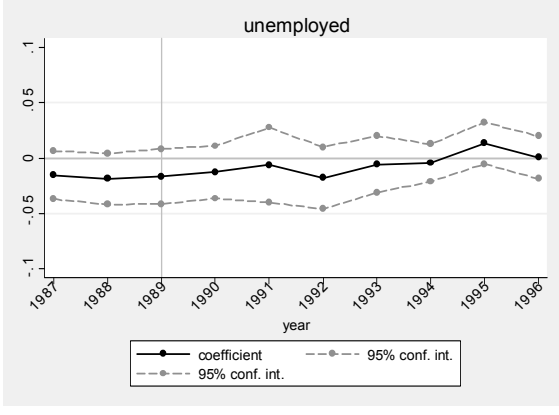
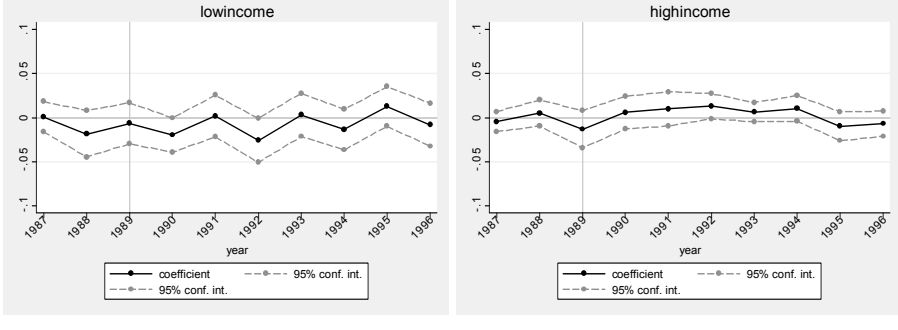


Figure 7. Sample of stayers and out-migrants. Left panel: Difference in probability of moving out of the treatment corridor between inhabitants with low and middle income, beyond any difference that also existed in the control corridors. Right panel: Difference in probability of moving out of the treatment corridor between inhabitants with high and middle income, beyond any difference that also existed in the control corridors.



6.3 Treatment intensity/Closeness to Uppsala city

The commuter train Upptåget improved commuting opportunities by public transit from all stations in the treatment corridor. However, while the train was substantially faster than bus connections, the decrease in travel time by public transit was not the same everywhere in the treatment corridor. For the train stations between Tierp and Uppsala, travel times by public transit to Uppsala were approximately cut in half, which in 1996 represented an absolute gain of between 17 and 35 minutes depending on station. From Tierp, on the other hand, Upptåget decreased travel time to Uppsala city by less than

30 % compared to bus in 1992 and by less than 40 % in 1996. Furthermore, the fastest public transit between Tierp and Uppsala was, both before and after the introduction of Upptåget, a long distance train that took about 40 minutes. It can therefore be argued that Upptåget had less of an effect on travel time by public transit between Tierp and Uppsala, even though it indeed increased the number of fast connections.

Likewise, it can be argued that Storvreta, which is the Upptåget station closest to Uppsala, might have been less intensively treated than the other stations. As explained in Section 3.1, the actual difference in travel time using bus or the commuter train Upptåget should have been less important.

Furthermore, the hypotheses from the model of LeRoy and Sonstelie (1983) is that the commuter train could have the effect of stretching the ring where low-income people use public transit further out along the commuter train line. However, it is possible that despite the commuter train, travel times by public transit from the Upptåget stations further away from Uppsala city were still too long to attract people depending on public transit.

In this section I therefore examine whether the baseline results are sensitive to excluding either Tierp or Storvreta from the analysis.

6.3.1 In-migration

The results with respect to in-migration, i.e., the estimation of equation 1 on the reduced in-migrants sample, are presented in Table 7 (with respect to employment status) and Table 8 (with respect to labor income). The left panel in each table shows the results when excluding in-migrants to Tierp and corresponding parts of the control corridors (i.e., to SAMS with a population center further than 46,000 meters from Uppsala city), and the right panel the results when excluding in-migrants to Storvreta and corresponding parts of the control corridors (i.e., to SAMS with a population center closer than 12,500 meters to Uppsala city). In each panel the columns correspond to column 1, column 2 and column 4 in Table 3 and Table 4. To save on space, only the coefficients of primary interest are reported, i.e., δ, β_1 and the coefficients in γ_1.

From Table 7 and Table 8 it can be seen that the overall effect of Upptåget on in-migration to the areas treated with commuter train access was larger closer to Uppsala city: Excluding in-migrants to Tierp (Storvreta) and corresponding parts of the control corridors, the introduction of Upptåget is related to a 5 (2.2) percentage point increase in the probability of locating in the treatment corridor for in-migrants, statistically significant at the 1 % (5 %) significance level.

The results presented in Table 7 and Table 8 further indicate that similar to what was found in the baseline analysis, the introduction of Upptåget had the same effect on the probability of locating in the treatment corridor for employed and unemployed in-migrants and for in-migrants with different labor income whether or not in-migrants to Tierp or Storvreta and the corre-

sponding parts of the control corridors are excluded. The results presented in Table 7 and Table 8 also confirm the findings from the baseline analysis, that the introduction of the commuter train had a larger effect on the probability of locating in the treatment corridor for in-migrants born in a non-western country than for in-migrants born in Sweden, but also indicate that the difference was larger closer to Uppsala city, i.e., when excluding in-migrants to Tierp and the corresponding control areas. On the other hand, the relative increase in the probability of locating in the treatment area for in-migrants with at most 10 years of education following the introduction of Upptåget found in the baseline analysis seems to be tied to the stations further from Uppsala city (i.e., when excluding in-migrants from Storvreta and the corresponding parts of the control corridors).

Finally, the results presented in Table 7 and Table 8, provide some evidence that the introduction of the commuter train had the effect to decrease the probability of locating in the treatment corridor for in-migrants aged 30-39 relative to the probability for in-migrants aged 18-29, both closer and further away from Uppsala city.

Table 7. Sample of in-migrants. Probability of locating in the treatment corridor. Employment status

	Without Tierp			Without Storvreta		
	(1)	(2)	(3)	(1)	(2)	(3)
post	0.050***	0.044***	0.054***	0.022**	0.018*	0.011
	(0.010)	(0.011)	(0.020)	(0.009)	(0.010)	(0.019)
unemployed#post		0.022	-0.011		0.005	-0.006
		(0.024)	(0.025)		(0.023)	(0.024)
female#post			-0.011			-0.007
			(0.019)			(0.019)
western#post			-0.025			-0.012
			(0.039)			(0.039)
non-western#post			0.309***			0.145**
			(0.057)			(0.060)
aged 30-39#post			-0.038*			-0.043*
			(0.023)			(0.023)
aged 40-49#post			-0.007			0.024
			(0.029)			(0.027)
aged 50-64#post			0.000			-0.029
			(0.033)			(0.032)
education =< 10 years#post			0.021			0.066***
			(0.023)			(0.022)
tertiary education#post			-0.036			0.016
			(0.027)			(0.027)
education info missing#post			-0.003			-0.031
			(0.047)			(0.048)
Dummy for employment status	no	yes	yes	no	yes	yes
Dummies for sex, age group, education level, and birth region	no	no	yes	no	no	yes
Constant	0.373***	0.373***	0.350***	0.344***	0.341***	0.352***
	(0.008)	(0.009)	(0.016)	(0.007)	(0.008)	(0.015)
Observations	11,889	11,889	11,889	11,839	11,839	11,839
R-squared	0.002	0.002	0.021	0.000	0.001	0.004

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8. Sample of in-migrants. Probability of locating in the treatment corridor. Labor income.

	Withouth Tierp			Without Storvreta		
	(1)	(2)	(3)	(1)	(2)	(3)
post	0.050***	0.048***	0.057***	0.022**	0.014	0.007
	(0.010)	(0.012)	(0.020)	(0.009)	(0.011)	(0.020)
low income #post		0.021	-0.013		0.022	0.011
		(0.024)	(0.025)		(0.023)	(0.024)
high income #post		-0.032	-0.024		0.016	0.021
		(0.030)	(0.032)		(0.028)	(0.031)
female#post			-0.012			-0.003
			(0.020)			(0.019)
western#post			-0.027			-0.012
			(0.039)			(0.039)
non-western#post			0.304***			0.141**
			(0.057)			(0.059)
aged 30-39#post			-0.036			-0.046**
			(0.023)			(0.023)
aged 40-49#post			-0.005			0.016
			(0.029)			(0.027)
aged 50-64#post			0.005			-0.036
			(0.034)			(0.032)
education =< 10 years#post			0.019			0.066***
			(0.023)			(0.022)
tertiary education#post			-0.033			0.014
			(0.027)			(0.027)
education info missing#post			-0.006			-0.031
			(0.047)			(0.048)
Dummies for labor income	no	yes	yes	no	yes	yes
Dummies for sex, age group, education level, and birth region	no	no	yes	no	no	yes
Constant	0.373***	0.369***	0.352***	0.344***	0.356***	0.367***
	(0.008)	(0.009)	(0.017)	(0.007)	(0.009)	(0.016)
Observations	11,889	11,889	11,889	11,839	11,839	11,839
R-squared	0.002	0.003	0.021	0.000	0.004	0.006

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

6.3.2 Out-migration

The results with respect to out-migration, i.e., the estimation of equation 2 on the reduced stayers and out-migrants sample, are presented in Table 9 (with respect to employment status) and Table 10 (with respect to labor income). The left panel in each table shows the results when excluding inhabitants in Tierp and the corresponding parts of the control corridors (i.e., ex-

cluding inhabitants in SAMS with a population center further than 46,000 meters from Uppsala city), and the right panel the results when excluding inhabitants in Storvreta and the corresponding parts of the control corridors (i.e., excluding inhabitants in SAMS with a population center closer than 12,500 meters to Uppsala city). In each panel the columns correspond to column 1, column 2 and column 4 in Table 5 and Table 6. To save on space, only the coefficients of primary interest are reported, i.e., φ, β_3 and the coefficients in γ_3.

Similar to what was found in the baseline analysis, the results presented in Table 9 and Table 10 suggest that the introduction of the commuter train Upptåget had no effect on the overall out-migration from the areas treated with commuter train access, whether or not Tierp or Storvreta and the corresponding control areas are excluded. In estimations where employment status is the only individual characteristic controlled for (Table 9, column 2), it appears that the train increased the probability of moving out of the treatment corridor for unemployed relative to the probability for employed, but taking other individual characteristics into account (Table 9, column 3), the point estimate becomes smaller and no longer significant at least at the 10 % significance level. Considering labor income, based on the richest specifications (Table 9, columns 3), the train seems to have increased the probability of moving out of the treatment corridor by about 1 percentage point more for people with high labor income than for people with lower labor income, but the effect is only statistically significant at least at the 10% significance level closer to Uppsala city, i.e., when excluding Tierp and the corresponding control areas. No other effects statistically significant at least at the 10 % significance level can be found.

Table 9. Sample of stayers and out-migrants. Probability of moving out. Employment status

	Without Tierp			Without Storvreta		
	(1)	(2)	(3)	(1)	(2)	(3)
treatment_corridor#post	0.009	0.003	0.002	0.004	0.001	-0.005
	(0.007)	(0.004)	(0.010)	(0.004)	(0.003)	(0.008)
uneployed#treatment_ corridor#post		0.034*	0.010		0.022*	0.012
		(0.020)	(0.010)		(0.011)	(0.010)
female# treatment_corridor#post			0.004			0.003
			(0.004)			(0.003)
western# treatment_corridor#post			0.007			-0.003
			(0.012)			(0.010)
non-western# treatment_corridor#post			0.016			0.021
			(0.034)			(0.033)
aged 30-39# treatment_corridor#post			-0.006			0.002
			(0.011)			(0.010)
aged 40-49# treatment_corridor#post			-0.002			0.004
			(0.011)			(0.009)
aged 50-64# treatment_corridor#post			-0.007			0.005
			(0.012)			(0.008)
education =< 10 years # treatment_corridor#post			0.003			0.000
			(0.004)			(0.005)
tertiary education# treatment_corridor #post			0.002			0.005
			(0.008)			(0.007)
education info missing # treatment_corridor#post			0.011			-0.027
			(0.042)			(0.034)
Dummy for treatment_corridor	yes	yes	yes	yes	yes	yes
Dummy for post	yes	yes	yes	yes	yes	yes
Dummy for being unemployed	no	yes	yes	no	yes	yes
Dummies for being unemployed #treatment_corridor	no	yes	yes	no	yes	yes
Dummies for being unemployed #post	no	yes	yes	no	yes	yes
Dummies for sex, age group, education level, and birth region	no	no	yes	no	no	yes
Dummy for just arrived non-western	no	no	yes	no	no	yes

(Dummies for sex, age group, education level, and birth region) #treatment_corridor	no	no	yes	no	no	yes
(Dummies for sex, age group, education level, and birth region)#post	no	no	yes	no	no	yes
Constant	0.059***	0.052***	0.118***	0.054***	0.048***	0.107***
	(0.005)	(0.004)	(0.009)	(0.005)	(0.004)	(0.007)
Observations	191,241	191,241	191,241	217,536	217,536	217,536
R-squared	0.001	0.010	0.061	0.000	0.008	0.051

Standard errors clustered on SAMS in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 10. Sample of stayers and out-migrants. Probability of moving out. Labor income

	Without Tierp			Without Storvreta		
	(1)	(2)	(3)	(1)	(2)	(3)
treatment_corridor#post	0.009	0.006	0.002	0.004	0.001	-0.006
	(0.007)	(0.005)	(0.011)	(0.004)	(0.004)	(0.009)
low income#treatment_corridor#post		0.024	0.005		0.011	0.004
		(0.022)	(0.011)		(0.011)	(0.010)
high income#treatment_corridor#post		-0.000	0.009*		0.005	0.008
		(0.007)	(0.005)		(0.006)	(0.005)
female# treatment_corridor#post			0.006			0.005
			(0.004)			(0.004)
western# treatment_corridor#post			0.007			-0.002
			(0.012)			(0.010)
non-western# treatment_corridor#post			0.018			0.026
			(0.035)			(0.034)
aged 30-39# treatment_areas#post			-0.008			0.000
			(0.012)			(0.010)
aged 40-49# treatment_corridor #post			-0.006			0.001
			(0.011)			(0.009)
aged 50-64# treatment_corridor #post			-0.009			0.003
			(0.012)			(0.008)
education =< 10 years # treatment_corridor #post			0.004			0.001
			(0.004)			(0.004)
tertiary education# treatment_corridor #post			0.001			0.003
			(0.009)			(0.008)
education info missing #			0.014			-0.026

treatment_corridor #post			(0.042)			(0.034)
Dummy for treatment_corridor	yes	yes	yes	yes	yes	yes
Dummy for post	yes	yes	yes	yes	yes	yes
Dummies for labor income	no	yes	yes	no	yes	yes
Dummies for labor income #treatment_corridor	no	yes	yes	no	yes	yes
Dummies for labor income #post	no	yes	yes	no	yes	yes
Dummies for sex, age group, education level, and birth region	no	no	yes	no	no	yes
Dummy for just arrived non-western	no	no	yes	no	no	yes
(Dummies for sex, age group, education level, and birth region) #treatment_corridor	no	no	yes	no	no	yes
(Dummies for sex, age group, education level, and birth region)#post	no	no	yes	no	no	yes
Constant	0.059***	0.057***	0.120***	0.054***	0.052***	0.109***
	(0.005)	(0.005)	(0.009)	(0.005)	(0.004)	(0.008)
Observations	191,241	191,241	191,241	217,536	217,536	217,536
R-squared	0.001	0.008	0.060	0.000	0.005	0.050

Standard errors clustered on SAMS in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

7 Conclusion

This paper looks at the role of public transit for residential sorting by studying how the introduction of a commuter train in one part of the Uppsala local labor market (Sweden) affected the migration patterns of people with different characteristics in the areas served. The studied case, *Upptåget*, was introduced in the early 1990s and connected locations north of Uppsala city to the local center and further to the greater Stockholm area. The focus is on whether there is sorting between individuals with different employment statuses and labor incomes, but I also consider sorting based on education level, age, sex, and birth region.

I find that the commuter train had a positive effect on overall in-migration to the areas served and no effect on overall out-migration, as could be expected since most inhabitants should not have experienced adverse income effects following the introduction of the train. With regards to sorting based on labor market status, I find no evidence of sorting based on employment status but some evidence that the train introduction increased the probability of moving out of the areas served for individuals with high labor incomes relative to the probability for individuals with lower income, in particular in

the treated areas closer to Uppsala city. At least in the short run, and similarly to what has been found in previous studies and with the hypotheses from LeRoy and Sonstelie (1983), marginal workers do not seem to be outbid from areas with good public transit, rather it is those with high labor income who leave.

Considering sorting along lines other than labor market status, the analysis suggests that Upptåget caused people born in non-western countries to be particularly attracted towards the areas served by the commuter train as compared to other similar areas, and that this cannot be explained by their employment status or labor income. Excluding the station closest to Uppsala city, it also appears that the commuter train caused people with at most 10 years of education to be particular attracted towards the areas served by the commuter train as compared to other similar areas, which again cannot be explained by employment status or labor income as defined in this study.

A part from effects working through employment status and labor income, the urban economic models presented in Section 2 do not provide explanations for residential sorting based on birth region, age or education level. A possible explanation for the in-migration patterns found could be that wealth is important for the purchase of a car, and low educated and people from non-western countries can be thought to have accumulated less wealth than others with the same current income or employment status making them more dependent on public transit. People from non-western countries could also face a higher fixed material cost for driving a car than others because they not only have to buy a car but for some also obtain a driving license, once again making them more dependent on public transit.[31]

The paper adds to the scarce evidence on the role of public transport in residential sorting. Studies on how other changes in local/regional rail transit infrastructure – in Sweden, elsewhere in Europe, or in other parts of the world – have affected the migration patterns of people with different characteristics would, however, be most welcome. If possible, future studies should include better information on the individual characteristics of those moving into, staying in, or moving out from the areas served, as well as better information on the housing construction and public services that have accompanied the changes in rail transit infrastructure, since this could also affect residential sorting. Studies regarding other types of public transit, for example express buses would also be interesting.

[31] At least one municipality in Sweden has as part of their labor market policy given specially help to refugees to obtain a driving license.

References

Arbetsgruppen Projekt Upptåget. (1988a). *Upptåget - En förstudie*.

Arbetsgruppen Projekt Upptåget. (1988b). *Upptåget - förslag till länspendel Uppsala-Tierp.*

Alonso, W. (1964). *Location and land use: toward a general theory of land rent*: Harvard University Press Cambridge, MA.

Åslund, O. (2000). *Immigrant settlement policies and subsequent migration*. Working Paper 2000:23, Department of Economics, Uppsala University.

Becker, G. S. (1965). A theory of the allocation of time. *Economic journal, 75*(299), 493-517.

Brueckner, J. K., & Rosenthal, S. S. (2009). Gentrification and neighborhood housing cycles: Will America's future downtowns be rich? *Review of Economics and Statistics, 91*(4), 725-743.

Brueckner, J. K., Thisse, J.-F., & Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review, 43*(1), 91-107.

Brueckner, J. K., & Zenou, Y. (2003). Space and unemployment: The Labor market effects of spatial mismatch. *Journal of Labor Economics, 21*(1), 242-262.

Edin, P.-A., Fredriksson, P., & Åslund, O. (2004). Settlement policies and the economic success of immigrants. *Journal of Population Economics, 17*(1), 133-155.

Edmark, K. (2009). Migration effects of welfare benefit reform. *Scandinavian Journal of Economics, 111*(3), 511-526.

Gibbons, S., & Machin, S. (2006). *Transport and labour market linkages: empirical evidence, implications for policy and scope for further UK research*: Commissioned for the Eddington Study.

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics, 63*(1), 1-24.

Gobillon, L., Selod, H., & Zenou, Y. (2007). The mechanisms of spatial mismatch. *Urban Studies, 44*(12), 2401-2427.

Ihlanfeldt, K. R. (2006). A primer on spatial mismatch within urban labor markets. In R. J. Arnott & D. P. McMillen (Eds.), *A Companion to Urban Economics* (pp. 404-417). Boston: Blackwell Publishing.

Ihlanfeldt, K. R., & Sjoquist, D. L. (1998). The spatial mismatch hypothesis: a review of recent studies and their implications for welfare reform. *Housing Policy Debate, 9*(4), 849-892.

Kahn, M. E. (2007). Gentrification trends in new transit oriented communities: Evidence from 14 cities that expanded and built rail transit systems. *Real Estate Economics, 35*(2), 155-182.

Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *Quarterly Journal of Economics, 82*(2), 175-197.

LeRoy, S. F., & Sonstelie, J. (1983). Paradise lost and regained: Transportation innovation, income and residential location. *Journal of Urban Economics, 13*(1), 67-89.

McKinnish, T., Walsh, R., & White, T. K. (2010). Who gentrifies low-income neighborhoods? *Journal of Urban Economics, 67*(2), 180-193.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review, 57*(2), 197-210.

Mortensen, D. T., & Pissarides, C. A. (1999). New developments in models of search in the labor market. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (Vol. 3, pp. 2567-2627). Amsterdam: Elsevier Science.

Muth, R. F. (1969). *Cities and housing*: University of Chicago press.
Pissarides, C. A. (2000). *Equilibrium unemployment theory*. Cambridge, MA: MIT press.
Ross, S. L., & Zenou, Y. (2008). Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory. *Regional Science and Urban Economics, 38*(5), 498-517.
Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review, 74*(3), 433-444.
Smith, T. E., & Zenou, Y. (2003). Spatial mismatch, search effort, and urban spatial structure. *Journal of Urban Economics, 54*(1), 129-156.
SOU. (2003:67). *Kollektivtrafik med människan i centrum*. Stockholm.
Statistics Sweden. (2005). *Geografin i statistiken - regionala indelningar i Sverige.* Meddelanden i samordningsfrågor för Sveriges officiella statistik 2005:2 [*Geography in statistics - regional divisions in Sweden.* Reports on Statistical Coordination for the Official Statistics of Sweden 2005:2]. Örebro.
Taylor, B. D., & Ong, P. M. (1995). Spatial mismatch or automobile mismatch? An examination of race, residence and commuting in US metropolitan areas. *Urban Studies, 32*(9), 1453-1473.
Upplands Lokaltrafik (UL). Time-tables 1985-2008.
Upplands Lokaltrafik (UL). (1988). Minutes board meeting February 11, 1988.
Uppsala County council. (1990). Documents County Council meeting September 24, 1990.
Wasmer, E., & Zenou, Y. (2002). Does city structure affect job search and welfare? *Journal of Urban Economics, 51*(3), 515-541.
Wasmer, E., & Zenou, Y. (2006). Equilibrium search unemployment with explicit spatial frictions. *Labour Economics, 13*(2), 143-165.
Wheaton, W. C. (1977). Income and urban residence: An analysis of consumer demand for location. *American Economic Review, 67*(4), 620-631.
Zenou, Y. (2002). How do firms redline workers? *Journal of Urban Economics, 52*(3), 391-408.
Zenou, Y. (2009). *Urban Labor Economics*. Cambridge: Cambridge University Press.
Zenou, Y., & Smith, T. E. (1995). Efficiency wages, involuntary unemployment and urban spatial structure. *Regional Science and Urban Economics, 25*(4), 547-573.

Appendix: Definition of some variables

Employment: Employment is based on the official annual employment statistics and refers to status during November each year. A person is classified as employed if he or she did payed work for at least one hour/week.

Labor income: Labor income refers to annual earnings from work, including self-employment and employer's income. A person is classified as having low (high) labor income if he or she had a labor income lower (higher) than the 20^{th} (80^{th}) percentile of the population 18-64 years in Uppsala city (i.e., SAMS with SAMS-number 3800000-3800167).

Western countries: Western countries include: Finland, Norway, Denmark, Iceland, Ireland, Great Britain, the Netherlands, Belgium, France, Monaco, Luxembourg, Lichtenstein, Spain, Portugal, Andorra, Germany, Switzerland, Italy, San Marino, Vatican City State, Austria, Greece, Canada, the USA, and the countries in Oceania.

Essay 3: Tenure Type Mix and Social Mix in Metropolitan Stockholm*

* I would like to thank Mats Wilhelmsson, Matz Dahlberg, Roger Andersson, Peter Fredriksson, and seminar participants at the 2008 ENHR New Housing Researchers Colloquium and the 2009 WRSA conference for helpful comments and discussions. I am grateful to Lena Magnusson Turner for help with the database GeoSweden.

1 Introduction

In many European countries, and also in other parts of the world, (e.g., in the US, Australia and Malaysia), policies to create neighborhoods with mixed housing is used/debated as a mean to obtain socially mixed neighborhoods and to tackle problems of social exclusion in disadvantaged neighborhoods.[1] The tools to obtain housing mix, the types of neighborhoods at which housing mix policies are directed, and the scale at which housing mix is aimed varies somewhat between different countries and time periods, as do the type of social mix and benefits hoped for from housing mix. As explained by Musterd and Andersson (2005, p.762), a general idea is however that "housing mix (a mix of housing types and tenure types) will create social mix (a mix of households according to their socioeconomic position) and that this will create better social opportunities for individuals".

In Sweden, socially mixed neighborhoods was stated a national housing policy goal in the mid-1970s, and housing mix a primary mean advocated to achieve it (see e.g., Holmqvist, 2009). There is a general feature to the goal in that it is partly motivated by the wish to counter overall residential segregation and to obtain social equality (see e.g., Bergsten & Holmqvist, 2007; Holmqvist, 2009). Bergsten and Holmqvist (2007), write that the general feature of the goal not least can be seen in that mixing strategies have been aimed at all types of neighborhoods and not only towards neighborhoods considered disadvantaged. The status of social mix as a national housing policy goal and the conditions for implementation have however been questioned (see e.g., Borevi, 2002; Holmqvist, 2009). Holmqvist (2009) concludes that although the policy goal has been rather consistent over time, the measures to implement the policy have become weaker and fewer, and, while the goal is formulated by the State, its implementation is largely depending on the ambitions of each municipality.[2] Here it can also be noted that in Sweden the debates about housing mix to a large extent have concerned tenure types: The tenure types of buildings are rather fixed in Sweden, where some buildings almost exclusively contain apartments inhabited by tenant-owners, other buildings exclusively contain apartments inhabited by renters, and private houses to a large extent are owner-occupied. To the extent that different population groups tend to be overrepresented in different tenure types, it thus seems reasonable to assume that the spatial distribution of tenure types could affect the spatial distribution of different population groups.

[1] For an overview of such policies in Europe see Musterd and Andersson (2005). For an overview of such policies in the UK as well as some information on such policies in other parts of the word see Graham, Manley, Hiscock, Boyle, and Doherty (2009). For overviews of other means debated/used to obtain socially mixed neighborhoods in different parts of the world, see e.g., Bergsten and Holmqvist (2007), Holmqvist (2009) and Nyström (2008).

[2] For an overview in English of social mix policies in Sweden, see (Holmqvist & Bergsten, 2009), an article mainly based on Bergsten and Holmqvist (2007) and Holmqvist (2009).

However, Musterd and Andersson (2005) point out that the debates about housing mix as a mean to create social mix and thereby better social opportunities for individuals are based on two crucial assumptions: that social mix really enhances individual opportunities and that there is a strong relation between social mix and housing mix. Musterd and Andersson (2005) further note that these assumptions are insufficiently tested.

The aim of the present paper is therefore to study the relation between housing mix and social mix, in the case of metropolitan Stockholm 1990-2008. More precisely I study: *i)* the extent to which tenure type mix have been achieved by measuring the degree of residential segregation between different tenure types; *ii)* the extent to which there is social mix by measuring the degree of residential segregation between different population groups – groups born in different parts of the world, income groups, age groups and family types; and, *iii)* the relation between housing mix and social mix by comparing the mix of population groups within different tenure types in the same residential area, i.e., the dependency in the spatial distribution of people by tenure type and by population group over metropolitan Stockholm.

The only previous study I have found that directly tests the relation between neighborhood housing mix and social mix is Musterd and Andersson (2005).[3,4] For 1995/1996, Musterd and Andersson (2005) look at the association between housing mix and social mix in just over 9000 neighborhoods (i.e., SAMS, see section 3.2), covering all of Sweden. Each neighborhood was classified into a housing mix category by the amount of mix between different housing types and ownership types and into different social (income and ethnic) mix categories. Musterd and Andersson (2005) then look at the correlation between the housing mix categories and the social mix categories, and at the percentage of low-income neighborhoods, neighborhoods with a high number of nationalities and neighborhoods with a high share of refugees that is to be found in each type of housing mix category. From this, Musterd and Andersson (2005) find that although homogeneous and heterogeneous areas with respect to housing mix are different from each other in terms of social and ethnic composition, the difference is not clear cut. In both types of areas a large share of low-income households as well as refugees and people

[3] Some studies on land-use policy also touch upon the relation between housing mix and social mix. Different types of land-use policies are thought to favor different types of housing development with respect to which and where tenure types and housing types are built, and in which price bracket. There are some studies on the impact of the land-use policies adopted by a city on the social composition of the city, see e.g., Ihlanfeldt (2004) for an overview of studies concerning cities in the US; and some studies on the impact of a region-wide land-use policy on the development of the overall residential segregation between social groups in the region, see e.g., Nelson, Dawkins, and Sanchez (2004), Galster and Cutsinger (2007), and Danermark and Jacobson (1989) for an early (and the only?) study of this type on Swedish urban areas. However, these studies do not directly test the relation between housing mix and social mix at the neighborhood level.

[4] For studies on the assumption that social mix enhances individual opportunities, see e.g., Galster (2007) for an overview and analysis of the Western European evidence base.

with non-Swedish nationality can be found. Musterd and Andersson (2005) therefore conclude that there is no clear relationship between housing mix and social mix. With reference to the findings in Musterd and Andersson (2005), Andersson however notes elsewhere (Andersson, 2008) that although the relation between housing mix and social mix is quite weak at the national level it might very well be much stronger in the larger cities, and call for studies that analyse the relation more in detail for cities of different size.

There is also some indirect evidence that indicates that tenure type mix might have little impact on social mix. Bråmå, Andersson, and Solid (2006) use maps to show the share of people with foreign background in different neighborhoods within each tenure type in Uppsala city and Stockholm municipality 2002. From the maps it is clear that there is a marked segregation between neighborhoods within each tenure type. In both cities there are neighborhoods where the share of residents with foreign background is markedly higher or markedly lower than in other neighborhoods. The geographical patterns shown in the maps are further remarkably similar for all tenure types. Enström Öst, Söderberg, and Wilhelmsson (2014), analyse whether there is segregation within the rental housing market in metropolitan Stockholm, a market with rent control. The study, on data from 2008, show that there is residential segregation within different tenure types and that while income segregation is significantly lower in the rent control segment than in the free market benchmark, the rental housing market is not less segregated than the free market benchmark with respect to education level, family type, age and ethnicity.

The contribution of the present paper is first to provide a direct measure of the relation between housing mix and social mix in one large urban area, metropolitan Stockholm. The present paper also covers a long time span, making it possible to follow the evolution over time. In addition, the paper adds to the small literature that measures residential segregation between people living in different tenure types.[5] Finally, the paper addresses two methodological issues present when studying the relation between housing mix and social mix by looking at the correlation between housing mix categories and income mix categories of residential areas: First, if the areas are

[5] There are many studies that measure residential segregation between population groups in Stockholm, the results from the present study is compared to the evidence base in Section 4. Studies that measure residential segregation between people living in different tenure types are however almost inexistent not only for Stockholm. The only study I have found that report a summary measure for residential segregation between people living in different tenure types is Li and Wu (2008). Li and Wu find that, for Shanghai year 2000, residential segregation between people living in different tenure types is higher than residential segregation between different types of rural immigrants and local residents and than residential segregation between people with different levels of education. Further, Meng and Hall (2006) measure the degree of residential segregation between different housing quality classes in a part of Lima, Peru, finding a high degree of segregation in housing quality.

unequal in population size, a large share of areas can be rather heterogeneous with regards to housing mix but homogenous with regards to social mix (as found by Musterd and Andersson, 2005), while most people live in areas that are heterogeneous in both regards, if the previous areas have small populations and the later areas have large populations; second, there are no guidelines to how the classification of areas into housing mix and social mix categories is to be done, and the choices could influence the results. The method in this paper is based on entropy measures, introduced from information theory to the study of racial mix and segregation by Theil and Finizza (1971) and Theil (1972). The way they are used do not require any classification of areas into social mix or housing mix categories, and takes into account the population size of areas. The method has previously been used by Miller and Quigley (1990), Hårsman and Quigley (1995) and Hårsman (2006) to measure the degree of dependency in the spatial distributions of people by ethnicity, income and household type in the San Francisco Bay Area/metropolitan Stockholm. The present paper thus adds the dimension tenure type.

I find that although the distribution of tenure types over municipalities and neighborhoods became more even over the studied period, in 2008 the average tenure type mix in the neighborhoods (i.e., SAMS) was still about 49 percent lower than the mix at the metropolitan level, and the average tenure type mix in the municipalities about 12 percent lower than the mix at the metropolitan level. The residential segregation was much lower between different population groups. While segregation between municipalities is modest for all studied population groups, a more important and increasing segregation between neighborhoods is found for people born in different parts of the world and for income groups. Further, I find that while the mix of family types was rather different between different tenure types in the same municipality over the whole studied period, this also came to be more and more the case with regards to income groups and birth region groups. The mix of different groups however tended to be similar within different tenure types in the same neighborhood (SAMS). While the entropy measures provide a purely descriptive picture, the findings thus suggest that tenure type mix could be more useful for creating social mix at the municipal level than for creating social mix at the neighborhood level.

The rest of the paper is organized as follows. In Section 2, the entropy concept and how it is used to measure residential segregation between tenure types and population groups, and dependency in the spatial distribution of people by tenure type and by population group, is explained. Section 3 describes the data, discusses the spatial divisions, tenure types and population groups used in the study, and provides an overview of the development of tenure types and population groups in metropolitan Stockholm 1990-2008. The main results are presented and discussed in Section 4 and Section 5 concludes.

2 Method

2.1 The entropy concept

To measure the degree of residential segregation between tenure types and between population groups, as well as the dependency in the spatial distribution of people by tenure type and by population group over metropolitan Stockholm, entropy measures are used. The entropy concept was introduced from information theory to the measurement of racial mix and segregation by Theil and Finizza (1971) and (Theil, 1972).[6]

Imagine an urban area with K residential areas (k = 1, …, K), J tenure types (j = 1, …, J) and I groups of people (i = 1, …, I). In this paper the ks denote in turn municipalities and neighborhoods (SAMS) and the is denote in turn groups based on birth region, family types, income and age. Let p_{ijk} be the share of the population in residential area k that belongs to group i and lives in tenure type j. Let

$$p_{i.k} = \sum_{j}^{J} p_{ijk} \tag{1}$$

be the share of the population in residential area k that belongs to group i. For concreteness, let the is be different family types (e.g., singles with children, singles without children, couples with children, couples without children) and the ks be neighborhoods. Then the family type entropy of neighborhood k, $H(i)_k$ is defined as

$$H(i)_k = \sum_{i}^{I} p_{i.k} \log\left(\frac{1}{p_{i.k}}\right).^{7} \tag{2}$$

The entropy of a neighborhood, $H(.)_k$, is a measure of its diversity or mix, here of the mix between different family types. The neighborhood's family type entropy is higher the more family types there are in the neighborhood and the more even the group shares.

In the same way, the tenure type entropy of neighborhood k, $H(j)_k$, is defined as

[6] The term entropy is originally from thermodynamics, where it refers to how much disorder, or mix, there is in a system.

[7] The definition of H requires the definition $0\times\log(1/0) \equiv \lim_{q\to 0}[q\log(1/q)]=0$.

$$H(j)_k = \sum_i^I p_{.jk} \log\left(\frac{1}{p_{.jk}}\right) \tag{3}$$

where

$$p_{.jk} = \sum_i^I p_{ijk} \tag{4}$$

is the share of the population in residential area k that lives in tenure type j.

Next, let w_k be the share of the population in the urban area as a whole that lives in neighborhood k. Then, the population weighted average family type entropy of all neighborhoods in the urban area is

$$\overline{H(i)_k} = \sum_k^K w_k H(i)_k = \sum_k^K w_k \left(\sum_i^I p_{i.k} \log\left(\frac{1}{p_{i.k}}\right)\right), \tag{5}$$

and the population weighted average tenure type entropy is

$$\overline{H(j)_k} = \sum_k^K w_k H(j)_k = \sum_k^K w_k \left(\sum_i^I p_{.jk} \log\left(\frac{1}{p_{.jk}}\right)\right). \tag{6}$$

2.2 Residential segregation

The entropy measure *H(.)$_k$-bar* above tells us the population weighted average mix of tenure types and/or population groups in the neighborhoods of an urban area. However, what is generally of interest is not the degree of mix in itself, but the mix in the neighborhoods compared to the mix in the urban area as a whole. For example, since the entropy increases with the number of groups, the average ethnic entropy of neighborhoods will be small in a city with few ethnic groups. This is however seldom considered as an urban problem.[8] What is usually considered as problematic is if ethnic or other population groups are distributed unevenly over the neighborhood, so that there is a high concentration of some groups in some neighborhoods and of other groups in other neighborhoods, i.e., if there is residential segregation.

[8] A very homogeneous city could however be a problem e.g., if diversity enhances creativity and economic growth. A city with e.g., only owner-occupied housing would possibly also be problematic, if not everybody can buy their housing, but there is on the other hand nothing intrinsically good about a city with 25 percent of each of four possible tenure types.

Let n_i be the number of people in the urban population that belongs to family type i, and N be the total urban population. Then, the share of the urban population that belongs to family type i is

$$p_{i,urban} = n_i / N \tag{7}$$

and the family type entropy at the urban level is

$$\max H(i) = \sum_{i}^{I} p_{i,urban} \log\left(\frac{1}{p_{i,urban}}\right). \tag{8}$$

The tenure type entropy at the urban level, max*H(j)* is defined analogously.

If each neighborhood have the same share of family types as the urban area as a whole, the average family type entropy of the neighborhoods, $H(i)_k$-*bar*, would be equal to max *H(i)*. If the proportional representations in the neighborhoods are different from one another and thus from the proportional representation at the urban level, the average family type entropy of the neighborhoods will be lower than the family type entropy at the urban level. The more different the neighborhoods are from one another, i.e., the higher the concentration of some groups in some neighborhoods and other groups in other neighborhoods, the smaller the average family type entropy compared to the family type entropy at the urban level. Normalizing the difference between the entropy at the urban level and the average entropy of the neighborhoods, and expressing it as a percentage, gives the "Information index of residential segregation" which is the measure used in this paper:

$$S(i) = \frac{\max H(i) - \overline{H(i)_k}}{\max H(i)} * 100 \tag{9}$$

The index takes it minimum value, 0, if the share of each family type in each neighborhood is the same as the share of each family type at the urban level, and its maximum value, 100, if the family types present in the urban area lives completely separated from one another in different neighborhoods. The residential segregation between tenure types, *S(j)*, is defined analogously.

2.3 Spatial dependency

To study the relation between housing mix and social mix a measure similar to *S(i)* described above is used, but instead of comparing the mix of population groups at the urban level to the mix of population groups within different residential areas, the measure compares the mix of population groups

within different tenure types in the same residential area, $(H_j(i))_k$, to the mix of the residential area as a whole, $H(i)_k$.

Let again the *i*s be different family types and the *k*s be neighborhoods. The population weighted average family type mix within the tenure types of all neighborhoods, hereafter called the conditional average family type entropy, is then

$$\overline{H_j(i)_k} = \sum_k^K w_k \left(H_j(i)\right)_k = \sum_k^K w_k \left[\sum_j^J p_{.jk} \left(\sum_i^I \frac{p_{ijk}}{p_{.jk}} \log\left(\frac{1}{p_{ijk}/p_{.jk}} \right) \right) \right] = \overline{H_j(i)_k} = \sum_k^K w_k \left(\sum_j^J \sum_i^I p_{ijk} \log\left(\frac{p_{.jk}}{p_{ijk}} \right) \right) \quad (10)$$

As can be seen from equation (10), within a neighborhood the mix within each tenure type is weighted by the share of the neighborhoods' population who live in that tenure type. A normalized measure of the average difference between the mix of family types within different tenure types in the same neighborhood, $(H_j(i))_k$-*bar*, and the mix within the neighborhoods as a whole, $H(i)_k$-*bar*, expressed as a percentage, is

$$\frac{\overline{H(i)_k} - \overline{H_j(i)_k}}{\overline{H(i)_k}} * 100. \quad (11)$$

Equation 11 is also a measure of dependency, in a statistical and not causal sense, in the distribution of people by family type (or some other type of *i*) and by tenure types, *j*, over the neighborhoods (or some other *k*) of an urban area. If the average conditional family type entropy, $(H_j(i))_k$-*bar*, is equal to the average unconditional ethnic entropy, $H(i)_k$-*bar*, the joint probability that an individual in neighborhood *k* belongs to family type *i* and lives in tenure type *j* is equal to the probability that an individual in neighborhood *k* belongs to family type *i* times the probability that an individual in neighborhood *k* lives in tenure type *j*, i.e., $\Pr(i \cap j) = \Pr(i) \times \Pr(j)$. That is, the distribution of people by family type and the distribution of people by tenure type over the urban area are statistically independent, and equation 11 takes its minimum value 0. If the average family type mix within each tenure type in the neighborhoods are different from one another, and hence from the average family type mix in the neighborhoods overall, $H_j(i)_k$-*bar* will be smaller than $H(i)_k$-*bar*, and $\Pr(i \cap j) \neq \Pr(i) \times \Pr(j)$. That is, the distribution of people by family type and the distribution of people by tenure type over the neighborhoods of the urban area are not independent and equation 11 is larger than 0. If the probability that an individual in neighborhood *k* belongs to family type *i* is

equal to the probability that the individual lives in tenure type *j*, $H_j(i)_k$-*bar* is 0 and Equation 11 takes its maximum value 100. That is, the closer equation 11 is to 100, the more (statistical) dependency is there in the spatial distribution of people by tenure type and by population group.

3 Data, spatial divisions, tenure types and population groups

In this section I present the data used for the analysis and discuss the spatial divisions, tenure types and population groups used. The section also provides an overview of the development of tenure types and population groups in metropolitan Stockholm 1990-2008.

3.1 Database and population

In this study I analyze social mix, tenure type mix, and the relation between the two in metropolitan Stockholm 1990-2008. Residential segregation is often thought to be more of an issue in larger cities since there are more scope for different groups to live in separate places and since residential segregation to a larger extent will affect for example school segregation. In smaller cities, different groups can live in different ends of the city but will still see each other in the grocery store, and the children will all go to the same and only school. Stockholm is the largest city of Sweden and metropolitan Stockholm house about 20 % of the total Swedish population, and an even larger share of the foreign born population.

For the study I rely on population-wide register data, compiled for research purposes by Statistics Sweden, and held by the Institute for Housing and Urban Research (IBF) in a database called GeoSweden. The database contain yearly data on *all* residents in Sweden from 1990 and onwards (it is continuously updated). The database contains information on the individual level on demographic variables (e.g., year of birth, gender, marital status, number of children, country of birth), as well as on income, employment and educational variables. Further, the database contains detailed information on the residential location of each individual. For selected years (1990, 1995, 2000, 2002, 2004, 2006 and 2008) the database also contains information on the real estate in which the individual lived.

For the study I use a sample of cross-sections containing all people 19 years or older[9] living in metropolitan Stockholm in the years 1990, 1995, 2000, 2006 and 2008.[10]

[9] The age refers to the age in the end the year.

[10] The initial idea was to look at every fifth year from 1990 and onwards. Given that there is no data on the real estate in which the individual lived in the end of the topical year 2005, year

The limit of 19 years is set since people usually finish upper secondary school in the year they turn 19 and few persons leave their parents' home before this. Further, people living in the tenure type "other" (see Section 3.3) are not included in the measures in the result section since they do not belong to the regular housing market and since the actual residence for parts of the population in this tenure type is unclear.

Table 1 shows the population development in metropolitan Stockholm 1990-2008.

Table 1. Total population and population 19 years or older in metropolitan Stockholm 1990-2008.

Year	Total population (in thousands)	Population 19 years or older (in thousands)
1990	1642	1276
1995	1726	1332
2000	1823	1410
2006	1918	1473
2008	1981	1523

3.2 Spatial divisions

The degree of residential segregation between tenure types and between population groups, as well as the dependency in the spatial distribution of people by tenure type and by population group, are calculated over two levels of aggregation: over the about 900 SAMS (Small Areas for Market Statistics) and over the 26 municipalities in metropolitan Stockholm. The SAMS classification was created by Statistics Sweden to satisfy demand for small area statistics from other users than municipalities.[11] The objective was to create fairly homogeneous residential areas of about 1000 inhabitants each. The homogeneity related to housing type, date of construction and tenure type (Musterd & Andersson, 2005). In metropolitan Stockholm, the SAMS were created from the previous classification "basområde" ("base area"), a real estate-based classification used for intra-municipal and sometimes regional planning and administration. A SAMS came to equal a "base area", except within Stockholm municipality where each SAMS came to equal several "base areas". When the SAMS classification was created, the intention was to update the classification as "base areas" changed, but for different reasons the SAMS-classification has remained unchanged a part from minor adjustments

2006 was chosen since it was the most recent year for which data was available at the time this study was started. When data for 2008 were made available, this year was added to keep the study as up-to-date as possible.

[11] The information in this paragraph is from Statistics Sweden (2005), when nothing else is indicated.

since it came into use in 1994.[12,13] The SAMS have been used frequently in Swedish studies as the formal division closest to neighborhoods.

It has been noted that if indices for residential segregation are computed over some predefined subareas, for example census tracts, SAMS or municipalities, there are two inherent issues: the checkerboard problem and the modifiable areal unit problem (MAUP).[14] The checkerboard problem is that the computed degree of residential segregation will be the same whether the subareas with relatively high concentrations of some groups are spatially clustered or evenly spread over the whole urban area; it is only the composition of each predefined subarea that counts, not their spatial relation to one another. The modifiable areal unit problem (MAUP) is constituted by two interrelated effects: a scale effect and a zoning effect. The scale effect refers to the fact that the computed degree of residential segregation will usually be smaller the larger (in population terms) the subareas over which residential segregation are computed, e.g., considering the same urban area, the degree of residential segregation computed over municipalities is usually smaller than the degree of residential segregation computed over census tracts or SAMS. The zoning effect refers to the fact that the computed degree of residential segregation will depend on where the boundaries between the subareas are drawn, i.e., how the "zoning" is done, even if the scale and number of subareas are fixed. If the predefined subareas are delimitated to be homogeneous with regards to some social or other factors the computed degree of residential segregation by these factor(s) will naturally be high initially but then usually fall if the same subarea boundaries are retained. Once again it is only the composition of each predefined subarea that counts, not the spatial relation between different groups of people regardless of which predefined subarea they live in. Reardon and O'Sullivan (2004 p.124), write "unless spatial subarea boundaries correspond to meaningful social boundaries – all measures of spatial and aspatial segregation that rely on population counts aggregated within subareas are sensitive to the definitions of the boundaries of these spatial subareas".

Several ways to deal with the checkerboard problem and the MAUP have been proposed (see e.g., Feitosa, Camara, Monteiro, Koschitzki, & Silva, 2007; O'Sullivan & Wong, 2007; Reardon & O'Sullivan, 2004). The present

[12] Although the SAMS-classification did not come into use before 1994, older information can be located within the existing classification by use of the more precise coordinates that real estates have in Sweden.

[13] The "base areas" have however not changed much either. Up to year 2000 the "base areas" were reviewed and revised every 5th year, in 2000 it was decided that a review and revision every tenth year would be sufficient, and since 1995, changes of boundaries are not made in order to enable the development over time to be studied (Regionplane- och trafikkontoret Stockholms läns landsting, 2008).

[14] The presentation of the checkerboard problem and the MAUP in this section mainly builds on the discussion of the issues in Reardon and O'Sullivan (2004). For earlier references on the checkerboard problem and the MAUP, the reader is referred to that article.

paper does not rely on any of these strategies. Instead I argue that SAMS and municipalities indeed correspond to meaningful social and political boundaries in the context of this paper.[15]

Regarding the SAMS, it can first be noted that since the SAMS classification in Stockholm builds on a previous classification used for intra-municipal and sometimes regional planning and administration, they should correspond to some lived reality, e.g., be related to which school children attend, infrastructure/communications, and other public services. Second, the political aim is not to create a completely smooth distribution of population groups or tenure types. In their interviews with municipal planning departments and housing companies about their perceptions of social mix and housing mix and how and where it was to be implemented, Bergsten and Holmqvist (2007), find that the level aimed at was the meso level "stadsdelsnivå" or "områdesnivå", i.e., a lower level than the municipality but a higher level than the staircase or the block. Some planning departments and housing companies even saw a mix on the level of the staircase or the block as undesirable. The SAMS should correspond rather well to this meso level.

Regarding the municipalities, their boundaries are highly relevant to the issue of housing mix and social mix. As noted in the introduction, in her dissertation Holmqvist (2009) shows that the implementation of the national housing policy goal of socially mixed neighborhoods largely has fallen on the municipalities. Although mixed housing has been advocated on a national level, housing policy at a national level in Sweden is limited to setting the financial and legal frames, while land use, housing provision, public housing, the implementation of the national directives etc. are responsibilities at the municipal level and at each municipality's discretion, within the legal and financial frames. This discretion can create imbalances between the municipalities in terms of the provision of different tenure types, and thus maybe in the distribution of population group, and there have been calls from municipalities with a large share of public rental housing for other municipalities to build more rental housing, see e.g., Lago et al. (2010).

3.3 Tenure types

GeoSweden does not contain direct information on which tenure type people live in, but by combining housing type and legal form of ownership, it is possible to approximately classify people into tenure types. The tenure types

[15] Further, not relying on predefined subareas but for example on density surfaces has its own problems, for example, people living close to one another in meters can be far from one another in terms of interaction if they are separated by a road or a railway, or if the children go to different schools. The SAMS on the other hand are rather coherent areas. Also, for the calculation of density surfaces, the researcher is still limited by the format in which data are delivered, i.e., most often aggregated on predefined subareas of some kind. GeoSweden could however be quite helpful in this respect since data is available for squares of 100m*100m

used in this paper are "public rental" housing, i.e., housing rented from a municipal housing company, "private rental" housing, "owner-occupied housing", housing in tenant-owner cooperatives, and "other" (for details on how the classification is done see Appendix). Row houses, semi-detached houses and detached houses can be part of tenant-owner cooperatives or be privately owned. Condominiums, i.e., privately owned apartments in apartment blocks, have on the other hand only been allowed since 2009 in Sweden and only in buildings produced for that purpose or that have not served for housing the last eight years.

From GeoSweden it is not possible to discern whether a person living in "owner-occupied" housing or a tenant-owner cooperative owns the dwelling he or she occupies or whether he or she rents it from the owner. Those who rent these types of housing are however few. Similarly, it is not possible to discern whether a person living in "private rental" actually rents a dwelling in a tenement house or owns a tenement house and live in one the apartments. Those who own a tenement house and live in one of the apartments are however few relative to the renters.

The group "others" is rather large: around 10 percent of the population in metropolitan Stockholm is classified into this category. The category includes farms regardless of legal form of ownership, real estates owned by the state, the church, directly by the municipality or by the County Council, estates of dead persons and those real estate for which data on legal form of ownership and house type is missing (generally, either both or none is known). About half of the persons classified into the category "other" cannot be connected to any real estate or live in a real estate for which data is missing.[16] As noted in section 3.1, people living in the tenure type "other" are not included in the measures in the result section since they do not belong to the regular housing market, and since the actual residence for some is unclear.

It should be noted that any measures related to tenure type in the present paper is based on the number of people living in different tenure types, not on the number of physical dwelling units of each type. The number of physical dwelling units is not available in GeoSweden since it is a database of individuals.

Table 2 shows the tenure type development in metropolitan Stockholm 1990-2008. In 1990 it was more common among the population to live in one of the owner tenures than in one of the rental tenures and the share of the population living in one of the owner tenures further increased 1990-2008. The increase was particularly large with respect to the share of the population living in tenant-owner cooperatives. From 1990-2008 the share of the population living in housing cooperatives increased by about 12 percentage points: from about 18 % in 1990 to about 30 % in 2008. Owner-occupied housing

[16] It is not possible to discern between the two since people who cannot be connected to a real estate are ascribed a fictive real estate number by Statistics Sweden.

however remained the most common tenure type 1990-2008, with about 30 % of the population in this tenure type over the whole period. The share of the population living in public rental housing decreased by about 10 percentage points from 1990-2008: from more than 26 % in 1990 to less than 17 % in 2008. The share of the population living in private rental housing saw a small increase 1990- 2000 followed by a decreased 2000-2008 to a lower level than in 1990. The large increase in the share of people living in tenant-owner cooperatives and the decrease in the share of people living in rental housing is likely in large part due to the many conversions of rental housing to tenant-owner cooperatives that took place in metropolitan Stockholm within the studied time period.

Table 2. Metropolitan Stockholm: Percentage of the population 19 years or older in each tenure type

year	Owner-occupied	Tenant-owner cooperative	Public rental	Private rental	other
1990	29.5	17.9	26.4	14	12.2
1995	29.6	21.4	24.9	14.4	9.7
2000	29.8	22.6	21.2	16	10.3
2006	30.6	28.3	19.1	11.9	10.1
2008	30.6	30.3	16.8	12.3	10

3.4 Population groups

In their interviews with municipal planning departments and housing companies, Bergsten and Holmqvist (2007) find that the understanding and practice of the social mix policy in Sweden have been rather consistent since it was introduced as a national housing policy goal in the middle of the 1970's, with social mix policy remaining a general policy for counteracting socioeconomic segregation rather than ethnic segregation and with age groups and family types as other categories frequently cited as desirable to mix. Therefore, four different categories of population groups are considered in this study: income groups, family types, age groups and also groups born in different parts of the world.

3.4.1 Family types

Four family types are considered in this study: (1) couples without children living at home, (2) couples with children living at home, (3) single persons with children living at home and (4) single persons without children living at home. Couples cohabiting who are not married or have registered partnership and who never had any common child are not classified as belonging to the same family but as belonging to category 3 or 4. The reason for this iis that GeoSweden only contain information on in which real estate a person

lives, not in which apartment.[17] All persons belonging to the same family in GeoSweden has the same family type, individuals living at home with their parents will thus be classified as belonging to group (2) or (3).

Table 3 shows the share of each family type in the total population and in each tenure type 1990-2008. The share of people belonging to each family type did not change much between 1990 and 2008. The largest group is singles without children living at home and the second largest couples with children living at home. The third largest group is couples without children living at home and the smallest group singles with children living at home.

In all studied years, couples without children living at home were over-represented in the two owner tenures and under-represented in the two rental tenures. Couples with children living at home were over-represented in owner-occupied housing and under-represented in the other tenure types. Singles with children living at home were over-represented in the rental tenures and under-represented in the owner-tenures. Finally, singles without children living at home were under-represented in owner-occupied housing and over-represented in all other tenure types.

[17] In 2010, the Swedish Tax Agency started to collect information on in which apartment people live which should improve future statistics.

Table 3. Metropolitan Stockholm: The share of family types in the population and in each tenure type

year	as % of	Couples NoChildren	Couples WithChildren	Singles WithChildren	Singles NoChildren
1990	**population>=19 years**	**20**	**33.7**	**8.4**	**37.9**
1990	owner-occupied	24.1	58.7	5.5	11.6
1990	tenant-owner cooperative	22.2	22.2	7.2	48.4
1990	public rental	16.6	25.7	12.4	45.3
1990	private rental	17.7	17.5	8.6	56.2
1990	Other	16.9	25.7	8.2	49.2
1995	**population>=19 years**	**19.7**	**32.5**	**8.5**	**39.2**
1995	owner-occupied	26.8	54.5	5.6	13.1
1995	tenant-owner cooperative	20.5	24.3	7.7	47.4
1995	public rental	14.9	24.8	12.5	47.7
1995	private rental	15.3	18.1	9	57.7
1995	Other	15.5	24.6	8.6	51.3
2000	population>=19 years	19	31.4	8.8	40.9
2000	owner-occupied	26.8	52.8	5.5	14.9
2000	tenant-owner cooperative	20.1	21.5	7.9	50.5
2000	public rental	13.3	24.6	13.3	48.8
2000	private rental	13.1	19.5	10.3	57.2
2000	Other	14.6	23.3	8.6	53.5
2006	**population>=19 years**	**18.4**	**32.5**	**9.1**	**40**
2006	owner-occupied	25.3	53.9	5.9	14.9
2006	tenant-owner cooperative	19.3	22.5	8.1	50
2006	public rental	12.3	25.2	14.1	48.4
2006	private rental	12.4	21	11.4	55.1
2006	Other	13.6	22.7	9.4	54.3
2008	**population>=19 years**	**18.1**	**33**	**9.2**	**39.7**
2008	owner-occupied	24.5	54.7	6.1	14.7
2008	tenant-owner cooperative	19	23.2	8.5	49.4
2008	public rental	12.3	25.1	14.2	48.5
2008	private rental	12	22	12.1	53.9
2008	Other	13.6	22.8	9.3	54.2

Note: All measures refer to the population 19 years or older.

3.4.2 Age groups

Four age groups are considered in this study: people 19-29 years, people 30-49 years, people 50-64 years, and people 65 years and older. The groups are thought to correspond generally to formation years, two periods when people are active on the labor market, whereof the first also generally correspond to family formation, and retirement years. Table 4 shows the share of each age group in the total population and in each tenure type 1990- 2008.

In all studied years the largest age group was the group aged 30-49 years. In 1990, the second largest group was the group aged 19-29 years but over the studied period the share in this group decreased and in 2008 this group was the smallest. The opposite development can be seen for the group aged 50-64 years, that was the smallest group in 1990 but the second largest in 2008. Finally, the share of the population belonging to the

group aged 65 years or older decreased somewhat over the studied period.

In all studied years, the youngest group was under-represented in owner-occupied housing and over-represented in the other tenure types. The group aged 30-49 years was all years over-represented in owner-occupied housing and in 1995, 2000 and 2006 also slightly over-represented in public rental housing. The group aged 50-64 years was all years over-represented in owner-occupied housing and under-represented in the other tenure types. Finally, the group aged 65 years or older was all years under-represented in owner-occupied housing and over-represented in tenant-owner cooperatives and in ”other”. In 1990, 1995 and 2000 the group was further over-represented in public rental and in 1990 and 2000 also over-represented in private rental housing.

Table 4. Metropolitan Stockholm: The share of age groups in the population and in each tenure type

year	as % of	19-29	30-49	50-64	65-
1990	**population>=19 years**	22	39.5	18.5	20.1
1990	owner-occupied	15.5	48.4	23.8	12.4
1990	tenant-owner cooperative	24.8	33.8	16.9	24.5
1990	public rental	25.4	37.6	16.3	20.7
1990	private rental	24.5	34.3	15.9	25.4
1990	Other	23.1	36.8	15.4	24.7
1995	**population>=19 years**	20.3	39.2	21	19.5
1995	owner-occupied	13.4	43	29.6	14
1995	tenant-owner cooperative	21.9	36.2	19.1	22.8
1995	public rental	24.1	39.4	16.5	20
1995	private rental	25.4	36.8	16.4	21.5
1995	Other	21.1	37	17.1	24.8
2000	population>=19 years	19.5	38.8	23.3	18.4
2000	owner-occupied	12.1	41.6	31.7	14.6
2000	tenant-owner cooperative	21.2	34.6	22.4	21.8
2000	public rental	22.4	40.7	18.3	18.5
2000	private rental	25.5	38.3	18.4	17.8
2000	other	22.1	36.7	19.1	22.1
2006	**population>=19 years**	17.8	39.8	24	18.4
2006	owner-occupied	10.6	43.8	29.6	16
2006	tenant-owner cooperative	19.1	36.8	22.8	21.3
2006	public rental	21.6	40.1	20.8	17.5
2006	private rental	23.5	39.1	20.6	16.8
2006	other	22.4	35.9	20.8	20.9
2008	**population>=19 years**	18.6	39.3	23.4	18.8
2008	owner-occupied	11.2	43.7	28.1	17
2008	tenant-owner cooperative	19.7	36.8	22	21.5
2008	public rental	22.3	39	21.2	17.5
2008	private rental	24.9	38.6	20.4	16.1
2008	other	23.5	35	20.4	21.1

Note: All measures refer to the population 19 years or older.

3.4.3 Income groups

Three income groups based on disposable income are considered in the study.[18] The income groups are: below the 20 percentile, between the 20 percentile and the 80 percentile and above the 80 percentile. The income percentiles refer to the population who were 19 years or older in the end of each year and who did not live in the tenure type "other". Table 5 shows the share of each income group in the total population (19 years or older and not in the tenure type "other") and in each tenure type 1990-2008.

The income groups are defined to contain the same share of the population each year. In all studied years the group with the lowest disposable income was over-represented in the two rental tenures and under-represented in the two owner tenures, and the over- and under-representation increased over time. The group with "middle" disposable income was also overrepresented in the rental tenures and under-represented in owner-occupied housing, but rather well represented in tenant-owner cooperatives. The group with the highest disposable income was over-represented in owner-occupied housing and, apart form 1990, in tenant-owner cooperatives. The group was underrepresented in rental housing, in particular in public rental housing.

[18] Disposable income is basically income from work and capital plus transfers and minus taxes. To be precise, for 1990, 1995 and 2000, it is the variable DispInk in the "LISA" database. For 2006 and 2008, to be consistent with the variable DispInk, it is the variables DispInk04-KapInk in the "LISA" database if KapInk is negative, and otherwise DispInk04. For more information see Statistics Sweden (2009).

Table 5. Metropolitan Stockholm: The share of income groups in the population and in each tenure type

year	as % of	<20 percentile	20<x<80 percentile	75 percentile<
1990	**population>=19 years**	20	60	20
1990	owner-occupied	17.2	50	32.8
1990	tenant-owner cooperative	19.2	60.9	19.9
1990	public rental	22.9	68.1	9
1990	private rental	21.3	64.8	13.9
1995	**population>=19 years**	20	60	20
1995	owner-occupied	17.7	50.7	31.6
1995	tenant-owner cooperative	17.9	61	21.1
1995	public rental	24	67.7	8.3
1995	private rental	20.6	64.7	14.6
2000	population>=19 years	20	60	20
2000	owner-occupied	16.3	53.7	30
2000	tenant-owner cooperative	18	59.1	22.9
2000	public rental	25.8	66.8	7.4
2000	private rental	22.1	64.2	13.7
2006	**population>=19 years**	20	60	20
2006	owner-occupied	15.1	53.9	31
2006	tenant-owner cooperative	17.5	60.1	22.4
2006	public rental	28.7	66.4	4.9
2006	private rental	24.4	65.5	10.1
2008	**population>=19 years**	20	60	20
2008	owner-occupied	14.5	54.3	31.2
2008	tenant-owner cooperative	16.9	60.5	22.6
2008	public rental	30.7	65.6	3.7
2008	private rental	26.7	65.2	8

Note: All measures refer to the population 19 years or older.

3.4.4 Birth region groups

Three birth region groups are considered in this study: people born in Sweden, people born in a western country and people born in a non-western country.[19] The reason for studying only three groups is that immigrant-dense neighborhoods in Sweden to a large extent are multi-ethnic, see e.g., Andersson (2007). That is, the most important part of "ethnic" residential segregation in Sweden is not between different ethnicities, but between native and foreign borns. I have however divided the foreign born group into people born in western and non-western countries, since people born in western countries generally have more financial resources than people born in non-western countries, and thus more possibilities on the housing market.

Table 6 shows the share of each birth region group in the total population and in each tenure type 1990-2008. The largest group is, not surprisingly, the

[19] Western countries include: Finland, Norway, Denmark, Iceland, Ireland, Great Britain, the Netherlands, Belgium, France, Monaco, Luxembourg, Lichtenstein, Spain, Portugal, Andorra, Germany, Switzerland, Italy, San Marino, Vatican City State, Austria, Greece, Canada, the USA, and the countries in Oceania.

group born in Sweden. The population share of the group however decreased over time, from 82.3 % of the population in 1990 to 75.9 % of the population in 2008. In 1990, the second largest group was the group of people born in a western country, 10 % of the population, while 7.7 % of the population was born in a non-western. From 1995, the group born in a non-western country has however been larger than the group born in a western country. In 2008, 16.5 % of the population in metropolitan Stockholm was born in a non-western country and 7.6 % in a western country.

In all studied years, the group born in Sweden was over-represented in the two owner tenures and under-represented in public rental housing. Before 2000, the group was also over-represented in private rental and other, while from 2000 the group was under-represented in these tenures. The group of people born in western countries was all years under-represented in the two owner tenures and over-represented in public rental housing. Before 2000 (2006) the group was further under-represented in private rental housing (other) and 2000 (2006) and onwards the opposite is the case. Finally, the group of people born in a non-western country was all years over-represented in rental housing. From 2006 the group was also over-represented in the type “other”.

Table 6. Metropolitan Stockholm: The share of birth region groups in the population and in each tenure type

year	as % of	Sweden	Western country	Non-western country
1990	**population>=19 years**	**82.3**	**10**	**7.7**
1990	owner-occupied	88.7	8.3	3
1990	tenant-owner cooperative	83.9	10.3	5.9
1990	public rental	72.7	12	15.3
1990	private rental	85.3	8.9	5.8
1990	Other	81.9	10.2	7.9
1995	**population>=19 years**	**80.6**	**9.2**	**10.2**
1995	owner-occupied	88.3	8	3.7
1995	tenant-owner cooperative	82.3	9.9	7.8
1995	public rental	68.3	10.4	21.3
1995	private rental	82.7	8.5	8.8
1995	Other	81.9	9.4	8.7
2000	**population>=19 years**	**79.3**	**8.6**	**12.1**
2000	owner-occupied	87.8	7.7	4.5
2000	tenant-owner cooperative	82.4	9.2	8.4
2000	public rental	65.5	9	25.6
2000	private rental	77.7	8.3	14
2000	Other	79	9	12
2006	**population>=19 years**	**77**	**7.8**	**15.2**
2006	owner-occupied	86.4	7.1	6.5
2006	tenant-owner cooperative	81.4	8.4	10.2
2006	public rental	59.3	7.8	32.8
2006	private rental	74	7.6	18.4
2006	Other	72.8	9	18.2
2008	**population>=19 years**	**75.9**	**7.6**	**16.5**
2008	owner-occupied	85.8	6.9	7.3
2008	tenant-owner cooperative	80.2	8.1	11.7
2008	public rental	57.7	7.6	34.7
2008	private rental	69.6	7.3	23.1
2008	Other	70.6	8.6	20.7

Note: All measures refer to the population 19 years or older.

4 Results

In this section I present the results from the calculations of the different entropy measures. In section 4.1, I present measures concerning the development of tenure type mix in metropolitan Stockholm 1990-2008, and in section 4.2, I present measures concerning the social mix as well as measures concerning the relation between tenure type mix and social mix.

4.1 Tenure type mix

Table 7 shows measures concerning the development of tenure type mix in metropolitan Stockholm 1990-2008. The first row shows the mix at the

urban level (*maxH(j)*); the second row the average mix within the municipalities ($H(j)_k$-*bar*, with *k* being municipalities); the third row the average mix within the SAMS ($H(j)_k$-*bar*, with *k* being SAMS); the fourth row the percentage difference between the mix at the urban level and the average mix within municipalities (*S(j)*, with *k* being municipalities), i.e., the degree of tenure type segregation between municipalities; and, the fifth row the percentage difference between the mix at the urban level and the average mix within SAMS (*S(j)*, with *k* being SAMS), i.e., the degree of tenure type segregation between SAMS.

Table 7. Tenure type mix in metropolitan Stockholm 1990-2008

	1990	1995	2000	2006	2008
Metropolitan level entropy	1.345	1.355	1.362	1.328	1.319
Average entropy within municipalities	1.178	1.196	1.196	1.171	1.162
Average entropy within SAMS	0.639	0.686	0.691	0.673	0.674
Segregation between municipalities	12.4	11.7	12.2	11.8	11.8
Segregation between SAMS	52.5	49.4	49.3	49.3	48.9

Note: All measures refer to the population 19 years or older living in owner-occupation, tenant-owner cooperatives, public rental and private rental.

From Table 7, row 1, it can be seen that at the metropolitan level, the tenure type mix increased from 1990 to 2000 but there after decreased again to be lower in 2008 than in 1990. The average tenure type mix at the municipal level followed a similar pattern (row 2). The average tenure type mix at the SAMS level also increased in the beginning of the studied period and decreased towards the end, but the mix in 2008 was larger than in 1990 (row 3). With regards to residential segregation, people living in different tenure types to a large extent also tended to live in different municipalities (row 4) and in different SAMS (row 5) over the whole period 1990-2008, although the distribution became a little more even over the period. Over the whole period, the average mix of tenure types within the municipalities was about 12 % lower than the mix at the metropolitan level, while the average mix of tenures types within the SAMS was about half the mix at the metropolitan level. The high value of tenure type segregation between SAMS in the beginning of the period, 52.5 %, was to be expected since the SAMS were created to be homogeneous with respect to housing type, date of construction and tenure type, but the value was still about 49 % in the end of the period.

4.2 Social mix and the correlation with tenure type mix

Table 8, Table 9, Table 10, and Table 11 show measures concerning the development of social mix in metropolitan Stockholm 1990-2008 with respect to family types, age groups, income groups and birth region groups. The tables also show measures concerning the relation between social mix

and tenure type mix. In each table, the first row shows the degree of family type/age/income group/birth region segregation between municipalities (*S(i)*, with *k* being municipalities); the second row shows the degree of family type/age/income group/birth region segregation between SAMS (*S(i)*, with *k* being SAMS); the third row shows the percentage difference between the average mix of population groups within the municipalities and the average mix of population groups within the tenure types of the municipalities (equation 11, with *k* being municipalities), i.e., the dependency in the spatial distribution of tenure types and population groups over municipalities; and, the fourth row shows the percentage difference between the average mix of population groups within the SAMS and the average mix of population groups within the tenure types of the SAMS, i.e., the dependency in the spatial distribution of tenure types and population groups over SAMS (equation 11, with *k* being SAMS).

4.2.1 Family types

From Table 8 it can be seen that over the whole period 1990-2008, the degree of family type segregation between municipalities in metropolitan Stockholm was rather low. In 1990, the average mix of family types within the municipalities in metropolitan Stockholm was only about 2.4 % lower than the mix at the metropolitan level, and in 2008 the difference had decreased to 1.7 %. Family type segregation was much higher between SAMS, although also decreasing over the period: in 1990, the average mix of family types within the SAMS was about 8.7 % lower than the mix at the metropolitan level, and in 2008 the difference had decreased to 6.8 %.

From Table 8 it can further be seen that although decreasing, the dependency in the spatial distribution of tenure types and family types over municipalities was rather high over the studied period. In 1990 (2008), the average mix of family types within the tenure types of the municipalities was about 6.4 (5.7) % lower than the average mix of family types within the municipalities as a whole. The dependency in the spatial distribution of tenure types and family types over SAMS was lower: over the whole period 1990-2008, the average mix of family types within the tenure types of the SAMS was about 3.2 % lower than the average mix of family types in the SAMS as a whole.

Table 8. Segregation between family types in metropolitan Stockholm 1990-2008

	1990	1995	2000	2006	2008
between municipalities	2.4	2.2	2	1.8	1.7
between SAMS	8.7	8	7.5	6.9	6.8
between tenure types within municipalities	6.4	5.8	5.9	5.7	5.7
between tenure types within SAMS	3.2	3.1	3.3	3.2	3.2

Note: All measures refer to the population 19 years or older living in owner-occupation, tenant-owner cooperatives, public rental and private rental.

4.2.2 Age groups

From Table 9 it can be seen that over the whole period 1990-2008, the degree of age group segregation between municipalities in metropolitan Stockholm was low. In 1990, the average mix of age groups within the municipalities in metropolitan Stockholm was only about 1.1 % lower than the mix at the metropolitan level, and in 2008 the difference had decreased to 0.4 %. The age group segregation was higher between SAMS, but still rather low and also decreasing over the period: in 1990, the average mix of age groups within the SAMS was about 4.1 % lower than the mix at the metropolitan level, and in 2008 the difference had decreased to 2.3 %. The order of magnitude of residential segregation between age groups in Table 9, and the findings that residential segregation between age groups decreased 1990-2008 is consistent with Biterman and Franzén (2006).[20]

From Table 9 it can further be seen that the dependency in the spatial distribution of tenure types and age groups over municipalities, as well as over SAMS, was low over the studied period. In 1990 (2008), the average mix of age groups within the tenure type segments of the municipalities was about 1.4 % (1.1 %) lower than the average mix of age groups within the municipalities as a whole. The numbers with respect to SAMS were only slightly larger.

Table 9. Segregation between age groups in metropolitan Stockholm 1990-2008

	1990	1995	2000	2006	2008
between municipalities	1.1	0.7	0.6	0.4	0.4
between SAMS	4.1	3.3	2.8	2.4	2.3
between tenure types within municipalities	1.4	1.3	1.3	1.1	1.1
between tenure types within SAMS	1.5	1.4	1.4	1.2	1.2

Note: All measures refer to the population 19 years or older living in owner-occupation, tenant-owner cooperatives, public rental and private rental.

4.2.3 Income groups

From Table 10 it can be seen that over the whole period 1990-2008, the degree of income segregation between municipalities in metropolitan Stockholm was rather low although increasing. In 1990 (2008), the average mix of income groups within the municipalities in metropolitan Stockholm was about 0.9 % (1.4%) lower than the mix at the metropolitan level. The degree of income segregation between SAMS was higher and also increasing, from 3.6 % in 1990 to 6.9 % in 2008. The order of magnitude of residential segregation between income groups in Table 10 is consistent with previous studies of income group segregation in metropolitan Stockholm 1990-2000 (-2006), e.g., Hårsman (2006) and Biterman and Franzén (2006, 2010). Regarding the

[20] Although comparable in level and directions, the results in the present study and previous studies cannot be directly compared since the groups and the residential areas are not defined in the same way.

change in residential segregation over the period, Biterman and Franzén (2006, 2010) find that residential segregation between income groups increased 1990-2006 although the increase was not straight. Hårsman (2006) on the other hand finds that residential segregation between age-income groups decreased from 1991 to 2001. The differences in the findings of Biterman and Franzén (2006, 2010) and the present study at the one hand and Hårsman (2006) on the other hand are likely due to the decreasing age segregation shown in Table 9 above and in Biterman and Franzén (2006).

From Table 10 it can further be seen that in the beginning of the period 1990-2008, the dependency in the spatial distribution of tenure types and income groups over municipalities and over SAMS was low. In 1990, the average mix of family types within the tenure types of the municipalities (SAMS) was about 3.1 % (1.5 %) lower than the average mix of income groups in the municipalities (SAMS) as a whole. The dependency however increased over the period. In 2008, the average mix of income groups within the tenure types of the municipalities was as much as 5.5 % lower than the average mix of income groups in the municipalities as a whole. For SAMS the number was 2.7 %.

Table 10. Segregation between income groups in metropolitan Stockholm 1990-2008

	1990	1995	2000	2006	2008
between municipalities	0.9	1.0	1.2	1.2	1.4
between SAMS	3.6	4.2	4.9	5.9	6.9
between tenure types within municipalities	3.1	3.1	3.4	4.6	5.5
between tenure types within SAMS	1.5	1.6	1.7	2.2	2.7

Note: All measures refer to the population 19 years or older living in owner-occupation, tenant-owner cooperatives, public rental and private rental.

4.2.4 Birth region groups

From Table 11 it can be seen that over the whole period 1990-2008, the degree of segregation between people born in Sweden, western countries and non-western countries over municipalities in metropolitan Stockholm was rather low although increasing. In 1990 (2008), the average birth region mix within the municipalities in metropolitan Stockholm was about 1.7 % (2.3 %) lower than the mix at the metropolitan level. The degree of segregation between these groups over SAMS was much higher and also increasing, from 8.1 % in 1990 to 11.9 % in 2008. The order of magnitude of residential segregation in Table 11 and the finding that it increased 1990-2000 (-2006) are consistent with previous studies of ethnic residential segregation in metropolitan Stockholm, e.g., Hårsman (2006); Andersson (2000) and Biterman and Franzén (2006, 2010).

From Table 11 it can further be seen that in the beginning of the period 1990-2008, the dependency in the spatial distribution of tenure types and

birth region groups over municipalities and over SAMS was rather low. In 1990, the average mix of birth region groups within the tenure types of the municipalities (SAMS) was about 3.3 % (1.4 %) lower than the average birth region mix in the municipalities (SAMS) as a whole. The dependency however increased over the period. In 2008, the average birth region mix within the tenure type segments of the municipalities was about 4.9 % lower than the average birth region mix in the municipalities as a whole. With regards to SAMS the corresponding number was 2.0.

Table 11. Segregation between birth region groups in metropolitan Stockholm 1990-2008

	1990	1995	2000	2006	2008
between municipalities	1.7	1.8	2.0	2.3	2.3
between SAMS	8.1	10.3	11.1	11.8	11.9
between tenure types within municipalities	3.3	4.2	4.7	5.3	4.9
between tenure types within SAMS	1.4	1.8	2.1	2.1	2.0

Note: All measures refer to the population 19 years or older living in owner-occupation, tenant-owner cooperatives, public rental and private rental.

5 Conclusions

In Sweden like in many other countries, policies to create neighborhoods with mixed housing have been advocated as a mean to obtain socially mixed neighborhoods. Little is however known about the relation between housing mix and social mix. In this paper I use entropy measures to study the distribution of tenure types, the distribution of population groups – birth region groups, income groups, age groups and family types – and the dependency in the distribution of people by tenure type and by population group, over metropolitan Stockholm 1990-2008. The dependency should be interpreted in a statistical, not casual, sense so the paper is purely descriptive. Further, the exact values found in this study are specific to the group definitions used. Other definitions would probably have given somewhat different values and possibly other tendencies. The order of magnitude and the direction of the development over time of residential segregation between different population groups found in this study are however in line with previous studies in the cases a comparison is possible.

With this said, I find that although the distribution of tenure types over municipalities and neighborhoods, i.e., SAMS, became more even over the studied period, the average tenure type mix in the SAMS was still about 49% lower than the mix at the metropolitan level in 2008, and the average tenure type mix in the municipalities almost 12 % lower than the mix at the metropolitan level. Residential segregation was much lower between different population groups. While segregation between municipalities is modest for

all studied population groups, a more important and increasing segregation between neighborhoods is found for people born in different parts of the world and for income groups. Further, I find that the mix of family types was rather different between different tenure types in the same municipality over the whole studied period. Over the period this further came to be more and more the case also with regards to income groups and birth region groups. The mix of different population groups however tended to be similar within different tenure types in the same neighborhood (SAMS). While the entropy measures provide a purely descriptive picture, the findings thus suggest that tenure type mix could be more useful for creating social mix at the municipal level than for creating social mix at the neighborhood level. This does not mean that there is no scope for tenure mix at the neighborhood level. Tenure type mix at the neighborhood level could for example allow people to stay in the same neighborhood when moving between tenure types, which might be desirable both for individuals and for the stability of neighborhoods.

For a more robust picture of the development of residential segregation between population groups and the dependency in the distribution of people by tenure type and by population group, the sensitivity of the results to the group definitions should be tested. Further studies could also look beyond tenure types, at the importance of factors such as apartment sizes and housing costs, for the population mix in neighborhoods and municipalities.

References

Andersson, R. (2000). Etnisk och socioekonomisk segregation i Sverige 1990–1998. [Ethnic and socioeconomic segregation in Sweden 1990–1998]. In J. Fritzell (Ed.), *Välfärdens förutsättningar: arbetsmarknad, demografi och segregation [The conditions of welfare: labour market, demography and segregation]. SOU 2000:37* (pp. 223–266). Stockholm: Fritzes.

Andersson, R. (2007). Ethnic residential segregation and integration processes in Sweden. In K. Schönwälder (Ed.), *Residential segregation and the integration of immigrants: Britain, the Netherlands and Sweden.* WZB Discussion Paper, No. SP IV 2007-602 (pp. 61-90).

Andersson, R. (2008). Neighbourhood effects and the welfare state. Towards a European research agenda? *Schmollers Jahrbuch* (Vol. 128, pp. 1-14). Berlin: Duncker & Humblot.

Bergsten, Z., & Holmqvist, E. (2007). *Att blanda?: En undersökning av planerares och allmännyttiga bostadsbolags syn på planering för en allsidig hushållssammansättning*. IBF Researach report, 2007:1, Uppsala university, Uppsala.

Biterman, D., & Franzén, E. (2006). Boendesegregation. In *Social rapport 2006* (pp. 185-240): Socialstyrelsen

Biterman, D., & Franzén, E. (2010). Boendesegregaton. In *Social rapport 2010* (pp. 175-226): Socialstyrelsen.

Borevi, K. (2002). *Välfärdsstaten i det mångkulturella samhället [The Welfare State in a Multicultural Society*. PhD Thesis, Uppsala university, Uppsala.

Bråmå, Å., Andersson, R., & Solid, D. (2006). *Bostadsmarknadens institutioner och grindvakter i den etniskt segmenterade staten – exemplen Stockholm och Uppsala*. Norrköping: Integrationsverket.
Danermark, B., & Jacobson, T. (1989). Local housing policy and residential segregation. *Housing, Theory and Society, 6*(4), 245-256.
Enström Öst, C., Söderberg, B., & Wilhelmsson, M. (2014). Household allocation and spatial distribution in a market under ("soft") rent control. *Journal of Policy Modeling, 36*, 353-372.
Feitosa, F. F., Camara, G., Monteiro, A. M. V., Koschitzki, T., & Silva, M. P. S. (2007). Global and local spatial indices of urban segregation. *International Journal of Geographical Information Science, 21*(3), 299-323.
Galster, G. (2007). Should policy makers strive for neighborhood social mix? An analysis of the Western European evidence base. *Housing Studies, 22*(4), 523-545.
Galster, G., & Cutsinger, J. (2007). Racial settlement and metropolitan land-use patterns: Does sprawl abet black-white segregation? *Urban Geography, 28*(6), 516-553.
Graham, E., Manley, D., Hiscock, R., Boyle, P., & Doherty, J. (2009). Mixing housing tenures: Is it good for social well-being? *Urban Studies, 46*(1), 139.
Hårsman, B. (2006). Ethnic diversity and spatial segregation in the Stockholm region. *Urban Studies, 43*(8), 1341-1364.
Hårsman, B., & Quigley, J. M. (1995). The spatial segregation of ethnic and demographic groups: comparative evidence from Stockholm and San Francisco. *Journal of Urban Economics, 37*(1), 1-16.
Holmqvist, E. (2009). *Politik och planering för ett blandat boende och minskad boendesegregation. - ett mål utan medel? [Policy and planning for social and housing mix and decreased housing segregation. - A goal without means?]*. Phd Thesis, Uppsala university, Uppsala.
Holmqvist, E., & Bergsten, Z. (2009). Swedish social mix policy: a general policy without an explicit ethnic focus. *Journal of Housing and the Built Environment, 24*(4), 477-490.
Ihlanfeldt, K. R. (2004). Exclusionary land-use regulations within suburban communities: a review of the evidence and policy prescriptions. *Urban Studies, 41*(2), 261-283.
Lago, A., Johansson, A., Ljungdell, A., Heglert, H., Nygren, J., & Berggren, K. (August 23, 2010). Vi bygger småhus om ni bygger hyresrätter, *Svenska dagbladet*.
Li, Z., & Wu, F. (2008). Tenure-based residential segregation in post-reform Chinese cities: a case study of Shanghai. *Transactions of the Institute of British Geographers, 33*(3), 404-419.
Meng, G., & Hall, G. B. (2006). Assessing housing quality in metropolitan Lima, Peru. *Journal of Housing and the Built Environment, 21*(4), 413-439.
Miller, V. P., & Quigley, J. M. (1990). Segregation by racial and demographic group: Evidence from the San Francisco bay area. *Urban Studies, 27*(1), 3-21.
Musterd, S., & Andersson, R. (2005). Housing mix, social mix, and social opportunities. *Urban Affairs Review, 40*(6), 761-790.
Nelson, A. C., Dawkins, C. J., & Sanchez, T. W. (2004). Urban containment and residential segregation: a preliminary investigation. *Urban Studies, 41*(2), 423.
Nyström, L. (2008). *Social Mix i några länder – Leder boendeintegration till integration i hela samhället?*. Boverket.
O'Sullivan, D., & Wong, D. W. S. (2007). A surface-based approach to measuring spatial segregation. *Geographical Analysis, 39*(2), 147-168.

Reardon, S. F., & O'Sullivan, D. (2004). Measures of spatial segregation. *Sociological Methodology, 34*(1), 121-162.
Regionplane- och trafikkontoret Stockholms läns landsting. (2008). *Basområdestypologier* (Rapport 2008:5). Stockholm.
Statistics Sweden. (2005). *Geografin i statistiken – regionala indelningar i Sverige.* Meddelanden i samordningsfrågor för Sveriges officiella statistik 2005:2 [*Geography in statistics – regional divisions in Sweden.* Reports on Statistical Coordination for the Official Statistics of Sweden 2005:2]. Örebro.
Statistics Sweden (2009). *Longitudinell Integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier (LISA) 1990–2007.* Arbetsmarknads- och utbildningsstatistik 2009:1. [*Integrated database for labour market research (LISA) 1990–2007*. Background Facts Labour and Education Statistics 2009:1]
Theil, H. (1972). *Statistical decomposition analysis: with applications in the social and administrative sciences*: North-Holland.
Theil, H., & Finizza, A. J. (1971). A note on the measurement of racial integration of schools by means of informational concepts. *Journal of Mathematical Sociology, 1*(2), 187-193.

Appendix: Classification of tenure types

<table>
<tr><th rowspan="2"></th><th colspan="10">House type</th></tr>
<tr><th>Legal form of ownership</th><th>Farm where the housing unit is worth < 50000 SEK</th><th>farm</th><th>Detached house, worth < 50000 SEK</th><th>Detached house, year round dwelling for 1-2 families</th><th>Detached house, holiday dwelling for 1-2 families</th><th>Several detached houses for > 2 families</th><th>Detached house with premises</th><th>Block of apartments, mainly dwellings</th><th>Block of apartments, dwellings and premises</th><th>Block of apartments, mainly premises</th></tr>
<tr><td>Unknown</td><td colspan="10" rowspan="4">other</td></tr>
<tr><td>The State</td></tr>
<tr><td>Municipality or County Council</td></tr>
<tr><td>The Swedish Church</td></tr>
<tr><td>Individual</td><td colspan="2" rowspan="6">other</td><td colspan="6">owner-occupied housing</td><td colspan="3">private rental</td></tr>
<tr><td>Estate of dead person</td><td colspan="6" rowspan="2">other</td><td colspan="3">other</td></tr>
<tr><td>Swedish joint-stock company</td><td colspan="3">private rental</td></tr>
<tr><td>Tenant-owner cooperative</td><td colspan="9">tenant-owner cooperative</td></tr>
<tr><td>Municipal housing company</td><td colspan="9">public rental</td></tr>
<tr><td>Other</td><td colspan="9">other</td></tr>
</table>

Essay 4: Immigration, New Religious Symbols, and the Dynamics of Neighborhoods*

Co-authored with Matz Dahlberg

* We thank Jean-Paul Carvalho, Gunnar Myrberg, Jon Fiva, Peter Nilsson, and seminar participants at UCLA, UC Irvine, and the Institute for Housing and Urban Research at Uppsala University for helpful comments and discussions. We are grateful to Zara Bergsten, Gustav Engström, Eva Jirner, Linna Martén, and Marcus Mohall who at various stages helped us with the geo-coded information in the data. Financial support from the Jan Wallander and Tom Hedelius foundations is gratefully acknowledged.

1 Introduction

Over the last decades immigration to western European countries has increased substantially, and in most EU 27 countries immigration is now the main driver of population growth (European Comission, 2011). In 2010, 9.4 percent of the population in EU 27 was foreign born and of those about two thirds were born outside EU 27 (Eurostat, 2011).[1]

The increased immigration has resulted in a culturally and religiously more diverse European population. This diversification manifests itself in several ways, but one visible way is through new features in the cityscape. A much debated example is the emergence of mosques and with them minarets and public calls to prayer. In Denmark, the recent unexpected political consent to a new mosque in Copenhagen, for example, stirred up strong feelings (see, e.g., The Economist, Aug 17, 2013).[2] In Switzerland, a constitutional amendment banning the construction of new minarets was even subjected to a national referendum in 2009 and approved.[3]

While there are mosques, often with minarets, in most (all?) western European countries, there are few countries where public calls to prayer occur.[4] Despite the, often heated, debates in Europe over mosques and public calls to prayer, little is known about how these new religious symbols are valued by the surrounding society and how they affect the dynamics of neighborhoods. One important question is, for example, if they are drivers of segregation.

The purpose of this paper is to examine the effects of public calls to prayer from a mosque in a Western country on neighborhood dynamics (in terms of house prices and migration behavior among native- and foreign-born individuals). We take advantage of an unexpected political decision that led

[1] There are, however, some notable differences between countries. In 2010, the share of the foreign-born population was 1.2 percent in Poland and 4.3 percent in Finland but 15.2 percent in Austria and 14.3 percent in Sweden (Eurostat, 2011).

[2] The Grand Mosque of Copenhagen is a planned mosque which will be the first purpose-built mosque in Denmark and one of the largest in Europe. Local *by*-laws prohibit public calls to prayer.

[3] Calls to prayer have been a frequent argument against minarets. This was also the case in the Swiss referendum against minarets (see e.g., the information provided on the Wikipedia page http://en.wikipedia.org/wiki/Swiss minaret referendum, 2009; consulted on May 5, 2015). There have also been physical attacks on mosques. In Sweden, for example, the mosque in Trollhättan was burnt down in 1993, and as late as January 2015 the Uppsala Mosque was the target of attempted arson

[4] To our knowledge, the only western European countries where public calls to prayer existed to some extent before 2013 were Austria (Vienna), Germany, the Netherlands and the UK, while in Norway permission for public calls to prayer no louder than 60 decibels every Friday has been granted for a mosque in Oslo but this has never been put into practice (Allievi, 2009; Sveriges Radio P1 "Människor och tro" [Swedish Radio P1 (public radio), "People and faith"], April 18, 2013).

the way to the first public calls to prayer from a mosque in Sweden (the Fittja Mosque in Botkyrka municipality in the Stockholm region). This allows us to examine the question by combining the hedonic price theory of house price capitalization[5] with a quasi-experimental approach, yielding a hedonic difference-in-difference estimator.[6]

The quasi-experiments are defined by specific dates when different decisions were made regarding public calls to prayer from Fittja Mosque. The first, highly unexpected, political decision came on September 25, 2012 and the first public call to prayer was made on April 26, 2013. By using high-frequency house price data that provide precise information on the date when a property is sold and that covers a period from January 1, 2011 to April 30, 2014, we are in a good position to estimate the effect of the call to prayer events on house prices. This estimate provides us with a measure of the marginal willingness to pay for public calls to prayer. Likewise, by using migration data with precise information about migration dates, we are able to estimate the effects of the call to prayer events on the migration and sorting pattern around the mosque.

Our study relates to three different strands of literature. First, it relates to the literature examining the effects of immigration or ethnic-racial mix on neighborhood dynamics (see e.g., Accetturo, Manaresi, Mocetti, & Olivieri, 2014; Boustan, 2010; Card, Mas, & Rothstein, 2008; Sá, 2014; Saiz, 2003, 2007; Saiz & Wachter, 2011).

Second, it relates to a large literature examining the effects of immigration or ethnic-racial mix on different outcomes related to the welfare-state, such as public goods provision (see e.g., Alesina, Baqir, & Easterly, 1999, 2000), natives' preferences for redistribution (see e.g., Alesina, Glaeser, Sacerdote, Durlauf, & Levy, 2001; Alesina & Glaeser, 2004; Dahlberg, Edmark, & Lundqvist, 2012; Luttmer, 2001); trust and participation (see e.g., Alesina &

[5] Hedonic price theory (see e.g., Rosen, 1974; Tinbergen, 1956) applied to housing explains the capitalization of local attributes in housing values, and housing prices have been used to estimate the valuation of different features that are not explicitly traded in their own market such as noise (see e.g., Pope, 2008a), pollution (see e.g., Greenstone & Gallagher, 2008), school quality (see e.g., Black, 1999), and crime (see e.g., Pope, 2008c).

[6] It can be noted that while quasi-experimental studies and difference-in-differences estimators are common in some areas of the economic literature, notably in the empirical labor literature, to deal with reversed causality/omitted variables, potential problems from these sources have been an issue in the majority of the traditionally applied hedonic house price models. Only in relatively recent literature has the hedonic house price approach been combined with a quasi-experimental approach when trying to value local attributes. Using information on exogenous changes in the level of local amenities, or in the available information about the level of local amenities, these studies compare housing prices before and after a change in places subjected to the change and in other places, or in places subjected to the change to different degrees. For a discussion about this, see Gibbons and Machin (2008). Examples of this type of studies are Gibbons and Machin (2005) in relation to a transport innovation, Pope (2008a, 2008b, 2008c) in relation to airport noise, flooding, and crime, and Fiva and Kirkebøen (2011) in relation to school quality. The present paper places itself in this vein of the hedonic literature.

La Ferrara, 2000, 2002; Andreoni, Payne, Smith, & Karp, 2011; Vigdor, 2004), and rise of anti-immigrant/right-wing extremism parties (see e.g., Halla, Wagner, & Zweimüller, 2012).

Third, it relates to a fairly small literature on how housing prices are affected by proximity to houses of worship (see e.g., Babawale & Adewunmi, 2011; Brandt, Maennig, & Richter, 2014; Carroll, Clauretie, & Jensen, 1996).

Regarding house prices, our findings indicate that the public calls to prayer made housing closer to the mosque relatively more expensive. Regarding sorting, we find no indications of native flight/native avoidance in the neighborhoods close to the mosque following the call to prayer events. Given the original character of Botkyrka municipality with a high share of immigrants, our findings are consistent with a story of local revitalization from public calls to prayer in neighborhoods where native-immigrant sorting has already taken place.

The rest of the paper is organized as follows: Section 2 contains information on immigration and mosques in Sweden and presents the political process preceding the first public call to prayer from a Swedish mosque. To get a sense of the extent to which the general public was informed about the call to prayer events, Section 3 discusses the media and Internet attention given to the specific events. Section 4 discusses what effects on neighborhood dynamics that could be expected from public calls to prayers. Section 5 presents the empirical analyses for the effects on house prices and Section 6 the empirical analyses for the effects on migration behavior. Section 7 concludes.

2 Sweden: Immigration, religion, mosques, and public calls to prayer

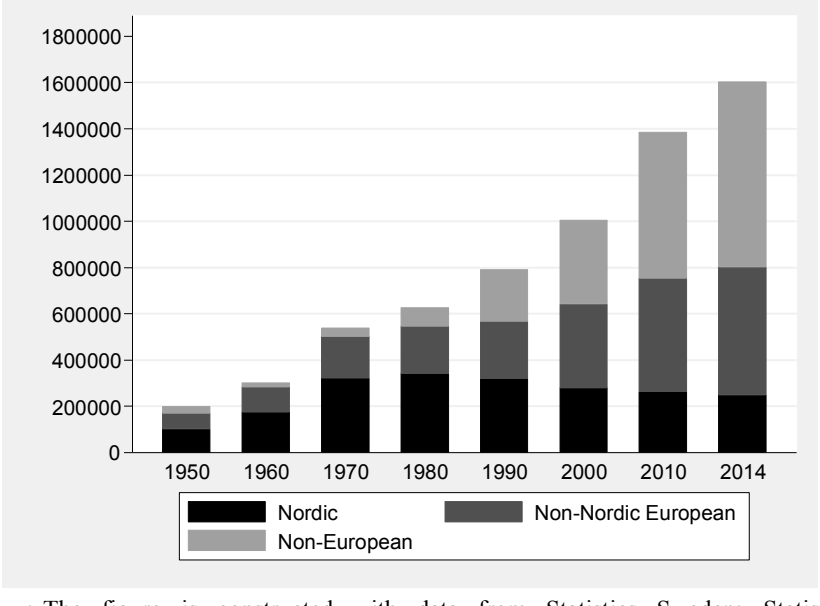
In this section we will briefly discuss, in turn, immigration to Sweden during the last decades, religion and mosques in Sweden, and the political process leading towards the first public call to prayer from a mosque in Sweden.

2.1 Immigration to Sweden

In Sweden, like in many other western European countries, the size and character of immigration have changed over the last decades. In 1970, less than seven percent of the Swedish population was foreign born (Statistics Sweden, Yearbook of Sweden 2012, table 4.30 "Population by country of birth"), and of those the large majority had arrived as labor immigrants from another Nordic or European country in the 1950s and 1960s. However, starting from the middle of the 1980s, Sweden has seen an important immigration of

refugees from other parts of the world. Consequently, there has been a drastic change over the last three to four decades in the number and origin of the foreign-born population in Sweden. The evolution and changing pattern of the foreign born-population in Sweden can be seen in *Figure 1*. The share of people born in the Nordic countries is decreasing over time while the share of people born in non-European countries is increasing. By the end of 2014, approximately 16 percent of the Swedish population was foreign-born and about half of these were born outside Europe (http://www.scb.se, i.e., the webpage of Statistics Sweden: Statistics database/Population/Population statistics/Foreign-born persons/Foreign-born persons in Sweden by country of birth, age and sex. Year 2000-2014, consulted on May 2, 2015). Among the ten most common countries of origin are Iraq, Iran, former Yugoslavia, Turkey, and Somalia; countries where Islam is an important religion.

Figure 1. Number of foreign born persons in Sweden by birth region, 1950-2014.



Source: The figure is constructed with data from Statistics Sweden: Statistical Yearbook of Sweden 1950 (table 55 "Persons born abroad, aliens and naturalized population"); Statistical Yearbook of Sweden 1960 (table 50 " Persons born abroad by country of birth and aliens by citizenship"); Statistical Yearbook of Sweden 2012 (table 4.30 "Population by country of birth"); http://www.scb.se (Statistics database/Population/Population statistics/Foreign-born persons/Foreign-born persons in Sweden by country of birth, age and sex. Year 2000-2014, consulted on May 4, 2015).

2.2 Religion in Sweden

There are no official statistics on religious beliefs in Sweden, but Christianity is the most prevalent religion. The Church of Sweden, separated from the

state since 2000, is an Evangelical Lutheran church of which around 70 percent of the population were members in 2011. Furthermore, the Swedish Commission for Government Support to Faith Communities (SST) estimates that in the same year there were about 110,000 persons belonging to a Muslim faith community in Sweden, which corresponded to just over 1 percent of the population. This makes Islam the second largest religion in Sweden.

2.3 Mosques in Sweden

The first building in Sweden constructed to be a mosque was the Nasir Mosque in Gothenburg taken into service by the Ahmadiyya Muslim community in 1976. Since then, five other mosque buildings have been constructed (year taken into use): Malmö Central Mosque (1984); Trollhättan's Mosque (1985); Uppsala Mosque (1995); Fittja Mosque (in the Stockholm region) (2007); and Gothenburg's Mosque (2011). In Stockholm there is moreover Stockholm's Mosque (2000), a power plant from the beginning of the 20th century redeveloped into a mosque. The mosques listed above are also those with minarets. There are also a number of mosques in buildings that were previously used by other faith communities for example in Gävle in a previous Methodist church and in Örebro in a building previously used by Jehovah's Witnesses. Finally, there are a number of more or less permanent Muslim places of worship in apartments and other premises.

2.4 Public calls to prayer in Sweden

The first public call to prayer from a Swedish mosque took place on April 26, 2013. The mosque in question was Fittja Mosque in Botkyrka municipality (Stockholm region). The call was preceded by a political process started by a citizen's proposal to Botkyrka municipality in January 2012, to allow public calls to prayers from Fittja Mosque once every Friday (or less frequently).[7] However, the development plan for the block where the mosque is built includes a safeguard that prohibits public calls to prayer. Furthermore, municipal regulations require permission from the police authority for messages through loudspeakers in public places.

Yet, in the second half of September 2012, pronouncements from the environmental health committee and the society construction committee in Botkyrka municipality unexpectedly expressed support for public calls to prayers. The environmental health committee recommended the municipal council to consider the citizen's proposal as answered, with the explanation that the committee did not consider public calls to prayers to be disturbing noise as long as they were kept below the noise limits and further that the

[7] The information in this section is from Botkyrka municipality's public documents on the case.

question about public calls to prayer was not a matter for the municipality but for the police authority. The environmental health committee further recommended that the municipal regulations should be reviewed so that public calls to prayer would not require permission from the police authority. The society construction committee also recommended the municipal council to consider the citizen's proposal as answered, with the explanation that the safeguard prohibiting public calls to prayer was to be considered a nullity with no support in existing law. According to the society construction committee, for such a safeguard to be valid it had to refer to a decibel level conflicting with the safeguards in environmental law (Miljöbalken) and not to the message transmitted.

February 21, 2013, Botkyrka municipal council decided in line with the committee pronouncements but also laid down that permission was required from the police authority.[8] The police authority in Stockholm County granted the permission for a period of one year on April 11, 2013. The permission was coupled with conditions relating to the direction of the loudspeakers (away from the closest housing), the establishment of a program for internal control of the calls, and the requirement to inform those living within one kilometer of the mosque before the first public call to prayer. The permission was later renewed.

At the time of writing, June 2015, none of the other mosques in Sweden had yet taken measures in view of public calls to prayer.

3 Public calls to prayer from Fittja Mosque: Events in focus

When studying the effects of public calls to prayers on neighborhood dynamics we focus on two events in the process towards the first public call to prayer from Fittja Mosque: the pronouncements from the municipal committees supporting public calls to prayer (September 25, 2012) and the actual permission from the police authority (April 11, 2013) shortly after followed by the first public call to prayer (April 26, 2013).[9]

[8] The decision by Botkyrka municipal council was also preceded by a pronouncement from the municipality's executive board (February 4, 2013) that was very similar to the council's decision. In addition, the decision by the municipal council was preceded by a dialogue with municipality residents through three focus groups. The subject of the dialogue was the position of religion in the public space. The outcomes of the focus groups were mainly that the municipality should continue working on the issue and that the dialogue should be broadened to include increased numbers of residents and perspectives.

[9] To be precise, September 25 was the date for the pronouncement from the society construction committee. The environmental health committee made their pronouncement as early as September 17, but it was not posted until September 26, 2012. Furthermore, the pronouncement received very little media attention both before and after it was posted. It thus seems reasonable to consider September 25, 2012, as the start of the event. Also, there are too few

The reason for focusing on the pronouncements from the municipal committees is that they gave reasons to believe that public calls to prayer would eventually become a reality. There was a broad political consensus in Botkyrka municipality that the municipality could not continue prohibiting public calls to prayer as such.[10] Furthermore, from the investigations of the society construction committee and the environmental health committee, it seemed uncertain whether there were any legal grounds at all, municipal or other, to prohibit public calls to prayer as such. In light of this, the decision of the Botkyrka municipal council, in February 2013, was expected. Nor did it have any practical consequences. The practical consequences instead appeared in connection with the second event we focus on; the permission for public calls to prayer from the police authority and the ensuing calls to prayer, April 2013.

For an event to have behavioral effects, in terms of a capitalization effect on house prices or in terms of an effect on migration decisions, it is important that people are informed about the issues at hand. We think this is likely to be the case, not least due to the widespread media coverage around the call to prayer events. The pronouncement from Botkyrka's society construction committee supporting public calls to prayer, which was made on Tuesday September 25, 2012, was the start of extensive media attention. On Thursday of the same week, the web-editions of all the major Swedish newspapers as well as national and local radio and television featured the news that Fittja Mosque could be the first Swedish mosque with public calls to prayer. On Friday, September 28, the news was in the paper versions of the major newspapers. The news was discussed on Internet forums and Swedish Wikipedia reported that Fittja Mosque was the first mosque in Sweden that had been granted permission for public calls to prayer from its minaret.[11] The first public call to prayer on April 26, 2013, was also covered by media; for example it was broadcast live on national television.

To obtain further indications that the public was well informed about the events, we first look at the daily number of visits to the Swedish Wikipedia page *Fittja moské* (Swedish for "Fittja Mosque") and, second look at weekly Google trends for the search term *böneutrop* (the Swedish word for "call to prayer").

---

observations in the data to separate effects from the police permission and from the actual start of the public calls to prayer, so these two events are studied as one starting on April 11, 2013.

[10] All six parties in the environmental health committee supported the committee's pronouncement. The only party out of six in the society construction committee that did not support the committee's pronouncement was the Christian Democrats, and the disagreement was more due to the form than the content or the practical implications of the pronouncement. The only party out of nine in Botkyrka municipal council that did not support the decision of February 21 2013, was the Sweden Democrats, a nationalistic party.

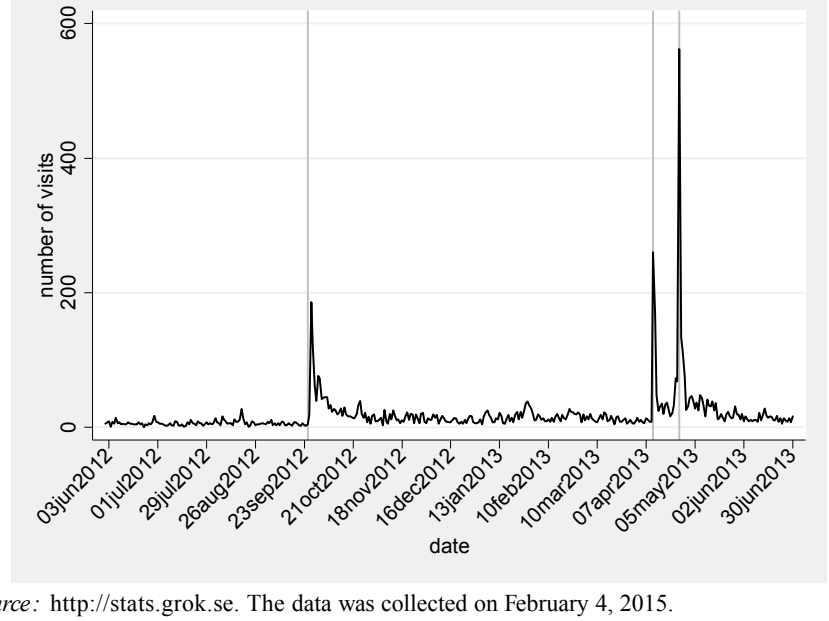
[11] This information was removed February 18, 2013. On April 26, 2013 information was added that the first public call to prayer had been conducted on that day.

From http://stats.grok.se/, where it is possible to get information on the number of daily visits to each Wikipedia page, we have obtained the number of visits to the Swedish Wikipedia page *Fittja moské* for the time period June 1, 2012, to June 30, 2013. From the data, presented in *Figure 2*, it can be seen that there are three clear peaks in the number of visits, coinciding with the September 2012 events, the police permission on April 11, 2013, and the first public call to prayer from Fittja Mosque on April 26, 2013.

From Google trends (http://www.google.com/trends/) it is possible to get a weekly index for the popularity of a specific search term. The index goes from 0 to 100, with 100 being the week with the highest popularity. The index is relative; Google trends do not provide the total number of searches. We obtained statistics for the search term *böneutrop* for the period June 3, 2012 to June 30, 2013. As is clear from the index for this search term, given in *Figure 3*, there are clear peaks at the same points in time as in *Figure 2*.

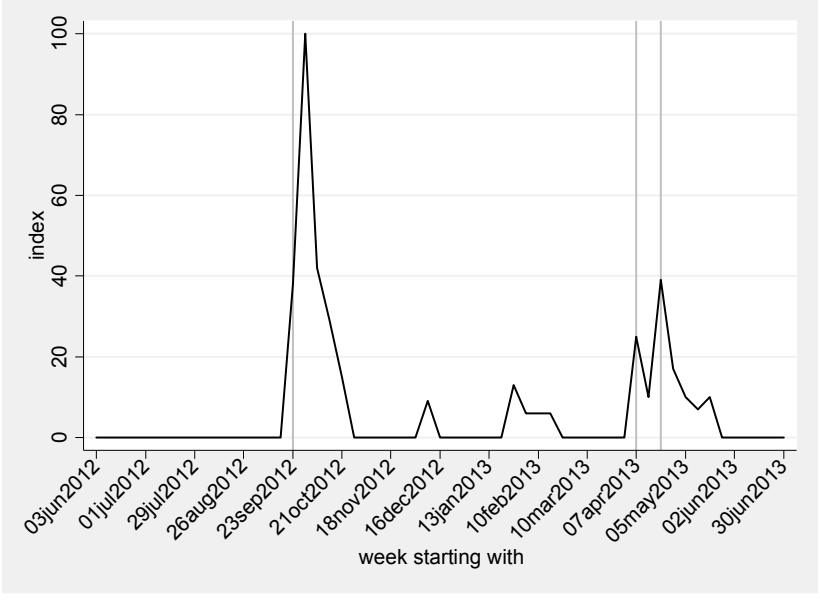
The peaks in *Figure 2* and *Figure 3* indicate that people were informed about the process related to public calls to prayers from Fittja Mosque, and also that the September 2012 and April 2013 events attracted the most interest. This strengthens the assumption that the events can cause capitalization effects in the housing market and affect migration patterns.

Figure 2. Daily number of visits to the Wikipedia page Fittja moské (Swedish for Fittja Mosque) for the period June 1, 2012-June 30, 2013.



Source: http://stats.grok.se. The data was collected on February 4, 2015.

Figure 3. Weekly Google trends for the search term böneutrop (Swedish for "call to prayer") for the period June 3, 2012-June 30, 2013.



Source: http://www.google.com/trends. The index-statistics were collected on February 4, 2015.

4 What effects should we expect from the public calls to prayer?

What effects should we expect from the public calls to prayer on house prices and migration patterns in the neighborhoods close to the mosque? To answer this we first consider the channels through which calls to prayer could have an effect on migration patterns and housing prices. There are several possible, not mutually exclusive, channels. One, and perhaps the one that first comes to most people's minds, is related to the sound of the public calls to prayer. The sound can be valued positively by some households, and negatively by others. However, we do not think this is an important channel in our case. There are few public calls to prayer (one every Friday at 1 pm), the loudspeakers are directed away from residential housing, and there are only a limited number of houses in the direct vicinity of the mosque.

Another, and we think, more likely channel for an effect from public calls to prayer on migration patterns and housing prices is the symbolic value of the calls. The calls to prayer can symbolize Islam and also create expectations that other expressions of Islam will follow in the neighborhoods close to the mosque. This can be valued differently by different

households, and it seems reasonable to assume that living close to the symbol of the public calls to prayer would generally be more positively valued by Muslims than non-Muslims.[12] This in turn can make people believe that the calls to prayer will attract Muslims to the neighborhoods close to the mosque and thus increase the share of Muslims in these neighborhoods. While this may increase the value of living close to the mosque for some Muslims if they have preferences for living with other who share their faith, it would decrease the value of living close to the Mosque for those who exhibit negative preferences towards living with Muslims or immigrants.[13] It is also possible that the media attention related to the public calls to prayer process informed the general public about the existence of the mosque itself, and that knowledge about the location of the mosque changed the valuation of proximity to it, in line with the above. Furthermore, the political support for public calls to prayer could be seen as an indication that Botkyrka as a municipality is welcoming toward ethnic and religious minorities, which could affect migration patterns in Botkyrka municipality as a whole, again in line with the above.[14]

In conclusion, it seems reasonable, if anything, to expect a relative increase in the share of Muslims, and – to the extent that Islam is more prevalent among immigrants than among natives – immigrants, in the neighbor-

[12] Bisin and Verdier (2000) build a model for the intergenerational transmission of ethnic and religious traits through family socialization and marital segregation decisions. Cultural transmission is modeled as an interaction between socialization inside the family and socialization outside the family. The socialization within the family for example takes place through parents spending time with their children, the choice of appropriate neighborhoods, schools, and acquaintances, and through attending religious service. Socialization outside the family occurs in society at large via imitation and learning from peers and role models. Bisin and Verdier (2000) show that this cultural transmission model produces different behavior for cultural minorities and majorities with respect to their efforts to marry monogamously and to socialize children to their own trait: minorities, all other things being equal and in equilibrium, have more highly segregated marriage markets, and more intensely exercise effort in directly socializing their children. In accordance with this model, one can hypothesize that for Muslims, who are a religious minority in Sweden, choosing a neighborhood close to the public calls to prayer could be important, while people belonging to the majority population (i.e., the Church of Sweden or with no strong religious belief) could be more indifferent about whether or not they live in the vicinity of public calls to prayers since their children with large probability will be socialized into the majority beliefs anyway. The following comment, from the person who wrote the citizen proposal on public calls to prayer, illustrates that public calls to prayers could be important for cultural transmission: "There should be public calls to prayer from a Mosque. If our children lived in a Muslim country they would hear calls to prayer five times a day. Now, they can at least hear them once a week." (Article on the website of the Swedish public service television company, April 29, 2013, http://www.svt.se/nyheter/regionalt/abc/premiar-forboneutrop, consulted on June 18, 2015).

[13] Of course, not all immigrants are Muslim, and not all Muslims are immigrants. However, to the extent that Islam is more prevalent among immigrants than among natives, an increased share of Muslims could be assumed to imply an increased share of immigrants.

[14] Descriptive studies (see e.g., Aldén, Hammarstedt, & Neuman, 2015; Bråmå, 2006) indicate that native avoidance/flight from neighborhoods with a marked increase in the share of foreign born inhabitants might be an issue also in Sweden.

hoods close to the mosque compared to other neighborhoods following the process leading to public calls to prayer.

The expected effect on house prices is less clear. Mirroring the argumentation in Saiz and Wachter (2011) with respect to neighborhoods with a growing immigrant population, even if Muslims/immigrants have preferences for living close to a mosque with public calls to prayer, this does not necessarily imply relatively higher prices close to the mosque as long as there are mobile non-Muslims/native price arbitrageurs. However, if non-Muslims/natives have preferences for *not* living close to a mosque with public calls to prayer, whether because of the calls themselves, what they symbolize or because of an expected increase in the Muslim/immigrant population close to the mosque, the calls to prayer may be associated with a relative negative impact on housing prices close to the mosque.

Looking to empirical studies, we are only aware of one on the effect of religious symbols other than churches on house prices; Brandt et al. (2014).[15] Brandt et al. (2014) use cross-sectional data to study the impact of houses of worship, including churches, mosques, synagogues, Buddhist and Hindu temples, on the prices of condominiums in Hamburg, Germany. Brandt et al. (2014) find a positive price premium for condominiums within 1000 meters of houses of worship and that the premium for mosques does not differ from the premium for houses of worship of other religions. Furthermore, they find no effect on condominium prices from church bell ringing.

Looking instead at empirical studies on the effect of increased immigration on relative house prices within urban areas, for example Saiz and Wachter (2011) for the US, Sá (2014) for the UK and Accetturo et al. (2014) for Italy, they find that the growth of a neighborhood's immigrant share is associated with relatively lower housing value appreciation. This is consistent with the idea that natives are willing to pay a premium for living in predominantly native areas. Saiz and Wachter (2011), however, also find that the association between growing immigrant density and relative housing value depreciation is stronger in neighborhoods where the population was initially predominantly white and in neighborhoods that were initially perceived as more valuable. As Saiz and Wachter (2011) argue, this is consistent with the view that in neighborhoods that were already minority-dense and poor, the marginal natives who still remained had lower willingness-to-pay for segregation. Thus, a growing immigrant share did not have a negative impact on relative housing values in these neighborhoods where socioeconomic sorting had already taken place, but rather served to revitalize the neighborhoods.

In relation to this it can be noted that prior to the process leading to the first public calls to prayers, Botkyrka municipality was the municipality

[15] For a review of the slightly larger literature on the effect of churches on house prices, see Brandt et al. (2014).

in the Stockholm region with the largest share of immigrants. In the end of 2011, 38 percent of the population in Botkyrka municipality was born abroad. In the northern part of Botkyrka, where Fittja Mosque is situated, the share was even higher, 56 percent. It is thus possible that the process leading to public calls to prayer will serve more to revitalize the neighborhoods close to the mosque than to scare away natives with preferences for segregation, resulting in an upward pressure on house prices.

Finally, it can be noted that there is a speculative component to housing. A household might be perfectly happy to live close to a mosque, public calls to prayer and Muslims, but if the household believes that a large share of the market discounts such location attributes, the household might be less willing to own housing close to the mosque following the process leading to public calls to prayers.

5 Effects on house prices

In this section we investigate the effects of the first public call to prayer from a mosque in Sweden on house prices.

Before presenting the baseline results in section 5.3, we present the data (section 5.1) and the econometric specification (section 5.2) used in the analyses. Section 5.4 presents placebo estimations, section 5.5 presents some sensitivity analyses, and section 5.6 presents some further analyses (effects on list prices and synthetic control estimations).

5.1 Data

Our house price data comes from Svensk Mäklarstatistik AB.[16] The data from Mäklarstatistik is based on information reported by real estate agents after the close of a sale. The data set includes the sales of dwellings in housing cooperatives (mainly apartments), privately owned houses and cottages and according to Mäklarstatistik covers about 80 percent of all housing sales made in Sweden.[17]

The data contain high-frequency (daily) house sales data with coordinates, housing characteristics (e.g., living area, number of rooms, plot area, year built, monthly fee, elevator, balcony), list (ad) prices and final (contract) price, list (ad) date, contract date, and date of possession. To the data

[16] Svensk Mäklarstatistik AB is an enterprise owned by two broker firms and two trade associations for brokers.

[17] Row houses, semi-detached houses and detached houses can be part of housing cooperatives or be privately owned. Condominiums, i.e., privately owned apartments in apartment blocks, have on the other hand only been allowed since 2009 in Sweden and only in buildings produced for that purpose or that have not served for housing the last eight years. The number of condominiums was still very small in the period under study in this paper.

provided by Mäklarstatistik, we add distance from each object to geo-coded urban data.

To improve on the precision of the coordinates provided by the real estate agents, for each sale we checked the coordinates provided by Google for the address in question. We use only observations for which the coordinates provided by real estate agents and Google are within 100 meters of each other. For those observations that are within 100 meters of each other, we use the coordinates given by Google.

The information in the data set makes it well suited for combining a quasi-experimental approach with traditional hedonic price theory (the housing sales are well defined in both time and space). We have data from January 1, 2011, to April 30, 2014.

For the baseline analysis, we extract two main samples from the house price data set. The first sample contains all apartments in housing cooperatives sold in Botkyrka municipality.[18] This sample will be used in a distance difference-in-differences specification (see next section) to test how the public calls to prayer process affected the relative valuation of housing at different distances from the mosque within Botkyrka municipality. *Figure 4* shows a map of Botkyrka, with the municipal border (marked by a thick grey line), water (light grey), built-up areas (halftone screen) and neighborhoods (SAMS[19], marked by solid black lines). As can be seen from the map, Fittja Mosque is situated in the northeast corner of Botkyrka.

The second sample contains all apartments in housing cooperatives sold in the northern part of Botkyrka or, more precisely, in the most northern of the areas marked by halftone screen in *Figure 4*, an area contained within approximately 4 km from the mosque. As can be seen from the figure, this

[18] We only consider apartments in housing cooperatives in this paper since there were few sales of privately owned houses in the neighborhoods close to Fittja Mosque in the period we study. Overall, in 2013, the dwellings in Botkyrka municipality consisted of 44 percent rental housing, 24 percent apartments in housing cooperatives, and 32 percent privately owned houses (http://www.scb.se , i.e., the webpage of Statistics Sweden: Statistics database/Housing, construction and building/Dwelling stock/Number of dwellings by region, type of building and tenure (including special housing) year 2013 - 2014. Consulted on April 29, 2015.

[19] SAMS, "Small Areas for Market Statistics", have been used frequently in Swedish studies as the formal division closest to neighborhoods. The SAMS-classification was created by Statistics Sweden to satisfy demand for small area statistics from other users than municipalities. The objective was to create fairly homogeneous residential areas of about 1,000 inhabitants each. In larger municipalities, the SAMS-classification is based on municipal subdivisions used for intra-municipal and sometimes regional planning and administration and in smaller municipalities it is based on election districts. The SAMS- classification came into use in 1994 and has remained unchanged since then apart from minor adjustments, for example to adapt the SAMS-borders to municipal borders. For more information, see Statistics Sweden (2005).

is a continuous and well delimited area. In the north and east the area is delimited by water and in the south and west by forests and fields.[20]

The original sample for Botkyrka (northern Botkyrka) contains 1821 (641) sales of apartments in housing cooperatives. After excluding sales where the Google coordinates differ by more than 100 meters from the coordinates provided by the real estate agents, newly built apartments – since they are sold at fixed prices – and apartments with missing or unusual contract price[21], living area[22], monthly fee[23], number of rooms[24], and/or with an unclear contract date[25], we are left with 1332 (523) sales of apartments in housing cooperatives. Table 1 shows summary statistics for the two samples described above.

[20] Since Botkyrka and its built-up areas are well-defined by natural borders, we have chosen to use only the housing sales in Botkyrka in the analyses. In the north, the municipality of Ekerö is disconnected from Botkyrka by a large body of water and there is little housing in the part of Ekerö closest to Fittja Mosque. Across the water to the east, the municipality of Huddinge is closer to Botkyrka than is Ekerö, but the number of sales in the neighborhoods close to Fittja Mosque were few in the period we study. Also, by using only observations from one municipality, Botkyrka, we get a more homogeneous study area (e.g., in terms of local politics). Further, it can be thought that the issue of public calls to prayer is more likely to affect the valuation of proximity to the mosque within the municipality formally concerned and that more symbolic aspects of public calls to prayer than the sound itself are limited by physical barriers such as water. We have nevertheless estimated distance specifications for northern Huddinge. From these estimations we find no significant effects of the public call to prayer events on house prices in Huddinge (these results are available on request). It is however hard to know whether this is due to the fact that there are no effects from the call to prayers on house prices in Huddinge or if it is due to increased uncertainty in the point estimates resulting from the small number of observations.

[21] Apartments sold for less than 1,000 SEK.

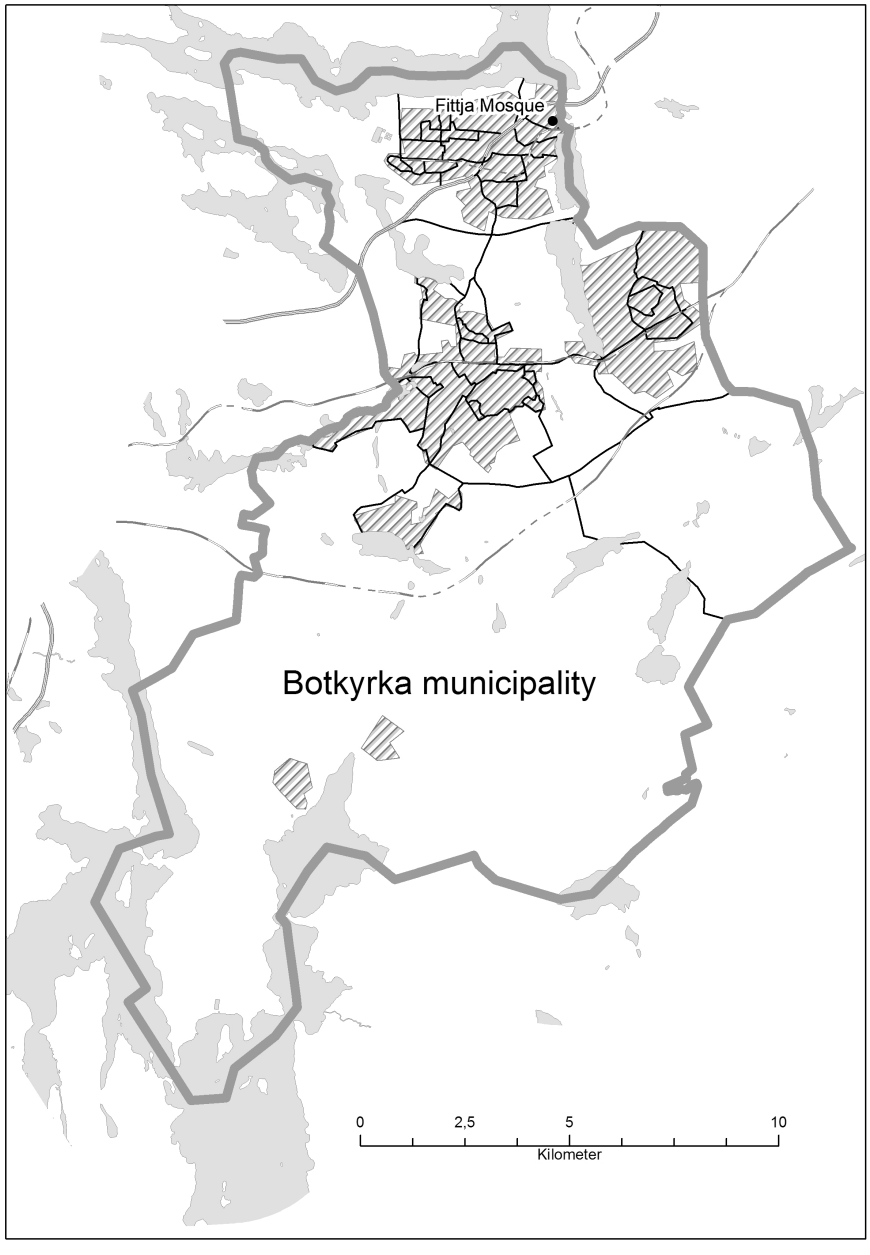
[22] Apartments with a living area smaller than or equal to 10 square meters or larger than 400 square meters.

[23] Apartments with a monthly fee of 0 or a monthly fee larger than 100,000 SEK.

[24] Apartments with 0 rooms or more than 20 rooms.

[25] Apartments with contract date before ad date.

Figure 4. Map of Botkyrka and Fittja Mosque



Note: Map created by Eva Jirner.

Table 1. Summary statistics house price samples

	Obs	Mean	Std. Dev.	Min	Max
Botkyrka municipality					
Contract price (SEK)	1332	1,166,411	391,151	250,000	3,775,000
Living area (square meters)	1332	73.1	22.7	24.0	169.0
Monthly fee (SEK)	1332	4,388	1,408	1,228	9,667
Number of rooms	1332	2.7	1.1	1.0	6.0
Distance to Fittja Mosque (1000 m)	1332	4.63	2.05	1.53	10.50
Distance to nearest public transit (1000 m)	1332	0.76	0.71	0.08	4.29
Distance to nearest state liquor store (1000 m)	1332	2.01	1.82	0.05	4.98
Building year	1028	1981	15	1957	2013
Northern Botkyrka					
Contract price (SEK)	523	1,028,614	299,043	400,000	2,150,000
Living area (square meters)	523	70.6	16.4	44.0	131.0
Monthly fee (SEK)	523	4,190	1,027	2,525	9,317
Number of rooms	523	2.6	0.9	1.0	5.0
Distance to Fittja Mosque (1000 m)	523	2.46	0.52	1.53	3.62
Distance to nearest public transit (1000 m)	523	0.45	0.16	0.11	0.91
Distance to nearest state liquor store (1000 m)	523	0.55	0.36	0.10	1.73
Building year	271	1980	10	1970	2004

5.2 Econometric specification

To examine the effects of public calls to prayer from Fittja Mosque on house prices, we adopt a quasi-experimental approach using a distance-defined difference-in-difference estimator. In the distance specification it is assumed that apartments at different distances from Fittja Mosque are treated differently, with properties close to the mosque being more treated. The quasi-experiments are the pronouncements from Botkyrka municipal committees supporting public calls to prayer (September 25, 2012) and the police permission to conduct public calls (April 11, 2013) followed shortly after by the first public call to prayer (April 26, 2013). We thus compare house prices close to and further away from Fittja Mosque before the September event, after the September event, and after the April events. This approach helps us to abstract from differences between places that are fixed over time, as well as from changes over time that are unrelated to the process surrounding the start of public calls to prayer from Fittja Mosque.

The distance specification is a semi-log hedonic price function that takes the following form in its richest specification:[26]

[26] Semi-log models are common in empirical hedonic studies (see e.g., Gibbons & Machin, 2005; Pope, 2008a, 2008b, 2008c). Moreover, studies that compare simulated "true" hedonic equilibria with various forms of estimated hedonic functions (Cropper, Deck, & McConnell, 1988; Kuminoff, Parmeter, & Pope, 2010) seem to indicate that a semi-log model could be suitable for estimating the implicit prices in a setting like the one in the present study.

$$\ln(p_{ijt}) = \beta_0 Dist_i + \beta_1 Dist_i \times D_{25Sept} + \beta_2 Dist_i \times D_{11April} + \gamma X_{it} + \varphi' Z_{it} + \mu_{time} + \lambda_j + \varepsilon_{ijt} \quad (1)$$

where $ln(p_{ijt})$ is the natural logarithm of the final (contract) price of apartment i in neighborhood j at day t. $Dist_i$ is a variable measuring the distance to Fittja Mosque from apartment i. The variable should control for the effect on house prices from proximity to the mosque itself as well as from omitted variables spatially correlated with the mosque. D_{25sept} is a dummy variable that takes the value 1 if apartment i was sold after September 25, 2012, and 0 otherwise. D_{25sept} captures any changes in selling prices following the pronouncement from the municipal committees supporting public calls to prayer. $D_{11April}$, which is a dummy variable that takes the value 1 if apartment i was sold after April 11, 2013 and 0 otherwise, captures any changes in selling prices following the actual permission from the police authority, followed shortly after by the first public call to prayer. X_{it} is a vector of structural characteristics for apartment i at day t. In the baseline specification the vector includes monthly fee, living area and number of rooms. Z_{it} is a vector of location attributes of apartment i at day t; the vector includes distance to nearest city center (proxied by distance to the nearest state liquor store[27]) and distance to nearest subway or commuter train station.[28] μ_{time} is a vector of time (month-by-year) fixed effects controlling for time variation in the data unrelated to the political process surrounding the public calls to prayer. To account for time-constant neighborhood effects, equation 1 is also augmented by neighborhood-fixed effects, λ_j, taking the value 1 if apartment i is situated in neighborhood (SAMS) j, and 0 otherwise. ε_{ijt} are the error terms.

Given the control variables, and under the assumptions of common time variation over the studied area – that there are no omitted variables systematically related to the September 2012 and April 2013 events under study that affect house prices – and that changes in house prices did not drive the timing of the events, β_1 should measure how the valuation of proximity to the mosque changed after the committee pronouncements supporting public calls to prayer and β_2 how the valuation of proximity to the mosque changed after the police authority gave permission for public calls to prayer.

The distance specification in equation 1 is similar to the one used by McMillen and McDonald (2004) to study the effect of a new rapid transit line from downtown Chicago to Midway Airport on house prices. It is also

[27] In Sweden, alcoholic beverages with more than 3.5% alcohol content are sold only at specific, state-owned, stores, which are typically located in a city or shopping center.

[28] Given the previous literature on the effect of proximity to religious buildings and especially churches, we considered including a variable controlling for proximity to churches. In Botkyrka municipality, however, such a variable turned out to be highly correlated with distance to nearest city center and nearest subway or commuter train station.

similar to the distance specification used by Aragón and Rud (2013) to study the local economic impact of a large gold mine in Northern Peru.

5.3 Baseline results

The baseline results are obtained from the estimation of equation 1 on data from Botkyrka, the municipality in which Fittja Mosque is located. In the first set of estimations, we use data from the whole of Botkyrka. In the second set of estimations, we restrict our data to include only observations from the northern part of Botkyrka, i.e., the part where the mosque is located (the most northern of the areas marked by halftone screen in *Figure 4*).

The results when we use data from the whole of Botkyrka are given in Table 2. In the first column, we control for apartment-specific attributes (i.e., monthly fee, living area, and number of rooms). In the second column, we add location-specific control variables to the specification (i.e., distance to the nearest subway station and distance to the nearest city center). In the third column, we finally add neighborhood-specific fixed effects to the specification to control for unobserved variables that might affect the house prices within a neighborhood. Time (month-by-year) fixed effects are included in all specifications. The standard errors are clustered on neighborhoods to allow for potential correlation in house prices within each neighborhood.

There are three main aspects to note from the results in Table 2. First, the point estimates on the difference-in-differences variables are stable over the different specifications (i.e., over the different columns). Second, the difference-in-differences estimates are, with one exception, statistically significant at least at the ten percent significance level. Third, the difference-in-differences estimates do also seem to be economically important; the point estimates indicate that the house prices increase by approximately 1.4–1.7 percent (depending on exact specification) *less* per kilometer away from the mosque after the September 2012-events and again by approximately 1.2-1.4 percent *less* per kilometer away after the April 2013-events, implying that it seems to have become relatively more expensive to live near the mosque after the events. Hence, there seems to be a willingness to pay for public calls to prayer.

Table 2. Distance specifications, whole of Botkyrka

	(1) ln(price)	(2) ln(price)	(3) ln(price)
distance_mosq	0.039**	0.036*	0.044
	(0.017)	(0.020)	(0.036)
Distance_Sep2012	-0.013	-0.014*	-0.017**
	(0.008)	(0.008)	(0.007)
Distance_Apr2013	-0.014**	-0.012**	-0.013**
	(0.005)	(0.004)	(0.005)
Month-by-year dummies	Yes	Yes	Yes
Apartment attributes	Yes	Yes	Yes
Location attributes	No	Yes	Yes
Neighborhood dummies	No	No	Yes
Observations	1,332	1,332	1,332
Adjusted R-squared	0.62	0.67	0.77

Standard errors clustered on neighborhoods (SAMS) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

The results when we only use data from the northern part of Botkyrka are given in Table 3. As can be seen from the table, estimating equation (1) but limiting the area analyzed to the part of Botkyrka closest to the mosque, we find results in line with the results when using data from the whole of Botkyrka. The point estimates on the difference-in-differences variables are still negative for the September 2012 and April 2013-events, but generally larger when using only data from the northern part of Botkyrka. For the two richest specifications (columns 2 and 3), the point estimate for the September 2012-event is significant at the ten percent level. The richest specification (column 3) also indicates a further statistically significant change in house prices for the April 2013-events (at the one percent significance level).

Taken together, the results in Table 2 and Table 3 indicate that, if anything, living close to Fittja Mosque became relatively more expensive following the public calls to prayer events, implying that the calls to prayer were positively valued.

Table 3. Distance specifications, Northern Botkyrka

	(1) ln(price)	(2) ln(price)	(3) ln(price)
distance_mosq	0.026	0.057*	0.091**
	(0.034)	(0.026)	(0.033)
Distance_Sep2012	-0.042	-0.045*	-0.037*
	(0.024)	(0.022)	(0.019)
Distance_Apr2013	-0.012	-0.022	-0.037***
	(0.017)	(0.020)	(0.010)
Month-by-year dummies	Yes	Yes	Yes
Apartment attributes	Yes	Yes	Yes
Location attributes	No	Yes	Yes
Neighborhood dummies	No	No	Yes
Observations	523	523	523
Adjusted R-squared	0.80	0.80	0.85

Standard errors clustered on neighborhoods (SAMS) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

5.4 Placebo results

To check the baseline model specifications, we run a placebo experiment in which we lag all call to prayer events one year. That is, we assume that the decision on September 25, 2012 took place on September 25, 2011 and that the events in April 2013 happened in April 2012. Then we re-estimate equation (1) with this false difference-in-differences specification; in this specification we would like to see small and insignificant point estimates on the false difference-in-differences variables.[29]

The placebo estimations are presented in Table 4. In the first two columns, data for the whole of Botkyrka is used and in the last two columns data only from the northern part of Botkyrka is used. Within each set of estimations, the first column controls for apartment- and location-specific attributes while the second column controls for apartment- and location-specific attributes as well as neighborhood-specific fixed effects. We get "difference-in-differences estimates" that are both insignificant and closer to zero (in five of the cases much closer to zero) than in the baseline analysis. This adds some trustworthiness to the baseline results: even though we have fewer observations in the placebo analysis, creating more uncertainty in the estimates, the point estimates close to zero in five of the cases are reassuring.

[29] No data after September 25, 2012, is used in the placebo analysis.

Table 4. Placebo estimations

	(1) ln(price) whole of Botyrka	(2) ln(price) whole of Botkyrka	(3) ln(price) northern Botkyrka	(4) ln(price) northern Botkyrka
distance_mosq	0.034	0.043	0.078**	0.136**
	(0.023)	(0.046)	(0.030)	(0.045)
Distance_Sep2011	0.000	-0.005	-0.007	-0.020
	(0.015)	(0.014)	(0.023)	(0.015)
Distance_Apr2012	0.002	-0.011	0.012	0.002
	(0.008)	(0.011)	(0.024)	(0.015)
Month-by-year dummies	Yes	Yes	Yes	Yes
Apartment attributes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes
Neighborhood dummies	No	Yes	No	Yes
Observations	630	630	249	249
Adjusted R-squared	0.63	0.77	0.73	0.79

Standard errors clustered on neighborhoods (SAMS) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

5.5 Sensitivity analyses

We will conduct two types of sensitivity analyses. First, we check how sensitive the baseline results are to measurement errors in the housing coordinates. Second, we will examine how controlling for the age of the building affects the results.

5.5.1 Measurement errors in coordinates

In the estimations presented thus far we have used only data for sales where the coordinates provided by real estate agents and the Google coordinates for the address in question are within 100 meters of each other. To further limit the risk of measurement errors in the coordinates, we here restrict the data to sales where the difference between the locations indicated by the two pairs of coordinates is less than 50 meters. The idea is that the more similar the coordinates provided by Google and a real estate agent, the more likely is it that they correspond to the real coordinates for the sale, as there is always a possibility that coordinates from Google as well as from real estate agents could be wrong. We lose some observations, but from the estimates in Table 5 it is clear that the baseline results are essentially unaffected.[30]

[30] In the first two columns, data from the whole of Botkyrka is used and in the last two columns, data from the northern part of Botkyrka is used.

Table 5. Sensitivity Analyses: Stricter measurement of housing coordinates

	(1) ln(price) whole of Botyrka	(2) ln(price) whole of Botkyrka	(3) ln(price) northern Botkyrka	(4) ln(price) northern Botkyrka
distance_mosq	0.034*	0.043	0.049*	0.097**
	(0.019)	(0.038)	(0.026)	(0.037)
Distance_Sep2011	-0.013	-0.017**	-0.045*	-0.038
	(0.008)	(0.008)	(0.024)	(0.020)
Distance_Apr2012	-0.012**	-0.012**	-0.014	-0.028*
	(0.005)	(0.005)	(0.020)	(0.012)
Month-by-year dummies	Yes	Yes	Yes	Yes
Apartment attributes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes
Neighborhood dummies	No	Yes	No	Yes
Observations	1,274	1,274	482	482
Adjusted R-squared	0.67	0.77	0.81	0.85

Standard errors clustered on neighborhoods (SAMS) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

5.5.2 Controlling for the buildings' construction period

In the estimations presented thus far we have controlled for an important but still rather limited set of housing attributes. The reason is that for many sales, information on other housing attributes, such as balcony and elevator, are missing. To see whether our results are sensitive to which housing attributes are controlled for, we include another potentially important housing characteristic among the regressors in equation (1) – time period for when the building was constructed. We create dummy variables for the following building periods: Housing built before 1966, between 1966 and 1975, between 1976 and 1985, between 1986 and 1995, between 1996 and 2005, and after 2005.

When including dummies for building period we lose about a fifth of the observations for the whole of Botkyrka (since construction year is missing or unclear for these observations). From the first two columns in Table 6, it can nevertheless be seen that, for the whole of Botkyrka, estimations including dummy variables for building periods give results similar to those obtained in the baseline specifications. The difference-in-differences point estimates have the same, negative, sign and are very similar in magnitude to those in the baseline. The April 2013-events are also statistically significant, at the five percent significance level, in the richest specification.

For northern Botkyrka, including dummies for building period when estimating equation (1) leads to the loss of almost half of the observations. As can be seen from the last two columns in Table 6, the difference-in-

differences point estimates have the same sign as in the baseline analyses and they are all statistically significant at least at the five percent significance level. The one difference compared with the baseline results is that the point estimates are larger in magnitude, with the exception of the estimate for the effect of the April 2013 events in the richest specification (last column), which is very similar to the corresponding estimate in the baseline specification.

The results in Table 6 indicate that even though we control for another important housing characteristic (and even though we lose a lot of observations by doing so), we still get results that lead us to the same type of conclusion as in the baseline analyses.

All in all, although the exact values and statistical significance of the difference-in-differences estimates for the September 2012 and April 2013 events vary somewhat, the results from the placebo and sensitivity analyses confirm the main message from the baseline estimations – within Botkyrka municipality, it seems as though living closer to Fittja Mosque has, if anything, become relatively more attractive following the public call to prayer events.

Table 6. Sensitivity analysis: Controlling for construction period

	(1) ln(price) whole of Botyrka	(2) ln(price) whole of Botkyrka	(3) ln(price) northern Botkyrka	(4) ln(price) northern Botkyrka
distance_mosq	0.030*	-0.016	-0.021	0.059
	(0.017)	(0.063)	(0.035)	(0.038)
Distance_Sep2011	-0.011	-0.012	-0.071**	-0.073**
	(0.009)	(0.008)	(0.021)	(0.021)
Distance_Apr2012	-0.008	-0.013**	-0.033**	-0.037***
	(0.005)	(0.005)	(0.011)	(0.010)
Month-by-year dummies	Yes	Yes	Yes	Yes
Apartment attributes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes
Neighborhood dummies	No	Yes	No	Yes
Observations	1,028	1,028	271	271
Adjusted R-squared	0.70	0.80	0.85	0.85

Standard errors clustered on neighborhoods (SAMS) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

5.6 Further results: Effects on list prices and synthetic control estimations

In this section we will conduct two further analyses. First, we estimate the effects of the public call to prayer events on list prices. Second, we estimate the effect on final, contracted, mean house prices in Botkyrka using the synthetic control method.

5.6.1 Effects on list prices

The final, contracted, prices give us information about how the market reacted to the public call to prayer events. However, it is also of interest to examine how the call to prayer events affected *expectations* about how the market would react. Do the effects on expectations match the actual market effects?

Since list prices can be considered as a measure of the expected final price level, where typically the owner and a real estate agent decide on the starting price, one way to investigate the effect on expectations is to look at the effect on list prices. To do this, we re-estimate equation 1 with list prices instead of final prices as outcome variable using a sample of our data containing all apartments advertised (and actually sold) between January 1, 2011, and April 30, 2014.[31]

The results, presented in Table 7, indicate that there were very similar effects from the call to prayer events on list prices and final prices. The point estimates are very similar to the ones in the baseline analysis, indicating that people expected the public calls to prayer to have a positive effect on housing prices in the vicinity of Fittja Mosque. It should however be noted that there is more uncertainty in the point estimates for northern Botkyrka when using list prices instead of final prices.

[31] Since our data is based on information reported by real estate agents after the close of a sale, the data set includes only advertisement information for properties that were actually sold.

Table 7. Distance specifications: Effects on list prices

	(1) ln(price) whole of Botyrka	(2) ln(price) whole of Botkyrka	(3) ln(price) northern Botkyrka	(4) ln(price) northern Botkyrka
distance_mosq	0.037	0.071*	0.062	0.086
	(0.021)	(0.040)	(0.042)	(0.066)
Distance_Sep2011ad	-0.012*	-0.019***	-0.030	-0.056*
	(0.006)	(0.006)	(0.025)	(0.028)
Distance_Apr2012ad	-0.017**	-0.017**	-0.027	-0.014
	(0.008)	(0.008)	(0.041)	(0.043)
Month-by-year dummies	Yes	Yes	Yes	Yes
Apartment attributes	Yes	Yes	Yes	Yes
Location attributes	Yes	Yes	Yes	Yes
Neighborhood dummies	No	Yes	No	Yes
Observations	1,106	1,106	410	410
Adjusted R-squared	0.63	0.75	0.75	0.81

Standard errors clustered on neighborhoods (SAMS) in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

5.6.2 Synthetic control estimations

In the previous sections we have used a distance specification to study the effect of the public call to prayer events on house prices within Botkyrka municipality. In this section we study the effect of the public call to prayer events on house prices in Botkyrka municipality as a whole. This is done using a synthetic control specification and a sample of data containing apartments in housing cooperatives sold in Stockholm County, except within Stockholm municipality and Huddinge municipality.[32] A map of Stockholm County and the municipalities it contains is provided in *Figure 5*.

The original sample for Stockholm County, excluding the municipalities of Huddinge and Stockholm, contains 38,415 sales of apartments in housing cooperatives. After excluding sales where the Google coordinates differ by more than 100 meters from the coordinates provided by the real estate agents, newly built apartments, and apartments with a missing or unusual contract price, living area, monthly fee, number of rooms and/or with unclear contract date (these are the same restrictions as imposed in the baseline analyses) we are left with 33,885 sales.

[32] The reason for excluding Stockholm municipality from the sample for the Stockholm County is that there is a mosque in Stockholm municipality, and housing within Stockholm municipality could thus potentially also be affected by the public calls to prayer process. The reason for excluding Huddinge municipality is that, as already mentioned, some parts of Huddinge are close to Fittja Mosque and could thus also be affected by the calls to prayer process.

Figure 5. Stockholm County and its municipalities

Note: Map created by Eva Jirner.

The synthetic control approach developed by Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010) is an extension of the traditional difference-in-differences approach. One advantage of the synthet-

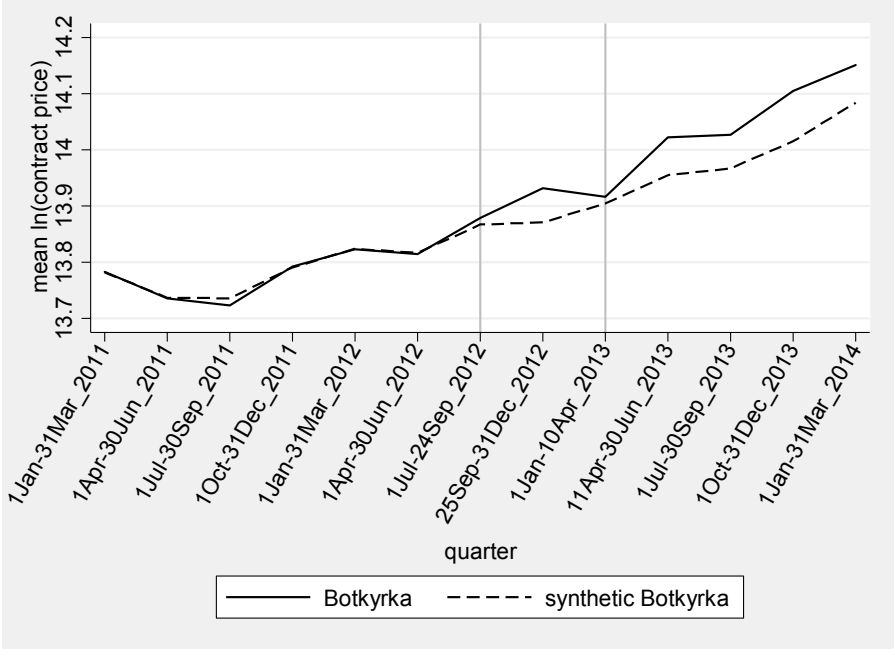
ic control approach is that the choice of control units is data-driven and relies only on pre-treatment data, thus reducing subjectivity in the choice of control units. Furthermore, the synthetic control approach makes inference possible in a setting like ours where there is only one treatment unit, i.e., when large sample inference techniques are not well suited. Also, in contrast to traditional difference-in-differences approaches, under some conditions the synthetic control approach allows the effects of confounding unobserved characteristics to vary over time (see Abadie et al., 2010).

In the synthetic control specification used in this paper, the effect of the September 2012 and April 2013 events on house prices in Botkyrka municipality are estimated by comparing the development of average house prices in the treatment unit Botkyrka municipality to the development of average house prices for a synthetic control group. The synthetic control group is constructed by a weighted combination of potential control units chosen to approximate Botkyrka municipality in terms of predictors of average house prices. In our case the potential control units are the other municipalities in Stockholm County (except for the municipalities of Huddinge and Stockholm) and the predictors are house prices, living area, monthly fee and number of rooms averaged over the whole pre-treatment period, share of foreign born in 2011, the tax base for 2011 as well as mean house prices in each quarter in the pre-treatment period. The development of average house prices for the resulting synthetic control group is an estimate of the counterfactual of what would have been observed for Botkyrka municipality without the treatment.

The trends for the mean of the natural logarithm of house prices for Botkyrka municipality and its synthetic counterpart are presented in *Figure 6*.[33] As is clear from the figure, the two price trends follow each other closely before the political announcement on September 25, 2012, indicating that the synthetic Botkyrka provides a good fit for the real Botkyrka. In the first quarter after the announcement, the average selling price in Botkyrka increases more than it does in its synthetic counterpart. After something that looks like a convergence pattern in the first quarter of 2013, there seems to be a clear and more persistent divergence in final house prices after the police decision and the first call to prayer in April 2013; in the four quarters following the April events, the house prices in Botkyrka are consistently above the house prices in synthetic Botkyrka.

[33] The municipalities included in synthetic Botkyrka are given in the Appendix, together with their respective weights. In the Appendix we also present a table with the values of prediction variables for Botkyrka and its synthetic counterpart.

Figure 6. House price trends, whole of Botkyrka: Synthetic control estimates on quarterly data for the period January 2011 - April 2014.

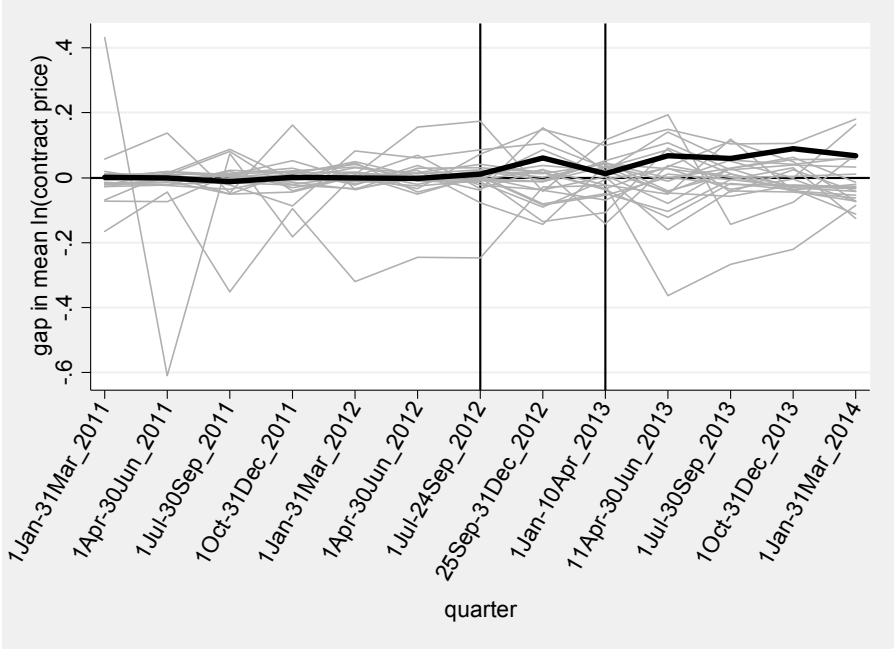


Note: The vertical lines indicate the September 2012 and April 2013 events

To get an indication of whether the positive price effect observed in Figure 6 has any statistical bearing, we estimate the synthetic control specification for every potential control municipality in Stockholm County with the purpose of assessing whether the effect estimated for Botkyrka municipality is large relative to the distribution of the effects estimated for the municipalities not exposed to treatment (the placebo estimations). From the placebo estimation, presented in *Figure 7*,[34] it seems as though it is only in the latter quarters of the period that we can talk about a potentially statistically significant effect; for that time period the estimated price gap (i.e., the estimated price effect) for Botkyrka is in the upper part of the gap size distribution, indicating a statistically significant price effect in Botkyrka.

[34] *Figure 7* presents the price gaps, i.e. the gap between the estimated mean price for a municipality and the estimated mean price for the municipality's synthetic version. The black line in *Figure 7* is the gap between the estimated mean price for Botkyrka and the estimated mean price for synthetic Botkyrka (i.e., the difference between the estimated price trends in *Figure 6*), and the grey lines are the gap between the estimated price trends in each of the placebo analyses.

Figure 7. House price gaps, whole of Botkyrka: Synthetic control estimates on quarterly data for the period January 2011 - April 2014.



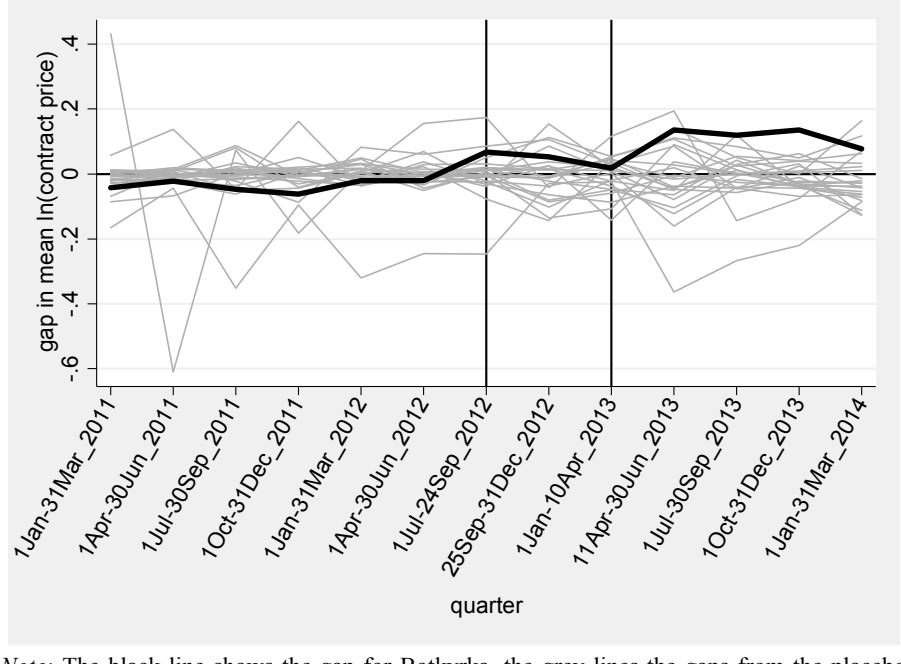
Note: The black line shows the gap for Botkyrka, the grey lines the gaps from the placebo estimations, and the vertical lines indicate the September 2012 and April 2013 events.

We also conduct a synthetic control analysis for the northern part of Botkyrka (the same geographic area as in the baseline analysis when examining the effects for the area closest to the mosque). All municipalities in Stockholm County except for Stockholm and Huddinge constitute the donor pool.[35] From the estimated gaps, presented in *Figure 8*, three things can be observed. First, even though "synthetic northern Botkyrka" is not a perfect fit for the real northern Botkyrka (the average house price is somewhat lower in the real northern Botkyrka than in synthetic northern Botkyrka in the pre-period), the overall pattern is very similar to that for the whole of Botkyrka (cf. *Figure 7*). Second, the shifts in the price gap for northern Botkyrka in September 2012 and April 2013 are larger (i.e., the "effects" on house prices are larger) than for the whole of Botkyrka (once again, cf. *Figure 7*). Third, there seems to be a statistically significant effect after the police decision and the first call to prayer in April 2013; there are few placebo estimates that are larger than the estimates for northern Botkyrka after April 2013.

[35] Of course, the observations from the southern part of Botkyrka do not form part of the donor pool either.

The indications from the synthetic control estimations of a positive house price effect from the public calls to prayer from Fittja Mosque in the whole of Botkyrka as well as in the northern part of Botkyrka are well in line with the results from the distance specifications in the baseline analysis. This strengthens our belief that if the public calls to prayer from Fittja Mosque have an effect on house prices, that effect is positive (i.e., house prices go up the closer one gets to the mosque).

Figure 8. House price gaps, northern Botkyrka: Synthetic control estimates on quarterly data for the period January 2011- April 2014.



Note: The black line shows the gap for Botkyrka, the grey lines the gaps from the placebo estimations, and the vertical lines indicate the September 2012 and April 2013 events.

6 Effects on migration behavior

In this section we examine the effects of the call to prayer events on migration of native- and foreign-born people into and out of neighborhoods close to Fittja Mosque. As explained in section 4, it seems reasonable, if anything, to expect a relative increase in the share of Muslims in the neighborhoods close to the mosque following the process leading to public calls to prayer. To the extent that Islam is more prevalent among the foreign born population than among natives, this could be mirrored by a relative increase in the share of foreign born people in these neighborhoods.

Before presenting the empirical results in section 6.3, we present the data (section 6.1), and the econometric specification (section 6.2) used in the analyses.

6.1 Data

Most available registers on population statistics in Sweden, including migration data, are quite infrequent (most often yearly). For our purpose we need more frequent data in order to be able to analyze if and how the migration patterns change around September 25, 2012, and April, 2013, in the neighborhoods close to Fittja Mosque. We have therefore asked Statistics Sweden to compile a data set that includes the following information for the years 2011, 2012, and 2013:

1. the number of individuals living in each neighborhood (SAMS) at the beginning of each year
2. the total number of individuals who each month moves into each neighborhood
3. the total number of individuals who each month moves out of each neighborhood
4. the total number of individuals born in Sweden who each month moves into each neighborhood
5. the total number of individuals born in Sweden who each month moves out of each neighborhood

From these data we are able to calculate the total number of individuals born abroad who each month move into each neighborhood, and the total number of individuals born abroad that each month move out of each neighborhood.[36]

While the frequency and information in this data set is better for our purpose than other available population data, there are some limitations. First, the data set includes only moves within Sweden. If the public call to prayer events affect the settlement of people arriving in Sweden directly from abroad, this will not be mirrored in our data. Likewise, if the public call to prayer events affect whether or not people choose to stay in Sweden, this will not be mirrored in our data. Even though we think this is a phenomenon of minor empirical importance, it can be worth mentioning that the estimations relating to the behavior of natives might be more reliable than the estimations relating to the behavior of foreign born people. Second, to prevent identification of individuals, for small neighborhoods Statistics Sweden

[36] An advantage with the migration data, compared with the house price data, is that we have all observations. While in the house price data we had only observations on properties sold, we now have all migrants, no matter whether they lived in a rental apartment, an owner-occupied apartment, or a house.

adjusts the values in cells with small values. Third, the data only make the distinction between foreign born and natives. It is possible that based on religion, there is also/instead, for example, sorting between different immigrant groups and between natives with parents born abroad and natives with parents born in Sweden.[37]

From the data set described above we use a sample containing observations for the neighborhoods in Botkyrka municipality. The original sample for Botkyrka municipality contains 36 month observations for each of 46 SAMS, i.e., 1656 month observations. Considering the problem of adjusted data, we exclude month observations that contain obvious adjustments, i.e., where the reported number of natives moving into (out of) a neighborhood is larger than the reported total number of individuals moving in (out). Table 8 shows summary statistics for the remaining sample for the whole of Botkyrka and for northern Botkyrka.

[37] We should later be able to look more closely into this using the database GeoSweden hosted by the Institute for Housing and Urban Research at Uppsala University. The database, which contains individual-based yearly register data covering the full population, is continuously updated and contains among other things information on residence, country of birth, and parents' country of birth. Direct study of the migration pattern of different religious groups is not possible since there is no official statistics on the religious beliefs of individuals.

Table 8. Summary statistics migration samples

	Obs	Mean	Std. Dev.	Min	Max
Botkyrka municipality					
Distance mosque (1000m)	1,632	4.5	2.9	0.2	16.0
Population	1,632	1,852	1,677	3.0	6,395
In-migration	1,632	17.8	18.6	0.0	97.0
In-migration natives	1,632	8.8	8.5	0.0	61.0
In-migration foreign	1,632	8.9	13.3	0.0	69.0
Out-migration	1,632	18.6	21.7	0.0	111.0
Out-migration natives	1,632	9.6	9.2	0.0	57.0
Out-migration foreign	1,632	9.0	14.8	0.0	79.0
Low education (%)	1,632	22.8	11.6	0.0	100.0
High education (%)	1,632	11.7	7.1	0.0	75.0
Mean income (SEK)	1,632	259,114	59,026	0.0	377,714
Northern Botkyrka					
Distance mosque (1000m)	752	2.1	1.0	0.2	4.4
Population	752	1,910	1,995	3.0	6,395
In-migration	752	19.1	22.3	0.0	97.0
In-migration natives	752	5.9	6.0	0.0	30.0
In-migration foreign	752	13.2	17.1	0.0	69.0
Out-migration	752	22.2	27.3	0.0	111.0
Out-migration natives	752	8.0	8.8	0.0	41.0
Out-migration foreign	752	14.2	19.3	0.0	79.0
Low education (%)	752	28.9	14.3	0.0	100.0
High education (%)	752	9.3	9.1	0.0	75.0
Mean income (SEK)	752	215,184	44,157	0.0	312,414

6.2 Econometric specification

When studying the effect of public calls to prayer on migration and sorting we use a similar type of distance specification as in the house price analysis. We examine whether the call to prayer process affected migration patterns differently in the neighborhoods close to and further away from the mosque. The assumption is that neighborhoods at different distances from Fittja Mosque are treated differently (with neighborhoods close to the mosque being more treated). The distance specification for migration takes the following form:

$$y_{jt} = \alpha_0 Dist_j + \alpha_1 Dist_j \times D_{Sept2012} + \alpha_2 Dist_j \times D_{April2013} + \gamma' Z_{jt} + \mu_t + \varepsilon_{jt} \quad (2)$$

where y_{jt} denotes the different migration outcomes in neighborhood j and month t. $Dist_j$ is a variable measuring the distance to Fittja Mosque from the population center of each neighborhood.[38] The variable should control for

[38] By using data from the database GeoSweden, hosted by the Institute for Housing and Urban Research at Uppsala University, we get information on housing coordinates for the full population, implying that we are able to calculate a neighborhood-specific coordinate

the effect on migration from proximity to the mosque itself as well as from omitted variables spatially correlated with the mosque. $D_{Sept2012}$ is a dummy variable that takes the value 1 for all months after September 2012, and 0 otherwise.[39] $D_{April2013}$ is a dummy variable that takes the value 1 for April 2013 and onwards, and 0 otherwise.[40] Z_{jt} is a vector of three neighborhood attributes; population size, share highly educated (more than high school), and share low-educated (less than high school) (these three variables are measured at the start of each year). μ_t is a vector of time (month-by-year) fixed effects controlling for time variation in the data unrelated to the political process surrounding the public calls to prayer and ε_{jt} denotes error terms.

Given the control variables, and under the assumptions of common time variation over the studied area – that there are no omitted variables systematically related to the September 2012 and April 2013 events under study that affect migration patterns – and that changes in migration patterns did not drive the timing of the events, α_1 and α_2 show how the call to prayer process affected the migration patterns in neighborhoods at different distances from the mosque.

6.3 Results

When presenting the migration results, we will do so separately for out-migration (section 6.3.1) and in-migration (section 6.3.2).

6.3.1 Results for out-migration

Table 9 shows the results from estimating equation (2) on data from the whole of Botkyrka municipality with respect to out-migration. The three outcome variables are total number of people moving out (column 1), number of foreign born people moving out (column 2), and number of natives moving out (column 3) of the neighborhoods. From the first row in the table it can be seen that before the public call to prayer events, total out- migration

based on the population density in each neighborhood. These statistics are calculated for 2010, which is the last year for which the information is provided in GeoSweden. This should, however, not be of any larger concern since this is a statistic that changes very slowly from one year to another.

[39] Since the call to prayer pronouncement took place on September 25, 2012, the moves that might have taken place in Botkyrka in the last days of September 2012 counted as if they took place before the pronouncement was actually made (since we have only monthly, not daily, observations on migration). However, we do not think this is of any major concern. There are probably fairly few moves made over a few days in Botkyrka, and it seems safe to assume that the moves actually made in the last days of September 2012 are unrelated to the call to prayer pronouncement.

[40] This means that the moves that took place after April 1 but before April 10 is counted as if they took place after the police decision on April 10. We do not consider this to be a problem from an econometric point of view. The number of moves are probably quite small and will not greatly affect the estimates. And if they have an impact, they work against finding an effect.

was relatively higher further away from Fittja Mosque. The out-migration of foreign born people was however relatively lower further away from the mosque. It is likely that this simply mirrors the fact that the foreign born population was smaller in the neighborhoods further away from the mosque (in the estimation we control for neighborhood total population but not neighborhood foreign born population). From the estimates for the time-interacted distance variable in Table 9 it seems like there were no effects of the public call to prayer events on relative out-migration of either total population, natives, or people born abroad at different distances from the mosque: the difference-in-differences estimates are all statistically insignificant.

From the results when we use only data from northern Botkyrka, given in Table 10, it can be seen that before the public call to prayer events, total out-migration of both foreign born and natives was relatively higher further away from Fittja Mosque. Regarding the public call to prayer events, the negative point estimate for distance to the mosque after September 2012, statistically significant at the 10 percent significance level, indicates that there was a relative decrease in total out-migration in neighborhoods further away from Fittja Mosque following the September 2012 events. When out-migration of natives and foreign born people are studied separately, however, the point estimates become insignificant. Furthermore, the estimates for distance to the mosque after April 2013 are all statistically insignificant. It thus seems that the public call to prayer events had little effect on relative out-migration at different distances from Fittja Mosque within northern Botkyrka. Furthermore, there is no evidence that any effect was different on the out-migration of natives and the out-migration of foreign born people.

Relating to the literature on "white flight", one can conclude from the results in Table 9 and Table 10 that the events related to the public calls to prayer from Fittja Mosque seem not to have induced any "flight" of natives from the neighborhoods close to the mosque.

Table 9. Effects on out-migration, whole of Botkyrka

	(1) Out-migration	(2) Out-migration foreign	(3) Out-migration native
distance_mosq	0.244**	-0.308***	0.552***
	(0.110)	(0.075)	(0.075)
Distance_Sep2012	-0.230	-0.238	0.008
	(0.196)	(0.167)	(0.121)
Distance_Apr2013	0.135	0.103	0.031
	(0.214)	(0.183)	(0.137)
Month-by-year dummies	Yes	Yes	Yes
Neighborhood attributes	Yes	Yes	Yes
Observations	1,632	1,632	1,632
Adjusted R-squared	0.85	0.77	0.71

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Out-migration in levels

Table 10. Effects on out-migration, northern Botkyrka

	(1) Out-migration	(2) Out-migration foreign	(3) Out-migration native
distance_mosq	1.580***	0.675*	0.905***
	(0.457)	(0.371)	(0.203)
Distance_Sep2012	-1.264*	-0.795	-0.469
	(0.746)	(0.529)	(0.433)
Distance_Apr2013	0.234	0.241	-0.007
	(0.819)	(0.600)	(0.464)
Month-by-year dummies	Yes	Yes	Yes
Neighborhood attributes	Yes	Yes	Yes
Observations	752	752	752
Adjusted R-squared	0.90	0.87	0.79

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Out-migration in levels

6.3.2 Results for in-migration

Table 11 shows the results from estimating equation (2) on data from the whole of Botkyrka municipality with respect to in-migration. The three outcome variables are total number of people moving in (column 1), number of foreign born people moving in (column 2), and number of natives moving in (column 3) to the neighborhoods. From the first row in the table it can be seen that, before the public call to prayer events, total in-migration was relatively higher further away from Fittja Mosque. The in-migration of foreign

born people was however relatively lower further away from the mosque. Similar to what was found for out-migration, it also seems that there was no effect of the public call to prayer events on relative in-migration at different distances from Fittja Mosque for either the total population, natives, or foreign born people: The difference-in-differences estimates presented in Table 12 are all statistically insignificant and small. The findings are similar when using data only from northern Botkyrka (Table 12).

Relating to the literature on "white avoidance" one can conclude from the results in Table 11 and Table 12 that the events related to the calls to prayer from Fittja Mosque seem not to have caused natives to avoid moving in to the neighborhoods close to the mosque.

Taken together, it does not seem as though the public call to prayer events affected the migration patterns of either total population, natives, or foreign born people within Botkyrka. The potential exception is the indication that within northern Botkyrka there was a relative decrease in total out-migration from neighborhoods further away from Fittja Mosque following the September 2012 events. There is however no evidence that the effect was different for native and foreign born out-migration. Thus, the public call to prayer events do not seem to have served as drivers of residential segregation between natives and foreign born people around Fittja Mosque, at least in the short run.

Table 11. Effects on in-migration, whole of Botkyrka

	(1)	(2)	(3)
	In-migration	In-migration foreign	In-migration native
distance_mosq	0.415***	-0.239***	0.655***
	(0.121)	(0.068)	(0.093)
Distance_Sep2012	-0.113	-0.016	-0.097
	(0.200)	(0.139)	(0.132)
Distance_Apr2013	-0.020	-0.051	0.030
	(0.221)	(0.154)	(0.150)
Month-by-year dummies	Yes	Yes	Yes
Neighborhood attributes	Yes	Yes	Yes
Observations	1,632	1,632	1,632
Adjusted R-squared	0.76	0.77	0.48

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
In-migration in levels

Table 12. Effects on in-migration, northern Botkyrka

	(1) In-migration	(2) In-migration foreign	(3) In-migration native
distance_mosq	1.313***	0.773***	0.540***
	(0.374)	(0.292)	(0.168)
Distance_Sep2012	-0.062	0.063	-0.125
	(0.685)	(0.575)	(0.297)
Distance_Apr2013	-0.853	-0.602	-0.251
	(0.799)	(0.637)	(0.349)
Month-by-year dummies	Yes	Yes	Yes
Neighborhood attributes	Yes	Yes	Yes
Observations	752	752	752
Adjusted R-squared	0.86	0.86	0.65

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
In-migration in levels

7 Conclusions

The last decades' immigration to Western Europe has resulted in a culturally and religiously more diverse population in Western European countries. Following this, new features have appeared in the cityscape. Using a quasi-experimental approach we study how one such feature, public calls to prayer, affects neighborhood dynamics (house prices and migration). The quasi-experiment is based on an unexpected political decisions that led the way to the first public call to prayer from a mosque in Sweden; the Fittja Mosque in Botkyrka municipality in the Stockholm region.

With regards to house prices, our results indicate that, if anything, the public call to prayer process increased house prices closer to the mosque. Although the point estimates and statistical significance vary somewhat over different specifications, all results point in the same direction: within Botkyrka municipality, the point estimates indicate that housing prices increased by 1.1-1.7 percent per kilometer closer to the mosque after municipal committee pronouncements supportive of public calls to prayer. This was followed by a further increase of approximately the same magnitude after the police permission to conduct public calls to prayer was given and the first public call to prayer was conducted. The point estimates indicate an even larger effect within the part of Botkyrka municipality closest to the mosque. We further find some evidence that house prices in Botkyrka municipality as a whole increased after the police permission and the first call to prayer.

With regards to migration we find no evidence that the public call to prayer process served as a driver of residential segregation between natives and people born abroad around Fittja Mosque.

Our findings are consistent with a story in which some people have a willingness to pay for the opportunity to more fully exert their religion which exerts an upward pressure on housing in the vicinity of a mosque with public calls to prayer. In the case we study, Fittja Mosque in Botkyrka municipality, it seems that there were few mobile price arbitrageurs, or people with preferences against symbols of Islam/Muslims/immigrants, who by moving out could keep house prices closer to the mosque down.

In Botkyrka municipality, the immigrant share of the population is much larger than in most other Swedish municipalities. It is thus possible that new features in the cityscape resulting from the last decades of foreign immigration could have a different effect on migration patterns and house prices in other places. This is an issue for further research. Another issue for further research is to look closer at whether the public call to prayer process affected sorting with respect not only to the native/foreign born dichotomy but with respect to country of birth/parents' country of birth.

Nevertheless, there seems to be a previously unmet willingness to pay for living close to a mosque with public calls to prayer, which may indicate that the possibility to fully exert one's religion/religious freedom is not entirely satisfied for everyone everywhere in today's Sweden.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association, 105*(490), 493-505.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review, 93*(1), 113-132.

Accetturo, A., Manaresi, F., Mocetti, S., & Olivieri, E. (2014). Don't Stand so close to me: the urban impact of immigration. *Regional Science and Urban Economics, 45*, 45-56.

Aldén, L., Hammarstedt, M., & Neuman, E. (2015). Ethnic segregation, tipping behavior, and native residential mobility. *International Migration Review, 49*(1), 36-69.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics, 114*(4), 1243-1284.

Alesina, A., Baqir, R., & Easterly, W. (2000). Redistributive public employment. *Journal of Urban Economics, 48*(2), 219-241.

Alesina, A., Glaeser, E., Sacerdote, B., Durlauf, S. N., & Levy, F. (2001). Why doesn't the United States have a European-style welfare state? *Brookings Papers on Economic Activity,* (2), 187-254.

Alesina, A., & Glaeser, E. L. (2004). Fighting poverty in the US and Europe: A world of difference: Rodolfo Debenedetti Lectures.

Alesina, A., & La Ferrara, E. (2000). Participation in heterogeneous communities. *Quarterly Journal of Economics, 115*(3), 847-904.
Alesina, A., & La Ferrara, E. (2002). Who trusts others? *Journal of Public Economics, 85*(2), 207-234.
Allievi, S. (2009). *Conflicts over mosques in Europe: Policy issues and trends*. London: NEF Initiative on Religion and Democracy in Europe.
Andreoni, J., Payne, A., Smith, J. D., & Karp, D. (2011). *Diversity and donations: The effect of religious and ethnic diversity on charitable giving*. Working Paper 17618, National Bureau of Economic Research.
Aragón, F. M., & Rud, J. P. (2013). Natural resources and local communities: Evidence from a Peruvian gold mine. *American Economic Journal: Economic Policy, 5*(2), 1-25.
Babawale, G. K., & Adewunmi, Y. (2011). The impact of neighbourhood churches on house prices. *Journal of Sustainable Development, 4*(1), 246-253.
Bisin, A., & Verdier, T. (2000). "Beyond the melting pot": Cultural transmission, marriage, and the evolution of ethnic and religious traits. *Quarterly Journal of Economics, 115*(3), 955-988.
Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics, 114*(2), 577-599.
Boustan, L. P. (2010). Was postwar suburbanization "White Flight"? Evidence from the black migration. Q*uarterly Journal of Economics, 125*(1), 417-443.
Bråmå, Å. (2006). 'White Flight'? The production and reproduction of immigrant concentration areas in Swedish cities, 1990-2000. *Urban studies, 43*(7), 1127-1146.
Brandt, S., Maennig, W., & Richter, F. (2014). Do houses of worship affect housing prices? Evidence from Germany. *Growth and Change, 45*(4), 549-574.
Card, D., Mas, A., & Rothstein, J. (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics*, 177-218.
Carroll, T. M., Clauretie, T. M., & Jensen, J. (1996). Living next to godliness: Residential property values and churches. *Journal of Real Estate Finance and Economics, 12*(3), 319-330.
Cropper, M. L., Deck, L. B., & McConnell, K. E. (1988). On the choice of functional form for hedonic price functions. *Review of Economics and Statistics, 70*(4), 668-675.
Dahlberg, M., Edmark, K., & Lundqvist, H. (2012). Ethnic diversity and preferences for redistribution. *Journal of Political Economy, 120*(1), 41-76.
European Comission. (2011). *Demography report 2010: Older, more numerous and diverse Europeans*.
Eurostat. (2011). *Statistics in focus 34/2011*.
Fiva, J. H., & Kirkebøen, L. J. (2011). Information shocks and the dynamics of the housing market. *Scandinavian Journal of Economics, 113*(3), 525-552.
Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics, 57*(1), 148-169.
Gibbons, S., & Machin, S. (2008). Valuing school quality, better transport, and lower crime: evidence from house prices. *Oxford Review of Economic Policy, 24*(1), 99-119.
Greenstone, M., & Gallagher, J. (2008). Does hazardous waste matter? Evidence from the housing market and the Superfund program. Q*uarterly Journal of Economics, 123*(3), 951-1003.

Halla, M., Wagner, A. F., & Zweimüller, J. (2012). *Does immigration into their neighborhoods incline voters toward the Extreme Right? The case of the Freedom Party of Austria*. Economics working papers 2012-05, Department of Economics, Johannes Kepler University Linz, Austria.
Kuminoff, N. V., Parmeter, C. F., & Pope, J. C. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities. *Journal of Environmental Economics and Management, 60*(3), 145-160.
Luttmer, E. F. (2001). Group loyalty and the taste for redistribution. *Journal of Political Economy, 109*(3), 500-528.
McMillen, D. P., & McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's midway line, 1983-1999. *Real Estate Economics, 32*(3), 463-486.
Pope, J. C. (2008a). Buyer information and the hedonic: the impact of a seller disclosure on the implicit price for airport noise. *Journal of Urban Economics, 63*(2), 498-516.
Pope, J. C. (2008b). Do seller disclosures affect property values? Buyer information and the hedonic model. *Land Economics, 84*(4), 551-572.
Pope, J. C. (2008c). Fear of crime and housing prices: Household reactions to sex offender registries. *Journal of Urban Economics, 64*(3), 601-614.
Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. J*ournal of Political Economy, 82*(1), 34-55.
Sá, F. (2014). Immigration and house prices in the UK. *Economic Journal*, n/a.
Saiz, A. (2003). Room in the kitchen for the melting pot: Immigration and rental prices. *Review of Economics and Statistics, 85*(3), 502-521.
Saiz, A. (2007). Immigration and housing rents in American cities. *Journal of Urban Economics, 61*(2), 345-371.
Saiz, A., & Wachter, S. (2011). Immigration and the neighborhood. *American Economic Journal: Economic Policy, 3*(2), 169-188.
Statistics Sweden. (2005). *Geografin i statistiken - regionala indelningar i Sverige.* Meddelanden i samordningsfrågor för Sveriges officiella statistik 2005:2 [*Geography in statistics - regional divisions in Sweden.* Reports on Statistical Coordination for the Official Statistics of Sweden 2005:2]. Örebro.
Tinbergen, J. (1956). On the theory of income distribution. *Weltwirtschaftliches Archiv, 77*, 155-175.
Vigdor, J. L. (2004). Community composition and collective action: Analyzing initial mail response to the 2000 census. *Review of Economics and Statistics, 86*(1), 303-312.

Appendix: Descriptive statistics related to the Synthetic control estimation for the whole of Botkyrka

In this appendix we present how the municipalities in the donor pool are weighted in the synthetic control estimation for the whole of Botkyrka (Table A1) and descriptive statistics for the balancing of the matching variables (Table A2).

Table A1. Donor pool and weights

Municipality	weight
UpplandsVäsby	0.000
Vallentuna	0.096
Österåker	0.258
Värmdö	0.000
Järfälla	0.000
Ekerö	0.000
Salem	0.026
Haninge	0.438
Tyresö	0.000
Upplands-Bro	0.000
Nykvarn	0.016
Täby	0.000
Danderyd	0.000
Sollentuna	0.000
Södertälje	0.000
Nacka	0.000
Sundbyberg	0.000
Solna	0.000
Lidingö	0.000
Vaxholm	0.000
Norrtälje	0.000
Sigtuna	0.138
Nynäshamn	0.028

Table A2. Balancing of matching variables

Variable	Botkyrka	Synthetic Botkyrka
ln(contract price)	13.79	13.79
living area (square meters)	72.55	71.24
monthly fee (SEK)	4,275	4,221
number of rooms	2.72	2.68
foreign born 2011 (%)	38.10	18.52
tax base 2011 (SEK)	32,360	38,164
ln(contract price) 1Jan-31Mar 2011	13.78	13.78
ln(contract price) 1 Apr-30Jun 2011	13.74	13.74
ln(contract price) 1Jul-30Sep 2011	13.72	13.74
ln(contract price) 1Oct-31Dec 2011	13.79	13.79
ln(contract price) 1Jan-31Mar 2012	13.82	13.82
ln(contract price) 1Apr-30Jun 2012	13.81	13.82
ln(contract price) 1Jul-24Sep 2012	13.88	13.87

Note: ln(contract price), living area, monthly fee, and number of rooms are averaged for the period January 1, 2011-September 24, 2012.