TECHNISCHE
UNIVERSITÄT
DARMSTADT

The Role of Amenities in the Location Decision of Households and Firms

Dissertation

Approved as fulfilling the requirements for the degree of

doctor rerum politicarum (Dr. rer. pol.)

by the

Department of Law and Economics of
Darmstadt University of Technology

by

Torsten Kristoffer Möller, M.A.
(born 19/09/1984 in Hannover)

First supervisor:	**Prof. Dr. Volker Nitsch**
Second supervisor:	**Prof. Dr. Jens Südekum**
Date of submission:	**17/10/2013**
Date of defense:	**21/05/2014**
Place of publication:	**Darmstadt**
Year of publication:	**2014**

D 17

Acknowledgements

At this point I want to thank several persons. My first supervisor and doctoral advisor Prof. Dr. Volker Nitsch discussed certain aspects related to many chapters of this dissertation during the last years. He provided helpful feedback which generated new ways of thinking. My second supervisor Prof. Dr. Jens Südekum who closely followed the development of my research over the last years, giving useful advices.

I very much would like to thank my family and friends who not only supported me during my PhD years.

A special thanks goes to my close colleague and friend Sevrin Waights for (not only) the endless ArcGIS sessions and inspiring research discussions. Furthermore, I would like to thank Dr. Gabriel Ahlfeldt for immediate feedback along the road and for making my visit at SERC at the London School of Economics and Political Science possible. I also want to thank Dr. Nicolai Wendland for his helpful comments and opinions.

I want to thank the Center for Metropolitan Studies at the Berlin University of Technology for the accommodation and great research environment. A special thanks goes to Elisabeth Asche and Professor Dr. Dorothee Brantz. For constant feedback, I would like to thank my colleagues Dr. Stefan Goldbach, Johannes Rode, Dr. Kristian Giessen, Felipe Carozzi and Mr Dell. Philip Savage provided excellent comments for issues related to linguistics.

Finally, I want to thank all conference participants who gave useful feedback on my research.

Contents

List of Figures

List of Tables

List of abbreviations

2SLS	Two Stage Least Squares
3SLS	Three Stage Least Square
ADF	Augmented Dickey Fuller
AMM	Alonso-Mills-Muth Model
CA	Conservation Area
CBD	Central Business District
CIRF	Cumulative Impulse Response Function
CRT	Chicago Rapid Transit Company
DD	Differences-in-difference
GDR	German Democratic Republic
GIS	Geographic Information System
GMM	Generalized Method of Moments
GSOEP	German Socioeconomic Panel
IHK	Chamber of Industry and Commerce
IIA	Independence of Irrelevant Alternatives
IRF	Impulse Response Function
IV	Instrumental Variable
MAUP	Modifiable Area Unit Problem
MLS	Major League Soccer
NBS	Nationwide Building Society
NCSA	National Center for Supercomputing Applications
NFL	Nation Football League (US)
NHGIS	National Historical Geographic Information System
NHL	Nation Hockey League
NIMBY	Not In My Backyard
OLS	Ordinary Least Square
PP	Philipps-Perron
PPML	Poisson Pseudo-Maximum Likelihood
PVAR	Panel Vector Auto Regression
RDD	Regression Discontinuity Design
SAR	Spatial Autoregression
SVAR	Structural Vector Auto Regression
TTWA	Travel to Work Areas

VIF	Variance Inflation Factors
VKT	Vehicle Kilometres Travelled

Zusammenfassung

Städte werden traditionellerweise als Produktionszentren gesehen: Unternehmen produzieren in Städten, da ihre Produktion dort aufgrund von Agglomerationseffekten produktiver ist. Menschen leben in Städten, da ihnen dort die Unternehmen Arbeits- und Verdienstmöglichkeiten bieten. Die Existenz von Städten sowie die damit eng verknüpfte Standortentscheidung von Haushalten und Unternehmen wird in der Regel aus Sicht dieser Produktionsverflechtungen erklärt. Räumliche Dichte wird mit Agglomerationsvorteilen auf der Produktionsseite und Agglomerationsnachteilen (*congestion*) auf der Konsumentenseite in Verbindung gesetzt. Seit einiger Zeit haben Stadtforscher zunehmend von dieser einseitigen Betrachtung Abstand genommen und Städte nicht nur als Produktions-, sondern auch als Konsumzentren betrachtet. Die Präferenzen von Arbeitern sind nachweislich heterogener geworden, das Humankapital sowie die Einkommen gewachsen. Unternehmen sind mobiler geworden, das Güterangebot diversifizierter (Brueckner et al., 1999; Kolko, 1999; Kotkin, 2000; Glaeser et al., 2001; Florida, 2002; Dalmazzo & de Blasio, 2011; Glaeser, 2011; Bauernschuster et al., 2012; Suedekum et al., 2012; Ahlfeldt, 2013). Diese Entwicklungen haben Arbeiter mehr Freizeit und Einkommen gebracht, das sie zum Konsumieren nutzen können. Daher sollten Arbeiter und Unternehmen nicht mehr nur länger auf klassische Produktionsanreize, sondern auch auf eine Reihe urbaner Annehmlichkeiten (*amenities*) reagieren.

Ziel dieser Dissertation ist eine detaillierte Untersuchung der “consumer city idea”. Motiviert von einem Mangel an empirischer Evidenz durchleuchtet die Arbeit verschiedene Aspekte der Rolle von Amenities bei der Standortwahl von Haushalten und Unternehmen, um auf diese Weise zu einem noch jungen Forschungsfeld beizutragen. Die Arbeit orientiert sich dabei an der Urban Amenity Klassifizierung von Glaeser et al. (2001). Diese Einteilung sowie verwandte Literatur werden detaillierter in Kapitel 2 behandelt. Ziel der umfangreichen Literaturzusammenfassung ist es, dem Leser eine Basis für die späteren Analysen zu geben. Außerdem wird auf die methodischen Entwicklungen im Forschungsfeld eingegangen, die sich vereinfacht als einen Wandel von der Herstellung simpler Korrelationszusammenhänge zu Kausalität beschreiben lassen.

Kapitel 3 untersucht, inwieweit kulturelle Annehmlichkeiten (*cultural amenities*) eine Rolle bei der Standwortwahl von Unternehmen spielen. Dabei werden *cultural amenities* als lokale, nicht-handelbare Güter und Dienstleistungen (wie Bars, Cafés etc.) definiert. Die Idee ist, dass innovative Dienstleistungsunternehmen hoch mobil sind und hauptsächlich von qualifizierter Arbeit als Inputfaktor abhängen. Gleichzeitig haben Hochqualifizierte und Kreative ein großes Interesse an einem sozial und kulturell abwechslungsreichen Umfeld (Florida, 2002). Hieraus ableitend stellt sich die Hypothese, dass Unternehmen, die bzgl. ihrer Standortentscheidung

ihren Mitarbeitern folgen, als Amenity Maximierer agieren. Ich teste die Hypothese empirisch mit Hilfe eines Location Choice Modells für Internet Start-ups in Berlin. Die Identifizierung des *cultural amenity* Effekts basiert auf dem Fall der Berliner Mauer, der als quasi-natürliches Experiment interpretiert wird. Amenities haben hiernach einen positiven Einfluss auf die Standortwahl und ziehen Unternehmen aus der Web-Branche an. Ein Anstieg der Dichte an *cultural amenities* um 1% führt zu einer erhöhten Wahrscheinlichkeit der Unternehmensansiedlung von 1,2%. Ein Vergleich mit anderen Dienstleistungsbranchen macht ferner deutlich, dass vor allem kreative Branchen von Amenities positiv beeinflusst werden, während die Schätzer für traditionelle Unternehmen nicht signifikant oder sogar negativ sind.

Kapitel 4 behandelt die Amenity Rolle in Bezug auf Ästhetik und die natürliche Schönheit eines Ortes am Beispiel von Kulturerbe. Der Ausweis von denkmalgeschützten Gebieten wird als Lösung eines Externalitätenproblems betrachtet, das für Hauseigentümer Nutzen im Sinne einer erhöhten Sicherheit bzgl. der Zukunft des Gebiets bedeutet, aber auch zusätzliche Kosten der Weiterentwicklung- und Baumöglichkeiten, d.h. Einschränkungen, mit sich bringt. Es wird ein simples theoretisches Modell entwickelt, nach dem das optimale Ausweisniveau so bestimmt wird, dass es den Nutzen der lokalen Eigentümer Pareto-maximiert. Das Modell impliziert, dass a) mit einer erhöhten Präferenz für historische Beschaffenheiten die Wahrscheinlichkeit eines Ausweis als Kulturerbe wächst, und dass sich b) marginale Neuausweise nicht signifikant in den Hauspreisen kapitalisieren. Die gewonnenen empirischen Ergebnisse entsprechen diesen Erwartungen.

In Kapitel 5 wird mit Transportgeschwindigkeit eine dritte Klasse von Urban Amenities nach der Definition von Glaeser et al. (2001) untersucht. Die bedeutende Rolle von Verkehr ist in der Ökonomie unumstritten. Die Untersuchung der Auswirkungen neuer Verkehrsprojekte gestaltet sich hingegen schwierig, da die Beziehung zwischen Verkehr und wirtschaftlicher Entwicklung von einem offenkundigen Simultanitätsproblem geprägt ist. Transportallokation ist nicht zufällig, sondern reagiert auf die Nachfrage nach Infrastruktur, da diese in der Regel immense Investitionen voraussetzt. Herkömmliche kausale Inferenz hat sich diesem Problem nur von einer einseitigen Betrachtung der Bereitstellung neuer Infrastruktur genähert. Die Arbeit schlägt daher eine in der Makroökonomie bewährte Methode vor, um die Struktur der sich gegenseitig beeinflussenden, endogenen Variablen zu untersuchen. Ich schätze bivariate Panel Vektorautoregressionsmodelle, die auf einzigartigen historischen Daten Berlins während einer hochdynamischen Periode, in der der Großteil des heutigen öffentlichen Nahverkehrs entwickelt wurde (1881-1935), basieren. Die Ergebnisse bestätigen die simultane Bestimmung von Transportinfrastruktur und urbaner Entwicklung. Ferner lässt sich eine Verdrängung von Haushalten durch Unternehmen in zentralen Lagen nach einem positiven Transportschock erkennen.

Kapitel 6 ergänzt die Untersuchung des Berliner Schienennahverkehrs durch die Anwendung derselben Panel VAR Methodik auf Chicago, Illinois, und der Entwicklung der Chicago Elevated über einen Zeitraum von über 100 Jahren (1910-2010). Die Untersuchung kann als zusätzlicher Robustheitstest sowie auch als eine Vergleichsstudie interpretiert werden. Die Ergebnisse entsprechen den Erkenntnissen aus der Berliner Analyse.

Die Untersuchung verschiedener Aspekte der “consumer city idea” im Rahmen dieser Arbeit macht die Bedeutung von Urban Amenities für das Verständnis, wie sich Haushalte und Unternehmen im Raum verhalten, wo und warum sie sich für einen Standort entscheiden, deutlich. Die räumliche Allokation von Personen sowie Unternehmen kann von mehr als nur wirtschaftlicher Aktivität erklärt werden. Neue und verbesserte empirische Methoden, eine zunehmende Verfügbarkeit räumlicher Daten (insbesondere “open data”) sowie verbesserte Softwarepakete zur Datenaufbereitung/-analyse (wie GIS) ermöglichen Untersuchungen, die uns zu verstehen helfen, wo wir leben und warum wir dort leben.

1. Introduction

Traditionally, cities have been regarded as centres of production: Firms produce goods in cities because agglomeration economies make them more productive. People live in cities because firms provide jobs and income. The existence of cities and closely related location decisions by households and firms have often been explained by these production linkages. Density is thought to offer agglomeration benefits on the production side but negative congestion effects on the consumption side. More recently, urban scholars have departed from this view and considered cities not only as centres of production but also of consumption (Glaeser et al., 2001). Workers have arguably become more heterogeneous in terms of taste, more educated and their incomes have risen. Firms have become more footloose and goods more diverse (Brueckner et al., 1999; Kolko, 1999; Kotkin, 2000; Glaeser et al., 2001; Florida, 2002; Dalmazzo & de Blasio, 2011; Glaeser, 2011; Bauernschuster et al., 2012; Suedekum et al., 2012; Ahlfeldt, 2013). These developments have left workers with more leisure time and income to spend on the consumption side so that workers and firms are no longer expected to only respond to classic production links but to a wide range of amenities.

This dissertation intends to shed further light on the consumer city idea. Motivated by a lack of empirical evidence, I contribute to this young field of research by investigating different aspects of the role of amenities in the location decision of households and firms. The work is structured around the classification of urban amenities as suggested by Glaeser et al. (2001). This classification as well as related literature is presented in Chapter 2. The comprehensive literature review is intended to provide a background for the analyses carried out in this work. Moreover, it shows the field's methodological development which is characterised by a move from correlations to establishing causality.

Chapter 3 investigates the role of cultural amenities in the location decision of firms. I define cultural amenities as localised goods and services, which is one of the four urban amenity categories defined by Glaeser et al. (2001). The idea is that innovative service firms are highly footloose and mainly rely on qualified labour as input factor. At the same time, highly qualified and "creative" individuals have a strong preference for a rich social and cultural life (Florida, 2002). It is therefore expected that firms, following its workers, act as amenity-maximising agents. I empirically test this hypothesis by estimating a location choice model for internet start-ups in Berlin. The identification of the cultural amenity effect is based on the fall of the Berlin Wall which is interpreted as a quasi-natural experiment. Amenities are found to positively impact on the location of web firms. A comparison with other service industries moreover suggests that amenities are significant to the location choice of creative sectors, whereas no effect can be observed for non-creative firms.

Chapter 4 is centred around the amenity role of aesthetics and physical setting, and on heritage preservation in particular. Heritage designation is considered to solve an externality problem thus providing benefits to home owners, in terms of a reduction of uncertainty regarding the future of an area, but at the costs of development restrictions. The chapter proposes a simple theory of the designation process, in which it is postulated that the optimal level of designation is chosen so as to Pareto-maximise the welfare of local owners. The implication of the model is that a) an increase in preferences for historic character should increase the likelihood of a designation, and b) new designations at the margin should not be associated with significant house price capitalisation effects. The empirical results are in line with these expectations.

In Chapter 5 a third type of urban amenities according to the Glaeser et al. (2001) definition is investigated in further detail, i.e. speed of transportation. Transport's important role in economics is beyond controversy. The estimation of the impact of a new infrastructure project is, however, not entirely straightforward as the relation between transport and economic development is plagued by a notorious simultaneity problem. The allocation of transport is not completely random and may respond to demand as infrastructure projects usually require large investment costs. Conventional causal inference has approached this problem by only focusing on the uni-directional effect of transport provision. I therefore propose a method, which is well established in macroeconomics, to explore the structure of mutually related endogenous variables. In particular, I run bivariate Panel VAR models using unique historical data for Berlin during a dynamic period when most of today's public rail network was established (1881-1935). Results do indeed suggest a simultaneously determined relation between transport and urban development.

Chapter 6 extends the previous analysis of the Berlin rail sector by applying the Panel VAR methodology to the city of Chicago, Illinois, and the development of the 'L' train over a period of over 100 years (1910-2010). The analysis can be interpreted as both an additional robustness test and a comparative study. Results are in line with the findings for Berlin.

The dissertation ends with the conclusion in Chapter 7, where I summarise the main findings and stress important contributions to the literature.

2. Literature review

This chapter reviews the recent literature on the role of amenities in the location decision of households and firms. The review is intended to provide a background for the analyses carried out in this work. The chapter begins with a very brief introduction to location theories and a quick review of the empirical workhorse models used in the literature. This is followed by a more detailed review of how the location choice is determined by different types of amenities.

2.1. Location theory and empirics

The question where people and firms locate and why they choose a particular location lies at the centre of location theory. Combining the fields of economics and geography, the question marks the origins of today's economic geography, or urban economics when dealing with cities (O'Sullivan, 2009) in particular. Von Thünen (1826) is one of the first to investigate this question from a more economic perspective. In his influential work "Der isolierte Staat" he introduces a spatial dimension by looking at the transport cost of different crops from surrounding fields to an exogenous market square. In his models, transport costs not only depend on geographical distance but also on a crop's weight and perishability. Assuming a (Ricardian) land rent, i.e. a renter's maximal willingness to pay for a unit of land, he derives concentric rings of land use around a town centre where the goods are assumed to be traded. This trivial but powerful model explains the location choice and land use allocation according to crops' transport costs to the market.

Von Thünen's early ideas were eventually picked up by Alonso (1964); Mills (1967, 1969); Muth (1969) who replace the exogenous town centre/market square by a central business district (CBD). This CBD is assumed to host all economic activity such that the city is characterised by a single employment centre. In this monocentric city model workers are assumed to commute to their jobs in the CBD. They face a trade-off between transport cost and space consumption and will choose their location based on optimised utility. At the same time, firms determine their location based on maximised profits; they require land for production and accessibility to markets. While different land use types were characterised by different crops in von Thünen's (1826) concentric ring model, it is residents and firms who compete for land in the monocentric city model. Final land use patterns are determined by residents' and firms' bid-rents. Firms are generally expected to outbid residents in the CBD to benefit from agglomeration economies. This widely used Alonso-Mills-Muth (AMM) Model lies at the core of urban economics and has often been adopted to address critical assumptions. Lucas & Rossi-Hansberg (2002) introduce for instance an endogenously determined CBD by incorporating agglomeration economies where firms are assumed to be more productive in a high employment surrounding.

Standard urban economic models share one common perspective on cities: A city is usually regarded as a place of production where land use patterns are determined only by the production side. Residential utility and hence location is predominantly defined by commuting opportunities and employment. More recently, economists have departed from this view. Especially applied research has increasingly investigated the role of consumption amenities. Workers have arguably become more heterogeneous in terms of taste, more educated and their incomes have risen. Firms have become more footloose and goods more diverse, such that workers and firms are no longer expected to respond not only to classic production links but to a wide range of amenities (Brueckner et al., 1999; Kolko, 1999; Kotkin, 2000; Glaeser et al., 2001; Florida, 2002; Dalmazzo & de Blasio, 2011; Glaeser, 2011; Bauernschuster et al., 2012; Suedekum et al., 2012; Ahlfeldt, 2013).

Glaeser et al. (2001) classify urban amenities into the following four categories[1]:

1. Localised goods and services
2. Aesthetics and physical setting
3. Public services
4. Speed in terms of transportation

Guided by this classification, the following sections provide an overview of the recent literature on the spatial allocation problem with respect to different types of urban amenities. The main questions which arise are: How do households and firms value amenities? Are amenities able to explain the location decision of residents and firms?

The greater part of the applied spatial allocation literature addressing these questions relies on a hedonic pricing approach which goes back to Rosen (1974). The approach allows for deriving the implicit price for specific characteristics of a composite good from the market even though only the composite is traded and not the sub-attributes. The idea is, that in equilibrium, the marginal benefit of an additional attribute (e.g. a flat endowed with a balcony) offsets the utility costs of the extra expenditure involved, assuming income and consumer preferences are given. Housing expenditures are then used to derive the monetary value of its observable attributes. The implicit price of an attribute can be determined by estimating how a marginal change in the attribute changes the housing expenditure (Gibbons & Machin, 2008). This revealed-preference method therefore indicates how much a certain amenity is valued by residents or firms. Following up on the hedonic pricing idea one might alternatively investigate the spatial distribution of population or migration flows. If an area is characterised by a high amenity endowment and hence by a higher (short-term) utility, people will relocate to this location since they are better off at the new place. Exploiting compensating differentials

---

[1] In a similar way, Brueckner et al. (1999) suggest the division into natural, historical and modern amenities where historical and modern amenities are theoretically linked by renovation to each other.

and the spatial equilibrium condition one can either directly use changes in population as an indicator for a region's attractiveness or calculate the differences in rents and wages as in the quality of life literature (Rosen, 1979; Roback, 1982; Blomquist et al., 1988; Greenwood et al., 1991; Gyourko & Tracy, 1991). Most of the research reviewed in the subsequent sections as well as my own work are based on these very briefly summarised methodological ideas.

2.2. Localised goods and services

Manufactured goods can be considered as national or even global goods in the sense that their consumption is hardly limited to a certain location. These tradable goods can be ordered online or in catalogs and are, thanks to a significant reduction in transport cost over the last decades, available to anyone anywhere. A lot of services like restaurants, theatres, concerts, museums or hair saloons are, however, only available locally (Glaeser et al., 2001). Their product cannot be shipped and consumers need to travel to the service providers. People or firms with preferences for the consumption of certain non-tradables will therefore locate in its proximity.

Since the provision of urban amenities like operas, bars or clubs involves high fixed costs a critical mass is needed, which is more easily reached in dense urban areas. Producers as well as consumers therefore benefit from agglomeration economies in dense areas. Assuming heterogeneous preferences and fixed cost in the provision of local services enables cities to agglomerate people with niche tastes which results in an even greater product variety. Special book stores or antique shops, Michelin star awarded restaurants or exotic cuisines are therefore expected to be found more in big cities than in the periphery. Theoretic models building on the Dixit-Stiglitz assumption of monopolistic competition demonstrate how firms and consumers co-locate, creating concentrated consumption clusters (Fujita, 1988; Glazer et al., 2003).

Empirical evidence on how local goods and services form part of the location decision process is still rare and predominantly based on correlations. Glaeser et al. (2001) have shown that US counties better equipped with local consumption amenities grew more quickly between 1977 and 1995. In their multivariate regression approach the presence of restaurants and concert venues in the base year significantly explains later population growth. Whilst the number of art museums is not significantly correlated with population growth, there is a negative association with the number of movie theatres and bowling alleys per capita. Glaeser et al. (2001) explain these results by the role education plays with respect to amenity appreciation/consumption, i.e heterogenous preferences. A positive correlation between population growth and restaurants is also reported for France (1975-1990).

By finding a positive bilateral correlation between the number of museums and US metropolitan area population in 1990, Glaeser & Gottlieb (2006) provide descriptive evidence that consumer amenities rely on high fixed costs and are therefore more likely to be found in bigger cities. Making use of the DDB Needham Life Style Survey from 1998, they furthermore show that city residents are significantly more likely to visit art museums, go to a bar, dine out, go to the movies or pop/rock/classical concerts. Conversely, home entertainment is negatively correlated. Larger cities are hence not only endowed with a higher level of amenities, their citizens also consume urban amenities more often. In a similar exercise, Borck (2007) assesses the importance of various kinds of consumption amenities for Germany. Based on the German Socioeconomic Panel (GSOEP) 1993-2003, his estimates yield an increased probability of dining out, going to the movies or a concert for people living in bigger cities. Introducing individual fixed effects, estimates stay robust for going to the cinema and to concerts. Similar findings for the US are reported by T. N. Clark (2004), who constructs an amenity index based on the number of operas, “Starbucks” coffee shops, juice bars, brew pubs, museums and whole-food stores as well as bicycle events. Their index is positively correlated with the change in population between 1980 and 1990 as well as 1990-2000. Moreover, there is a positive association between the number of high tech patents (1975-1995) and consumption amenities.

Large cities are not only equipped with a high number of restaurants and museums but are often also home to big sports teams. Residents have the opportunity to visit home games and identify themselves with the team. Following a compensating differential approach, Carlino & Coulson (2004) estimate people’s indirect willingness to pay for having a professional NFL franchise in their neighbourhood. Their fixed effects estimates (on a city level) indicate that NFL franchise raises annual rents by about 8% in central cities. But not all local amenities are bound to large cities. Golf courses are for instance rather located in less dense areas. Investigating house transaction prices for a suburban area of Rancho Bernardo, California, Do & Grudnitski (1995) estimate a golf course location premium of 7.6%.

With a particular focus on the amenity role of restaurants, Schiff (2013) investigates the aforementioned hypothesis that localised service and product variety increases with city size. He constructs indices for cuisine variety based on a sample of 127,000 restaurants across 726 US cities. He finds that in the top quartile by land area, a one standard deviation increase in log population is correlated with a 57% rise in the number of unique cuisines. Moreover, holding population constant, a one standard deviation decrease of land area (increase in density) raises cuisine count by 10%. A positive relation between the number of restaurants and population has also been found by Waldfogel (2008). He additionally provides evidence on preference externalities, i.e., specific types of restaurants are more likely to be found in certain neighbourhoods. For instance, Chinese neighbourhoods are characterised by a high density of Chinese restaurants.

However, it is difficult to establish any causal relation between endogenous consumption amenities and the location of firms and households. To overcome this problem, Falck et al. (2011) propose an identification strategy based on a quasi-natural experiment in German history. Using proximity to baroque opera houses, they exploit the fact that opera houses were historically mainly built due to prestigious reasons and do not symbolise economic power. According to their causal inference, cultural amenities explain the distribution of high human capital employees between German city districts. A different approach to capturing the role and the value of urban amenities is proposed by Ahlfeldt (2013). He exploits a novel dataset of geo-tagged photos uploaded to online communities in order to capture “urbanity” - a composed measure of urban amenities. The definition of urbanity is therefore not restricted to localised consumption goods but is also expected to capture aesthetics and the architectural beauty of a city. His estimates for Berlin and London yield an indirect utility elasticity with respect to urbanity of 1% and demonstrate the important role amenities play in cities.

Summing up, urban consumption amenities tend to have a positive effect on the location of households. People move to amenity-rich areas and are willing to pay for the consumption of local services. However, this group of amenities seems to be understudied, most probably due to its endogenous nature.

2.3. Aesthetics and physical setting

When people think of a particular location, its physical setting or aesthetic appearance is probably what comes first to mind: California is generally associated with a mild climate and long beaches, New York with its skyline in Manhattan and Rome is known for its long history as expressed by its historic building stock. Architectural design or climate conditions are therefore often considered as a determining factor in the spatial allocation of households and firms. This section reviews the recent literature on how (i) natural amenities and (ii) how pure beauty, architectural design as well as heritage conservation are valued by people.

2.3.1. Natural amenities

Natural amenities, like the proximity to the nearest coast line, an elevated location or a mild climate might be considered as positive natural amenities by households, depending on their preferences. Applying a workhorse model in urban economics, Mahan et al. (2000) use hedonic regression techniques to estimate the amenity value of wetlands in Portland, Oregon. They find that property prices rise with increasing distance to the nearest wetland and prices go up by US-$ 436 per 1,000 feet. Quite opposite empirical results are reported by Wu et al. (2004) who find a positive price effect of being located closer to wetlands. Portland's inhabitants are moreover willing to pay higher house prices for locations with more open space and which are closer to parks and lakes. Elevation is also positively correlated with house prices. Anderson & West (2006) make further use of the hedonic price approach when examining the amenity role of open space in the twin cities of Minneapolis and St. Paul. Their location fixed effect estimates yield a positive amenity role of open spaces which is moreover higher in denser neighbourhoods. Furthermore, the valuation of proximity to open space depends on specific neighbourhood characteristics. Neighbourhoods characterised by high incomes, high crime and child rates value open spaces higher than average. Positive natural amenity values are also found within a nationwide study of one million housing transactions for England between 1996 and 2008 (Gibbons et al., 2011). Gardens, green spaces, areas of water as well as broadleaved woodland, coniferous woodland, enclosed farmland and freshwater and flood plain locations raise house prices at a ward level. The Travel to Work Area (TTWA) fixed effect estimates show, moreover, that house prices are higher in proximity to rivers, National Parks as well as National Trust sites.

Another strand of literature suggests that "people vote by their feet" (Tiebout, 1956). These mainly regional economic analyses try to capture the demand of certain location attributes by examining where people migrate to or what places experience the strongest growth in population. Rappaport (2007) estimates partial correlation between population growth and local climate variables on a US county level. He not only finds a positive correlation between the daily maximum temperature in January and population growth but, amongst others, also a positive climate association with growth in employment, number of elderly people and number of college graduates. Moreover, he observes a change in preferences over the last decades. People are more likely to follow nice weather, an observation in line with the consumer city idea. Similar empirical analyses have also been carried out for Europe. Using functional urban areas for the EU-12 countries between 1980 and 2000, Cheshire & Magrini (2006) also find that "weather matters". However, they conclude that climate amenities only affect population growth on a national scale but not between European countries. More empirical evidence for Europe is provided by Rodríguez-Pose & Ketterer (2012). Instead of population growth they use net migration on a NUTS1/NUTS2 level as dependent variable. Based on Hausman-Taylor estimations they find that migrants have a preference for milder climates.

A similar preference for warm coastal areas but based on a different approach is also found by Chen & Rosenthal (2008). Following the quality of life literature (Rosen, 1979; Roback, 1982; Blomquist et al., 1988; Gyourko & Tracy, 1991), they first develop a set of quality of business and of life indicators and then match them with US census information (1970-2000). According to their estimates, young highly-qualified move to areas with a good business environment whereas retirees move away from economic activity to high-amenity places in terms of nature and beauty.

2.3.2. Beauty and architectural design

How architecture or a city's "beauty" is perceived is a matter of taste and therefore not easy to generalise and value. To solve this problem Florida et al. (2011) make use of information on community satisfaction from a large-scale survey covering 28,000 people and 8,000 communities in the US. They find a strong positive correlation between a ranking measure of community satisfaction and physical setting and pure beauty.

Another approach to revealing residential preferences which is somehow similar to the use of survey data is to focus on referenda for individual partly public building projects. Sports stadia have become a very popular research subject in this regard. One example is the study by Coates & Humphreys (2006), who investigate the outcome of the referenda held on the (re-)construction of the Lambeau Field in Green Bay, Wisconsin (professional American football stadium), the Compaq Center and on a newly proposed arena (both professional basketball arenas) in Houston, Texas. Precincts in close distance to Lambeau Field (Green Bay) and the newly proposed Basketball area in Houston cast a significantly larger share of positive votes, whereas proximity has no particular impact for the existing arena in Houston. Matching the referendum data with Census tract information, Coates & Humphreys (2006) conclude for Green Bay that urban precincts, precincts with a high share of renters and of white collar jobs show stronger support for the stadium. In Houston, precincts characterised by a high number of renters are conversely more likely to be against the basketball arenas, whereas a higher share of blacks, college graduates and a higher median family income are positively correlated with positive votes. Ahlfeldt & Maennig (2012) find contrasting results when exploiting the 2001 referendum on the Allianz Arena (professional football) in Munich, Germany. Making use of a spatial autoregressive (SAR) model their estimates show that people living in close distance to the new grounds on average oppose the construction whereas at the overall city level the referendum was positive. They regard the results as being in line with the NIMBY (Not In My Backyard) hypothesis.[2] Dehring et al. (2008) extend the purely referendum-based approach by additionally using data on house prices. They first estimate a hedonic price function with variables capturing the proximity to the potentially new home stadium of NFL's Dallas Cowboys in Arlington, Texas. They particularly look at pre-referendum events in this step to capture potential signalling effects on homeowners. In a second step they match the market signals with the results from a 2004 referendum on the stadium. The authors provide evidence for the homevoter hypothesis; support for the stadium rises by between 0.9% and 1.2% for every US-$ 1,000 increase in house prices.

There are numerous purely hedonic analyses of sport facilities and arena architecture in particular. Feng & Humphreys (2008) for instance estimate a spatial lag model, using a contiguity as well as a distance-based spatial weight matrix, to obtain residential willingness to pay for major sport stadia. Their analysis is based on 10,000 transaction prices on family housing for 2000. They examine the price effect of two different facilities in Columbus, Ohio: the Nationwide Arena, which is home of NHL's Blue Jackets, and the Crew Stadium, a major league soccer (MLS) stadium. Their Spatial-2SLS approach with spatially lagged explanatory variables as instrumental variables (IV) yields that housing values increase by 1.75% for each 10% decrease in distance to the facility. The total willingness to pay sums up to US-$ 222.54 millions for the Nationwide Arena and US-$ 35.7 millions for the Crew Stadium. Ahlfeldt & Maennig (2010a) also make use of hedonic regression techniques in their analysis of three multi-functional sports arenas situated in the municipality of Prenzlauer Berg, Berlin, Germany; namely the Max-Schmeling Arena and Velodrom/Swimming Arena. As the arenas were built as part of the application for the Olympics 2000, special attention was paid to its architectural designs. Ahlfeldt & Maennig (2010a) find that the Velodrom has a significant positive effect on standard land values which decreases with distance. Whilst the Max-Schmeling-Arena has more ambiguous effects, the authors identify a total impact radius of three km for the Berlin sport arenas. Difference-in-difference estimates carried out in a follow-up study (Ahlfeldt & Maennig, 2009)[3]

---

[2] A NIMBY attitude describes residential opposition to a proposed new development (e.g. an airport) in close proximity to their homes. However, they often agree to the need for this new development, but not in their neighbourhood ("backyard").

[3] Even though the initially reviewed paper was published later, it was submitted earlier.

also yield a positive land value growth but this is only regarded as a short-run novelty effect. Similar price responses are also found for London's 2012 Olympic stadium. Property prices inside host boroughs rose by 2.1-3.3% after the announcement of the city's successful bid to host the mega event (Kavetsos, 2012). Based on year and location fixed effects price estimates are about 5% higher in a three miles radius around the Olympic stadium. Also examining external price effects of sports stadia in London but focusing on football, Ahlfeldt & Kavetsos (2013) find large effects on surrounding property prices originating from the New Wembley and Arsenal's new Emirates Stadium. With an impact area of 3-5 km, their difference-in-difference estimates are in range of the aforementioned studies. In particular they observe an increase in prices of up to 15% in close proximity to the New Wembley. Property prices go up by 1.7% for each distance decrease of 10% to the new Arsenal grounds.

Even though stadia provide a good example of iconic design, a large body of the literature tries to investigate architectural amenity effects of functionally less specific buildings, often making use of landmarks, price awarded or officially designated buildings. Professional expertise for measuring architectural quality is used, for instance, by Vandell & Lane (1989) as well as by Gat (1998). They both conclude that design impacts on the rents of office buildings in Boston and Cambridge and in Tel Aviv, Israel, respectively. Moorhouse & Smith (1994) take another approach and restrict their analysis to an arguably homogeneous neighbourhood in Boston's South End which is characterised by numerous Victorian style row houses. This set-up enables them to estimate the premium of differentiating design features which range between 11 and 20% of the price. Ahlfeldt & Mastro (2012) follow a similar approach by restricting their case-study to 24 residential buildings in Oak Park, Illinois designed by US architect Frank Lloyd Wright. They estimate a price premium of 8.5% within a radius of 50 metres. Moreover, they find that the premium decreases steeply with greater distance to Wright buildings and ranges around 5% within 50-250 m. The effects become significantly weaker and eventually diminish with even greater distance.

Cultural heritage is another example a location's beauty and physical appearance. Heritage designation has become a popular research subject, probably due its role as an important but controversially discussed planning policy. Following the definition of the UK Planning Act 1990, conservation areas are identified by a "special architectural or historic interest, the character or appearance of which is desirable to preserve or to enhance" (Section 69). One of the early studies which estimates the price effect of being located inside a historical district of Baltimore was carried out by Ford (1989). She observes a positive correlation between listed transaction prices for the years 1980-1985 and historic district location. A positive internal price effect has also been found by Schaeffer & Millerick (1991) for the Chicago neighbourhoods of Beverly Hills and Morgan Park. The effect seems, however, to depend on whether local or national authorities implemented the designation. While nationally designated areas experience positive price effects, local designation status is associated with decreasing prices. The authors explain the different outcomes by the fact that national designators are more likely to identify buildings or areas of a wider importance. In a study on Philadelphia's CBD, Asabere et al. (1994) also find a negative impact of locally designated areas, looking at apartment sale prices between 1980 and 1991. They estimate a negative premium on apartment prices by as much as 24%. In a similar study on owner occupied houses (1986-1990), Asabere & Huffman (1994) report, however, a premium of 26% if houses where located in federally certified historic districts. These findings stress the fact that the price effect might depend on the designator being in line with Schaeffer & Millerick (1991).

There exists a long list of hedonic price analyses which mainly contribute to the previously reviewed literature by expanding estimations to different cities and exploiting different price datasets. Deodhar (2004) for example estimate the internal premium for Sidney's upper north shore exploiting property sales between 1999 and 2000. On average property prices experience a premium of 12%. A comparable premium of 16% is is found for single-family residences in San Diego (2000-2006). Narwold et al. (2008) note that the premium paid exceeds the pure capitalisation due to tax savings, suggesting that built heritage generates an additional value. In Abilene, Texas, the estimated internal heritage premium is 17.6%. House prices inside a census tract moreover rise by 0.14% per additional designation (Coulson & Leichenko, 2001). Examining a sample of nine cities in Texas, Leichenko et al. (2001) find significant price premia ranging between 5 and 20%. A range of 14-23% is reported by Coulson & Lahr (2005) for historic designation effects on property values in Memphis, Tennessee, for a period between 1998 and 2002.

More recent analyses are on the one hand not only interested in the internal but also in the external effect of heritage designation. Houses in close proximity to preserved historic areas, presumingly with a direct view, are expected to experience a price increase as they indirectly benefit from the designation without being subject to any cost such as not being able to alter their façades. On the other hand, recent urban economic heritage literatures tries to address endogeneity concerns such as omitted variable biases. Noonan (2007), for instance, makes use of repeat-sales for Chicago between 1990 and 1999. This approach enables him to take differences in order to eliminate time-invariant characteristics which might be correlated with the variables of interest. His SAR estimates yield an internal price effect of 2% for each additional landmark inside a block. In a follow-up study, Noonan & Krupka (2011) introduce instrumental variables which are defined as interactions of historic quality and neighbourhood demographics. They find that standard OLS overestimates the effects, suggesting that districts have been systematically selected into the designation program. Instrumented estimates are even negative while they conclude that the overall policy effect for Chicago is close to zero. Noonan & Krupka (2011) also investigate the external effect and create mutually exclusive distance rings around designated areas. Their results suggest an external effect in close proximity. However, the ring estimates might likely suffer from multi collinearity. There also exists a number of heritage studies for Europe. Ahlfeldt & Maennig (2010b) for instance focus on condominium apartment sales in Berlin (2007). Computing a landmark density measure for neighbourhoods defined by an area of 600 metres, they estimate a marginal price effect of 0.10% per landmark. They test for external premia by separately including a distance measure to the closest landmark as well as a gravity-based accessibility indicator into the regression model. The distance bands yield an external premium of 2.8% for apartments within 50 metres of a landmark. The effect diminishes with greater distance and lies at around 1.4% for a distance bin of 50-100 metres. A steeply declining external premium is also found for the Dutch city of Zaanstad. Exploiting a panel of transaction data between 1985 and 2007 and estimating a SAR model, Lazrak et al. (2013) find a premium of 0.28% within a 50-metre radius of listed heritage areas. On average, a house sells for a price which is 26.4% higher if being located inside a conserved neighbourhood.

Koster et al. (2012) and van Duijn & Rouwendal (2013) furthermore try to investigate whether heritage appreciation depends on socio-economic attributes. Following a semi-parametric regression discontinuity approach, Koster et al. (2012) ask whether rich households sort themselves into amenity-rich city centres. Exploring a house sales database which covers approximately 75% of all owner-occupied housing transactions of the Netherlands between 2002-2009, they find a price difference of 5% at the boundary. They moreover observe that richer, older and higher qualified households have a higher willingness to pay for living inside a conservation area. Similar results for the Netherlands are found by van Duijn & Rouwendal (2013). Their residential sorting model accounts for the fact that residents are not only able to consume heritage in their own but also in neighbouring municipalities. They provide additional evidence on residential sorting into designated neighbourhoods where the highly educated have the highest marginal willingness to pay. Final simulation exercises indicate that the price of a standard house in Amsterdam would decrease by 17% in the absence of cultural heritage, and by 8% in Utrecht respectively. Finally, Ahlfeldt, Möller et al. (2013)[4] examine the economics of heritage conservation for England. They develop a theoretic model of the designation process in which they postulate that the optimal level of designation is chosen to Pareto-maximize the welfare of local owners. The implications of the model are twofold: An increase in historic preferences, proxied by education, is expected to raise the likelihood of additional heritage designation. This is tested and confirmed by IV Tobit estimations using UK census data between 1991 and 2011. Secondly the model implies that new designations at the margin should not be associated with significant house price capitalisation effects. Standard as well as spatial difference-in-difference estimations of the benefits (externalities) and cost (restrictiveness) of heritage conservation to local homeowners are in line with these expectations.

To sum up, recent literature has shown that people are well aware of their surroundings and appreciate the amenity value of heritage, climate or architectural design. Most architectural features not only have an internal but also an external effect where the range of the spill-overs might vary quite strongly. New construction projects are evaluated and supported by home owners if they expect an increase in house prices, following the home voter analysis. People might however also be against projects in their close proximity as suggested by the NIMBY literature. Spatial planning policies are therefore often introduced to solve coordination problems inherent to free markets. Recent literature, specifically looking into the role of heritage,

---

[4] The paper is part of this thesis and is discussed in greater detail in Chapter 4.

has found evidence for the availability of heterogeneous preferences with respect to heritage appreciation, like the urban phenomena of gentrification.

2.4. Public services

A third type of urban amenities as categorised by Glaeser et al. (2001) is summarised under the category of public services. This is obviously a very broadly defined category. The following review is therefore restricted to schooling as well as crime - two amenities which are among the most popular in applied urban economic research.

2.4.1. Schooling

The hedonic house price approach is also popular when it comes to the valuation of school quality. Especially in countries where household location determines what school a pupil attends (catchment-areas) the hedonic price approach provides a way of estimating the parent's willingness to pay for their children's education. A lot of educational researchers have followed this approach and improved and refined the methodology over time, moving from correlation to causality. The median premium of most of these studies is about 4% and an inter-quartile range of 4% whereas refined boundary discontinuity yields a median figure of 3.5% with an inter-quartile range of 1.3% (Gibbons & Machin, 2008; Machin, 2011). According to Nguyen-Hoang & Yinger (2011), the majority of the analyses were carried out for the United States (36), followed by the United Kingdom (6), France (2) and Norway (2). The subsequent paragraphs briefly review a selection of hedonic price regressions to assess the value of school quality.

Looking into the percentage of students reaching 9th grade proficiency in Ohio 2000/2001, Brasington & Haurin (2006) observe a positive association between house prices and school performance. Prices go up by 7.6% if school performance increases by one standard deviation. Cheshire & Sheppard (2004) follow a traditional hedonic approach, attempting to capture as many (un-)observables as possible by applying a wide range of neighbourhood and socio economic controls. Their estimates yield a premium of 9.8% for a one standard deviation increase in primary school test performance. L. Rosenthal (2003) follows an instrumental variable strategy to establish a causal relation between house prices and school quality. Inspections by the Office of Standards in Education are used as a source of exogeneity. Her estimates indicate a dwelling price elasticity of 5% with respect to exam performance based on data for England between 1995 and 1998.

Figlio & Lucas (2004) generate a large panel dataset using repeat real estate transactions between the mid-1980's and 2002 for Florida. Exploiting the rich panel information and applying a wide range of fixed effects as well as neighbourhood-year interactions they detect a 10% premium for schools which received an "A" as a grade in each year. Panel estimation techniques to deal with endogeneity are also used by Clapp et al. (2008) who examine the relation between property prices and school district attributes such as 8th grade math test scores for Connecticut (1994-2004). Their estimate yield a 1.3 to 1.4% price increase for a positive one standard deviation change in math scores. Bogart & Cromwell (2000) make use of a school redistriction in Shaker Heights, Ohio (1987). Their difference-in-difference estimates based on transactions between 1983 and 1994 indicate that a disruption of neighbourhood schools lowers house values by 9.9%.

Discontinuity designs using administrative boundaries have become a very popular tool in the more recent literature. The idea is to compare households which are in close proximity to each other and are hence assumed to be subject to the same observable and unobservable area effects but at the same time are on two sides of an administrative boundary. This boundary can, for example, be a school attendance zone which determines the treatment. Households are assumed to differ only in terms of access to school quality as common neighbourhood factors are eliminated. The boundary discontinuity approach has, amongst others, been used by Fack & Grenet (2010) to establish a causal link between middle-school test results and house prices. Looking into data for Paris (1997-2003), they find that a one standard deviation increase in test performance leads to a rise in house prices by 2%. With 3.5%, the price effect is found to be significantly stronger for high-school test results in the Australian Capital Territory between 2003 and 2005 (Davidoff & Leigh, 2008). Similar results are also reported for the London area (1996-2001) using the proportion of students reaching the target grade in primary school English, science, and maths tests by Gibbons & Machin (2006) who extend the discontinuity approach by using changes over time within geographical school clusters and by instrumenting school performance with salient school characteristics. A one standard deviation increase in performance causes house prices to rise by 3.8%. More positive causal links based on the boundary discontinuity approach are found for Boston, Massachusetts (1993-1995) using elementary math and reading scores (Black, 1999), for British primary schools relying on grade data from maths, science, and English tests between 1996 and 1999 (Gibbons & Machin, 2003) as well as for elementary schools in Mecklenburg County, North Carolina, exploiting math and reading scores between 1994 and 2001 (Kane et al., 2005). Machin & Salvanes (2007) make use of a quasi-experimental change in education politics in Oslo 1997, which enables school choice independent from living in a certain catchment area. Their discontinuity estimates yield a house price premium of about 2-4% for a one standard deviation increase in average marks before the reform. In a recent paper, Gibbons et al. (2013) introduce several methodological improvements to the discontinuity approach like the matching of identical properties across boundaries, the inclusion of spatial trends and boundary effects or the re-weighting of transactions that are closest to district boundaries. They observe prices going up by about 3% in response to a one standard deviation change in school average value-added based on age-7 to age-11 test scores as well as to prior achievement using a sample covering the whole of England.

A slightly different focus is used in the study by Gibbons & Silva (2008) when looking into the relation between urban density and school performance. Exploiting the compulsory switch from Primary to Secondary Education and using census information for more than 1.2 million pupils in England they find significant but small benefits on education in dense urban areas.

The authors explain their results by higher competition and greater school choice in dense urban areas.

2.4.2. Crime

One often-cited reason for the renewed interest in cities and urban resurgence is the drop in crime rates over the last decades. Between the 1960's and 1980's New York City was characterised by historically high crime rates and workers in big cities had to be compensated for crime disamenities by higher wages (Glaeser & Gottlieb, 2006). In fact, examining 127 US cities, Cullen & Levitt (1999) find that one additional reported crime per capita reduces population by 1%. Their analysis is based on data covering the 1970's till the 1980's. They address endogeneity concerns by instrumenting city crime rates by lagged changes in the punitiveness of state criminal justice system.

Crime rates are generally found to be higher in cities compared to smaller towns and rural areas. Looking at the correlation between different types of crime and city population, Glaeser & Sacerdote (1999) report an overall elasticity of about 0.16 where the elasticity for murder is estimated to be twice as large (0.32). In their subsequent analysis, Glaeser & Sacerdote (1999) try to understand what determines higher crime rates in cities. They argue that higher pecuniary benefits, lower probabilities of arrest and recognition as well as the presence of more female-headed households are among the most important factors in explaining high urban crime rates.

Similar to school amenities, there is a long tradition in valuing the role of crime in urban economics. A straightforward approach to measuring the costs of crime is pursued by Brand & Price (2000), who sum up a crime-related-costs-reported-in-victimisation survey. Based on the British Crime Survey, they derive average costs for different crimes like burglary (£2,300), robbery (£5,000) or homicide (at least £1 million). These calculations ignore any psychological costs not taking into account the fear of crime or its risk. This might be one of the reasons why most analyses on assessing the costs of crime rely on hedonic house price estimations. In contrast to schooling, where attendance-zones can be exploited in a boundary discontinuity design, for crime analyses it is difficult to eliminate unobservable neighbourhood effects. Most estimates can therefore not be interpreted as describing a causal link.

One of the early hedonic analyses was carried out by Thaler (1978) who examines crime data from the Rochester Police Department in the state of New York of 1971. His estimates yield a price reduction of 3% followed by a one standard deviation increase in property crime rates. A negative relation is also found by Hellman & Naroff (1979). They estimate a house price elasticity of -0.63 with respect to (total) crimes for the city of Boston, Massachusetts. D. E. Clark & Cosgrove (1990) apply a two-stage hedonic model for public safety, making use of the 1980 Public Use Microdata Sample for the US. A 10% increase in their public safety index is associated with a rise in monthly rents by 1.3%, additionally providing evidence for the expected disamenity character of crime. A positive relation between the number of crimes and property prices have, against expectations, been found by Lynch & Rasmussen (2001) for data on Jacksonville, Florida (1994/95). They stress the importance of properly weighing a crime measure used in the estimations. Their improved measure, where they weigh crime offences by the cost of crime to victims, then yields a negative elasticity of -0.05. Bowes & Ihlanfeldt (2001) investigate the effect of crime on house prices in the context of proximity to rail stations. They regard crime as a negative underlying factor of locations close to rail stations due to a better access to a neighbourhood for outsiders. Using sale prices of single-family homes in the City of Atlanta and DeKalb County they observe a reduction in house prices by 3% per additional crime per acre and year.

Schwartz et al. (2003) study the price effect of crime for New York City covering a period ranging from 1976 to 1998. They first of all observe significant reductions in crime rates over the years as initially stated. For New York City, murder rate fell by 69%, property crime by 56% and violent crime by 53% between 1988 and 1998. They address the problems arising from unobserved location effects by using repeat-sales, which enables them to difference out any unobserved time-invariant neighbourhood characteristics. Their panel estimates yield an elasticity of violent crimes of 0.15. Finally, Gibbons (2004) approaches the endogeneity problem by estimating non-parametric models jointly with instrumental variable techniques. In a standard fixed effect fashion he uses deviations from the local spatial average of the variable. Instrumentation is built on the difference between spatially lagged values of crime rates, offences reported on non-residential properties as well as alcohol consumption proxied by distance to nearest public house or wine bar, depending on the specification. Gibbons (2004) estimates a 10% reduction in London house prices (1999/2000) for a one standard deviation increase crime density (incidents per square kilometre). Moreover, high incidence burglary does not impact on house prices, whereas small incidences like graffiti, vandalism or damage to property induce a negative price effect. He explains the difference by house buyers not being able to directly observe the first type of crime while vandalism might be a sign of an unstable neighbourhood.

The reviewed studies have shown that public amenities have a significant impact on the surrounding neighbourhood. Particularly focusing on the positive amenity of good schooling and the disamenity of crime, the value might differ over space and time. Overall, parents are willing to pay a certain premium for good schooling which can be estimated from house prices. Residents also value a safe neighbourhood and high crime rates are generally associated with a negative premium. Similar to the local endowment of private goods and services, public services are subject to notable endogeneity, challenging empirical researchers who seek to obtain unbiased estimates.

2.5. Transportation

The disciplines of economic geography in general and urban economics in particular originate from questions regarding the costs of moving goods; crops, as in von Thünen (1826) and people, as in Alonso (1964); Mills (1967, 1969); Muth (1969). Glaeser et al. (2001) emphasise the “speed” of urban transportation in their categorisation of urban amenities. In fact, the majority of transport studies examine the impact of reduced travel time due to new railroads or new highways on house prices and population growth.[5]

Early hedonic price study estimates are usually based on comparisons between house prices before and after a new line was constructed, being rather descriptive and far from causal inference. Dewees (1976) for instance cross-sectionally estimates a hedonic price function using property sales for Toronto in 1961 and 1971. His analysis focuses on a 13-mile-long stretch where streetcar service was replaced by a new subway line. He finds an increase in the slope of the price function which declines perpendicularly from the tracks and diminishes at around 1/3 mile from a station. Another early analysis of Toronto finds positive price effects, too, making use of 385 housing transactions in 1978. Bajic (1983) estimates that households using the new Spadina Subway Line save up to 34 hours in commuting time per year, an implied value of US-$ 120. Reviewing twenty years of the Bay Area Rapid Transit system (BART) in San Francisco, Cervero & Landis (1995, 1997) generally find positive price effects: Locations close to a BART station experience, on average, a premium of 12% in contrast to freeway junctions; the premium diminishes by US-$ 2 per meter distance from a station. Bowes & Ihlanfeldt (2001), who consider positive as well as negative underling factors of being closely located to a rail transit station, find mixed results for their analysis of sales prices in the City of Atlanta and DeKalb County. They observe negative effects in close proximity to stations which become positive at a distance of about 1/2 mile. Armstrong & Rodríguez (2006) distinguish between four municipalities with commuter rail service access in Eastern Massachusetts and three municipalities without access. Their comparison relies on a total of 1,860 transactions of single-family detached properties in 1992/1993 and yields a transport premium of 9.6-10.1%. Their spatial hedonic price estimates further indicate a reduction in prices by 1.6% per additional minute of travel time to a station.

The recent research focus has been on more carefully solving the identification problem, stating that infrastructure supply and demand are determined simultaneously: Planners have certain expectations and infrastructure is not only built based on past but also on future/expected urban development. McMillen & McDonald (2004) exploit a large dataset of repeat sales between 1983 and 1999 which enables them to differentiate out time-invariant characteristics. Examining the effect of a new rapid line from downtown Chicago to Midway Airport, they observe significant anticipation effects. The improved accessibility is already capitalised into house prices six years prior to the official end of the construction works. House price gradients with respect to proximity to the nearest station decrease from -4.2% per mile before 1987 to a slope of -19.4% between 1991 and 1996, indicating the positive price effects induced by new transport opportunities.

---

[5] The focus of this review is placed on railway- and highway transportation since it is intended to provide a background for the rail-bound transport analyses carried out in Chapters 5 and 6. However, there exist a long list of literature on other transport hubs such as airports (Feitelson et al., 1996; Carlsson et al., 2004; J. Nelson, 2004; Ahlfeldt & Maennig, 2007; Nitsch, 2009). See for instance Ahlfeldt & Maennig (2007) who estimate a land value discount of 5-9% within a distance of 5 km along the air corridor of Berlin's former downtown airport Tempelhof due to noise pollution. Investigating the disamenity effect of noise using voting data from a referendum on the closure of the same airport, where the closure is interpreted as natural experiment, Nitsch (2009) finds that airport noise is surprisingly weak in explaining voting behaviour.

Another attempt to deal with the simultaneity problem while predicting transport effects is to directly assume exogeneity in a quasi-experimental setting and to use a difference-in-difference estimator. Baum-Snow & Kahn (2000) investigate the expansion of urban transit for five US cities between 1980 and 1990. Their difference-in-difference estimator yields an increases in mean prices of US-$ 4,972 when moving from 3 km to 1 km with respect to distance to nearest transit. Another prominent example, which moreover uses data on a fine spatial level, is Gibbons & Machin (2005), who interpret the extension of London's Jubilee Line and the Docklands Light Railway during the 1990s as a quasi-experiment. Their difference-in-difference approach yields a positive response of house prices to improved access and a reduction in travel time in particular. House prices increase by 1-4% per 1 km reduction in proximity to stations, depending on the specification. This translates into an upward shift of 7-20% in response to a reduction in travel time by one standard deviation. Quasi-experiments have also been used by Michaels (2008), observing a trade increase in rural US counties caused by the construction of the Interstate Highway System, as well as by Ahlfeldt & Feddersen (2011) who find a positive economic effect in two rural cities in Germany which were connected to the high speed railway network. Their estimates yield a GDP elasticity with respect to market accessibility of 0.25-0.30. The exogenous variation of the two latter studies is based on the assumption that the rural locations were not connected on purpose but by accident, as the planners intended to connect only major cities. Another paper which applies a difference-and-difference estimator and moreover a similar dataset to the analysis carried out in Chapter 5 is Ahlfeldt et al. (2011). Exploiting a highly disaggregated dataset for Berlin covering a period between 1881 and 1914, they use the inauguration of the city's first metro line in 1902 as a quasi-experiment to disentangle the positive accessibility effects from negative noise disamenities. Their hedonic land value estimates yield an accessibility premium of 11% per km, and negative noise effect which declines by 0.15% per additional decibel.

One strand of literature opened up by Baum-Snow (2007) follows an instrumental variable approach using historic infrastructure plans to generate exogenous variation. By instrumenting recent highway development using a 1947 national interstate highway plan, he establishes a causal relation between the provision of new highways in the US and suburbanisation. Baum-Snow (2007) concludes that the construction of a new highway which passes through a central city depopulates the centre by 18%. Duranton & Turner (2011) extend the instrumenting literature, introducing plans of historical major rail roads (1898) and expedition routes between 1835 and 1850. They predict an increase in vehicle kilometres travelled (VKT) following the provision of new roads. Duranton & Turner (2012) further find empirical evidence for a positive population as well as a positive employment response caused by an increase in urban road stock following a similar IV approach. These results are replicated for Spanish highways and cities by Holl & Viladecans-Marsal (2011) and for Japan by Hsu & Zhang (2011). Baum-Snow et al. (2012) carried out similar analyses by investigating decentralisation trends in China. They particularly stress the importance of infrastructure in countries which are still developing and invest highly into their infrastructure. They find population elasticities for radial highway construction comparable to those reported for the western hemisphere. A few of the aforementioned studies also investigate the effect of improvements in the public transport network. Baum-Snow et al. (2012) show that railroads have a significant effect on industry location in China. Public transportation, however, has no effect on VTK in the US (Duranton & Turner, 2011) whereas urban population grows with a city's stock of large buses (Duranton & Turner, 2012).

The quasi-experimental and the IV approach of circumventing the simultaneity problem share a common idea: They both generate exogenous variation in the transport provision to separate supply and demand. What follows is an investigation of the isolated supply side effect of transport on the allocation of land use; a new line or road is built and the urban structure adjusts to the new situation. So far, demand side driven reactions of transport infrastructure have been largely ignored by economists. However, as initially argued, it is very likely that planners take future urban development into account when planning new roads and lines, for instance to respond to residential commuting needs. Firms could also request (lobby for) a better access to workers and customers. There is very little literature on the demand-side driven relation between transport and development. Levinson & Karamalaputi (2003) represent an exception. Their analysis yields a positive effect of population on highway lane expansion in Minnesota (1978-1998).

However, the process must neither be exclusively supply nor exclusively demand side driven but could also work simultaneously. Levinson (2008) calls this joint process of infrastructure and land development co-development. In this situation, residents and firms incorporate the changes of the transport network in their location decision. At the same time rail extensions depend on the location of economic agents. Levinson (2008) empirically tests for a co-development process using a sample of 33 London boroughs for a time period between 1871 and 2001. He first tests whether a change in the transport network Granger causes population to rise (fall) in the periphery (CBD). In a rather ad-hoc approach, he then reverses the estimation equation to see whether population affects station density, too. The analysis confirms the hypothesised co-development process for London. In a follow-up study on the twin cities Minneapolis and St. Paul, Xie & Levinson (2010) could, however, find only evidence for a supply side driven relationship. Unfortunately, Levinson's approach does not allow for a simultaneous adjustment of transport and land development. If the two variables influence each other one should account for simultaneous effects in order to obtain unbiased results as well as robust standard errors. Cervero & Hansen (2002) apply a three-stage least square (3SLS) approach to simultaneously estimate induced demand and induced supply effects. Applying and testing a wide range of potential instruments, they eventually conclude that road supply has been both a cause of and a reaction to vehicle miles traveled for a sample of 34 California urban counties (1976-1997).

Overall, the transport literature finds mostly positive effects of new projects on economic activity, population growth or on house prices. For households and firms, geography and physical distance matter for transporting goods, for commuting as well as for market accessibility.

Recent literature has tried to strengthen the causal link between the provision of transport and economic activity, where the focus has been on a rather supply-side, unidirectional relation.

2.6. Firm location choice

Amenities might not only attract people but also determine the location of firms. So far, the majority of firm location studies has concentrated on the production side and “hard” economic factors like access to resources and consumers, localisation economies as well as tax levels, economic activity and FDI (Guimarães et al., 2000; Kolko, 2001; List, 2001; Figueiredo et al., 2002; Guimarães et al., 2003; Crozet et al., 2004; Holl, 2004; Isaksen, 2004; S. Rosenthal & Strange, 2005; Devereux et al., 2007; Arzaghi & Henderson, 2008; Hong, 2009; Claussen et al., 2010; Brülhart et al., 2012). However, firms, especially in the service sector, might be attracted to the same type of urban amenities as households. At the end of the day, it is people who work and run companies who themselves consider cities not only as a place of production but of consumption (Glaeser et al., 2001).

The significant reduction in transport and hence shipping costs as well as an increase in the economic importance of knowledge-based companies, with neither a capital- nor a land intensive production, have made firms more flexible in deciding on where to manage their company and where to produce and sell their products. Service firms, especially, are becoming increasingly footloose thanks to portable computers and internet services. These firms and their employees are theoretically able to live anywhere (Kotkin, 2000). As qualified labour becomes the most important (and sole) input for service firms, these companies increasingly depend on the location of their (potential) employees. This would imply that firms follow their workers and not the other way around (Kolko, 1999). If service firms follow skilled labour, they act as amenity maximisers when deciding where to locate. Gottlieb (1995) identifies two potential channels for this relation. On the one hand, it is a company’s executive or an entrepreneur who maximises his own utility with respect to urban amenities and moves his firm to an amenity-rich location. On the other hand, workers might drive the firm location by either being irreplacable or accepting lower wages as a compensation for living in a pleasant neighbourhood. In his subsequent empirical analysis of questionnaires, Gottlieb (1995) finds that preference for amenities was explicitly pronounced in high-technology firms. In a reduced form model where he regresses employment in engineering and management services on classic firm location determinants and a wide range of urban amenities he finds (descriptive) evidence for the amenity maximising firm hypothesis (Gottlieb, 1995). A similar approach was followed by Kolko (1999) who restricts his analysis to natural amenities as a source of exogenous variation. According to his results, natural amenities like proximity to the nearest coast and climate variables do not significantly predict the location of firms. An early study looking into the effect of architectural design on office space in downtown Chicago is Hough & Kratz (1983). Architectural significance is determined as being a historic landmark in the case of an older building and by being awarded by the Chicago American Institute of Architects if it is a newer one. Hough & Kratz (1983) find a rent premium for “new” but not for “old” architecture. Examining the role of rather endogenous amenities, N. Lee & Nathan (2010) make use of the 2007 London Annual Business Survey to investigate the effect of cultural diversity on firm innovation. Even though their cross-sectional estimates do not permit a causal interpretation, they conclude that culturally driven amenities can become an important economic asset for a city.

Summing up, the link between urban amenities and firm location is highly understudied in economics. Early analyses yield diverse results but generally consider amenities as an important economic asset.

The literature review has shown that even though the consumer city idea was recently (re-)vitalised by Glaeser et al. (2001) amenities have been studied for a long time. Early hedonic regression analyses try to assess the value of architectural design, proximity to a transport hub or the effect of high crime rates. The more recent development can be described by a diversification in terms of research questions as well as by methodological improvements. In particular, this review demonstrates the tendency in applied empirical work of moving from pure correlations towards causal inference. The following chapters are intended to build on the reviewed studies by investigating different aspects of the amenity role in the location decision process of households and firms.

3. Culturally clustered or in the cloud? Location choice of internet firms in Berlin

3.1. Motivation

In the past, manufacturing firm location was characterised by classic/first nature type location factors like natural advantages, cheap land and labour, or later physical infrastructure. Today's knowledge-based economy, however, is based on the idea of generating and quickly spreading innovation. IT companies for instance have neither a capital- nor a land intensive production. They are highly footloose thanks to portable computers and wireless internet. Due to these technological improvements as well as a significant reduction of travel and transport costs over the last decades New Economy firms and its employees are theoretically able to live/work anywhere (Kotkin, 2000). As qualified labour becomes the most important (and sole) input for service firms these companies increasingly depend on the location of their (potential) employees. This would imply that firms follow their workers and not the other way around (Kolko, 1999).

Highly qualified and "creative" individuals have a strong preference for a rich social and cultural life (Florida, 2002). According to social science, these creative heads can be assigned to a new social milieu which has evolved over the last years. They have been labeled "movers and shakers" ("Experimentalisten"), the unconventional creative avant-garde, the new Bohemia (Sinus Sociovision GmbH, 2011). Members of this milieu are very individualistic, digitally networked and highly mobile in geographical as well as in mental scope. I consider this milieu as the driver of a currently observable start-up boom in Berlin and expect them to be highly attracted by a distinct provision of urban amenities.

As reviewed in Section 2.2, the provision of urban amenities like theatres, bars or clubs involves high fixed costs and a critical mass is needed which is easier reached in dense urban areas. Cities have therefore been more and more regarded as a place of consumption than of production (Glaeser et al., 2001). Cities are not only endowed with a higher level of amenities, their citizens also consume urban amenities more often (Glaeser & Gottlieb, 2006).

If service firms follow skilled labour those firms act as amenity-maximising agents when deciding where to locate (Gottlieb, 1995). Amenities can therefore become an important economic asset for a city. Even though the important role of amenities is highly accepted in the urban economic literature, most amenities tested empirically do not explain the whole story. Quality of life indices based on compensating differentials implicitly control for amenities but do not allow for the determination of distinct effects (Gabriel & Rosenthal, 2004; Chen & Rosenthal, 2008). Measures of local amenities like distance from a major coast and average annual precipitation (Kolko, 1999) or other climate amenities like July/January temperature (Glaeser et al., 2010) have definitely the advantage of being purely exogenous but ignore the discussion on urban consumption amenities. It is questionable whether these amenities are able to attract a young footloose generation – the movers and shakers – founding and working for internet firms. As economic conditions and technology change, society changes as well.

Measures which might be more appropriate are, for instance, the cuisine variety a location offers (Schiff, 2013), or music nodes and clubs (Ahlfeldt, 2011a). However, since urban amenities are man-made, they are highly endogenous. Estimates are therefore most likely subject to severe omitted variable biases. This might be a reason why there are very few attempts to include endogenous amenities in econometric analyses.

Motivated by this lack of empirical evidence, I contribute to the literature of firm location and consumer cities by testing the amenity-oriented firm location hypothesis: Knowledge-based service industries locate at urban amenity-rich places. In particular, I concentrate on local service/consumption goods like restaurants, bars or theatres. Throughout the chapter I label the composite of local consumption goods as cultural amenities.

I test the stated hypothesis empirically by looking at the rise of the internet industry in Berlin over the last years. First of all, internet firms provide a perfect example of the knowledge-based service sector which is highly footloose. Secondly, potential labourers as well as the firms' entrepreneurs can be characterised as relatively young, highly qualified and somehow creative individuals who are expected to be attracted by urban/cultural amenities. Thirdly, limiting the analysis to start-ups enables the assumption of taking the existing economic environment as given. The location choice is expected to be unconstrained by previous firm decisions (S. Rosenthal & Strange, 2003). And finally and most importantly, I use the sudden fall of the Berlin Wall as a source of exogenous variation. Nowadays, Berlin is globally known as having an open, creative and artistic environment which is regarded as fertile ground for innovation. A specific subculture has evolved in the aftermath of German reunification which still strongly affects today's cultural scene. The subcultural development originates from the open, chaotic and tolerant environment after the fall of the Berlin Wall. The "wild east" with its political vacuum and abandoned places became home to artists and creatives (Schwannhäußer, 2007). I make use of this very particular subcultural development and use proximity to the former Wall as well as squat density as instrumental variables for the presumably endogenous current endowment of cultural amenities. The applied instruments are assumed to affect the location of internet start-ups only indirectly via the cultural amenity channel conditional on a large set of controls like land values or centrality. The exclusionary restriction is strengthened by the time dimension and the fact that the internet was not used commercially/by the general public during the fall of the Iron Curtain but only became popular at the end of the 1990s.[1]

The next section (Section 3.2) provides an overview of the development of the internet industry in Berlin as well as the city's cultural development. I state more reasons for using Berlin as a case study. I also provide arguments in favour of an intra-urban analysis. Section 3.3 introduces a footloose start-up model to motivate the empirical strategy which is outlined in Section 3.4. Section 3.5 provides an overview of the data used, followed by the discussion of the empirical results (Section 3.6). Previewing my results, I conclude in the final Section 3.7 that cultural amenities indeed affect the internet start-up location.

3.2. Internet industry in Berlin

Today, the "Nerd Revolution" (tip, 2011) describes the growing number of internet start-ups founded not by business students but by computer developers. Berlin seems to provide a hub function, at least in Germany, for this latest development, sometimes even compared to Silicon Valley. There are more than 5,700 firms with over 50,000 employees working in the IT and communication sector (Berlin Business Location Center, 2012b). A lot of international investors, mainly venture capitalists and business angels, not only visit Berlin but move to the city to financially support and collaborate with local start-ups. Moreover, experts predict further growth of this fairly young sector and even expect that the next Facebook will come from Berlin.

To get an idea of the movement's origins this section sums up Berlin's recent history with respect to the main research questions and provides arguments in favour of an intra-urban analysis.

3.2.1. (Sub-) Cultural rise after re-unification

The fall of the Berlin Wall in 1989, which had run through the heart of pre-WWII Berlin, reshaped the city's geography. Former border locations like West Berlin's Kreuzberg as well as today's Mitte and Prenzlauer Berg in the East were all of a sudden in the new geographical centre of the city, causing a re-newed interest in the historical CBD. Looking at rents, Ahlfeldt et al. (2012), for instance, observe a re-emergence of the former rent gradient towards East Berlin's district of Mitte.

Due to the fall of the Iron Curtain Berlin's population suddenly rose from 2.1 million (West) or 1.3 million (East) in 1989 to 3.4 million. This implies a sudden increase in economic mass and market size. Accessibility to a wide range of physical amenities (parks, water bodies), social amenities (friends and family) and cultural amenities also experienced a strong rise. A higher number of residents decreases the cost of provision of certain cultural and public goods.

---

[1] In fact, in 1993 the National Center for Supercomputing Applications (NCSA) at the University of Illinois Urbana-Champaign released Mosaic - the first browser which made the internet available for the general public (Vetter et al., 1994).

This is of special interest for service industries providing local non-tradable goods since they are characterised by high consumer transport costs as well as by a required critical mass due to high fixed costs (Schiff, 2013). Assuming that customer's willingness to travel to e.g. restaurants is described by a steep spatial decay (i.e. they are not willing travel far), these places will cluster in central areas additionally allowing for a greater variety (Fujita, 1988; Glazer et al., 2003).

Due to its history and its renewed status as German capital[2], Berlin is by definition a tourist magnet. Tourists have an additional interest in services like restaurants, bars and theatres, especially in the historic CBD. Hence, the reborn historical centre offers new potentials for services due to the improved accessibility.

Owing to underinvestments of the local GDR government in historical built-up structure, a lot of East Berlin buildings were abandoned, rents were low. Empty houses, the political vacuum and the new tolerant, open environment drew in artists as well as squatters[3] (e.g. in Prenzlauer Berg or Mainzer Straße in Friedrichshain) and eventually students. Abandoned warehouses and industry complexes provided free/open space for artists and cultural events. A lot of techno music clubs were established in empty buildings. Night life was young and vivid. Curfew did not exist (and still does not). This pioneering development has been increasingly commercialised over the years: In the summer of 1999 the techno parade “Love Parade” attracted more than 1.5 million visitors. Electronic music clubs like Berghain located in the district of Friedrichshain ranked as number one techno club in the world in 2009 (DJ Mag, 2009). Nowadays, the city attracts easyjetters (Rapp, 2009) from all over Europe. However, this young, mobile and often highly skilled generation do not always return to their home countries but stay in Berlin, settle and look for jobs.

Amongst four universities, eleven technical colleges, a great number of research institutes, Europe's largest fibre glass network as well as a wide range of sector specific exhibitions (e.g. Berlin Web Week, Droidcon, re:publica, Social Media Week etc.), it is the quality of life which is an often quoted argument for start-ups locating in Berlin. As O'Leary, partner at the venture capitalist Earlybird, puts it:

> “There is no other place in the world where I can find such a bunch of creativity and freedom.”

His company as well as fellow venture capital funds invested more than in 136 mio. Euros during the first three quarters of 2012.

3.2.2. Berlin discovers the internet

By mid-1999 the German internet industry was lagging behind the US economy by five years (McGrane, 2000). It was exactly in that year that the German internet economy kicked off after the Samwer brothers sold their first German internet start-up to a US company. After having experienced the work and management environment in Silicon Valley, in 1999 the three brothers

---

[2] On June 20th, 1991 the German Parliament decided to move the capital of reunified Germany from Bonn back to Berlin.

[3] Who yet might be considered as the pioneers of today's gentrification (Clay, 1979; Friedrichs, 2000).

moved back to Cologne, Germany, and subsequently founded the online auction house alando.de in a backyard in Berlin-Kreuzberg. Only six months later they sold the company for US-$43 million to eBay. This can be regarded as the start of the Berlin internet economy. From that moment on Berlin transformed itself into Germany's Mecca for young internet entrepreneurs.

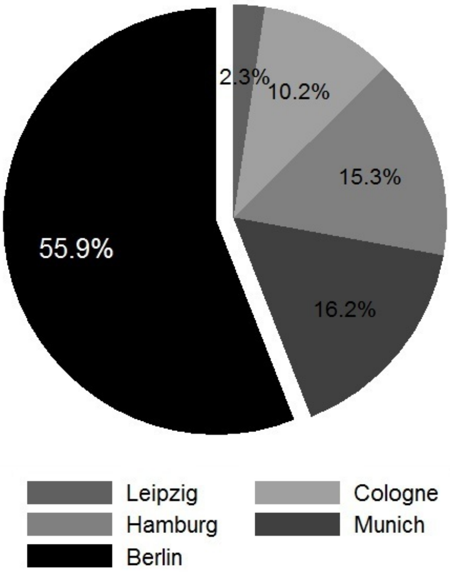
Quickly, agglomeration economies came into play. From the very beginning, the Samwer brothers tried to establish strong linkages within the Berlin founder community. Start-up Lounges, weekly breakfast rounds and seminars were supposed to foster the exchange of ideas and experience regarding the founding process. Following classic Marshallian externalities, spillovers and a highly specialised labour market support the development of the local internet industry. In a sector which is characterised by mainly young companies bearing a high risk to failure, the exchange of experience is of even greater importance compared to "mature" industries.

Additionally, the young sector was spurred by important financial as well as technological developments: The introduction of the "Neuer Markt" (1997) - German equivalent to US Nasdaq - made it easier for the new start-ups to raise capital from venture capitalists. Moreover, internet became cheaper, faster and, with the introduction of Apple's iBook in 1999, even footloose. The iBook was the first portable computer with integrated wireless network (WiFi) which not only allowed for saving costs on local cable network infrastructure but also from that moment on programmers were able to work from anywhere. Companies like SoundCloud for instance even started off in bars (Sankt Oberholz), enjoying the social environment and saving on renting office space. Silicon Valley's garages are Berlin's bars.

According to the US technology magazine Wired (McGrane, 2000), the very first internet start-ups settled in Berlin mainly for two reasons: (i) cheap rents in centrally located backyards, and (ii) a cultural scenery and night life which was as vivid and unshaped as the entrepreneurs themselves; both being the outcome of the reunification process.

During the past years an increasing number of internet start-ups settled in Berlin, making the city the nation's biggest home to internet firms. According to the online database provided by the start-up network Gründerszene (2013), Berlin, with more than 504 internet start-ups, is by far the sector's most important city, followed by Hamburg (138) and Munich (146, see also Figure 3.1 on page 28). Despite it's leading position, it's still Berlin where the Chamber of Industry and Commerce (IHK) recognizes the strongest growth in innovative web firms. To sum up, the birth of the internet economy does not look like being the result of an historic accident (Krugman, 2010). Recent anecdotal evidence instead tells us that the initial firm births are highly linked to Berlin's rich endowment of very distinct cultural amenities. The first-movers are then expected to be followed by new start-ups which on the one hand also want to benefit from amenities and on the other hand from agglomeration economies of the newly created internet cluster. Given the above described development, Berlin serves as a perfect city to empirically test the stated firm amenity maximiser hypothesis.

Figure 3.1.: Share of internet start-ups by city.



Notes: Data are extracted from Gründerszene (2013).

3.2.3. An intra-urban analysis

So far, most research on the determinants of firm location has been carried out on a regional or metropolitan level. There are substantially fewer intra-urban analyses. An exception is represented by S. Rosenthal & Strange (2005) as well as Arzaghi & Henderson (2008), who both use census tract level data from New York City. Within-city analyses, however, provide interesting insights when it comes to the assessment of location factors. First of all, the availability of highly disaggregated data on a city level implies a high geographic variation compared to an analysis which is based on a country's variation in regions or provinces, as this number is usually comparably small. Secondly, there might be a lot of location factors which only affect locations at a very close distance. Especially when thinking of cultural amenities, it is reasonable to assume that their influence diminishes with a steep decay. Thirdly, using highly disaggregated data allows for including location fixed effects on a larger aggregation level to control for unobservables. And finally, as the chapter's empirical approach builds on a conditional logit model, it must be ensured that the assumption of independence of irrelevant alternatives holds. An entrepreneur must theoretically be able to take all locations for setting-up his firm into consideration; a reasonable assumption in an intra-urban setting. After having made a case for using Berlin for an intra-urban analysis, I present a model of a footloose start-up.

3.3. Footloose start-up model

This section introduces a model of a footloose start-up in order to derive an estimable equation. It is based on the firm model introduced by Crozet et al. (2004), which has also been used by Brülhart et al. (2012).

The idea is to derive a profit function which describes a firm's profitability depending on its location. The firm location choice model assumes an investor setting up a new firm. The founder then decides on a firm location given a set of alternatives. The profit function consists of factors varying over location i and across sectors j. Quantity is set to be the strategic variable of a representative firm. Suppose that consumer's demand (=firm's supply) relies on a Cobb-Douglas utility function and is given by:

$$Q_{ij} = \frac{\alpha_j m_i^{\gamma_j}}{p_{ij}^{\delta_j}}, \tag{3.1}$$

with α_j as the share of income spent on the particular good (of sector j), m_i denotes the (exogenous) income of the consumers at location i, γ_j is the income elasticity and δ_j the price elasticity for sector j. The demand is satisfied by the firms at a price p_{ij}. This is a simplifying assumption, since in the world of internet start-ups not only workers and firms are footloose but also consumers, i.e. demand. Now suppose that firms have identical production costs when producing in the same location; a reasonable assumption for internet industries. Following that assumption, individual firms' quantities will be equal.

$$Q_{ij} = N_{ij} q_{ij}, \tag{3.2}$$

where N_{ij} is the number of firms on the market. Ignoring any taxes, a representative firm's profit function producing and selling at location i is given by:

$$\pi_{ij} = (p_{ij} - c_{ij}) q_{ij}, \tag{3.3}$$

where c_{ij} is a unit production cost function. It is now possible to derive the total equilibrium quantity Q_{ij}^* as well as the equilibrium price p_{ij}^*:[4]

$$Q_{ij}^* = N_{ij} q_{ij} = \frac{\alpha_j m_i^{\gamma_j}}{N_{ij}^{\delta_j}} \left(\frac{N_{ij}\delta_j - 1}{c_{ij}\delta_j} \right)^{\delta_j} \tag{3.4}$$

$$p_{ij}^* = \frac{N_{ij}\delta_j}{N_{ij}\delta_j - 1} c_{ij} \tag{3.5}$$

The following profit function can be obtained when plugging in p_{ij} and q_{ij}:

$$\pi_{ij} = \frac{\alpha_j m_i^{\gamma_j}}{N_{ij}^{\delta_j + 1}} \left(\frac{\delta_j}{N_{ij}\delta_j - 1} c_{ij} \right)^{(1-\delta_j)} \tag{3.6}$$

Assuming a price elasticity greater than one, $\delta_j > 1$, profits increase with consumers' expenditure/market size and decrease with production costs and number of active firms/competition.

---

[4] See Crozet et al. (2004) for a more detailed derivation.

Let firms be price takers on the input market and let the unit cost be defined as function of the number of firms (N_{ij}), wages (w_{ij}), which both vary over location and sector, location variant rents (r_i) and location invariant capital cost (k):

$$c_{ij} = N_{ij}^{-\theta_j^N} w_{ij}^{\theta_j^w} r_i^{\theta_j^r} k^{\theta_j^k}, \tag{3.7}$$

where θ_j^N, θ_j^w, θ_j^r and θ_j^k denote the respective input shares. Unit costs rise with wages, office rents and capital cost and fall with the number of other firms due to agglomeration economies. The wage is determined by vector E_j, a composite of (unobservable) worker individual specific characteristics like education, work experience etc., and a location variant amenity shifter A_i, capturing the stock of cultural amenities surrounding location i. For simplicity, I assume that E_j is location invariant and identical within sectors:[5]

$$w_{ij} = E_j A_i^{-\tau_j}, \text{ with } \tau_j > 0, \tag{3.8}$$

where τ_j describes how strongly amenities are capitalised into wages and varies across sector j. Wages might decline with the endowment of amenities for two reasons: (i) workers (including the entrepreneur) are willing to work at lower wages if they get compensated by amenities (Rosen, 1979; Roback, 1982; Blomquist et al., 1988; Gyourko & Tracy, 1991; Gottlieb, 1995), and (ii) spill-overs due to face-to-face contact in bars, coffee shops etc.[6] The latter idea is closely related to Storper & Venables (2004) who consider the face-to-face contact as a key element of urban concentration. Especially creative industries require the exchange of ideas and information. Urban amenities like bars provide an external location to hold meetings. A third-party location might be preferred due to a lack of office space (especially for young start-ups), its neutral character or due to the preference for a more relaxed, creative, stimulating environment. Substituting wages into the unit cost leaves us with:

$$c_{ij} = N_{ij}^{-\theta_j^N} (E_j A_i^{-\tau_j})^{\theta_j^w} r_i^{\theta_j^r} k^{\theta_j^k} \tag{3.9}$$

Plugging the unit cost into the maximised profit function and assuming a sufficiently large number of firms yield the following expression:

$$\pi_{ij} = \alpha_j m_i^{\gamma_j} N_{ij}^{\theta_j^N(\delta_j-1)-2} r_i^{\theta_j^r(1-\delta_j)} (E_j A_i^{-\tau_j})^{\theta_j^w(1-\delta_j)} k^{\theta_j^k(1-\delta_j)} {\frac{\delta_j}{\delta_j - 1}}^{1-\delta_j} \tag{3.10}$$

Log-linearizing the maximised profit function results in:

$$\begin{aligned} \ln \pi_{ij} = {} & \ln \alpha_j + \gamma_j \ln m_i + (\theta_j^N(\delta_j - 1) - 2) \ln N_{ij} + \theta_j^w(1-\delta_j) \ln E_j - \tau_j \theta_j^w (1-\delta_j) \ln A_i \\ & + \theta_j^r(1-\delta_j) \ln r_i + \theta_j^k(1-\delta_j) \ln k + (1-\delta_j) \ln \frac{\delta_j}{1-\delta_j} \end{aligned} \tag{3.11}$$

---

[5] Admittedly a strong assumption but reasonable when taking into account the relatively high homogeneity of the sector (all young, IT specialists etc.).

[6] The relation is more indirect where it is assumed that spill-overs boost worker productivity making labour cheaper in relative terms.

Assuming homogeneous sectors and mobile firms, profits are equal at every location, spatial equilibrium then requires amenities to be capitalised into wages and rents. The location choice is independent of any amenity endowment. Empirically, amenities and rents should have no significant effect.

$$\frac{\partial \pi_i}{\partial A_i} = 0 \tag{3.12}$$

However, assuming firm heterogeneity, especially with respect to amenity appreciation ($\tau_j > 0$), the results are expected to differ over sectors. Since I am particularly focussing on internet start-ups the following key hypothesis can be formulated:

$$\frac{\partial \pi_{ij}}{\partial A_i} > 0 \tag{3.13}$$

A footloose start-up acts as amenity maximiser when deciding on a firm location if profits rise with cultural amenities.

3.4. Empirical approach

3.4.1. General estimation approach

Based on the log-linearized version of the profit function developed in Section 3.3 the following estimable equation can be formulated:

$$\ln \pi_{ij} = \beta_0 + \beta_{1i} \ln N_{ij} + \beta_{2i} \ln A_i + \beta_{3i} \ln r_i + \beta_{4i} \ln G_i + \beta_{5j} + \ln v_{ij} \tag{3.14}$$

Capital cost and income are dropped as both factors are spatially not restricted to the city of Berlin. Potential consumers are web users all around the world, making it impossible to control for their income. Employee characteristics E_j are also not included as regressors for the above stated reasons. G_i stands for a number of controls which are going to be discussed in the data section (Section 3.5). β_{5j} are sector fixed effects, absorbing sector specific unobservables. The equation can be estimated by a conditional logit model when the added stochastic term $\ln v_{ij}$ is assumed to follow an i.i.d. extreme-value type 1 distribution. As I only test the firm model for internet start-ups which all belong to the same sector, the estimable equation needs to be slightly adopted, where sector fixed effects are absorbed by the constant and N_i denotes the number of firms in all sectors.

$$\ln \pi_i = \beta_0 + \beta_{1i} \ln N_i + \beta_{2i} \ln A_i + \beta_{3i} \ln r_i + \beta_{4i} \ln G_i + \ln v_i \tag{3.15}$$

The conditional logit model serves as a well-established econometric framework when it comes to the estimation of firm location decisions. It is based on McFadden's (1974) random utility maximisation which was adapted to a random profit maximisation problem by Carlton (1983). Consider an investor or entrepreneur j who chooses a location i out of a set of spatial choices

I for setting up a new firm. The profit π_{ij} the entrepreneur j derives at location i is composed by a deterministic U_{ij} and a stochastic term ε_{ij}:

$$\pi_{ij} = U_{ij} + \varepsilon_{ij} \tag{3.16}$$

Location i will be prefered over k if:

$$\pi_{ij} > \pi_{ik}, \forall k,\ k \neq i \tag{3.17}$$

The probability that location i is chosen by the entrepreneur is given by:

$$P_{ij} = \text{Prob}(\pi_{ij} > \pi_{ik}), \forall k,\ k \neq i \tag{3.18}$$

Assuming independently distributed error terms and additionally following a Weibull distribution results in the conditional logit formulation

$$P_{ij} = \frac{\exp(U_{ij})}{\sum_{k=1}^{K} \exp(U_{ik})}, \tag{3.19}$$

where the deterministic component U_{ij} is assumed to be a linear combination of explanatory variables.

In the past, conditional logit models could not consider the full set of location choices when the set was large. To avoid cumbersome estimations Guimarães et al. (2000) used smaller choice sets which were selected randomly. The size of choice sets increases with the fineness of the spatial level, such as statistical blocks, as in this work. To be able to use all information and allow for the replicability of the results, Guimarães et al. (2003) have shown that it is possible to obtain equivalent coefficients for the conditional logit model when estimating it using a Poisson count model. By assuming that each location decision is determined by a vector of choice-specific attributes which are common to groups of individuals (or in this case of firms), the log-likelihood function of the conditional logit model is identical to the Poisson log-likelihood up to a constant. It is therefore possible to estimate the profit function using a Poisson model with the number of firms in each location n_i as dependent variable.

$$E(n_i) = \lambda_i = \exp(\beta_0 + \beta_{1i} N_i + \beta_{2i} A_i + \beta_{3i} r_i + \beta_{4i} G_i) \tag{3.20}$$

The conditional logit model relies on the independence of irrelevant alternatives (IIA) assumption. This means that consistent estimates require the stochastic terms to be independent across locations. The location decision between two alternatives is not allowed to change when a third alternative location is added or changed. An entrepreneur must therefore theoretically be able to compare all locations available in the choice set. The finer the spatial level, the more alternatives there are, increasing the likelihood of violating the IAA. This is in line with Figueiredo et al. (2002), who argue in their paper on location decisions of Portuguese entrepreneurs that entrepreneurs choose firm locations close to where they live. They know the area better and finding a new location implies additional search costs. That is another reason for investigating the location choice problem in an intra-city framework. I assume that within a city an entrepreneur is able to compare all potential locations.

Another violation of the IIA assumption might occur when there are unobserved location characteristics that are spatially correlated. I therefore include location fixed effects to control for any spatially-fixed unobservables by adding a set of location dummies d_v (Brülhart et al., 2012).

As previously established, the above derived profit function can be estimated using a Poisson model. The Poisson estimator, however, relies on the strong assumption that the conditional mean equals the conditional variance, VAR(Y|X) = E(Y|X). In practice this assumption is often violated and the data at hand suffer from overdispersion, i.e., the variance exceeds the expected value. Very often there is also a larger number of zeros, as described by the Poisson distribution. I therefore weaken the Poisson assumption and apply a Poisson Pseudo-Maximum Likelihood (PPML) estimator as originally suggested by McCullagh & Nelder (1989) and later by Santos Silva & Tenreyro (2006).

A PPML estimator requires two specifications: the functional form of the conditional expectation E(Y|X) and of the conditional variance VAR(Y|X). The conditional mean is defined as above (now plus voting precinct dummies d_v):

$$E(n_i \mid N_i A_i r_i G_i d_v) = \exp(\beta_0 + \beta_{1i} N_i + \beta_{2i} A_i + \beta_{3i} r_i + \beta_{4i} G_i + \beta_{5i} d_v) \tag{3.21}$$

Assuming the conditional variance to be proportional to the conditional mean, VAR(Y|X) $\propto$ E(Y|X), it is possible to estimate $\tilde{\beta}$ by solving the following set of first-order-conditions:

$$\begin{aligned}\sum_{n=1}^{N} [n_i - \exp(\tilde{\beta_0} + \tilde{\beta_{1i}} N_i + \tilde{\beta_{2i}} A_i + \tilde{\beta_{3i}} r_i \\ + \tilde{\beta_{4i}} G_i + \tilde{\beta_{5i}} d_v)] N_i A_i r_i G_i d_v = 0\end{aligned} \tag{3.22}$$

$\tilde{\beta}$s are a Generalized Method of Moments (GMM) estimators and consistent when the conditional mean is correctly specified. If the assumption about the proportional relation between conditional expectation and variance is violated, the standard errors of the estimates are inefficient, whereas the estimated coefficients are not affected. All inference has therefore been based on Eicker-White robust standard errors.

The way the weights have been defined, the PML estimator is numerically equal to the Poisson pseudo-maximum likelihood (PPML) estimator. Therefore I obtain consistent estimates based on a Poisson likelihood function without requiring the dependent variable to be made of integers (Gourieroux et al., 1984). Building on large sample asymptotic, the PPML approach has been proven to be efficient and robust (Gourieroux et al., 1984; Santos Silva & Tenreyro, 2006, 2011).

3.4.2. Identification

The inclusion of cultural amenities in the empirical model raises obvious endogeneity concerns mainly because their existence highly depends on demand from economic subjects. There are two potential types of endogeneity. Firstly, estimates might suffer from a simultaneity bias. It becomes difficult to disentangle whether cultural amenities attract firms or whether causality runs the other way around. Secondly, the likelihood of unobservables in the error term which affect both internet start-ups and amenities is very high. Therefore identification becomes crucial.

The suggested identification strategy to deal with the risen endogeneity concerns is twofold. First of all, I control for location fixed effects by adding a set of location dummies. Due to the spatial scope of the expected unobservable fixed effects, the geographic bodies of the location controls must be sufficiently fine. I use voting precincts from 2008 to control for fixed effects. There are 1,201 precincts for 15,937 statistical blocks. Voting precincts are by definition supposed to reflect homogeneity in terms of demographics (Berliner Parlament, 2008). They have previously been used by Ahlfeldt (2013) as unit of analysis to represent a self-contained neighbourhood. The voting precincts are therefore expected to soak up any unobservable fixed effects. As there are only about 600 start-ups distributed over the whole of Berlin fixed effects are restricted to voting precincts with at least five firms.

I secondly follow an instrumental variable strategy. I make use of the fall of the Berlin Wall and interpret it as a quasi-natural experiment. The historic event was not foreseen by any market players[7] and can therefore be regarded as an exogenous shock (Redding & Sturm, 2008; Redding et al., 2011; Ahlfeldt et al., 2012). I exploit German reunification as the source of exogenous variation from which I derive a set of instrumental variables. In particular and most preferably, I use distance to the former Berlin Wall to instrument cultural amenities. The idea is that proximity to the former border explains the spatial endowment of current cultural amenities sufficiently well. Municipalities like Prenzlauer Berg, Mitte, Friedrichshain and Kreuzberg which were originally located in the periphery of either East or West Berlin all of a sudden became central locations experiencing a huge accessibility shock (Redding & Sturm, 2008). There is ethnological evidence that a specific subculture has evolved in the aftermath of German reunification, predominantly in the “wild east” with its political vacuum and abandoned places (Schwannhäußer, 2007). Former border areas became home to artists, creatives, students and squatters. Bars and clubs opened. The identifying assumption is that proximity to the Wall has no direct effect on the location choice of internet start-ups, only indirectly via the amenity channel. Identification is conditional on a large set of control variables, where the inclusion of a historic CBD dummy as well as land values from 1992 are of special importance. They control for the link between proximity to the Wall and today’s cultural amenities neither being driven by centrality nor lower rents after re-unification. The exclusionary restriction is backed-up by the time dimension and the fact that the internet was

---

[7] According to opinion polls, less than ten percent of the West German population expected the re-unification (Herdegen, 1992).

not used by the general public at the beginning of the 1990s. The IV PPML estimator then solves the following first order condition (Tenreyro, 2007):

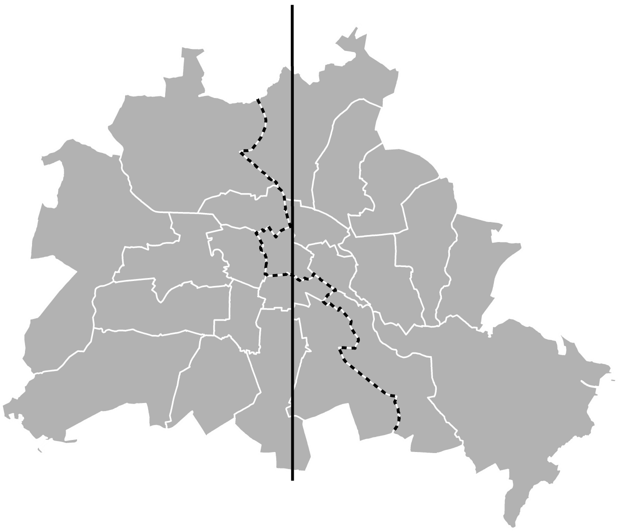
$$\sum_{n=1}^{N} \left[n_i - \exp(Y_i \bar{\beta}) \right] z_i = 0 \tag{3.23}$$

For robustness reasons I suggest a second set of instruments which is admittedly weaker in terms of the exclusionary restriction. Following the previous line of argumentation, I use the location of (i) squatted buildings since 1987 as well as (ii) historical cultural amenities for 1998/1999 and (iii) for 1936 to instrument the current level of amenities. Firstly, squatters reflect the immigration into the new, open, tolerant areas. Together with artists and students they are considered to be the pioneers of the gentrification (Clay, 1979; Friedrichs, 2000). They are the first ones to open (sub-)cultural bars and clubs. However, they develop the area for themselves without any intentions of making the area a hip place which would drive up rents in the long-run. Secondly, 1998/99 cultural amenities directly capture the young, open techno scene. The idea is that today's cultural life originates from a subculture which developed during the 1990s provoked by German reunification. And finally, the historic amenities from 1936 are motivated by the idea that there is some path dependency in the development of amusement areas. Neighbourhoods which were known for their endowment of bars and for their nightlife in 1936 are expected to be still equipped with urban amenities today. Even though I consider this latter set of instrumental variables as weaker in terms of the exclusionary restriction, they at least allow for circumventing simultaneity. Squatted houses and 1998/99 cultural amenities are a result of the reunification years and no direct link to internet firms can be established. By that time, the number of internet users was still very small; mobile computers and wireless internet connection scarcely available.

While the exclusionary restriction is fairly reasonable in terms of the time dimension, one might be worried about the spatial dimension. Potential concerns are that the Berlin Wall only captures centrality or/and the distance measure is correlated with unobservables driving the estimates. Moreover, one might question the randomness of the location of the wall. To address these concerns I employ the instrumental strategy conditional on controls, whereas centrality should in particular be captured by a dummy indicating whether a block lies inside the historic CBD from 1933. Additionally, transport controls, an East Berlin dummy, spatial trends and, above all, the aforementioned voting precinct fixed effects are expected to pick up any remaining centrality forces determining firm location. I run a number of additional robustness checks, where I compare distance from a block to the CBD (centrality) to cultural amenities in terms of explaining start-up location.

Berlin's distribution into four sectors after World War II was decided during the Conference of Yalta held February 4-11, 1945. It is a result of negotiations between Roosevelt, Churchill and Stalin. The sector's border and hence the Berlin Wall followed old definitions of Gross Berlin (1920). If East and West only wanted to split the city in half, the border would always run through the centre and thus would always be correlated with centrality. If start-up location

Figure 3.2.: Placebo Wall



Notes: The placebo Wall is indicated by the solid line, the dotted line marks the actual Berlin Wall.

was not determined via the amenity channel by the Wall but only by centrality or any other unobservables correlated with the instrument, a placebo Wall which vertically cuts the city in half should also give significant estimates. I therefore re-run the benchmark model using a placebo Wall as indicated by Figure 3.2 on page 36. In alternative fixed effect specifications I use an old district definition as well as municipality location fixed effects which both share the same borders with the former Wall. As firm location is only explained by within district/municipality variation a potential unobservable effect is additionally hindered.

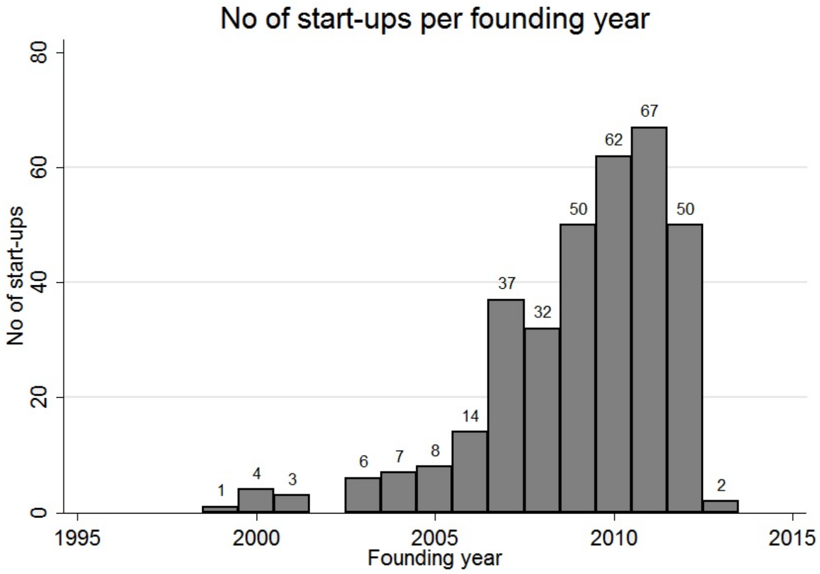
To support the case that firms/sectors are heterogeneous and therefore rely on different sets of location factors, I re-run the benchmark specification for a set of counter-factual service firms. The idea is that cultural amenities more strongly drive the location of young, creative web start-ups than of other knowledge-based service industries. These counter-factual models can also be interpreted as additional robustness tests with respect to centrality or unobservables which might be correlated with distance to Wall.

3.5. Data

3.5.1. Dependent variable

To determine the location factors of internet start-ups I use the number of web firms per statistical housing block as dependent variable. The regressand's count data character further encourages the Poisson estimation approach. The statistical blocks are the unit of analysis. The city is structured into 15,937 of these statistical housing blocks. The firm data originate from two sources: As primary source, I extracted firm information of all firms listed in the online database provided by Gründerszene (2013). Gründerszene is a magazine as well as an

Figure 3.3.: Number of internet start-ups per founding year



Notes: Information on the founding year are only available for about 58% of the total sample.

online platform for the German web economy and its start-ups which was founded in 2006.[8] The firm addresses were geocoded and processed in a geographic information system (GIS) environment. As a second source, I used the Berlin start-up map which maps Berlin Web 2.0 start-ups. It is accessible via the Berlin Business Location Center (2012a), a public business promoter and location marketing office owned by the state of Berlin. The data from the two different sources were merged and double entries deleted. The sample represents a total of 600 internet start-ups listed in April 2013. 345 of these firms have information on their founding date. As indicated by Figure 3.3 on page 37, first internet firms started settling in Berlin at the end of the 1990's, whereas the development took off around 2007.

3.5.2. Cultural amenities

Data on current cultural amenities were taken from OpenStreetMap (2013). It is argued that the potential self-selection by uploading spatial data to OpenStreetMap reflects people's perception of their surroundings and, contrary to causing biases, reveals preferences. There might be a bias, however, if web entrepreneurs themselves report amenities. This would further support the application of an IV estimator. I take into account mainstream as well as subcultural amenities. Cultural amenities include bars/pubs, cinemas, theatres, clubs, operas, beer gardens, cafés, restaurants and art places.

Not the actual number of amenities is of people's interest but the mass of cultural amenities they are surrounded by. A potential amenity indicator should therefore be able to capture the number of amenities within a certain proximity, whereas amenities nearby should get a stronger

---

[8] Even though there might be a number of already more mature firms, I follow the Gründerszene classification and consider all firms listed as start-ups, and thus as a homogeneous group of young web bound firms.

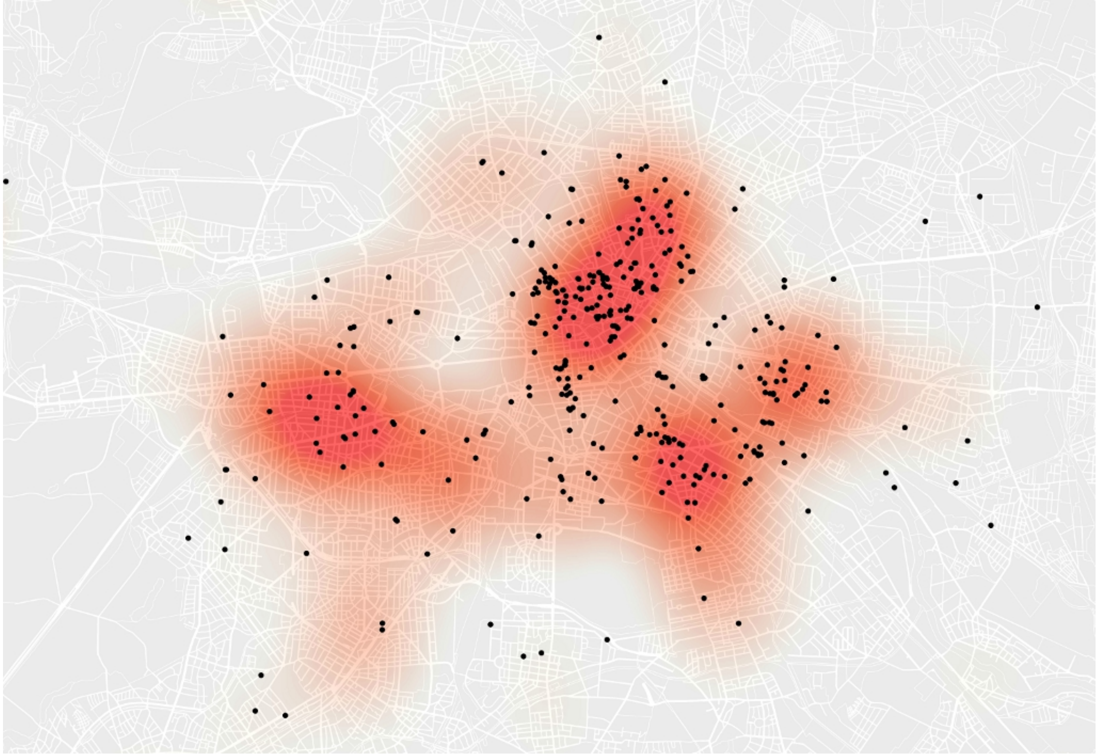
weight. Since the definition of the amenity measure already implies a certain assumption and hence affects the results, I briefly discuss three different measures I use. First of all, I compute kernel densities around each point representing a cultural amenity (Silverman, 1986), applying a radius of 2 km. This radius goes back to Gibbons & Machin (2005) who predict a distance of 2 km as being the maximum distance people are willing to walk to the nearest station and has already been used in the context of urban amenities by Ahlfeldt, Möller et al. (2013). Even though the density measure fulfills the above stated requirements, estimate interpretation is rather abstract and not intuitive. I therefore secondly employ a gravity based accessibility measure as suggested by (Fujita & Ogawa, 1982):

$$A(i) = \sum_{l} A_l e^{-bd(i,l)},\ i \neq l \text{ and } d_{ll} = \frac{1}{3}\frac{\sqrt{Area_i}}{\pi}, \tag{3.24}$$

where the access to cultural amenities in block i, $A(i)$, is defined by the number of other amenities at all other surrounding locations l spatially discounted by a decay parameter b and $d(i,l)$ a measure of distance between i and l. As public transport plays a major role in moving people in big and dense cities, Euclidean distances only provide a rough estimation of proximity to other firms. However, replicating the transport network places a strong weight on the location of public transport stations, which most likely picks up correlated unobservables. I therefore stick to the straight line distances between block centroids. Assuming that start-ups are only attracted by amenities in their close neighbourhood, I apply a distance decay parameter of two which is supposed to capture walking speed (Ahlfeldt, 2011b). For robustness tests I thirdly create buffer rings of various radiuses around each block centroid and use the number of amenities which fall inside a ring.

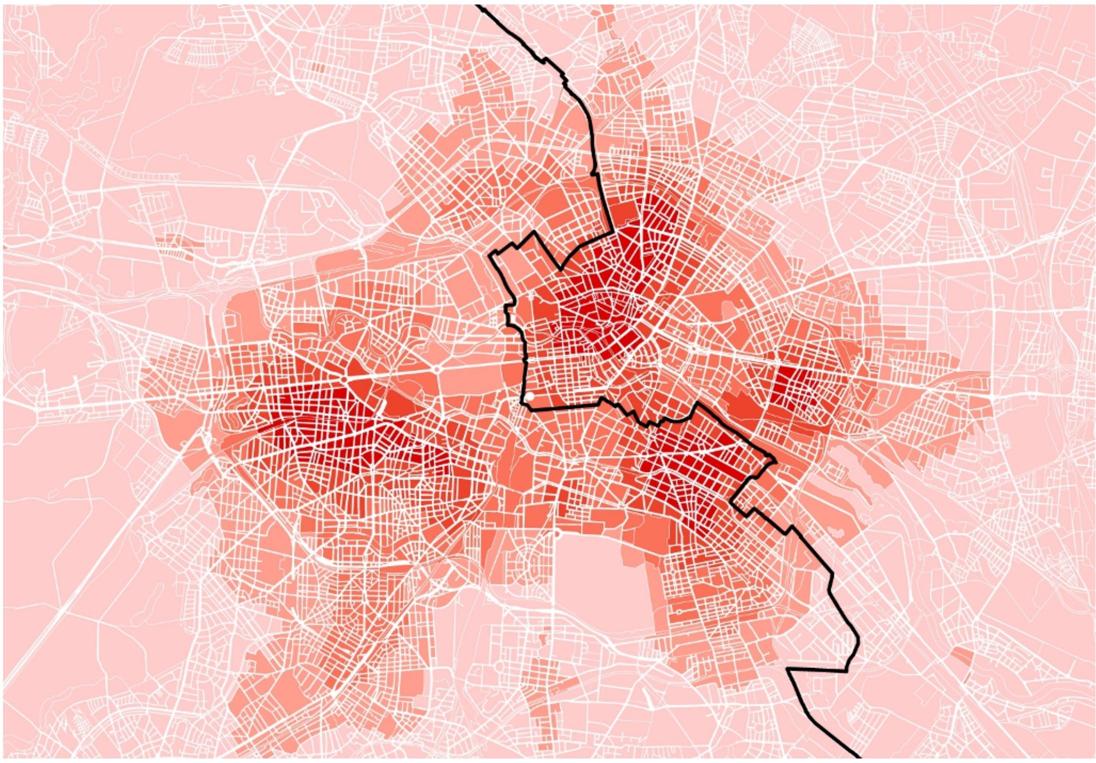
To sum up, each of the measures suggested comes with certain advantages and disadvantages. The application of all three of them helps to get a better understanding of the forces at work. Their application is hence regarded as a robustness check, controlling whether the estimates are independent of the measure chosen. The distribution of internet start-ups (points) and density of cultural amenities is illustrated by Figure 3.4 on page 39. Proximity to the Berlin Wall is computed for every block centroid. I calculate straight line distances as well as a potentiality measure similar to the access to cultural amenities, as indicated by equation (3.24). To motivate the first stage, illustrative evidence for today's endowment of former border locations with cultural amenities is given by Figure 3.5 on page 39, jointly showing the Berlin Wall and access to cultural amenities. A similar map motivating the application of squat density as a secondary instrument is shown in the appendix (Figure A.1 on page 126). For the secondary set of instruments I use historical cultural amenities of 1998/1999 and squatted houses since 1987. Squatted buildings are taken from Hausbesetzungs Geschichte Berlin (2010). For 1998/99, mainstream bars, clubs, theatres and restaurants were extracted from Siebenhaar et al. (1998), a guide book especially designed for young people on behalf of the state of Berlin. A detailed and ethnological analysis of Berlin's subculture and the origins of the Berlin techno underground scene is provided by Schwannhäußer (2007). She refers to a website Verblichene

Figure 3.4.: Distribution of start-ups and cultural amenities



Notes: Black dots denote the location of web start-ups based on Gründerszene (2013) and Berlin Business Location Center (2012a). Red amenity clouds represent the amenity density measure with a radius of 2 km (Silverman, 1986), with dark red indicating a high amenity density.

Figure 3.5.: Berlin Wall and access to cultural amenities



Notes: Black solid line denotes the Berlin Wall. Cultural amenities are aggregated at a statistical block level using the accessibility measure from equation (3.24) with dark red indicating a high accessibility to amenities.

Locations (1999) listing the locations of subcultural Berlin and its events before the gentrification process kicked in. Another list of historical amenities from 1936 was compiled from the guide book Baedeker (1936). The collection is significantly smaller and contains theatres, operas, operettas, vaudevilles, cabarets and cinemas. The extracted data were geocoded and processed as described above.

3.5.3. Control variables

According to the empirical specification (3.22) the number of internet start-ups is not only determined by cultural amenities but also by the number of other firms inside a block, the rent as well as a set of control variables. The number of firms is proxied by the total employment inside a block (i.e. across all sectors). The variable can be considered to capture localised general agglomeration/urbanization economies (Arzaghi & Henderson, 2008) as a large number of employees suggests a strong economic activity. The coefficient is expected to be positive if localised agglomeration economies positively impact on internet firms. A negative coefficient would reflect the competitive aspect of being closely located to other firms. As noted above, internet firms offer online services and users are not required to be physically close. In contrast to offline firms I do not control for any other market potentiality in the classic sense. Anecdotal evidence sees low rents as one of the main drivers of the Berlin web 2.0 boom. I use standard land values from 1992 per square metre as measured by the German committee of valuation experts (Gutachterausschuss für Grundstückswerte in Berlin, 1992). Data were spatially smoothed using an inverse distance weighting. I particularly include data from 1992 to reflect the spatial pattern in land prices right after re-unification to make sure that the link between instruments and amenities is not confounded with, e.g., lower prices in former border regions. In an additional robustness specification, I replace 1992 land values by residential rent data from 2010 (Immobilien Scout, 2012) to assess today's role of rents for the location decision of start-ups.

I additionally control for further location factors which might determine the location of young internet firms. I control for the number of immigrants per block. Areas characterised by migration are expected to attract young entrepreneurs as they are signal of cultural variety and tolerance. Data come from the statistical office Berlin Brandenburg (Amt für Statistik Berlin-Brandenburg, 2011a,b). Berlin is home of a large number of knowledge-creating and –spreading institutions. Young start-ups are often founded as spin-offs of universities. I therefore expect a positive relation between firm location and proximity to universities and research institutes. To test this I calculate Euclidean distances between all statistical blocks and research institutes/universities in a GIS environment. Among cultural amenities, entrepreneurs might also be attracted by natural amenities. Proximities to water bodies and green spaces are therefore computed. Additionally, young founders might also have a need for exercising after work. I hence control for the number sport facilities inside a block. These facilities include gyms, outdoor sport fields, swimming baths, and tennis courts (Gelbe Seiten Deutschland, 2012; OpenStreetMap, 2013). Especially in a capital like Berlin, historical districts can be of special interest. I therefore include a dummy variable to see whether a firm is located inside the historical CBD based on a definition from the historian Leyden (1933). Above all the dummy is supposed to control for centrality. As initially stated, the former eastern part of Berlin has especially attracted the movers and shakers, indicating a need to include a dummy variable which indicates whether a firm is located in the former east. In a recent review on the spatial concentration of entrepreneurship, Chatterji et al. (2013) note that entrepreneurship clusters might evolve from urban revitalisation projects. After re-unification 22 renewal areas were implemented to increase housing quality in Berlin. I therefore control for being located in one of the areas by adding an urban renewal dummy. Data are gratefully provided by Ahlfeldt, Maennig & Richter (2013), who test whether there are any positive housing externalities originating from the planning instrument. Transport accessibility is generally another important factor for the location of firms. For internet start-ups, however, transport serves more for commuting than in terms of market accessibility, since the output is usually a service good which is consumed or ordered “online”. I therefore control for accessibility to public transport infrastructure by including kernel density measures of 2 km (Silverman, 1986) for bus, trams, the underground and light rail network (BVG, 2006). Moreover, I control for the disamenity effect of noise originating from trains, underground trains on overground tracks as well as tram and street noise. The data are taken from maps published by the Berlin Senate Department for Urban Development (2007) which indicates the level of noise on a highly disaggregated 10x10 meter grid. To control for spatial trends I also add X and Y coordinates to the estimation model.

I note that there are numerous co-variates, such as the number of sport facilities or of migrants inside a block which ignore any spatial relation to the surroundings. For example, a block might very well be located inside a multi-cultural neighbourhood even though the block’s number of migrants is low. However, the inclusion of various measures relying on the same functional specification might cause multi-collinearity among regressors, which results in biased estimation results (Thill & Kim, 2005). The benchmark specification was therefore tested for multi-collinearity by computing variance inflation factors (VIF) (Kennedy, 2003).

3.5.4. Placebo firms

The selection of branches of other knowledge-based service industries is based on an overview provided by Eickelpasch et al. (2009) who analyse development perspectives for the service sector in eastern Germany. I hence rerun the benchmark model of the internet start-ups for consultancies, lawyers, insurance companies, financial advisors, agencies, engineering offices, publishers and architects. I consult the yellow pages for Berlin (Gelbe Seiten Deutschland, 2012) to obtain the postal addresses of all service firms. The data were processed in the same manner as the start-up information. Exemplary, agencies as well as of financial advisories are mapped jointly with cultural amenities in the appendix (Figure A.2 on page 126).

3.6. Results

3.6.1. Internet start-ups

To test the implications of the footloose start-up model I begin the analysis by estimating the regression model as outlined by equation (3.22). Table 3.1 on page 42 reports the key estimates for five different specifications (see Table A.1 on page 128 for extended table).

I initially abstract from equation (3.22) by ignoring rents (from 1992) and other firms and by only focusing on the the effect of cultural amenities on the location of internet start-ups (column 1). I use the log of the amenities to facilitate the interpretation.[9] The transformation reduces the total number of observation by 87 to 15,850 remaining blocks. Cultural amenities significantly (at a 1% level) influence the location choice of internet start-ups. In particular, a 1% increase in amenity density raises the probability of a firm locating inside a block by about 1.5%. Adding the employment and rent variable raises the attractive force of cultural amenities only slightly (column 2). Employment positively affects the location of internet start-ups, indicating the presence of localised agglomeration economies. In line with general intuition, rents have a negative impact on the firm location.

Table 3.1.: Estimation results: Footloose start-up model

	(1) # start-up	(2) # start-up	(3) # start-up	(4) # start-up	(5) # start-up
log cult amen.	1.547***	1.640***	0.885***	0.671***	1.195***
	(0.098)	(0.118)	(0.190)	(0.157)	(0.307)
employment		4E-4***	4E-4***	4E-4***	4E-4***
		(0.000)	(0.000)	(0.000)	(0.000)
rent		-5E-5**	-3E-5	-3E-5	-3E-5
		(0.000)	(0.000)	(0.000)	(0.000)
Controls	No	No	Yes	Yes	Yes
FE	No	No	No	Yes	Yes
IV	No	No	No	No	Yes
N	15850	15850	15850	15850	15850
OVERIDP					0.236
F (first)					1,775.682

Notes: Clustered standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01, Instruments: distance to Wall, squat density, OVERID (OVERIDP) denotes Hansen's J statistic of the overidentification test (and its p-value).

To control for alternative explanations, I add a set of control variables in specification (3). The likelihood of internet start-ups locating at a block due to the endowment with cultural amenities is reduced to 0.9%. Moving on to the estimates of the additional set of controls, migration positively affects the location of young web firms (see Table A.1 on page 128). This is very much in line with the creative class defined by Florida (2002) being highly attracted

---

[9] I also estimate the non-logarithmised density of cultural amenities as a robustness test, see Section 3.6.2.

by a tolerant surrounding. The migration effect stays robust in all specifications. As internet start-ups are knowledge-based firms, exchange with research institutions might be important. However, proximity to research institutes is positively correlated, i.e. the further away from a research institute the more likely a start-up location, throughout all specifications except the specification in column (3) where distance to research institutes is insignificant. A similar negative relationship is found for proximity to universities. It was said that venture capitalists move to Berlin as well and co-locate close to their recipients. This is found to be highly significant. Estimates indicate that the probability of a start-up locating inside a block decreases by between 0.2% and 0.3% per km distance from venture capitalists, depending on the specification. Of the two natural amenities, distance to nearest water bodies and to green space, only the first one is significant and positively attracts web firms. The provision of sport facilities is also found to positively affect firm location whereas exercising can be seen as a complement to cultural amenities in terms of leisure consumption. One unfortunate drawback of the sports measure is that it only takes into account the number of facilities inside a block due to the aforementioned multi-collinearity concerns. The transport controls are mainly insignificant with a few exceptions: Bus stops have a positive effect on firm location in specification (3). In the more demanding specifications (4) and (5), light rail stations attract start-ups. Noise disamenities have mixed effects. Noise originating from the underground running on overground tracks and trams positively affect start-up location. More intuitively, noise caused by the light rail system, trains and simple streets noise drive firms out. One explanation could be that trams and the two underground lines U1 and U2, which run on overground tracks, make relatively less noise than for instance normal trains. As indicated by the East Berlin dummy, new web firms are more likely to start a business in the former East, as suspected earlier. Moreover, start-ups are not particularly attracted by the historic CBD/centrality. There is no evidence that urban renewal zones attract entrepreneurs either: Once the instruments are introduced, the urban renewal dummy becomes insignificant. Additional controls for spatial trends (by the X-/Y-Coordinates) are all insignificant. Column (4) introduces location fixed effects at the voting precinct level. The cultural amenity coefficient decreases to 0.67. Finally, instruments (proximity to wall, squat density) for the endogenous amenity variable are introduced in specification (5). Cultural amenities continue to have a positive impact on the location decision of internet start-ups. A 1% increase in amenity density causes the likelihood of a firm location to rise by 1.2% still with a significance level of 1%. The economic significance is slightly smaller than in the baseline specification (1). The uninstrumented PPML estimates (3) and (4) are biased downwards. There might for instance be a general unobservable conservative attitude of people involved in the urban development process (planners, residents etc.) which slows down the creation of amenities (loud bars and clubs) and young start-ups. This attitude would underestimate the causal relationship between amenities and firm location and lead to an underestimation of the real effect if not addressed properly. The control variables are in line with general expectations indicating that start-ups are attracted by economic activity (employment, proximity to capital), tolerant (migration) and pleasant locations (proximity to water, noise disamenities). I consider this last specification as the most demanding one and will refer to it as benchmark model for the subsequent analysis.

Instrument validity relies on two requirements: (i) instruments need to be valid, i.e. uncorrelated with the error term, and (ii) relevant in terms of prediction power, so they require a high correlation with the endogenous regressors. The first requirement can generally not be tested. However, when the model is overidentified and there are more instruments than endogenous variables one can perform a test of overidentifying restrictions. As I instrument current cultural amenities using two different instruments, I test the null hypothesis that the applied instruments are jointly valid assuming that at least one instrument is exogenous. The computed Hansen's J statistic (OVERID) and its p-values (OVERIDP) do not reject the validity of the instruments. The evaluation of an instrument's strength is based on the F-statistic of the first stage regression. Stock et al. (2002) argue that the F-statistic should be greater than ten for a set of instruments to be relevant. The benchmark first-stage regression passes this threshold. The full first-stage estimates are illustrated by Table A.2 on page 130, where column (1) refers to the benchmark specification and the remaining columns belong to robustness checks discussed in the next section.

3.6.2. Robustness

To ensure the robustness of the benchmark results I first test alternative start-up model specifications. I then test the amenity-maximiser hypothesis for eight alternative service sectors and compare the outcomes to the benchmark estimates.

To address the concerns that proximity to Wall only reflects centrality or is correlated with unobservables driving the results, I perform a few robustness exercises which are shown in Table 3.2 on page 45. The full table is in the appendix (Table A.3 on page 132). In column (1) I control for centrality by including distance to CBD[10] instead of the historic CBD dummy in the uninstrumented PPML model. Unfortunately, it is not possible to include the centrality control variable in the IV model due to reasons of convergence. Distance to CBD is insignificant and the cultural amenity coefficient is almost identical to the estimates of the PPML model without the distance to CBD variable (column (3) Table 3.1 on page 42). Specification (2) uses distance to CBD and not to Wall as instrument. Centrality does not significantly attract web start-ups via the amenity channel. Proximity to a placebo wall which cuts the city vertically in two halves also yields insignificant estimates (column 3). Estimating the benchmark model with alternative location fixed effects that share borders with the Berlin Wall provides, however, a positive and significant amenity effect. Cultural amenities explain start-up location slightly weaker than in the benchmark model when using district fixed effects (column 4) and slightly stronger when using municipality fixed effects (column 5).

---

[10] Following standard practice for Berlin, I define the CBD to be at the intersection of Friedrich- and Leipzigerstraße (Ahlfeldt & Wendland, 2011).

Table 3.2.: Estimation results: Robustness exercises (1)

	(1) # start-ups	(2) # start-ups	(3) # start-ups	(4) # start-ups	(5) # start-ups
log cult dens	0.673***	0.573	-0.644	1.121***	1.507***
	(0.160)	(1.035)	(1.446)	(0.405)	(0.459)
dist to CBD	0.007				
	(0.078)				
employment	4E-4***	4E-4***	4E-4***	5E-4***	4E-4***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
rent	-4E-5	-4E-5	-3E-5	-1E-5	-3E-5
	(0.000)	(0000)	(0.000)	(0.000)	(0.000)
Controls	Yes	Yes	Yes	Yes	Yes
FE	voting	voting	voting	district	municipality
IV1		dist to CBD	placebo Wall	dist Wall	dist Wall
IV2				d. squat	p. squat
N	15850	15850	15850	15850	15850
OVERID		0.000	0.000	1.360	0.002
OVERIDP				0.244	0.968

Notes: Clustered standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01. OVERID (OVERIDP) denotes Hansen's J statistic of the over-identification test (p-value).

To ensure that the results are not driven by the applied indicators I re-run the benchmark specification using different measures of cultural amenity endowment and different instruments (Table 3.3 on page 47, extended Table A.4 on page 135). In the first column I restrict the IV estimation by only using proximity to Berlin Wall as an instrument. The amenity coefficient is insignificant with fixed effects, probably because they soak up too much variation. Column (1) illustrates the estimates without location dummies where amenities are significant at a 10% level. I then use the non-logarithmised density of current cultural amenities (column 2) which allows me to run the model on the full sample of 15,937 statistical blocks. The amenity density variable is highly significant and still positively affects the location choice of web firms. In columns (3)-(5) I capture the stock of current cultural amenities by a gravity-based potentiality measure with a distance decay of two which is supposed to capture walking speed (as suggested by Ahlfeldt (2011b)). The inherent assumption is that entrepreneurs are only attracted by amenities in their close neighbourhood. I first instrument current amenities using the benchmark instruments distance to Berlin Wall and squat density (specification 3). I then use distance to Wall and squat potentiality (specification 4) and both Berlin Wall and squat potentiality (specification 5) as instrumental variables. The instrumented cultural amenity estimates are all positive and significant at the 1% level. Finally, I make use of the historical cultural amenities. I create buffer rings of several distances around a block centroid and count the number of cultural amenities that fall inside a ring. This follows Arzaghi & Henderson (2008), who apply a similar indicator to capture the access to nearby advertising agencies as agglomeration measure. I define five (column 6) and four (column 7) rings moving out in increments of 500 metres up to 2,000 metres whereas specification (6) has an additional ring of 250 metres. The ring approach comes with the advantage of being intuitive to interpret. However, the variables might quite likely suffer from unobservables fixed effects in the error terms despite the use of voting precincts dummies and instruments. For data reasons, I am additionally only able to use the supposedly weakest set of instruments, the count of historical cultural amenities (1998/99) inside a ring. Moreover, the definition of the blocks is relatively heterogeneous compared to the definition of New York City census tracts (Arzaghi & Henderson, 2008). There might exist large blocks with the smallest buffer rings around the centroid still inside the block. Estimates should therefore be interpreted with particular caution. Estimation results for the two models are very similar. One additional cultural amenity in a ring between 500 and 1000 metres around a block centroid increases the probability of a firm location by 0.12%. In contrast, in the neighbouring ring (1000-1500m) the effect is negative and the likelihood decreases by 0.06-0.07% and the increases by 0.05-0.06% again (1500-2000m) which might indicate the presence of multi-collinearity. In the specification reported in column (8), I use a density measure of historical cultural amenities of 1936 as an instrument. Amenity estimates turn out to be insignificant. This is most likely due to the fact that the sample of 1936 amenities is very small. Column (9) makes use of the historic cultural amenities from 1998/99 again but this time as a standard kernel density measure with a radius of 2 km. Jointly instrumenting contemporary amenity density with the historic ones and proximity to Wall yields a positive amenity effect on the location of web firms. Their likelihood of locating inside a particular housing block increases by almost 1% if amenity density goes up by 1%. I eventually replace the 1992 land values with rent data from 2010 (column 10). The rent coefficient stays insignificant where the amenity effect is slightly larger than in the benchmark model.

Table 3.3.: Estimation results: Robustness exercises (2)

	(1) start-ups	(2) start-ups	(3) start-ups	(4) start-ups	(5) start-ups	(6) start-ups	(7) start-ups	(8) start-ups	(9) start-ups	(10) start-ups
cult amenity	1.570*	0.021***	0.014***	0.032**	0.031***	0.019	0.000	1.600	0.987**	1.590***
	(0.844)	(0.006)	(0.004)	(0.013)	(0.011)	(0.015)	(0.006)	(1.269)	(0.377)	(0.431)
ring 250m						0.019				
						(0.015)				
ring 500m						-0.007	2E-4			
						(0.011)	(0.006)			
ring 1000m						0.012***	0.012**			
						(0.004)	(0.005)			
ring 1500m						-0.006**	-0.007***			
						(0.002)	(0.003)			
ring 2000m						0.005**	0.006***			
						(0.002)	(0.002)			
employment	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	4E-4***	5E-4
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
land value	-3E-5	-3E-5	-2E-5	-1E-5	-3E-5	-3E-5	-2E-5	-2E-5	-3E-5	0.056
1992/rent	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.059)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Rent 2010
FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cult (X)	ln cult dens	cult dens	cult pot	cult pot	cult pot	rings	rings	ln cult dens	ln cult dens	ln cult dens
IV1	dist Wall	dist Wall	dist Wall	dist Wall	p. Wall	hist rings	hist rings	cult 1936	cult 1998	dist Wall
IV2		d. squat	d. squat	p. squat	p. squat				dist Wall	d. squat
N	15850	15937	15937	15937	15937	15937	15937	15850	15850	15850
OVERID	0.000	5.526	6.942	7.049	3.219	0.000	0.000	0.000	1.043	0.012
OVERIDP		0.019	0.008	0.008	0.073				0.307	0.911

Notes: Standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01, OVERID (OVERIDP) denotes Hansen's J statistic of the overidentification test (and its p-value). IVs: dens denotes a density measure, dist a Euclidean distance measure, pot a potentiality measure, rings the amenity rings. Cult stands for the the cultural amenity to be instrumented. Column (10) includes rents from 2010 (Immobilien Scout, 2012) instead of 1992 land values.

All instruments fulfill the overidentifying restrictions indicating their statistical validity. Referring to the first stage regressions (Table A.2 on page 130) all instruments are also sufficiently strong. The instruments are all individually significant and have the expected coefficients. One exception is shown in column (5), where distance to Wall is positively correlated to the gravity-based cultural amenity measure, probably due to multi-collinearity. Interestingly, 1936 cultural amenities are negatively correlated with today's amenity endowment. This would imply that amenity rich areas have changed over time and amenities moved to other locations within the city. In fact, according to anecdotal evidence the entertainment industry was located more in the Western area around Kurfürstendamm during the 1930s (Leyden, 1933). Moreover, as stated before, the 1936 sample is very small and its application as an instrument not a very good robustness test. Nonetheless, the overall results of the robustness checks using alternative measures provide evidence for the attractive role cultural amenities play for the location choice of internet start-ups. However, I consider none of the models to be as good as the benchmark specification with respect to the identification.

3.6.3. Placebo firms

The work tries to establish an empirical link between cultural amenities and firm location. It was argued that internet start-ups provide a perfect example for a footloose and knowledge-based firm. However, there might be other service sectors affected by an area's endowment with cultural amenities. I therefore re-estimate the benchmark model using eight alternative (placebo) service firms instead of the original internet start-ups. I assume that these firms face no cost of moving and re-adjust their location when attracted by other places which offer a better set of location factors. This, admittedly, is a simplifying assumption. However, compared to manufacturing industries, the moving of service firms usually only involves the relocation of office equipment. Moreover, it is important to note that the alternative service firms are (only) used as a placebo test to establish further robustness. Estimates are reported in Table 3.4 on page 49 (extended in Table A.5 on page 137).

Table 3.4.: Estimation results: Placebo firms

	(1) architects	(2) consultancies	(3) engineering	(4) insurance	(5) law	(6) publisher	(7) finance	(8) agencies
log cult dens	0.110	0.567***	-0.187	-0.221	-0.147	0.234	-0.551*	0.836**
	(0.196)	(0.216)	(0.336)	(0.465)	(0.176)	(0.376)	(0.320)	(0.333)
employment	3E-4***	3E-4***	3E-4***	3E-4***	3E-4***	3E-4***	3E-4***	4E-4***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
rent	-5E-6	1E-4***	5E-5*	8E-5**	8E-5***	6E-5*	1E-4***	4E-5
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	15850	15850	15850	15850	15850	15850	15850	15850
OVERID	5.102	0.677	0.001	2.167	2.510	3.117	0.014	0.241
OVERIDP	0.024	0.410	0.974	0.141	0.113	0.077	0.906	0.624

Notes: Standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01. Instruments: distance to Wall, squat density, OVERID (OVERIDP) denotes Hansen's J statistic of the overidentification test (and its p-value). There are at least ten law firms per voting precincts (instead of five).

Cultural amenities now only have an effect on the location of three out of the eight industries. In particular, cultural amenities only significantly determine the location of agencies as well as consultancies. A 1% increase in density raises the probability of an agency locating at a location by 0.84%, for consultancies the probability is 0.57%. The estimated probabilities are smaller than for internet start-ups. In contrast, firms offering financial services experience on average a negative effect and seem to get driven out by amenities. Their likelihood decreases by 0.55%.[11] I cannot observe any statistically significant effects for architects, engineering offices, law firms, publishers and insurance companies.

The interpretation of the remaining variables is limited as they are not at the centre of this research. The majority of the coefficient estimates are comparable to the benchmark model using internet start-ups as dependent variable. I therefore briefly report the most striking differences. Land value estimates yield a diverse but rather positive effect on firm location (insignificant for insurance companies, architects and agencies). Service firm location is generally independent of the proximity to research institutions, an intuitive result considering the research un-intensity of the firm selection. Law firms are the only type of service firms positively affected by proximity to universities. Interestingly, sport facilities positively affect all firms except engineering offices, whereas the picture of the transport role becomes rather mixed again. The importance of light rail stations stands out and affects almost all firms except agencies. There is a significant tendency to locate in former East Berlin, except for insurance companies and publishers. In contrast to web companies, the probability of locating inside an urban renewal area is higher for all firms but publishers.

Finally turning to the validity of the instruments, six out of the eight model specifications pass the test of overidentification. For architects as well as for publishers the null hypothesis of joint instrument validity must be rejected. The overidentification test must to be interpreted with care since it relies on the assumption that at least one instrument is exogenous. The test therefore only serves as a rough indicator for the validity of the applied instruments. First-stage regressions are reported in Table A.6 on page 139. Theoretically, the first stage regression models for the latter eight service industries should be equivalent to the one for the internet start-up model. Practically however, the inclusion of voting precinct dummies slightly varies due to the distinct distribution of firms over the city. All F-statistics confirm the IVs' relevance for all models.

Complementing the main analysis of internet start-up location with the estimates for alternative service industries allows us to draw the following conclusions: First of all, the comparison serves as an additional robustness check in terms of the amenity effect being driven only by centrality. If the amenity variable was highly correlated with centrality, one would expect mature and financially more potent service firms to outbid young internet start-ups in the centre. Secondly, the results indicate that the endowment of an area with cultural amenities cannot be regarded as a generalizable location determinant. The footloose start-up model as built above cannot be applied to any economic sector. In fact, it is only internet start-ups as well as

---

[11] The positive effect of amenities on agencies and the negative one on financial advisories is also reflected by Figure A.2 on page 126.

agencies and thus rather creative industries which are positively affected by the cultural amenities measure. The remaining firms conversely do not incorporate amenities into their location choice. Companies in particular offering financial or juridical advice are even driven out by amenities. These remaining service industries can be classified into rather conservative sectors. The movers and shakers are more likely to be found among web firms and agencies than among financial service firms. This first group is literally the "creative class" (Florida, 2002). Thirdly, the inter-sectoral comparison provides evidence on the existence of an urban "buzz" (Storper & Venables, 2004). Concentrated urban areas are characterised by a specific atmosphere (e.g. originating from cultural amenities or tolerance) which only affects very specific industries. Even if the applied instrumental variables were correlated with unobservables I was unable to control for, there would be some urban forces which particularly attract creative firms.

3.7. Summary

It was argued that especially young, highly-qualified and creative individuals are attracted by urban amenities. Knowledge-based service firms highly dependent on qualified labour are therefore expected to act as amenity-maximising agents and to locate in amenity-rich areas.

I test this hypothesis by looking at the evolving internet start-up sector in Berlin which serves an example of knowledge-based service firms. Following an instrumental variable approach which makes use of the fall of the Berlin Wall as a quasi natural experiment, I try to fill the gap of missing studies empirically assessing the role of urban amenities. It was shown that cultural amenities positively impact on the location of start-ups; a one-percent increase in amenity density raises the probability of a start-up location by about 1.2%. These results are proven to be robust by estimating various specifications in terms of amenity measures and instruments applied.

It was also shown that the results do not generally apply to all service types. Conservative service sectors like law or financial firms are not found to be affected by an area's endowment with cultural amenities. It is more creative branches like agencies and, above all, internet firms which act as amenity maximisers. The chapter therefore additionally provides evidence on the existence of an urban "buzz" (Storper & Venables, 2004). Concentrated urban areas are characterised by a specific atmosphere (e.g. originating from cultural amenities or tolerance) which affects very specific industries and not others. I find that these affected industries are closely related to the creative class.

Entrepreneurs are generally regarded as highly beneficial for a country's economy, both by economists as well as by politicians. They create new jobs, promote innovation and economic growth. Especially the IT and software sector is considered to be a key sector with great potentials. Hence, there are lot of different political initiatives to support entrepreneurs such as providing cheap office space, developing cheap credit programs or offering workshops on how to found a company. The results of this chapter enable a different perspective on how to promote entrepreneurs. It was shown that cultural amenities play an important role in attracting start-ups. This suggests an implementation of cultural-political initiatives in economic policy. Even though subcultural diversity might not be anticipated as economically beneficial in the short-run, its destruction might, however, stop attracting footloose creative heads in the long-run. Moreover, the results stand in contrast to artificially created science and technology parks in the periphery. Even though these parks are equipped with appealing incentives like cheap rents or access to the public transport network, it is not very likely that young innovative firms will relocate to the periphery but stay in developed, central and amenity-rich areas.

4. Game of zones: The economics of conservation areas [1]

4.1. Motivation

One of the key motivations for a variety of spatial planning policies is how to solve coordination problems inherent to free markets. Wherever non-traded positive or negative non-pecuniary externalities exist, prices no longer provide efficient signals to market actors. In such a situation individually rational decisions may be collectively irrational which implies that it is theoretically possible to improve welfare by means of regulatory policies. Among such policies historic preservation that aims at the protection of historic buildings with a particular aesthetic, cultural or historic value, occupies a leading position in terms of the rigidity of the related regulations as well as the complexity of related social and private costs and benefits. The policy is controversial because the preservation of socially desirable buildings comes at the cost of restricting individual property rights. On the one hand, the policy would not be equitable if individual owners bore the cost of a presumed social welfare improvement. On the other hand, it can be argued that by imposing binding standards the policy helps to overcome a coordination problem among homeowners. Since owners can no longer "free ride" on the character of nearby buildings while making inappropriate changes to their own properties the policy helps to solve a so-called prisoner's dilemma and eventually benefits the owners (Holman & Ahlfeldt, 2012).

With this contribution we provide a framework to empirically analyse the practice of preservation policy and its impact on the utility of local homeowners. We develop a simple model world in which we distinguish between a *heritage effect*, which can be internal or external, i.e. the effect of the appearance of a historic building on the perceived value of the house itself (internal) or nearby houses (external), and a *policy effect*, which results from the legal treatment of the designation policy. We argue that with positive heritage effects, the policy benefits the owners by removing uncertainty regarding the future of the neighborhood, i.e. the presence of the heritage effect. These benefits are opposed by the costs of regulation (in the form of development restrictions and maintenance obligations) so that the net effect of the policy effect is ambiguous. Our theoretical framework predicts positive, but diminishing returns to designation. We consider the policy (locally) Pareto-efficient if the designation share is maximized under the condition that the benefits of designation do not exceed the costs for any owner in the neighbourhood.

---

[1] This chapter is mainly based on a revised version of Ahlfeldt, Möller et al. (2013)

From the theoretical framework we derive two empirical specifications that allow us to test the nature and (local) welfare impact of the preservation policy. Firstly, provided that the planner behaves as an agent of the owners, new designation will result from increases in the local preferences for heritage. Secondly, a Pareto-optimal designation policy implies that at the margin, the costs and benefits of designation will offset each other, resulting in a zero impact of designation on the value of designated properties. At all other locations in a neighbourhood the effect will be positive. We test these implications using two different empirical approaches. Firstly, we identify a causal effect of changes in neighbourhood composition, what we define as gentrification, on the likelihood of designations using a tobit IV approach. Secondly, we use a hybrid difference-in-differences (DD) and regression discontinuity design (RDD) identification strategy to estimate the causal effect of new designations on the market value of properties. Our analysis is based on the whole of England, making use of 1 million property transactions from 1995 to 2010 and of about 8,000 designated conservation areas, of which 915 have been designated in the same observation period. We also make use of ward level education data from the UK census for 1991, 2001, and 2011 in order to analyse the effect of changing neighbourhood characteristics on the designation status. Previewing our results we find that that an increase in the local share of residents holding a university or college degree leads to an expansion of the designated area. The property price effect inside newly designated conservation areas turns out not to be statistically distinguishable from zero. We find evidence that the effect just outside the conservation area boundary is positive and significant. These results are in line with a Pareto-optimal designation policy at the local neighbourhood level, which can thus be argued to solve a coordination problem among homeowners (and landlords) within a neighbourhood. We emphasize that it is not possible to conclude from these results that the policy is globally Pareto-optimal since excessive historic preservation on a wide scale may lead to adverse welfare impacts through supply restrictions as argued, for example, by Glaeser (2011).

Our analysis of the conservation area designation process adds to a growing body of literature on the political economy of housing markets, which implicitly or explicitly assumes that property owners are able to influence political outcomes in their own interest (Ahlfeldt, 2011a; Ahlfeldt & Maennig, 2013; Brunner & Sonstelie, 2003; Brunner et al., 2001; Cellini et al., 2010; Dehring et al., 2008; Fischel, 2001a,b; Hilber & Mayer, 2009; Oates, 1969). We also contribute to a literature investigating the costs and benefits of spatially targeted policies that aim at improving neighbourhood quality (Cheshire & Hilber, 2008; Cheshire et al., 2011; Hilber & Vermeulen, 2010; Rossi-Hansberg et al., 2010) as well as research that has looked into the value amenities add to neighbourhoods and cities more generally (Ahlfeldt et al., 2012; Brueckner et al., 1999; Cheshire & Sheppard, 1995; Glaeser et al., 2001). As reviewed in Section 2.3.2, there is also a growing body of literature that has investigated property price effects of designation policies, mostly focused on the U.S. (Asabere et al., 1989, 1994; Asabere & Huffman, 1994; Coulson & Lahr, 2005; Coulson & Leichenko, 2001; Glaeser, 2011; Leichenko et al., 2001; Noonan & Krupka, 2011; Schaeffer & Millerick, 1991). The key contribution of this study is to provide insights into the political economy of conservation area designation and to examine whether the outcome is Pareto-efficient for local home-owners. We also make a number of more specific, though still important contributions. Firstly, the theoretical framework we develop lends a structure to the designation process that helps to interpret the existing evidence that has typically been derived from ad-hoc empirical models. Secondly, our analysis of conservation area effects on property prices is one of the few rigorous analysis of this kind available for Europe (Ahlfeldt & Maennig, 2010b; Koster et al., 2012; Lazrak et al., 2013) and the first to analyse England. It is unique in terms of size and spatial detail of the data set and special in its focus on spatial modelling of heritage externalities. Thirdly, our differences-in-differences analysis of designation effects on property prices is the only study along with Koster et al. (2012) that uses a quasi-experimental research design to separate the policy effect of designation from correlated location effects. It is unique in using particularly carefully selected control groups. Fourthly, we make use of a novel combination of RDD and DD approaches to identify the policy effects on outcome trends and discontinuities from quasi-experimental variation, which could be applied more generally to program evaluations. Fifthly, we provide the first empirical analysis of the determinants of heritage designation. More generally, we establish a novel connection between the spatial outcome of a political bargaining process and one of the most striking contemporary urban phenomena: gentrification.

The structure of this chapter is as follows. The next Section 4.2 introduces our theoretical model of heritage designations and the institutional setting. Section 4.3 our empirical strategy. A presentation and discussion of our empirical results is in Section 4.5. The last section summarizes the main findings (Section 4.6).

4.2. Theory and context

4.2.1. Theoretical framework

We assume that a linear neighbourhood exists along a spatial dimension x on the interval $[0, 1]$. Each parcel of land at point x is occupied by a housing structure which is endowed with $h(x)$ units of internal heritage. The aggregate of the distribution of internal heritage gives the heritage character (external heritage) H of the neighbourhood at any point in time. Owners care about their *initial endowment* of internal heritage $h(x)$, which is under their full control, and the *long run* external heritage, which may be damaged by their neighbours' property (re)developments. Such redevelopments occur in the long run with a probability of $(1 - \pi)$ where $0 \leq \pi < 1$ is the 'preservation probability' in the absence of conservation policies. The effect of conservation areas is to increase the preservation probability to 1 for parcels of land within their boundaries.[2] Therefore, long-run external heritage depends on both the internal heritage distribution and the level of designation.

---

[2] Our argument does not depend on the assumption of full preservation probability, only that preservation is *more likely* inside conservation areas.

Within the neighbourhood, the initial internal heritage monotonically decreases in x. The theoretical argument does not depend on the functional form. For simplicity we assume $h(x)$ to be a linear function of the heritage endowment at the neighbourhood's centre (h_0):

$$h(x) = h_0(1 - x) \tag{4.1}$$

One way to rationalize this distribution is to assume a neighbourhood that grew outwards from its historical centre (at $x = 0$) until the neighbourhood limit (at $x = 1$) and an internal heritage that strictly depends on the age of the housing unit.[3]

To protect the neighbourhood heritage, a planner can choose to designate a conservation area that covers all locations in the neighbourhood from the historical centre up to a point $x = D$ and hence, a share $0 \leq D \leq 1$ of the neighbourhood. Since heritage is monotonically decreasing in x it is always rational to start designating at $x = 0$. By affecting the preservation probability, the designation share D determines the external heritage amount to be expected in the long run. The expected long-run external heritage derived from undesignated locations ($x > D$) corresponds to the integral of the distribution of internal heritage multiplied by the preservation probability, $\int_D^1 \pi h(x)\mathrm{d}$. This is added to the amount derived from designated locations ($x \leq D$), which is simply the integral of the internal heritage as the preservation probability is equal to one, $\int_0^D h(x)\mathrm{d}$.

$$E[H \mid D] = \int_0^D h(x)\mathrm{d}x + \int_D^1 \pi h(x)\mathrm{d}x \tag{4.2}$$

$$E[H \mid D] = h_0(1 - \frac{D}{2})D + \frac{\pi}{2}h_0(1 - D)^2 \tag{4.3}$$

The expected external heritage integral $E[H \mid D]$ is indicated by the whole grey-shaded area in Figure 4.1 on page 57. The expected amount of external heritage saved by the preservation policy is illustrated as the black-dotted area $\check{H}$ which denotes the difference in (expected) external heritage between a scenario with no designation and a scenario with a designation share D. This amount is:

$$\check{H} = h_0(1 - \pi)(1 - \frac{D}{2})D \tag{4.4}$$

As evident from the partial derivatives, the amount of external heritage saved by the policy increases with designation share but at a decreasing rate:

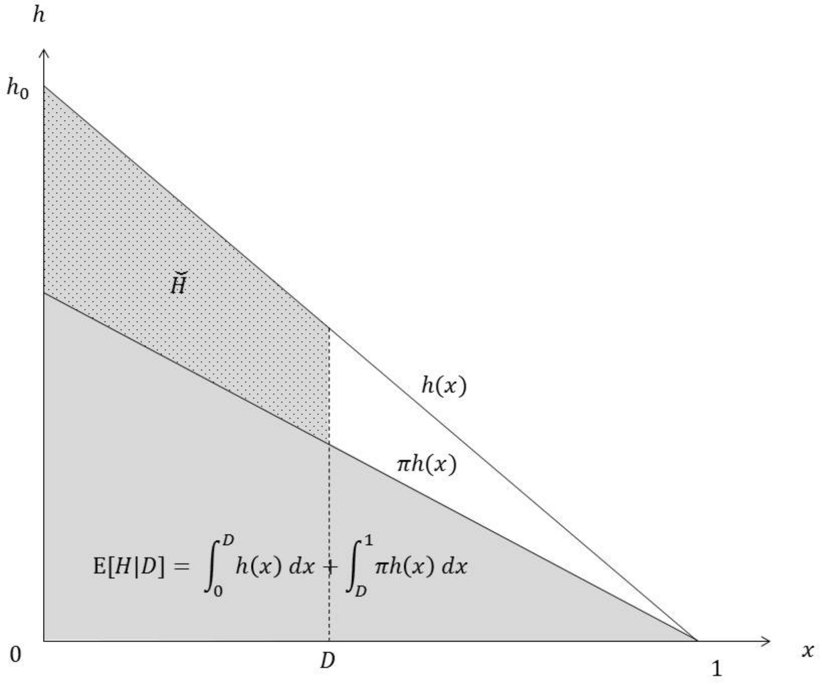
$$\frac{\partial \check{H}}{\partial D} = \frac{\partial E[H \mid D]}{\partial D} = h_0(1 - D)(1 - \pi) > 0 \tag{4.5}$$

$$\frac{\partial^2 \check{H}}{\partial D^2} = \frac{\partial^2 E[H \mid D]}{\partial D^2} = -h_0(1 - \pi) < 1 \tag{4.6}$$

---

[3] Alternatively, x can simply be interpreted as the rank of a property in the heritage distribution.

Figure 4.1.: Expected heritage distribution with partial designation



Notes: The function $h(x)$ gives the internal heritage at each location in the neighbourhood. The expected external heritage is equal to the grey-shaded area and is the integral of $h(x)$ up to the designation share plus the integral of π times this $h(x)$ from the designation share until the neighbourhood limit at $x = 1$. The stippled area marked $\check{H}$ is the amount of expected external heritage preserved by the policy.

The partial derivatives of $\check{H}$ (which are the same as of H) with respect to D establish a central stylized fact of our theory: There are diminishing returns to designation.

To link the distribution of heritage in the neighbourhood to the utility U of an individual residing at x we define a utility function:

$$U(x) = A(x)X^{\delta}L^{1-\delta} \tag{4.7}$$

where X is a composite consumption good and L is housing space. The Cobb-Douglas form is motivated by the empirical observation that housing expenditure shares tend to be relatively constant across geographies and population groups (Davis & Ortalo-Magne, 2011). $A(x)$ is a composite amenities term:

$$A(x) = a(x)e^{\phi h(x)}e^{\gamma E[H|D]}e^{-c\tilde{D}(x)} \tag{4.8}$$

where a is a further composite indicator of m non-heritage amenities,[4] $h(x)$ is the internal heritage endowment (i.e. heritage character of the specific housing unit), ϕ is the internal heritage preference parameter, $E[H \mid D]$ is external heritage (i.e. expected heritage of surrounding units, which depends on the designation policy) and is conditional on designation share as defined

---

[4] Non-heritage amenities are given by: $a = b\Pi_m a_m^{\rho_m}$ where the different amenity levels are denoted a_m and are given a collective scaling factor b and individual parameters ρ_m.

above, γ is the external heritage preference parameter, and c represents the costs of designation policies, which arise from the development restrictions imposed inside conservation areas. The cost to an individual is $e^{-c\tilde{D}(x)}$ and depends on the local designation status $\tilde{D}(x)$, a binary function of x, which takes the value of one if $x \leq D$ and zero otherwise.

We assume a social planner seeking a Pareto-efficient designation share, which in the model implies maximizing the designation share (and the external heritage effects) on the condition that by designation utility is not reduced at any location in the neighbourhood.

The positive marginal utility effect of designation at any location in the neighbourhood is given by:

$$\frac{dU(x)}{dD} = \frac{\partial U}{\partial E[H \mid D]} \frac{\partial E[H \mid D]}{\partial D} = \gamma U(x) h_0 (1-d)(1-\pi) \tag{4.9}$$

The negative utility effect to an owner of a property changing designation status from zero to one is:

$$\frac{dU(x)}{d\tilde{D}} = \frac{\partial U}{\partial \tilde{D}(x)} = -cU(x) \tag{4.10}$$

By setting the social marginal benefit equal to the private marginal cost of an affected owner the planner finds the Pareto-efficient designation share D^* by solving for D:

$$D^* = 1 - \frac{c}{(1-\pi)\gamma h_0} \tag{4.11}$$

Based on the resulting efficiency condition we can derive some useful comparative statics (see also Figure B.1 on page 142 in the Appendix). The (Pareto) optimal designation share is greater when people have a greater taste for external heritage γ or where there is altogether more heritage (determined by the heritage endowment at the neighbourhood centre h_0, and implicitly the age of the neighbourhood):

$$\frac{\partial D^*}{\partial \gamma} > 0 \tag{4.12}$$

$$\frac{\partial D^*}{\partial h_0} > 0 \tag{4.13}$$

There is less optimal designation when the preservation probability π (if left undesignated) increases or if the cost of designation increases:

$$\frac{\partial D^*}{\partial \pi} < 0 \tag{4.14}$$

$$\frac{\partial D^*}{\partial c} < 0 \tag{4.15}$$

These theoretical implications are in line with intuition and can be transformed into empirically testable hypotheses in principle. However, the heritage at the neighbourhood centre h_0, the preservation probability π and the costs to owners of conservation policies c are all difficult to observe in reality. For that reason we will concentrate on testing the first comparative statics implication about taste for heritage (proxied by education level of the local population) in the empirical section.

To develop a testable hypothesis on whether the efficiency condition is fulfilled, i.e. the planner sets $D = D^*$, we incorporate capitalization effects in the next step. We first assume that individuals maximize their utility defined above subject to a budget constraint: $W = X + \theta(x)L$, where $\theta(x)$ is a housing bid-rent. Furthermore we assume spatial equilibrium such that all locations offer the same level of utility $\bar{U}$ which we set equal to one:

$$U(x) = A(x)[\delta W]^{\delta}[(1-\delta)\frac{W}{\theta}]^{1-\delta} = \bar{U} = 1 \tag{4.16}$$

This can be rearranged to give the spatial equilibrium bid-rents for a representative individual:

$$\theta(x) = (1-\delta)[\delta^{\delta} W a(x) e^{\varphi h(x)} e^{\gamma E[H|D]} e^{-c\tilde{D}(x)}]^{\frac{1}{1-\delta}} \tag{4.17}$$

In keeping with intuition, the bid-rent increases in the expected external heritage, which depends on the designation share D and the internal heritage endowment $h(x)$ and decreases in the designation cost, which is locally constrained to $x \leq D$ as defined above.

The spatial equilibrium condition can be used to derive the marginal effect of an increase in designation share on rents in the neighbourhood. At all locations in the city a marginal increase in designation share D triggers a positive effect on rent through an increase in expected external heritage. At the margin, in addition, the change in designation status $\tilde{D}$ also creates a cost.

$$\frac{d\theta(x)}{dD} = \begin{cases} \frac{\partial\theta(x)}{\partial E[H|D]}\frac{\partial E[H|D]}{\partial D} + \frac{\partial\theta(x)}{\partial\tilde{D}(x)}d\tilde{D}(x) & \text{if } x = D \\ \frac{\partial\theta(x)}{\partial E[H|D]}\frac{\partial E[H|D]}{\partial D} & \text{if } x \neq D \end{cases} \tag{4.18}$$

Substituting in the Pareto optimal designation share $D = D^*$ derived above we get:

$$\frac{d\theta(x)}{dD} = \begin{cases} \frac{\theta(x)}{1-\delta}[\gamma h_0(1-1+\frac{c}{(1-\pi)\gamma h_0})(1-\pi) - c] = 0 & \text{if } x = D \\ \frac{\theta(x)}{1-\delta}[\gamma h_0(1-1+\frac{c}{(1-\pi)\gamma h_0})(1-\pi) - c] = \frac{\theta(x)}{1-\delta} & \text{if } x \neq D \end{cases} \tag{4.19}$$

The two conditions directly translate into two testable hypotheses. If the designation process is in reality Pareto optimal, we expect the marginal effect of designation on housing rents to be zero at newly designated locations and to be positive at all other locations in the neighbourhood. Likewise, an excessive or restrictive designation policy will be associated with negative or positive marginal designation effects.

Assuming that the preservation probability (if undesignated) and the preservation costs are held constant our theory predicts that, in equilibrium, (Pareto optimal) designations occur as a result of an increase in the benefits associated with (external) heritage. Such increases in benefits will occur mechanically over time if the internal (and thus the external) heritage depends on housing age. The effective benefits will also increase as a result of neighbourhood turnover, if the in-migrating residents have larger heritage preferences than the incumbents. Designation then becomes a collateral effect of 'gentrification'. The older the conservation area, the greater the accrued benefits of designation may be.

Contrary to the assumption in our theory there is evidence suggesting that heritage externalities (Ahlfeldt & Maennig, 2010b; Holman & Ahlfeldt, 2012) or housing externalities more generally (Rossi-Hansberg et al., 2010) decline quite steeply in distance. The implication is that at the centre of a conservation area, where the effective external heritage is largest, the marginal designation benefit will be larger than at the margin. We justify our simplified theory on the grounds that most conservation areas are small in reality even compared to the narrow scope of housing externalities. Moreover, we allow designation effects to vary in distance to the conservation area boundary and provide estimates of designation effects at the boundary, the critical point for Pareto-efficiency as the policy benefits are presumably at their lowest.

4.2.2. Institutional context

In England, the designation of conservation areas started in 1967 and continues today under the provisions 69 and 70 of the Planning Act 1990 (Listed Buildings and Conservation Areas).[5] Conservation areas are those that have been identified as having "special architectural or historic interest, the character or appearance of which is desirable to preserve or to enhance" (Section 69). The Planning Policy Guidance Note 15 (PPG15) states that a conservation area "may form groups of buildings, open spaces, trees, historic street patterns, village greens or features of historic or archaeological interest. It is the character of the areas rather than individual buildings that conservation areas seek to enhance". Conservation areas are designated on the grounds of local and regional criteria. After the designation, the Local Authority has more control over minor developments and the demolition of buildings (Bortrill, 2005). However, the protection an area receives when it is designated a conservation area is determined at the national level to reflect the wider interests of society.

In 2011 there were around 9,800 conservation areas in England. Conservation areas vary in character and size. Many have strong historical links, for example an architectural style associated with a certain period. Besides these characteristics, designation is made based on softer benefits said to have emanated from conservation area designation including: the creation of a unique sense of place-based identity, encouraging community cohesion, and promoting regeneration (HM Government, 2010). This 'instrumentalisation' of conservation policy, which seeks to en-compass heritage values, economic values and public policy outcomes, has been identified as a key shift in the English policy context (Pendlebury, 2009; Strange & Whitney, 2003). This is reflective of the notion of heritage not as a single definable entity, but as s

---

[5] However, the first legislation to protect the historic environment was enacted in 1882 when the Ancient Monuments Protection Act was passed to protect a small number of designated ancient monuments. More statutory measures came into force in the ensuing years, but it was the passage of the Ancient Monuments Consolidation and Amendment Act in 1913 that set out a more comprehensive legislative framework for the protection of ancient monuments.

political, social, cultural and economic "bundle of processes" (Avrami et al., 2000, cited in Pendlebury (2009) p. 7).

In combination with bottom-up schemes leading to designation (e.g. community-led designation), the complex heritage preservation agenda which pursues a multitude of objectives and the institutional setting with responsibilities shared across several institutional layers creates significant scope for organized interest groups like property owners to influence the outcome of a political bargaining process.

4.3. Empirical strategy

4.3.1. Designation process

The first potentially testable implications of our theoretical model are the partial derivatives (4.12) to (4.15). As mentioned in the theory section it is difficult to find feasible proxies for the variables π, c and h_0. We therefore concentrate on testing the first of these conditions, i.e., the 'taste' for heritage γ has a positive effect on optimal designation share D^* in a neighbourhood. We adopt the common assertion that the demand for urban consumption amenities increases in education and income (Brueckner et al., 1999; Carlino & Saiz, 2008; Glaeser & Gottlieb, 2006; Shapiro, 2006; van Duijn & Rouwendal, 2013). In particular, we assume that the preference for heritage γ_n in a neighbourhood n is related to the share of people in the neighbourhood who hold a higher education certificate (DEG_i)[6] with the following functional form:

$$\gamma_{nt} = DEG_{nt}^{\vartheta} e^{-\varepsilon_{nt}} \tag{4.20}$$

where $\vartheta > 0$ such that the relationship is positive. Since the purpose of our empirical exercise is to evaluate the causal impact of changes in heritage preferences on designation status - and not the causal impact of education on heritage preference - it is sufficient to assume that ϑ captures a correlation between education and heritage preferences. ε_{nt} is a random disturbance term capturing determinants of heritage preferences that are not correlated with education. Rearranging the Pareto-efficient designation share equation (4.11), substituting the education degree proxy relationship and taking logs we arrive at the following empirical specification:

$$\log(1 - D_{nt}) = \alpha - \vartheta \log(DEG_{nt}) - \omega_n + \varepsilon_{nt} \tag{4.21}$$

$$\text{where } \alpha = \log(1 - \pi) - \log(c) \text{ and } \omega_n = \log(h_{0n}) + l_n \tag{4.22}$$

The n subscripts correspond to the individual 'neighbourhoods' of our theoretical model and we choose to represent these empirically as UK Census wards. Wards are the smallest geographical areas that are comparable between 1991 and 2011 Censuses. Subscript t stands for time periods

---

[6] We also use income as a proxy for a subsample of our data set - results are reported in the appendix (Section B.5.1).

for which we use the Census years of 1991 and 2011. All idiosyncratic time-invariant location components l_n (location-specific determinants of designation not modelled in our theory) and the unobserved heritage endowment h_{0n} of a neighbourhood n as captured by ω_n as well as the preservation probability π and the costs to owners of conservation policies are removed by taking first-differences:

$$\Delta \log (1 - D_n) = \Delta\alpha - \vartheta\Delta \log (DEG_n) + \Delta\varepsilon_n \tag{4.23}$$

Our estimation equation now depicts that a neighbourhood change reflected in a positive change in (log) educational degree share causes the (logged) share of non-designated land on the left-hand side to decrease. This is just another way of saying that a positive change in educational degree leads to a higher designation share, although the transformation is non-linear. Note that we implicitly assume that we are in equilibrium in the sense that all areas that should be designated at t are in fact designated. To support the case, we estimate our model using a long difference between 1991 and 2011, which is more than two decades after the start of the policy and the initial wave of designations. Results for the smaller differences between 1991-2001, and 2001-2011 respectively, are reported in the appendix (Section B.5.1).

Equation (4.23) evidently follows from a stylized model world. In the empirical implementation we add a number of covariates to control for alternative determinants of designation. The on-going designation is then only determined by the local changes in preferences and the steady aging of buildings and the effects on heritage, which are differentiated out. To control for the contagion effects in designation we add the initial (1991) designation share. A number of variables are added to account for heterogeneity in the net benefits of designation and abilities to express (collective) opinions in a political bargaining that may influence the designation decision. These include the initial (1991) degree share, the homeownership rate, and the household size (both in initial shares and changes). We alter the baseline model in a number of robustness checks to account for institutional heterogeneity at the TTWA level, neighbourhood appreciation trends and, to the extent possible, the historic and physical quality of the housing stock.

In practice, however, it is difficult to control for all determinants of designation that are external to our model. One particular concern is that areas can be designated if the heritage is threatened by poor maintenance in a declining neighbourhood. Such derelict is likely to be negatively correlated with our explanatory variable and is unlikely to be fully captured by the control variables we have at hand. At the same time, the policy itself could make it more likely that educated people are attracted to designated areas due to a different valuation of uncertainty (reverse causality). Since an OLS estimation of equation (4.23) can result in a significant bias in either direction we make use of instrumental variables z_n, which predict changes in education, $\rho(z_n, \Delta \log DEG_n) \neq 0$, but must be conditionally uncorrelated with the differenced error term, $\rho(z_n, \Delta\varepsilon_{nt}) = 0$. We argue that rail station (in London additionally Tube station) density as well as effective employment accessibility (both time-invariant in levels) are good predictors of neighbourhood gentrification (Florida, 2002; Glaeser et al., 2001).[7] We also argue that it is unlikely that these level variables directly impact on the likelihood of designation conditional on the unobserved heritage endowment in the fixed effects ω_n.

Another empirical concern is that, theoretically, a decrease in preferences for heritage must provoke a reduction of the designated area. The abolishment of conservation areas, however, is extremely rare in England (as in most institutional contexts) so our data is left-censored (we do not observe increases in the share of non-designated land). We therefore take the model to the data using a tobit approach:

$$Y_n^* = \Delta\alpha - \vartheta\Delta\log\left(DEG_n\right) + \Delta\varepsilon_n, \quad \Delta\varepsilon_n \sim N(0,\sigma^2) \tag{4.24}$$

$$Y_n = \begin{cases} Y_n^*, & \text{if } Y_n^* = \Delta\log\left(1 - D_n\right) \quad < 0 \\ 0, & \text{if } Y_n^* \qquad\qquad\qquad\quad \geq 0 \end{cases} \tag{4.25}$$

4.3.2. Pareto optimality

To test whether the designation share in practice is set at the (locally) Pareto-optimal level (D^*) we estimate the effect of the event of designation on property prices within and surrounding conservation areas. In its essence our quasi-experimental methods are a derivative of the established difference-in-differences (DD) methodology (Bertrand et al., 2004). We draw elements of the increasingly popular regression discontinuity designs (RDD) (Imbens & Lemieux, 2008), however, to relax the DD assumptions of homogeneous trends and a singular treatment date to separate smooth variation (e.g. externalities) and discontinuities (e.g. policy zones) in treatment effects from correlated unobservables.

Difference-in-differences

We define a group of 912 ‘treated’ conservation areas as those that were designated between the years 1996 and 2010 to ensure we observe property transactions both before and after the designation date. Our counterfactuals are established via various control groups of housing units that are similar to the treated units but are themselves not treated. These control groups are discussed in more detail in the results section and in the appendix (Section B.3.2).

Our baseline DD model takes the following form:

$$p_{it} = \beta^I I_i + \beta^E E_i + \beta^{IPost}(I_i \times Post_{it}) + \beta^{EPost}(E_i \times Post_{it}) + X_i^{'}\mu + f_n + Y_t + \epsilon_{it} \tag{4.26}$$

---

[7] Our measure of effective employment accessibility aggregates employment in surrounding regions weighted by distance. We use exponential distance weights that are popular in the theoretical (Fujita & Ogawa, 1982; Rossi-Hansberg et al., 2010) and empirical literature (Ahlfeldt et al., 2012; Ahlfeldt & Wendland, 2013) and the decay parameter estimate provided by Ahlfeldt (in press). Transport infrastructure is captured by a kernel density measure (Silverman, 1986) with a radius of 2 km which is considered to be the maximum distance people are willing to walk (Gibbons & Machin, 2005).

where p_{it} is the natural logarithm of the transaction price for property i in time period t, I_i is a dummy variable equal to one if the observation is internal to a treated conservation area, E_i indicates observations external to the treated CA. While our standard models use a buffer area of 500m we also experiment with various alternative spatial specifications. $Post_{it}$ is a dummy variable indicating whether the transaction year t is equal to or greater than the designation year, X_i is a vector of controls for property, neighbourhood and environmental characteristics, f_n is a set of n location fixed effects and Y_t are year effects. The β^{IPost} and β^{EPost} parameters give the difference-in-differences estimates of the designation effect on the properties within and just outside a conservation area. We show in Section B.3.2 that β^{IPost} is equal to the net marginal policy (designation costs and benefits) effect while β^{EPost} reflects the pure (albeit spatially discounted) policy benefit.

Temporal regression discontinuity design of differences (RDD-DD)

The standard DD specification (4.26) identifies the policy treatment effect under some arguably restrictive assumptions. Firstly, the treatment and control groups follow the same trend before and after the treatment. Secondly, the treatment occurs at a singular and a priori known date and affects the level (and not the trend) of the outcome variable. These assumptions are evidently violated if the outcome variable does not respond immediately to the treatment, e.g., because of costly arbitrage, or in anticipation of the treatment, for example because of an investment motive by buyers (Ahlfeldt & Kavetsos, 2013). In our case, a positive pre-trend can also be associated with the gentrification that causes designation according to our theoretical model, a reverse causality problem.

To address these limitations of the standard DD we refine the model to accommodate differences in trends across the treatment and the control group. We borrow the functional form from the RDD literature where a (temporal) treatment effect is identified as an instant adjustment - a discontinuity - conditional on higher order polynomial (pre- and post-) trends, which are assumed to be unrelated to the treatment (Bento et al., 2010). In our regression discontinuity design of differences (RDD-DD) we combine an RDD-type polynomial specification of trends with the control group-based counterfactual from the DD. It is therefore possible to attribute pre- and post-trends to the treatment as long as it is credible to assume that treatment and control groups would have followed the same trend in the absence of the treatment. It is notable that even if this assumption is violated the RDD-DD (unlike the standard RDD) will at least remove macro-economic shocks from the treatment effect by taking differences from the control group. This improves identification so long as the control group remains unaffected by the treatment. Our RDD-DD with linear trends takes the following form:

$$\begin{aligned} p_{it} &= \beta^{I} I_i + \beta^{IYD}(I_i \times YD_{it}) + \beta^{E} E_i + \beta^{EYD}(E_i \times YD_{it}) + \beta^{IPost}(I_i \times Post_{it}) \\ &+ \beta^{IPostYD}(I_i \times Post_{it} \times YD_{it}) + \beta^{EPost}(E_i \times Post_{it}) \\ &+ \beta^{EPostYD}(E_i \times Post_{it} \times YD_{it}) + X_i'\mu + f_n + Y_t + \epsilon_{it} \end{aligned} \tag{4.27}$$

where YD_{it} is the number of years since the designation date, with the pre-designation years having negative values. As in the RDD, the polynomial degree of the trend can be increased subject to sufficient degrees of freedom. We make use of a quadratic trend specification and evaluate the fit of the parametric polynomial function using a semi-parametric version of (4.27) that replaces the YD_{it} variables with full sets of years-since-designation effects (details in Section B.3.2).

A significant 'dis-in-diff' parameter (β^{IPost} or β^{EPost}) can be entirely attributed to the treatment even under the existence of complex relative trends that are unrelated to the treatment or may even have caused the treatment as the comparison is made just before and just after the treatment date. Under the assumption of homogeneous counterfactual trends the significant pre-trend parameters (β^{IYD} or β^{EYD}) describe the anticipation effects. Significant post-trend parameters ($\beta^{IPostYD}$ or $\beta^{EPostYD}$) then indicate changes in relative trends after the treatment. In conjunction, the 'dis-in-diff' and the pre- and post-trend parameters describe the full temporal structure of the treatment effect. As a program evaluation tool that is applicable to a variety of event studies, the RDD-DD thus naturally comes with a stronger test (dis-in-diff) and a weaker test (trends) of whether there exists an effect of the treatment.

Spatial regression discontinuity design of difference-in-differences (RDD-DD)

In contrast to our theory, in reality there most likely exists a spatial decay to the heritage externalities. This decay implies that the external heritage effect should be stronger at the centre of the conservation area than at the boundaries. The policy benefit, which is a transformation of the external heritage effect, should also be greater at the centre of the newly designated conservation area. Likewise, the predicted positive policy effects just outside the boundary should be decaying in distance to the conservation area (CA) boundary. At the CA boundary there may be a discontinuity as the cost of the policy ends abruptly at the boundary, whereas potential externalities decay smoothly across it. The combination of trends and discontinuities potentially caused by the treatment resembles the temporal identification problem just described and will be addressed by a similar combination of RDD and DD tools. Essentially, we use the RDD tools to capture how the difference (before and after) in the differences (treatment vs. control) of property prices varies along the (internal and external) distances from the CA boundary. Unlike in the standard (spatial) RDD, unobserved time-invariant spatial effects can be held constant due to the availability of spatiotemporal variation. In our spatial RDD-DD model it is therefore possible to attribute spatial trends (with respect to distance to the CA boundary) as well as a discontinuity (at the CA boundary) to the treatment provided that the spatial trends are uncorrelated with unobserved temporal trends. The spatial RDD-DD we estimate takes the following form:[8]

$$\begin{aligned}p_{it} = {} & \beta^{I}T_i + \beta^{ID}(T_i \times D_i) + \beta^{IPost}(T_i \times Post_{it}) + \beta^{IDPost}(T_i \times D_i \times Post_{it}) \\ & + \beta^{E}O_i + \beta^{OD}(O_i \times D_i) + \beta^{OPost}(O_i \times Post_{it}) + \beta^{ODPost}(O_i \times D_i \times Post_{it}) \\ & + X_i^{'}\mu + f_n + Y_t + \epsilon_{it}\end{aligned} \tag{4.28}$$

where D_i is the distance from the property to the conservation area boundary (internal distances are negative values), O_i indicates properties outside a treated conservation area and T_i indicates the conservation area that is nearest to a property that is treated at any point of the study period. In order to fully explore the extent of spatial externalities O_i indicates a larger area outside CAs rather than just within 500 m as indicated by E_i in previous models. Specifically, the empirical analysis uses properties within 1,400 m of the treated conservation area. As with the temporal RDD-DD specification we also estimate an expanded model specification in which we allow for quadratic distance trends and semi-nonparametric specifications replacing the distance variable with some distance bin effects. The coefficient β^{IPost} gives the intercept of the internal effect (i.e. the internal effect at the boundary) and β^{IDPost} estimates how this changes with respect to internal distance. Jointly, these terms capture the net policy costs and benefits of designation for internal treated areas. A zero β^{IPost} coefficient would be reflective of a zero effect at the boundary and would be in line with the optimality condition derived in the theory section. A negative β^{IDPost} would be in line with the existence of policy benefits (due to increased preservation probability) that spillover with decay. The parameters β^{OPost} and β^{ODPost} allow for a spatial discontinuity treatment effect at the boundary and heterogeneity in spatial trends inside and outside the treated areas. As with β^{IDPost}, a jointly negative $\beta^{IDPost} + \beta^{ODPost}$ would be in line with the decaying policy benefits external to the conservation area. The discontinuity at the border is measured by the external intercept term β^{OPost}. A statistically positive estimate would indicate a cost to the policy. A jointly positive effect of $\beta^{IPost} + \beta^{OPost}$ would in turn indicate the existence of policy benefits.

4.4. Data

We have compiled two distinct data sets for the two stages of the empirical analysis. Both data sets make use of data provided by English Heritage. These include a precise GIS map of 8,167 conservation areas in England, the Conservation Areas Survey containing information on community support and risk status (average condition, vulnerability and trajectory of a conservation) and a complete register of listed buildings.

For the analysis of the determinants of designation we use UK census wards as a unit of analysis. Shares of designated land within each Census ward are computed in a Geograph-

---

[8] In models with historical CAs as control groups the following terms are also included $\beta^{CD}(C_i \times D_i) + \beta^{EC}EC_i + \beta^{ECD}(EC_i \times D_i)$, where C_i indicates internal to control CA and EC_i external to control CA. This ensures that spatial effects are estimated conditional on the spatial trends in control CA.

ical Information Systems environment. Various ward level data on educational level, average household size and homeownership status and vacancy rate were obtained from the UK Census. Any changes in ward boundaries between the years were corrected for using the online conversion tool GeoConvert. For robustness tests we also collected a measure of the ward's average income (Experian). The instrumental variables station density and employment potential are regenerated data that stem from nomis (workplace employment) and the Ordinance Survey (rail stations).

For the analysis of the capitalization effects of designation we use transactions data related to mortgages granted by the Nationwide Building Society (NBS) between 1995 and 2010. The data for England comprise 1,088,446 observations and include the price paid for individual housing units along with detailed property characteristics. These characteristics include floor space (m^2), the type of property (detached, semi-detached, flat, bungalow or terraced), the date of construction, the number of bedrooms and bathrooms, garage or parking facilities and the type of heating. There is also some buyer information including the type of mortgage (freehold or leasehold) and whether they are a first-time buyer. Importantly, the transaction data includes the full UK postcode of the property sold allowing it to be assigned to grid-reference coordinates.

With this information it is possible within GIS to calculate distances to conservation area borders and to determine whether the property lies inside or outside of these borders. Furthermore, it is possible to calculate distances and other spatial measures (such as densities) for the amenities and environmental characteristics such as National Parks, as well as natural features like lakes, rivers and coastline. The postcode reference also allows a merger of transactions and various household characteristics (median income and ethnic composition) from the UK census, natural land cover and land use, various amenities such as access to employment opportunities, cultural and entertainment establishments and school quality. A more detailed description of all the data used is in the appendix (Section B.4).

4.5. Empirical results

4.5.1. Designation process

Table 4.1 on page 70 reports the results of our tobit model of the designation process defined in equation (4.24). The non-instrumented baseline model is in column (1). As predicted by our theory, increases in educational levels that are presumably correlated with heritage preferences are associated with reductions in the share of non-designated land. More precisely, an increase in the degree share by 1% is associated with a 0.12% reduction in the share of non-designated land. This decrease corresponds to an $0.12\% \times (1 - \bar{D}_{t-1})/\bar{D}_{t-1} = 2.61\%$ increase in the share of designated land for a ward with the mean of the positive initial designation shares $\bar{D}_{t-1} = 4.4\%$. The effect substantially increases once we instrument the change in degree share using rail station density and employment potential (column 2). This increase is in line with unobserved (positive) deterioration trends that a) increase the likelihood of designation and b) are negatively correlated with changes in degree share. Introducing the instruments, the effect of a 1% increase in degree share on the share of non-designated land increases to 0.52%, which for a ward with the mean initial designation share $\bar{D}_{t-1}$ corresponds to an increase in the designated land share of about 11%. While we have argued that our estimates are supposed to reflect a causal estimate of gentrification (proxied by degree shares) on designation probabilities and not necessarily a causal effect of degree share on designation share, a parameter estimate of $\hat{\vartheta} = 0.52$ is at least indicative of heritage preferences increasing relatively steeply in education.

In a series of robustness checks columns (3) to (5) of the same table provide variations of the benchmark model (2). We add TTWA effects to control for unobserved institutional heterogeneity in column (3). Column (4) adds a measure of property price appreciation, which we obtain from ward-level regressions of log property prices on a time trend (and property controls, see the appendix for details). In this specification we control for a potentially positive correlation between owners' risk aversion and the value of their properties - typically their largest assets. This is a potentially important control since a larger risk aversion increases the benefit from a policy that increases certainty regarding the future of the neighbourhood and, thus, potentially increases the optimal designation share. It is a demanding control since positive price trends are potentially endogenous to changes in neighbourhood composition and may thus absorb some of the gentrification effect on designation. The price trends are indeed positively, though not statistically significantly, associated with increases in the share of designate land. Adding controls capturing vacancy trends and levels, the density of listed buildings and some risk and vulnerability assessments from the Conservation Areas Survey tend to increase the education effect (column 5).

Across all specifications we find that, besides positive changes in designation share, high initial levels of degree shares are positively correlated with increases in the share of designated land. While high initial and positive changes in homeownership rate, *ceteris paribus*, are associated with less designation, it is notable that the (positive) impact of neighbourhood change on designations shares (interaction term) is particularly large in high homeownership areas (see column 6). This is in line with a political economy literature that suggests that homeowners tend to form well-organized interest groups (Ahlfeldt & Maennig, 2013; Brunner & Sonstelie, 2003; Dehring et al., 2008; Fischel, 2001a). Aside from the uninstrumented model (1), the results in Table 4.1 suggest contagion effects in designation, i.e. designated land shares tend to increase where shares were initially high.

Table 4.1.: Designation process

	(1) Tobit	(2) IV Tobit	(3) IV Tobit	(4) IV Tobit	(5) IV Tobit	(6) IV Tobit
	Δ log non designation share$_t$					
Δ log degree share$_t(\vartheta)$	-0.116***	-0.519***	-0.587***	-0.528***	-0.560***	-0.513***
	(0.019)	(0.061)	(0.105)	(0.062)	(0.061)	(0.060)
log degree share$_{t-1}$	-0.127***	-0.276***	-0.337***	-0.280***	-0.289***	-0.269***
	(0.010)	(0.024)	(0.046)	(0.024)	-(0.025)	(0.023)
log designation share$_{t-1}$	-0.004	-0.022*	-0.027**	-0.035***	-0.026*	-0.025**
	(0.012)	(0.013)	(0.013)	(0.013)	(0.033)	(0.013)
Δ log homeownership$_t$	0.189***	0.263***	0.319***	0.262***	0.259***	0.255***
	(0.025)	(0.029)	(0.042)	(0.029)	(0.033)	(0.028)
log homeownership$_{t-1}$	0.127***	0.057***	0.091***	0.053***	0.042***	0.153***
	(0.015)	(0.018)	(0.018)	(0.018)	(0.018)	(0.040)
Δ log aver. household	0.042	0.031	0.006	0.028	-0.009	0.047
size$_t$	(0.037)	(0.040)	(0.046)	(0.039)	(0.040)	(0.040)
log aver. household	0.073*	-0.016	-0.122*	-0.011	-0.075	-0.013
size$_{t-1}$	(0.046)	(0.049)	(0.066)	(0.049)	(0.051)	(0.049)
log price trend				-0.011		
				(0.021)		
Δ log vacancy rate$_t$					-0.021***	
					(0.006)	
log vacancy rate$_{t-1}$					-0.022***	
					(0.010)	
log listed building density					1E-4	
					(0.004)	
aver. condition score					-0.070***	
(1 best, 4 worst)					(0.019)	
aver. vulnerability score					-0.045***	
(1 low, 8 high)					(0.017)	
aver. trajectory score (-2 im-					0.043	
proving, +2 deteriorating)					(0.036)	

Δ log degree share$_t \times$						-0.201***
homeownership$_{t-1}$						(0.085)
Constant	0.013	0.052	0.231*	0.018	0.039	0.096**
	(0.046)	(0.048)	(0.119)	(0.067)	(0.065)	(0.051)
TTWA Effects	No	No	Yes	No	No	No
Chi2		350.753	634.960	368.036	475.892	354.198
ExogP		0.000	0.000	0.000	0.000	0.000
Overid		0.017		0.100	0.073	0.009
OveridP		0.897		0.752	0.787	0.926
Observations	7965	7965	7965	7965	7965	7965

Notes: $\Delta \log$ non designation share$_t$ denotes the share of non designated land as defined by (4.23). See the data section for a description of control variables. IVs are station density and employment potential in all models except model (1). Model (4) includes a dummy variable indicating 60 wards for which no price trend could be computed due to insufficient transactions. Standard errors in parentheses and clustered on fixed effects in (3). *p< 0.05, **p< 0.01, ***p< 0.001. Chi2 is the model Wald χ^2 statistic, ExogP the p-value of the Wald exogeneity test, Overid (OveridP) denotes the test of overidentifying restrictions and its p-value.

Further robustness

Our IVs comfortably pass the typical statistical tests. They are overidentified, i.e. the instruments are jointly valid assuming that at least one instrument is exogenous, and relevant in terms of prediction power (see first stages in Table B.5 on page 157). We have also experimented with four alternative sets of IVs. We have moreover split up the 1991–2011 long difference into two shorter differences (1991–2001 and 2001–2011), used the change in income as a proxy for heritage preferences (for 2001–2011) and run the baseline model in OLS keeping only observations with positive changes in shares of designated land. The results are presented in the appendix (Section B.4) and support those discussed here.

4.5.2. Pareto optimality

Difference-in-Differences

Table 4.2 on page 73 shows the results from an estimation of the standard DD equation (4.26) for different selections of control groups and fixed effects. Each model includes controls for property, location, and neighbourhood characteristics, year effects and location fixed effects to hold unobserved time-invariant effects constant. Column (1) is a naïve DD using the mean price trend of all properties located beyond 500 m of a treated conservation area as a counterfactual. Columns (2) to (7) provide more credible counterfactuals by restricting the control group to properties that are presumably similar to the treated properties. Column (2), with ward fixed effects, and (3), with nearest CA fixed effects, provide a spatial matching by restricting the sample to properties within 2 km of a treated CA, where many unobserved location characteristics are likely to be similar. In column (4) we impose the additional restriction that properties in the control group must fall within 500 m of the boundaries of a historically designated conservation area (before 1996), which increases the likelihood of unobserved property characteristics being similar. While areas that are designated at any point in time are likely to share many similarities, the diminishing returns to designation in our theoretical framework also imply that heritage-richer areas should generally be designated first. To evaluate whether the designation date of the treated conservation areas, relative to those on the control group, influences the DD estimate, we define CA designated 1996–2002 as a treatment group and form control groups based on CAs designated just before (1987–1994) or right after (2003–2010) in columns (5) and (6). In column (7), finally, we use environmental, property and neighbourhood characteristics to estimate the propensity of being in a treated (1996–2010) CA over a historical (<1996) CA. Then the treated CAs are matched to their 'nearest-neighbour', i.e. the most similar non-treated CA, based on the estimated propensity score (Rosenbaum & Rubin, 1983). A fixed effect is defined for each treated CA and its nearest-neighbour control CA such that the treatment effect is estimated by the direct comparison between the treated CA and its nearest-neighbour.

We anticipate that the strength of the counterfactual increases as we match the treatment and control group based on proximity (2 and 3), proximity and qualifying for designation (4, 5 and

6) and qualifying for designation and a combination of various observable characteristics (7). As the credibility of the counterfactual increases, the statistical significance of the treatment effect tends to decrease. Benchmarked against the nationwide property price trend both the internal effect (Inside × Post) and the external effect (Within 500 m × Post) are significant at the 5% level. The magnitudes of these effects are of similar size, implying a 2.8% premium for houses inside newly designated conservation areas and a 2.3% premium outside. The spatial matching (2 and 3) renders the internal treatment effect insignificant (2 and 3). With further refinements in the matching procedure the external effect also becomes insignificant. Table 4.2 results, thus, suggest that designation does not lead to significant property price adjustments. Evidence is weak for positive (policy) spillovers to nearby areas.

Table 4.2.: Conservation area premium – designation effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				log property transaction price			
Inside treated CA ×	0.028***	0.014	0.014	0.003	-0.024	-0.077	-0.003
Post designation	(0.009)	(0.009)	(0.010)	(0.012)	(0.070)	(0.111)	(0.015)
Within 500 m buffer of	0.023***	0.013***	0.012***	0.004	0.012	-0.005	-0.005
treated CA × Post des.	(0.004)	(0.004)	(0.005)	(0.006)	(0.027)	(0.022)	(0.010)
Inside treated CA	-0.043***	-0.038***	-0.048***	-0.037***	-0.062	0.029	-0.024
	(0.009)	(0.009)	(0.010)	(0.012)	(0.057)	(0.108)	(0.021)
Within 500 m buffer of	-0.010**	-0.004	-0.011**	-0.005	0.003	0.006	-0.002
treated CA	(0.004)	(0.004)	(0.005)	(0.005)	(0.030)	(0.023)	(0.013)
Hedonic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Neighbourhood controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward effects	Yes	Yes	No	No	No	No	No
Nearest treated CA effects	No	No	Yes	Yes	Yes	Yes	No
Matched CA effects	No	No	No	No	No	No	Yes
Treatment group: CAs designated	1996-2010	1996-2010	1996-2010	1996-2010	1996-2002	1996-2002	1996-2010
Control group	Full England sample	Within 2 km of treated CA	Within 2 km of treated CA	Within 500 m of CA designated before 1996 & within 2 km of treated CA	Within 500 m of CA designated before 1987-95 & within 2 km of treated CA	Within 500 m of CA designated before 2003-10 & within 2 km of treated CA	Within 500 m of pre-1996 CA matched on propensity score
R^2	0.921	0.922	0.915	0.915	0.861	0.864	0.909
AIC	-587375	-156426	-130469	-67046	-5409	-8476	-41184
Observation	1,088k	302k	302k	178k	21k	32k	133k

Notes: Standard errors in parentheses are clustered on location fixed effects. Conservation area control groups in columns (4)-(7) have separate fixed effects for the areas inside and outside a conservation area. *p<0.10, **p<0.05, ***p<0.01.

Temporal RDD-DD

Table 4.3 on page 76 illustrates the results of the estimation of the (temporal) RDD-DD outlined in equation (4.27). We present the results of a variety of models that feature linear (1–5) and quadratic (6–10) trends and several of the control groups utilized in Table 4.3. One important finding across these specifications is that the external (Within 500m × Post) 'dis-in-diff' parameter estimate is significant in four of 10 specifications at the 5% level and in one half of the specifications at the 10% level, whereas, the internal (Inside × Post) parameter is only significant in one specification at the 10% level (column 8). This suggests primarily that there exists a significant treatment effect exactly at the treatment date only for the external area. This interpretation is in line with the predictions of our theoretical model. Another finding illustrated by Table 4.3 is the positive change in the internal price trend after a CA has been designated (Inside treated CA × Post designation × Years designated). The change in trend, which is significant at the 5% level in seven of the 10 models, may be regarded as evidence for a cumulative internal effect of the designation policy. There is also a faster appreciation in the external area post-designation that is significant in four of the 10 models. In short, the temporal RDD-DD has confirmed that designation policy causes no immediate effect inside the conservation area but shows instead that it increases the speed of price appreciation over time. The RDD-DD has also uncovered that areas external to the conservation area receive an immediate shift in prices at the designation date in line with our theoretical hypothesis.

Figure 4.2 on page 77 provides a graphical illustration of the predicted effect of being in the treatment group over the control group against years-since-designation. A horizontal red line is drawn at the mean of the pre-treatment effects in order to illustrate the differences between the RDD-DD results and those of the standard DD. The positive impact of designation on (relative) price trends suggested by the RDD-DD (black lines) is supported by the functionally more flexible semi-parametric estimates for the 'years-since-designation bins' (grey dots).[9] However, the post-treatment effects are never statistically distinguished from the pre-period mean, which is in line with the DD estimates.

Figure 4.3 on page 78 provides an analogical illustration for the external treatment effect, i.e. the spillovers onto areas adjacent to the designated CAs. Again, the post-period estimates do not deviate significantly from the pre-period mean. However, the top-left panel illustrates a large discontinuity at the treatment date that is statistically significant in Table 4.3 on page 76. As with the internal effects, there is a positive trend shift post-designation.

---

[9] Confidence bands for the semi-parametric 'bins' model are presented in the appendix (Section B.5.2).

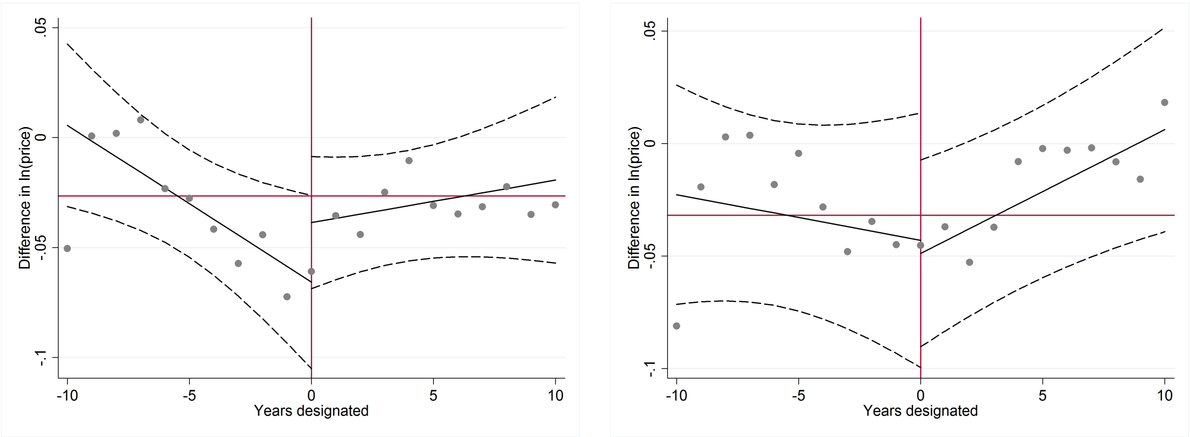
Table 4.3.: Regression discontinuity design of differences between treatment and control (RDD-DD)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	log property transaction price									
Inside treated CA ×	0.015	0.022	0.024	0.027	-0.006	0.023	0.033	0.038*	0.036	0.020
Post designation	(0.015)	(0.015)	(0.015)	(0.017)	(0.018)	(0.023)	(0.021)	(0.023)	(0.024)	(0.024)
Within 500m buffer of	0.006	0.013*	0.015**	0.020**	-0.007	0.013	0.017**	0.022**	0.017	0.009
treated CA × Post des.	(0.007)	(0.007)	(0.007)	(0.008)	(0.003)	(0.008)	(0.008)	(0.009)	(0.010)	(0.014)
Inside treated CA ×	0.000	-0.004	-0.004	-0.007**	-0.002	-0.010	-0.016*	-0.019*	-0.019*	-0.020*
Years designated	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.010)	(0.009)	(0.010)	(0.010)	(0.011)
Inside treated CA ×						-0.001	-0.001	-0.001	-0.001	-0.002*
Years designated^2						(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Inside treated CA × post	0.003	0.007**	0.008**	0.009**	0.008*	0.020	0.026**	0.032**	0.031**	0.031
designation × Years des.	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.014)	(0.012)	(0.013)	(0.013)	(0.014)
Inside treated CA × post						0.000	0.000	0.000	0.000	0.001
designation ×Years des.^{2}						(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Within 500m of treated	0.002	-0.002*	-0.002*	-0.004***	-0.001	-0.001	-0.004	-0.007*	-0.004	-0.009
CA × Years des.	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)
Within 500m of treated						-0.000	-0.000	-0.000	0.000	-0.001
CA ×Years des.^{2}						(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Within 500m of treated CA ×	0.001	0.004***	0.004***	0.005***	0.003	0.003	0.007	0.011**	0.008	0.009
post des. × Years des.	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.005)	(0.005)	(0.005)	(0.006)	(0.010)
Within 500m of treated CA ×						0.000	0.000	0.000	-0.000	0.001
post des. ×Years des.^{2}						(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Hedonic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Neighbourhood controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward effects	Yes	Yes	No	No	No	Yes	Yes	No	No	No
Nearest treated CA effects	No	No	Yes	Yes	No	No	No	Yes	Yes	No
Matched CA effects	No	No	No	No	Yes	No	No	No	No	Yes
Control group as in prev. Tab., column	(1)	(2)	(3)	(4)	(7)	(1)	(2)	(3)	(4)	(7)

R^2	0.920	0.921	0.912	0.914	0.907	0.920	0.921	0.912	0.914	0.907
AIC	-547688	-147818	-120160	-64425	-39321	-548078	-147839	-120191	-64467	-39329
Observation	995k	277k	277k	164k	123k	995k	277k	277k	164k	123k

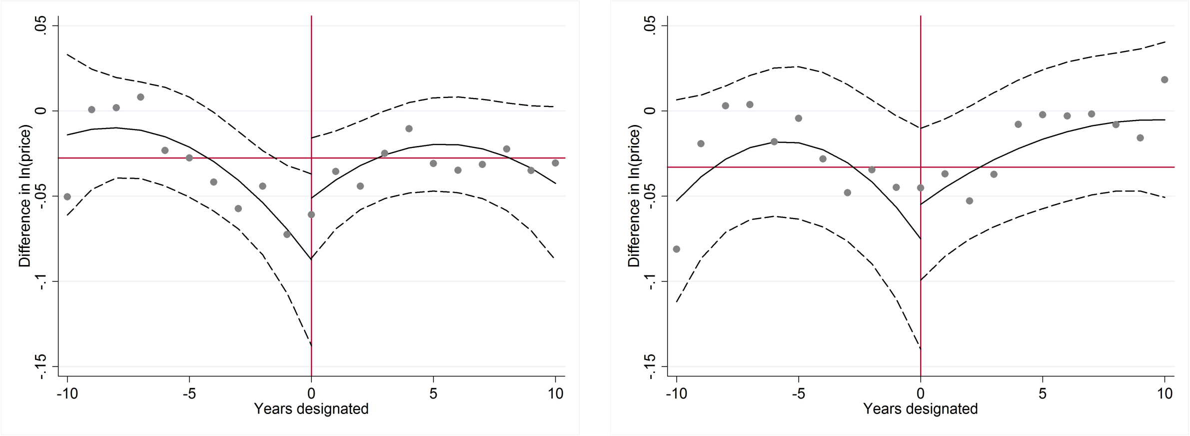
Notes: Standard errors in parentheses are clustered on location fixed effects. Conservation area control groups in columns (4)-(7) have separate fixed effects for the areas inside and outside a conservation area. Observations dropped if years designated falls outside of range -10/+10 years. $^*p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Figure 4.2.: RDD-DD internal estimates



(a) Nearest treated CA (linear trends)
Tab. 4.3, column (4)

(b) Matched CA (linear trends)
Tab. 4.3, column (5)



(c) Nearest treated CA (quadratic trends)
Tab. 4.3, column (9)

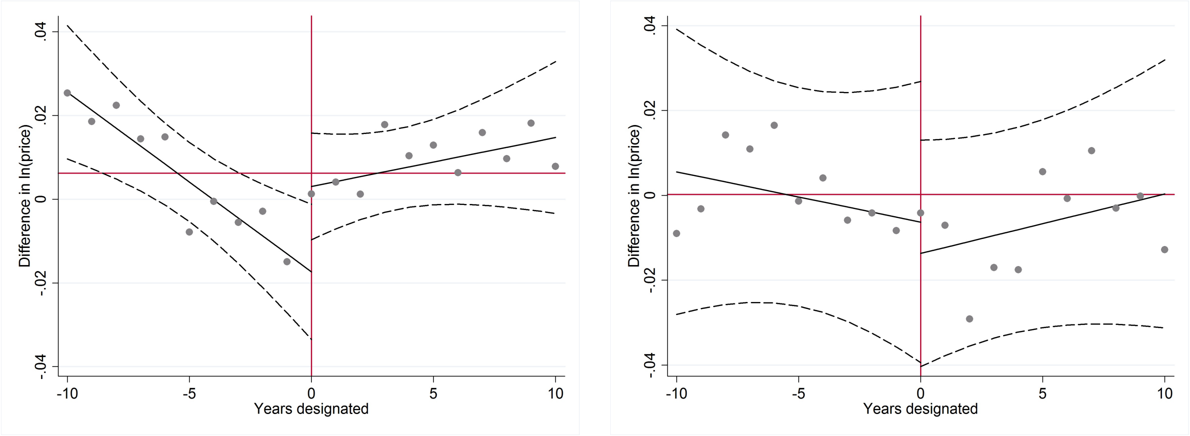
(d) Matched CA (quadratic trends)
Tab. 4.3, column (10)

Note: The solid lines are graphical illustrations of the parametric estimates presented in Table 4.3 on page 76 and estimated using equation (4.27). The dashed lines indicate the 95% CI which are calculated using standard errors of multiplicative interaction terms presented by Aiken et al. (1991). The grey dots plot the point estimates of ‘years-since-designation bins’ effects obtained from separate regression described and presented in more detail in the appendix (Section B.5.2). The horizontal red line illustrates the mean of the pre-treatment estimates.

Spatial RDD-DD

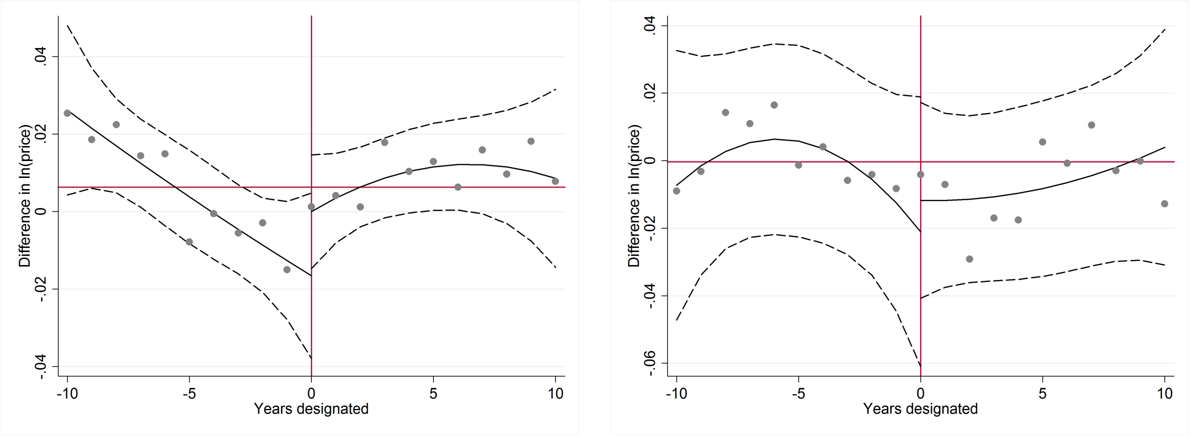
Table 4.4 on page 80 shows the results of the estimation of the (spatial) RDD-DD model outlined in equation (4.28). As with the temporal RDD-DD, we present the results of a variety of models that feature linear (1–5) and quadratic (6–10) trends and several of the control groups utilized in Table 4.2. One interesting and consistent feature of Table 4.4 is that the positive dis-continuity coefficient (Outside × Post) matches the expected (positive) sign under the existence of a policy cost inside. However, the parameter is statistically insignificant in all models.

Figure 4.3.: RDD-DD external estimates



(a) Pre-1996 CA within 2 km (linear trends)
Tab. 4.3, column (4)

(b) Matched pre-1996 CA (linear trends)
Tab. 4.3, column (5)



(c) Pre-1996 CA within 2 km (quadratic trends)
Tab. 4.3, column (9)

(d) Matched pre-1996 CA (quadratic trends)
Tab. 4.3, column (10)

Note: The solid lines are graphical illustrations of the parametric estimates presented in Table 4.3 on page 76 and estimated using equation (4.27). The dashed lines indicate the 95% CI which are calculated using standard errors of multiplicative interaction terms presented Aiken et al. (1991). The grey dots plot the point estimates of 'years-since-designation bins' effects obtained from separate regression described and presented in more detail in the appendix (Section B.5.2). The horizontal red line illustrates the mean of the pre-treatment estimates.

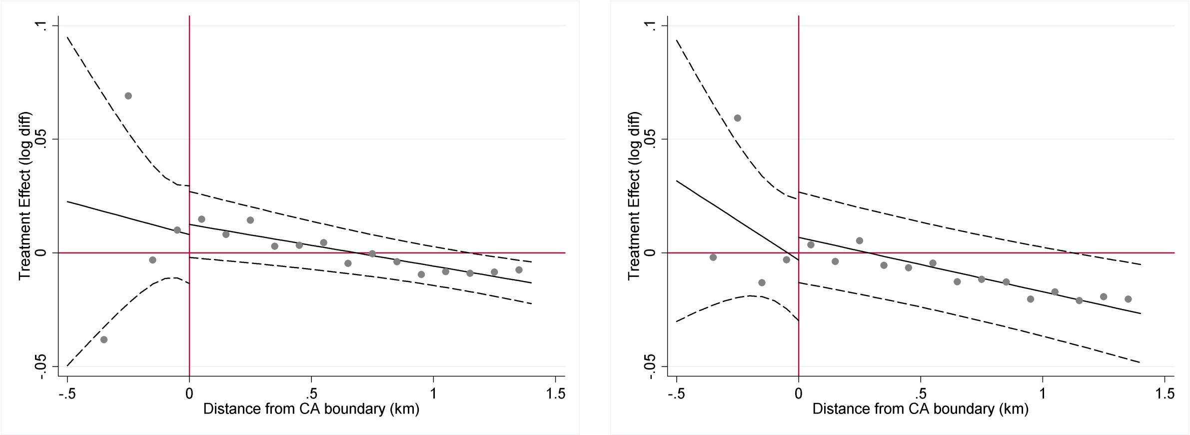
Table 4.4.: Spatial regression discontinuity design of difference-in-differences (RDD-DD)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	log property transaction price									
Within 1.4km of treat.	0.027***	0.014	0.012	0.008	-0.003	0.026**	0.014	0.012	0.008	-0.005
CA × post des.	(0.010)	(0.010)	(0.011)	(0.011)	(0.014)	(0.011)	(0.012)	(0.012)	(0.012)	(0.015)
Within 1.4km of treat. CA ×	-0.057	-0.032	-0.030	-0.029	-0.070	-0.096	-0.046	-0.040	-0.040	-0.118
dist. to boundary × post des.	(0.081)	(0.075)	(0.080)	(0.077)	(0.068)	(0.156)	(0.154)	(0.162)	(0.157)	(0.143)
Within 1.4km of treat. CA ×						-0.059	-0.017	-0.018	-0.017	-0.099
dist. to boundary^2× post des.						(0.132)	(0.131)	(0.140)	(0.136)	(0.130)
Outside treated CA ×	0.004	0.005	0.005	0.004	0.010	0.009	0.009	0.008	0.009	0.016
post design.	(0.010)	(0.010)	(0.010)	(0.009)	(0.011)	(0.012)	(0.012)	(0.011)	(0.011)	(0.012)
Outside treated CA × dist.	0.039	0.016	0.013	0.011	0.046	0.064	0.014	0.013	0.004	0.080
to boundary × post des.	(0.081)	(0.075)	(0.080)	(0.078)	(0.069)	(0.157)	(0.155)	(0.163)	(0.159)	(0.145)
Outside treated CA × dist.						0.070	0.028	0.025	0.029	0.109
to boundary^2× post des.						(0.133)	(0.132)	(0.140)	(0.136)	(0.130)
Hedonic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Neighbourhood controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward effects	Yes	Yes	No	No	No	Yes	Yes	No	No	No
Nearest treated CA effects	No	No	Yes	Yes	No	No	No	Yes	Yes	No
Matched CA effects	No	No	No	No	Yes	No	No	No	No	Yes
Control group	Full England sample	Within 2 km of treated CA	Within 2 km of treated CA	Within 1.4 km of CA designated before 1996 & within 2km of treaded CA	Within 1.4 km pre-1996 CA matched on propensity score	Full England sample	Within 2 km of treated CA	Within 2 km of treated CA	Within 1.4 km of CA designated before 1996 & within 2km of treaded CA	Within 1.4 km pre-1996 CA matched on propensity score

R^2	0.921	0.922	0.915	0.914	0.905	0.921	0.922	0.915	0.914	0.921
AIC	-587538	-156448	-130478	-118076	-101076	-587533	-156444	-130478	-118074	-587538
Observation	1,088k	302k	302k	281k	327k	1,088k	302k	302k	281k	327k

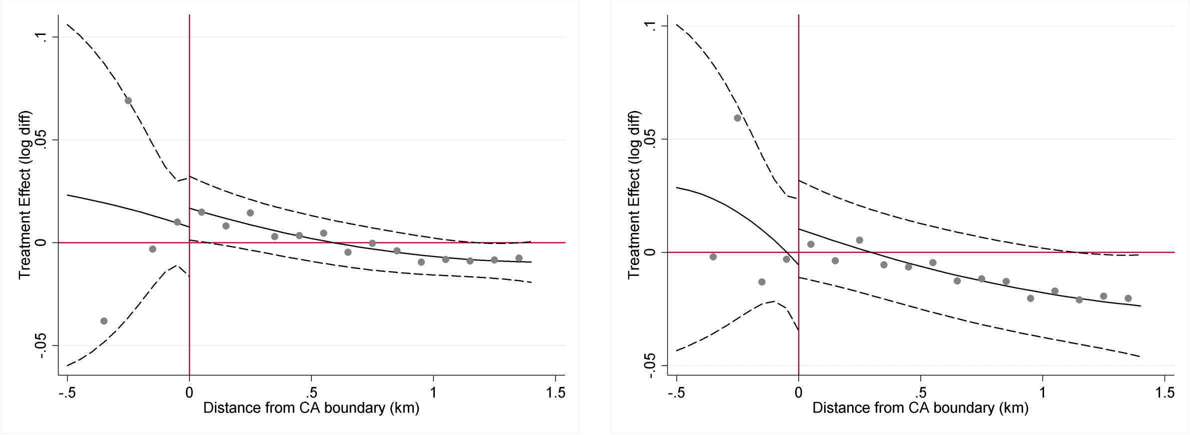
Notes: Standard errors in parentheses are clustered on the location fixed effects. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Figure 4.4.: RDD-DD spatial treatment effects



(a) Pre-1996 CA within 2 km (linear trends)
Tab. 4.4, column (4)

(b) Matched pre-1996 CA (linear trends)
Tab. 4.4, column (5)



(c) Pre-1996 CA within 2 km (quadratic trends)
Tab. 4.4, column (9)

(d) Matched pre-1996 CA (quadratic trends)
Tab. 4.4, column (10)

Note: The solid lines are graphical illustrations of the parametric estimates presented in Table 4.4 on page 80 and estimated using equation (4.28). The dashed lines indicate the 95% CI which are calculated using standard errors of multiplicative interaction terms presented by Aiken et al. (1991).

We have argued that the model predictions for capitalization effects under a (locally) efficient designation policy and a spatial decay in heritage externalities hold at the conservation area boundary, i.e. we expect a zero effect just inside and a positive effect just outside the boundary. Figure 4.4 on page 81 illustrates the joint effect of the parametric estimates reported in Table 4.4 on page 80 at varying (internal and external) distances from the CA boundary. With the control group of historical CAs within 2 km of the treatment CA (left panels) we find a positive capitalization effect just inside and outside the boundary, which is in line with the baseline DD result in Table 4.2 on page 73, column (4). Moreover, the treatment effect increases toward the centre for the CA and decreases in external distance to the boundary until it becomes zero at around 700 m. This distance is in line with existing evidence on a relatively steep decay in heritage and housing externalities (Ahlfeldt & Maennig, 2010b; Lazrak et al., 2013; Rossi-Hansberg et al., 2010). However, the effect is statistically indistinguishable from zero at almost all distances. The single exception is a significant (at 5% level) 1.6% effect just outside the CA in the quadratic model. While the effect is only significant within 100 m of the CA, this is precisely where we expect a positive effect in a world with spatial decay in heritage (housing) externalities. In the context of the model the lower and not statistically significant effect just inside the CA indicates the presence of a cost that compensates for some of the benefits associated with designation.

With the control group of matched CAs (right panels) the treatment effect just inside the CA boundary is remarkably close to zero. The joint effect just outside the boundary is positive, although not statistically significant. Briefly summarized, the spatial RDD-DD model suggests that across the treated CAs owners - at least on average - are not harmed by designation. There is some evidence that owners just outside a conservation area receive some benefit.

4.6. Summary

Historic preservation policies are among the most restrictive planning policies used to overcome coordination problems in the housing market internationally. These policies aim at increasing social welfare at the cost of constraining individual property rights. From the perspective of owners of properties in conservation areas, the policy may help to solve a collective action problem, preventing owners from free riding on the heritage character of nearby buildings while inappropriately altering their own property. If property owners value the heritage character of nearby buildings and can influence the designation process they will seek out a (local) level of designation where the marginal costs of designation equate the marginal benefits. An increase in the marginal benefit of designation will lead to an increase in designation activity. If the policy is Pareto-efficient, additional designations in a neighbourhood will not lead to an adverse impact on those being designated.

We provide evidence that is supportive of this scenario using two empirical approaches that follow from a simple model of (locally) efficient conservation area designation. First, we present a neighbourhood level IV tobit analysis that reveals a positive impact of an increase in degree share, which is presumably (positively) correlated with heritage preferences, on the share of designated land. Gentrification, by increasing the value of neighbourhood stability to local owners, can cause designation. Second, we combine the strengths of difference-in-differences (DD) and regression discontinuity designs (RDD) to estimate the capitalization effect of designation on newly designated areas as well as spillovers to adjacent areas. This RDD-DD methodology qualifies more generally as a useful tool for program evaluations where a treatment is suspected to lead to an impact on (spatial or temporal) trends and discontinuities. Within newly designated conservation areas we find no significant short-run effects of designation and some evidence for positive capitalization effects in the long run. There is some evidence for positive spillovers onto properties just outside.

These results are in line with a Pareto-efficient designation policy, at least from the perspective of the local owners. Either, the policy is deliberately Pareto-maximizing local owner welfare or, as suggested in the literature on the political economy of housing markets, homeowners are able to successfully influence the outcome of local policies in their interest. In any case, it is important to note that our results do not imply that the policy is necessarily welfare-enhancing on a wider geographic scale. Depending on the excessiveness of the policy and the general restrictiveness of the planning system, historic preservation may constrain housing supply and generate welfare losses. The net-welfare effect to a wider housing market area is an interesting and important question that we leave to future research.

5. Chicken or egg? Transport and urban development in Berlin

5.1. Motivation

As reviewed in Section 2.5, infrastructure and especially mass transit play a major role in economics and especially in urban economics. Transport induces welfare effects due to increased accessibility and reduced shipping costs and impacts on the location of households and firms. Estimating the impact of transport infrastructure is, however, not entirely straightforward since the relation between transport and economic activity is plagued by a notorious simultaneity problem. While the view that an ease of access to other locations within a region should have a positive impact on the attractiveness of a location is unchallenged, it is also hard to believe that the allocation of transport is completely random. Due to the demand required to recover large investment costs, new infrastructure is more likely to connect economically successful places. Recent research has approached this problem by only focusing on the uni-directional effect of transport provision, building causal inference on quasi-experimental and instrumental variable strategies.

However, it is very difficult to address the simultaneity problem entirely with conventional models. It is challenging to assess to which extent the reverse causation is an empirically relevant phenomenon because instruments for this direction are particularly difficult to find. Moreover, it is *a priori* unclear whether the planner responds to economic activity broadly defined – and reflected in land values – or absolute demand from potential passengers – reflected in population. Effectively, the relation can be described by a complex structure of various endogenous variables whose mutual causal links are difficult to model with the standard tool sets of causal inference. Theory offers some guidance on the potential directions and temporal sequence of the dependencies, but ultimately fails in producing predictions that would allow us to describe the whole dynamic process by a causal relationship of X on Y. A method that is well established to explore the structure of mutually related endogenous variables is the structural vector auto regression (SVAR) model. Commonly used in macroeconomics where isolated shocks are hard to observe and models consist of many endogenous variables, VAR models have become a well-established method to describe complex dynamic relations. The method can be adopted to incorporate spatial information in a Panel VAR (PVAR), where the regions form a cross-sectional dimension. This method is proposed as an informative tool to go beyond the state of knowledge that has been achieved in the transport economics literature on the uni-directional effects of transport on land (or property) value.

Unfortunately, as with most applications of empirical methods, the benefits do not come without costs. Causality in a narrow sense is an issue, and while the PVAR strength is to model the simultaneous relationships of the most important variables conditional on each other, it is vulnerable to variables that belong to the (unknown) full model, but are omitted from the analysis. Also it is difficult to parameterize the strength of the (mutual) dependencies analysed. Acknowledging the strength and weaknesses, I believe that the novel application of the established PVAR method to a well-known estimation problem has the potential to complement existing uni-directional causal research by providing a better picture of the general dynamics.

In line with the macroeconomic approach, a dynamic study area is required characterised by a battery of shocks. These criteria can typically be found in massively growing developing countries and their metropolises. However, reliable data for a spatially fine level and a sufficiently long time period is difficult to find for these cities. I therefore look into the history of a by now developed city and employ a unique historic dataset for Berlin between 1881 and 1935 for the analysis. This period is characterised by great infrastructure projects as well as a strong growth in population, which enables me to empirically address the research question. I try to fill a gap of studies fully exploiting changes in the transport infrastructure, most probably due to a scarcity of projects, as noted by Gibbons & Machin (2008). In line with Baum-Snow et al. (2012) I argue that the analysis of public infrastructure during that time is of special interest as the city was industrialising, economically developing and characterised by huge transport investments somehow comparable to the process developing countries experience today. Moreover automobiles can largely be ignored as transport mode.

The analysis follows a twofold approach: Firstly, I estimate the interaction between population and transport. Secondly, the analysis is extended by looking into the relation between land values and transport. I separately estimate the empirical model for a core and a peripheral sample in order to allow for potential outbidding. The estimates suggest a simultaneous relation between transport and urban development. Planners, however, follow only economic activity but not residents. I find evidence for an outbidding of residents due to transport innovations. The results are proven robust to price anticipation effects.

The remainder of this chapter is structured as follows. The next Section 5.2 provides a historical overview of the Berlin transport sector. This is followed by an introduction of our historical dataset in Section 5.3 as well as of the PVAR methodology in Section 5.4. In Section 5.5, the empirical findings of the benchmark models as well as of additional robustness tests are reported and interpreted. Section 5.6 briefly summarizes and comments on the main findings.

5.2. Historical background

The public rail network in Berlin is made up of two different modes, namely a rapid transit system (“S-Bahn”) and the underground (“U-Bahn”). This section provides a brief overview of its historical development.

5.2.1. S-Bahn

Berlin’s rapid transit system as it is known today is a result of combining various suburban lines (“Vorortsbahn”), the original city line (“Stadtbahn”) and the circular line (“Ringbahn”) in 1930. Therefore, there are various reasons and purposes why, where and how the S-Bahn was developed over the years (Klünner, 1985; Gottwaldt, 1994; Kiebert, 2004, 2008).

The suburban lines connected Berlin with surrounding cities and its suburbs. Especially the early lines originate from long-distance connections to major cities like Potsdam (“Stammbahn” 1838), Hamburg (“Hamburger Bahn” 1846) or Dresden (“Dresdner Bahn” 1887). Initially, the long-distance lines had to share their tracks with the new upcoming suburban lines. In 1891 a new tariff system for local mass transit was introduced, pushing up the passenger numbers by about 30% and the suburban lines increasingly started to run on their own tracks. The majority of these lines were developed by public companies and planned by the government. For instance, the “Ostbahn”, which was intended to go through the Prussian regions of Pommern and East Prussia, was built in order to develop the periphery along the tracks. The “Görlizer Bahn” (1866/67) or the “Wetzlarer Bahn” (also “Canon Train”), linking Berlin with Metz at the French border, were planned by the military in order to move troops more rapidly. Later on, new lines were directly built for local mass transit in order to improve access from the periphery, like the North-South connection (1934-39). However, private developers like J.A.W. Carsten, who financed the station “Lichterfelde” (1868) in order to sell his newly established country estates in that area, intervened in the expansion of the S-Bahn network as well. The electronics company Siemens further financially supported the exploitation of the section between Fürstenbrunn and Siemensstadt (1905) in order to improve the commuting situation for its workers. Moreover, Spindler Brothers, who ran a laundry and drying factory in Köpenick at the Eastern border of the city, were strongly in favour of building a transport line between Schöneweide and Spindlersfelde (1891). Hence, the suburban lines were driven by both public and private interests.

The city line went from Stralau-Rummelsburg to Westkreuz, Halensee and was built in 1882. This East-West connection running through the historical city centre was planned to decongest the traffic between Berlin’s terminal stations. The tracks were mainly built on land owned by the government and the project was carried out publicly.

The first sections of the circular line Moabit-Gesundbrunnen-Potsdamer-Ringbahnhof and Moabit-Charlottenburg-(Westend)-Grunewald-Tempelhof were opened in 1881, and 1882 respectively. The circular line was financed by the state of Prussia but run by the Niederschlesisch-Märkische Eisenbahn, a public company owned by Prussia. The idea behind the circular line was to connect radian lines extending from the centre with each other and the important terminal stations. Various parts of the new line were built on undeveloped land and thus outside the city border. Or, as Elkins & Hofmeister (1988) state: "The actual position of the ring line was a compromise between the desire to maximise utilization by being as close as possible to the core of the city and the desire to minimize land-acquisition costs by avoiding areas of existing urban development" (p. 114).

Like the circular line, many other lines of the light rail system were extended into undeveloped areas, connecting Berlin with other villages. Only the East-West and North-South connections went through the city centre. New villages were founded close to the new lines, such as "Glienicke an der Nordbahn". Companies like AEG or Borsigwerke in Tegel built new factories in close proximity to the new stations (e.g. "Kremmener Bahn"). Even though a few rapid transit lines were developed at the request of the private sector, most of the lines were developed by the public sector. In the 1880s the majority of the long-distance lines, which were closely related to the rise of the suburban lines, were nationalised. However, most of the nationalised lines were still run independently. They had their own management as well as their own trains/coaches. From 1920 on, all lines were eventually nationalised under the "Reichseisenbahn".

5.2.2. U-Bahn

The underground was developed about a third of a century later than the rapid transit system. The first line was opened in 1902 and went from Stralauer Tor (later Warschauer Brücke) to Potsdamer Platz and then to Zoo. The first underground was constructed on elevated tracks since the Berlin government was afraid of damaging its newly installed drainage system. The project was pushed forward by the company "Siemens & Halske", which as far back as 1891 proposed a densely linked network, connecting the historic city centre to its surrounding municipalities. The new line was eventually developed by the "Hochbahngesellschaft" a company jointly founded by Siemens & Halske and Deutsche Bank as the main funder. While the line's Eastern section to Nollendorf Platz was built on viaducts, the city of Charlottenburg successfully negotiated the tracks to run under ground when passing though its territory. Not obscuring the view of the prominent church "Kaiser-Wilhelm-Gedächtniskirche" was one of Charlottenburg's reasons for the changed routing. In the West (Westend), the line was extended into an undeveloped area where Deutsche Bank owned land. Driven by financial speculations the bank was expecting rising land rents due to improved access. As a result of the newly established connection, Western Charlottenburg turned into an attractive business area. The extension of the first underground line leading into central Berlin was initially hampered by the tram operator "Große-Berliner-Straßenbahn" being afraid of losing its monopolistic role in that area. Despite the concerns, the line eventually went via Mohrenstraße and Spittelmarkt through the city centre (Gottwaldt, 1994).

Especially the municipalities in the South West had great interest in developing their undeveloped land. They competed for wealthy citizens by turning it into attractive residential areas. The underground played a crucial role in developing these areas. The city of

Schöneberg ("Schöneberger Linie" 1910) even planned and financed its own line between Nollendorfplatz and Hauptstraße (today Innsbrucker Platz) in order to develop its Western territory. As the "Hochbahngesellschaft" did not expect any profits from the new line it was completely planned as a public enterprise. The area through which Schöneberger Linie ran changed significantly. Individually designed stations were built at prominent squares. A similar approach was followed by the villages of Wilmersdorf and Dahlem. Newly planned country estates and academic institutes were supposed to benefit from improved access by constructing the "Wilmersdorf-Dahlemer U-Bahn" (1913). The line was divided into three sections regarding ownership: While the section between Wittenbergplatz and Nürnberger Platz belonged to the Hochbahngesellschaft, Nürnberger Platz-Breitenbachplatz was owned by the city of Wilmersdorf and Breitenbachpatz-Thielplatz by Domäne Dahlem. The line was extended to the lake "Krumme Lanke" in 1929. This extension was mainly financed by the land speculator and private developer Adolf Sommerfeld in order to connect his newly established residential quarters in Dahlem. Moreover, he wanted to improve the access to the surrounding woods, establishing them as recreational areas (Kurpjuweit & Meyer-Kronthaler, 2009).

In contrast to the S-Bahn network, Berlin's underground was intended to serve local mass transit from the beginning. The lines were extended into more central areas. Moreover, the network was developed later than the rapid transit system; the technology was superior, allowing for an underground system, and planners as well as investors had already gained experience by evaluating the effects of the S-Bahn. Anecdotal evidence suggests that the rise of the U-Bahn was mainly driven by the idea of developing new land in close proximity to the historical core (especially in the South West). Public as well as private planners competed for wealthy citizens and increasing land rents.

Summing up, even though the link between transport and land development is not completely clear when analysing the history of Berlin's transport system, the majority of the projects and newly constructed lines seemed to lead the development in an area and not the other way around.

5.3. Data

This section provides a brief overview of the historic data used in the main analysis. The data differ in time as well as in spatial coverage and level of detail. Since spatial aggregation/disaggregation might result in a loss of information or biased estimates (modifiable area unit problem, MAUP) I decided to work with distinct panels. I end up with three samples: (i) I estimate the interaction between transport and population at a municipality ("Ortsteile") level of 93 municipalities using data for every five years from 1870 to 1935 (14 time periods), (ii) the relation between transport and land values is estimated on a block level covering 2,481 blocks of an inner sample of historical land value maps between 1881 and 1914 (7 time periods), as well as on (iii) a grid square level (length of 150 m, 12,596 grid squares) making use of the full extent of the historical sources which are further explained in the next paragraphs (see also Table 5.1 on page 90).[1]

Table 5.1.: Twofold approach - three samples

Panel	Development indicator	Unit of analysis	Spatial coverage	Time period
1	Population	Municipality	Gross Berlin	1870-1935
2	Land values	Historical city block	Inner Müller sample	1881-1914
3	Land values	Grid (150 m)	Müller sample	1881-1914

Table 5.2.: Number of stations

Year	S-Bahn	U-Bahn	Total
1870	18	0	18
1875	31	0	31
1880	55	0	55
1885	64	0	64
1890	65	0	65
1895	88	0	88
1900	96	0	96
1905	103	15	118
1910	109	28	137
1915	109	46	155
1920	112	46	158
1925	113	60	173
1930	120	94	214
1935	127	94	221

To compute rail densities, the historic transport network of Berlin has been reconstructed for the entire observation period (1870-1935/36) in a geographic information system (GIS) environment (Schomacker, 2009; Mauruszat, 2010; Straschewski, 2011). Similar to Levinson (2008) and in line with the density measures used in the previous analyses, station densities for the combined network are computed using a kernel with a radius of 2 km (Silverman, 1986). By choosing this radius I follow Gibbons & Machin (2005) who estimate a distance of 2 km as the maximum distance people are willing to walk to the nearest station. By definition, the density measure reflects the degree of network concentration. It is high in areas which are characterised by a high number of stations in close distance, and low vice versa. The measure does not, however, incorporate the number of lines which run through a station, the travel time/speed and the frequency for which I do not have data for either. I compute a joint station density since the two rail systems cannot be regarded as substitutes for each other: The rapid

---

[1] To address any concerns regarding the modifiable area unit problem, the grid sample approach has additionally been carried out for a 300 m grid sample (Appendix Section C.2).

Figure 5.1.: Combined rail network in 1870, 1900 and 1935

Notes: The lines indicate the combined rail network, i.e. the rapid transit system and the underground lines that were added up to each respective year.

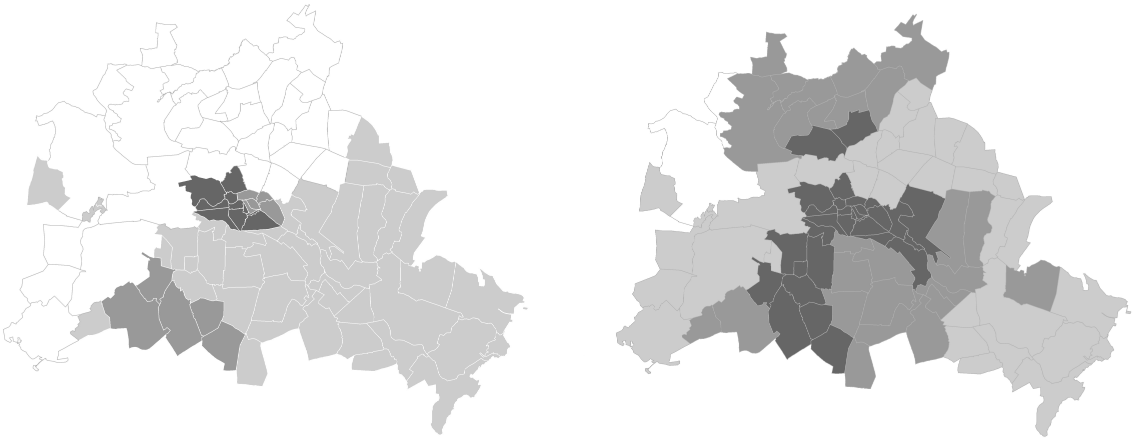
transit system was developed earlier and originates from inter-city connections. Conversely, the underground was built as public transport mode from the very beginning and runs through more central areas.

Figure 5.1 on page 91 and Figure 5.2 on page 92 provide an overview of the development of the total rail network during the observation period. In 1870, there existed only seven lines connecting surrounding cities with Berlin, whereby each line had its own terminus station. The lines did not reach the very centre of the city and only a few stations were built along the tracks. Turning to 1935, one can observe a massive development of the transport network. By that time, the circular line as well as the East-West and North-South connections were merged with the suburban lines. From 1902 onwards, underground tracks were added to the rail network. One can also observe the strong growth in accessibility of the South West up to 1935. The total number of stations increased from 18 in 1870 to 221 in 1935 (see also Table 5.2 on page 90).

The population data for Berlin refer to 93 municipalities ("Ortsteile") as they were structured in 1935. The data were collected from the Statistical Yearbook of Berlin (Statistisches Amt der Stadt Berlin, 1920) and Leyden (1933) and aggregated to the municipality level. The observation area covers the entire Greater Berlin. Between 1870 and 1935 population grew strongly from more than 900,000 up to about 4,200,000 inhabitants. Figure 5.3 on page 93 illustrates the population density over the observation period. Apart from a general increase for the whole of Berlin, a slight decentralisation pattern becomes visible.

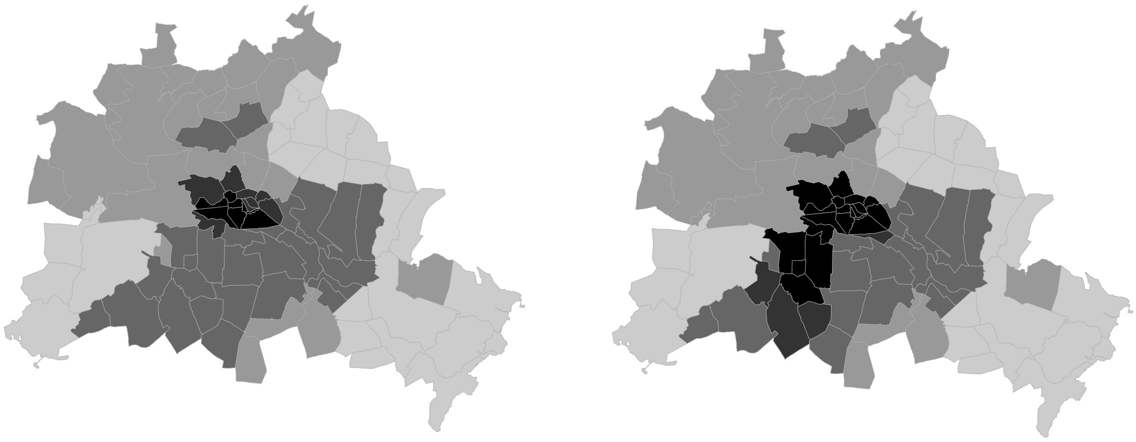
The technician Gustav Müller (1881-1914) produced maps of land values at the plot level for Berlin between 1881 and 1914. The historical maps were georeferenced in a GIS environment

Figure 5.2.: Rail station density on municipality level



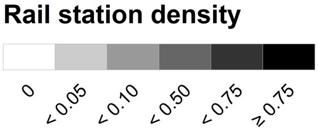
(a) 1870

(b) 1890



(c) 1910

(d) 1935



(e) Legend

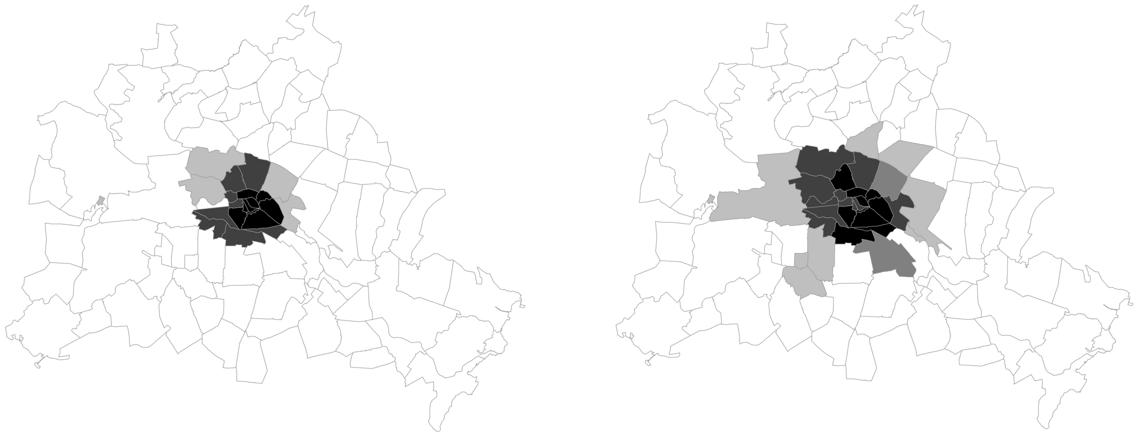
Notes: Station densities were computed using a kernel radius of 2 km and a quadratic kernel function (Silverman, 1986).

and land values were extracted for seven time periods at a plot level.[2] The values were aggregated at the 1910 block level for inner Berlin. Figure 5.4 on page 94 illustrates the development of the land values in Reichsmark between 1890 and 1914. I additionally aggregated the complete coverage of the historical maps at a grid square level, where each grid square has a length of 150 m, i.e. an area of 0.0225 square km (Figure 5.5 on page 96). There are several rather technical advantages of using grid squares. First of all, they do not imply a density bias. That means if for instance an area was made up of a high number of smaller blocks, this specific

---

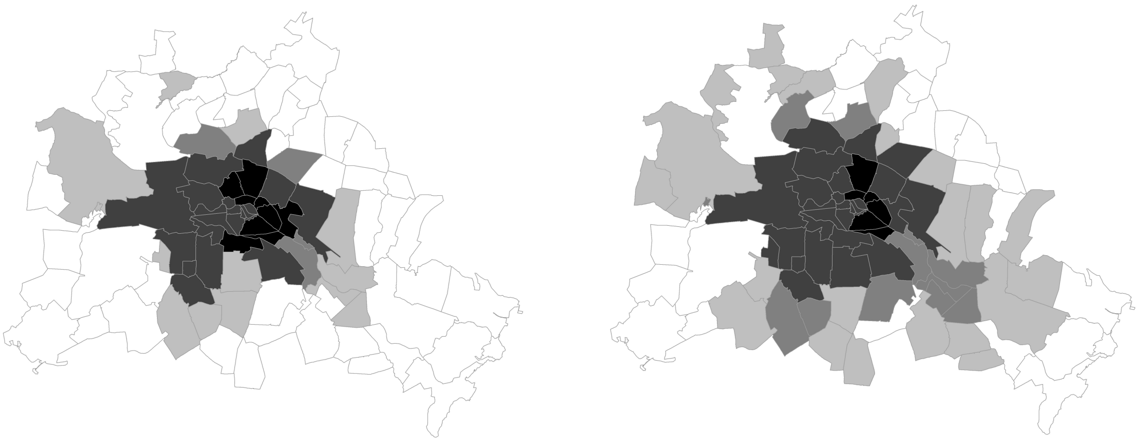
[2] The land value data were extracted for 1881, 1890, 1896, 1900, 1904, 1910 and 1914. Data for 1885 are unfortunately not available.

Figure 5.3.: Population density on municipality level



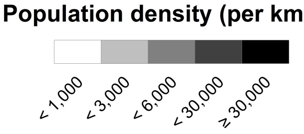
(a) 1870

(b) 1890



(c) 1910

(d) 1935



(e) Legend

Notes: Population are extracted from Statistisches Amt der Stadt Berlin (1920) and Leyden (1933).

area would be weighted more strongly in the econometric analysis due to the higher number of observations. Furthermore, the grid approach yields a balanced panel over space and time. However, one cannot deny the fact that grids provide only an abstract picture of reality. One also needs to pay attention to not generating too many artificial observations when creating the variables/distributing the available spatial information on the grid squares. The size of the grid squares is chosen to represent an average sized housing block. The Müller maps were used previously by Ahlfeldt & Wendland (2011); Ahlfeldt et al. (2011); Ahlfeldt & Wendland (2013).

Figure 5.4.: Land values on block level in 1890 and 1914

(a) 1890 (b) Legend

(c) 1914

Notes: Land values (in Reichsmark) were extracted from Müller (1881-1914).

In the analysis, I distinguish between a core and peripheral region to capture the presumably distinct behaviour of economic agents in the city. The idea is explained in greater detail in the next section. I follow the definition of “Berlin City” introduced by Leyden (1933) to determine the boundaries of the core.[3]

Tables 5.3, 5.4 and 5.5 provide the summary statistics for the different samples.

Table 5.3.: Municipality sample summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
railDens	overall	0.211	0.354	0	1.995	N=1302
	between		0.268	0	0.834	n=93
	within		0.233	-0.503	1.372	T=14

---

[3] Throughout this and throughout the next chapter the terms core and CBD are used synonymously.

population	overall	29743.77	56296.1	0	354.684	N=1302
	between		47686.4	0	216.328.3	n=93
	within		30298.21	-146107.2	221449.6	T=14

Table 5.4.: Block sample summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
railDens	overall	0.389	0.247	0	1.248	N=15034
	between		0.145	0	0.735	n=2486
	within		0.207	-0.209	1.096	T=6.047
land value	overall	159.523	239.052	0	2348	N=15034
	between		199.025	0	1502.143	n=2486
	within		117.688	-1043.33	1522.384	T=6.047

Table 5.5.: Grid (150m) sample summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
railDens	overall	0.219	0.251	0	1.769	N=88172
	between		0.193	0	0.856	n=12596
	within		0.160	-0.529	1.272	T=7
land value	overall	81.839	159.151	1	2180	N=52254
	between		118.221	1	1665.429	n=12596
	within		65.084	-1103.589	1335.422	T=4.15

5.4. Methodology

Panel vector autoregression is an adaption of the vector autoregression (VAR) methodology (Sims, 1980). A VAR model consists of a system of equations which are estimated simultaneously. Each variable in this system is explained by its own lags and lagged values of the other variables (Gravier-Rymaszewska, 2012). VARs have become very popular in applied empirical research mainly because they treat all variables as being endogenous and independent and allow for a simultaneous determination. Causality is allowed to run in any direction, from transport to urban development and from urban development to transport infrastructure (Drakos & Konstantinou, 2011). Since the observation period does not allow for any identification relying on the idea of natural experiments but is characterised by a battery of shocks - such as in a more macroeconomic set-up - the PVAR approach allows for setting up a model with only a small set of assumptions in order to interpret the impact of transport/development shocks.

Figure 5.5.: Land values on grid level (in Reichsmark) in 1890 and 1914

(a) 1890

(b) 1914

Notes: Land values (in Reichsmark) were extracted from Müller (1881-1914).

The panel VAR incorporates spatial information by using municipalities as a cross-sectional dimension. Exploiting rich panel data sets, PVARs extend the standard VARs model by allowing for unobserved individual heterogeneity. Even short panels improve asymptotic results as the sampling properties do not depend on the number of time-series observations but on cross-sectional observations (Gilchrist & Himmelberg, 1999).

PVARs are still more popular in macro- than in microeconomic research. They have for instance been applied in analyses of monetary policy and investment behaviour (Love & Zicchino, 2006; Assenmacher-Wesche & Gerlach, 2008), the supply of development aid (Osei et al., 2005; M'Amanja & Morrissey, 2006; Gillanders, 2011; Gravier-Rymaszewska, 2012) or security economics (Drakos & Konstantinou, 2011). An extensive overview of PVAR models used in macroeconomics and finance is provided by Canova & Ciccarelli (2013). In urban economics PVARs are not yet in common use. Exceptions are Miller & Peng (2006) who look, for example, into US housing price volatility as well as C.-I. Lee (2007), who investigates the question of whether the provision of public rental housing crowds out private investment. However, to the author's knowledge this is the first work applying PVAR techniques in an intra-urban context.

Building on Holtz-Eakin et al. (1988) and Canova & Ciccarelli (2004) the PVAR is specified as follows:

$$\begin{gathered} y_{i,t} = A_0 a_{i,t} + M_1 y_{i,t-1} + ... + M_p y_{i,t-p} + u_{i,t}, \\ (i = 1, ..., N; t = 1, ..., T) \end{gathered} \tag{5.1}$$

$$u_{i,t} = \mu_i + v_t + \varepsilon_{i,t} \tag{5.2}$$

where $y_{i,t}$ is a $\kappa x 1$ vector of κ panel data variables, the M's are $\kappa x \kappa$ coefficient matrices of the lagged variables $y_{i,t}$, p denotes the number of lags and $a_{i,t}$ is a vector of deterministic terms (linear trend, dummy or a constant) with the associated parameter matrix A_0. The unobserved individual effect μ_i, the time fixed effect v_t and the disturbance term $\varepsilon_{i,t}$ jointly compose the error process u_t. It is assumed that u_t has zero mean, i.e. $E(u_t) = 0$, independent u_t's and a time invariant covariance matrix. As the panels are rather short (due to the availability of the historic data) lag length p is a priori determined to be one.

I begin by estimating the reduced form VARs using system GMM (Arellano & Bover, 1995). Writing the reduced form as a system of single equations gives the following equations: For the municipality sample equations (5.3) and (5.4) and for the block/grid samples equations (5.5) and (5.6) where the error process is defined as above:

$$R_{i,t} = \delta_1 + \beta_{11} Pop_{i,t-1} + \beta_{12} R_{i,t-1} + u_{1i,t} \tag{5.3}$$

$$Pop_{i,t} = \delta_2 + \beta_{21} Pop_{i,t-1} + \beta_{22} R_{i,t-1} + u_{2i,t} \tag{5.4}$$

$$R_{j,t} = \vartheta_1 + \alpha_{11} LV_{j,t-1} + \alpha_{12} R_{j,t-1} + u_{1j,t} \tag{5.5}$$

$$LV_{j,t} = \vartheta_2 + \alpha_{21} LV_{j,t-1} + \alpha_{22} R_{j,t-1} + u_{2j,t} \tag{5.6}$$

with $R_{i,t}$ denoting rail station density in municipality i (statistical block/grid square j) at time t, Pop population and LV land values.

After having estimated the unknown parameters of the reduced form and computed the moving average representation of the VAR model (Wold decomposition), the impulse response functions (IRF) can be derived. The IRFs indicate how a variable reacts to a unit innovation in the disturbance term in period t holding all shocks constant. In particular, I am interested in the reaction of urban development to transport shocks and vice versa. The importance of that particular shock is determined by decomposing the variance. However, the IRFs do not allow for a structural interpretation since the models have not been identified. The variables only respond to the reduced form disturbances but not to the structural innovations. One therefore needs to isolate the shocks in the IRF by orthogonalising the residuals. The identifying restriction imposes a recursive ordering of causality (Choleski decomposition): The earlier a variable appears in the system the more exogenous it presumably is.

I adopt the identifying restriction (recursive order) by assuming that (i) transport is not affected by a contemporaneous population shock, whereas population is subject to contemporaneous rail shocks, and for the secondary set of panels that (ii) transport is not affected by a contemporaneous shock in land values, whereas land values responds to contemporaneous rail shocks. The idea is that new stations are subject to physical constraints/“time-to-build effects” (Love & Zicchino, 2006; Kilian, 2011) and do no instantly respond to population or land value shocks. There might also be information delays (Inoue et al., 2009) which impose a reaction time of one period in the system. In contrast, population and land values are assumed to react instantaneously to contemporaneous transport shocks. Assuming that markets incorporate all realizations of relevant outcomes (weak form efficient market hypothesis), prices in particular are expected to adjust immediately to transport shocks. Moreover, the construction of new lines is usually made public in advance so that residents/firms have time to adjust their location according to their preference for accessibility. A stricter interpretation of the (strong) efficient market hypothesis would even imply that prices adjust directly after the construction of a new line has been announced. Investors anticipate a future rise in land values and the values increase even before the station has officially been inaugurated. To control for anticipation effects I re-estimate the model using announced station density instead of actual station density.

The estimation model controls for individual heterogeneity by including individual fixed effects μ_i. Applying standard mean-differencing procedures, however, generates biased estimates. This is because the fixed effects are correlated with the regressors due to the auto-correlated dependent variables (Arellano & Bond, 1991; Arellano & Bover, 1995; Blundell & Bond, 1998). I therefore use forward-mean-differencing (also Helmert transformation) to eliminate the fixed effects. This way it is possible to remove the mean of all future observations available for each municipality/block-time pair.[4] The Helmert transformation preserves the orthogonality between the variables and their lags which is essential for the use of lags as instruments in the system GMM estimation (Arellano & Bover, 1995).[5] I time-demean the series by subtracting

---

[4] See also Gilchrist & Himmelberg (1999), p. 257 f. for details.

[5] I adopt the STATA routines `pvar` and `helm` by Inessa Love who developed the programs for her paper Love & Zicchino (2006). The original programs are available at http://go.worldbank.org/E96NEWM7L0.

the mean of each variable computed for each spatial unit-year pair. This final transformation controls for time fixed effects ϑ_t. The PVAR is estimated in logs. Since I look at the relative effects (log differences), it does not matter whether population enters in levels or densities.

The analysis of the impulse response functions requires an estimate of their confidence bands. Standard errors for the IRFs are obtained by Monte Carlo simulations: I generate random draws of the VAR coefficients and recalculate the impulse responses using the estimated covariance matrix of the errors and the estimated coefficients. Repeating this procedure several times allows for the generation of 5th and 95th percentiles of the distribution. These percentiles are then used as confidence intervals for the IRFs (Love & Zicchino, 2006).

The empirical analysis follows a twofold approach: Firstly, I look into the interaction between population and transport, secondly between land values and transport. However, this approach only explains how people/land values react to changes in the transport network. These results would hardly allow us to derive any adjustments in the actual land use pattern, i.e. it is not possible to draw any insights on how firms react and whether, for instance, an outbidding of residents occurs. Since I am unfortunately not in the possession of historic firm data, an alternative strategy is required. I distinguish between a core region and the periphery as defined by Leyden (1933) and estimate the model for the two spatial samples separately. Assuming that businesses strongly depend on agglomeration economies in order to use land productively, they will pay a higher price for land than residents would do. They are therefore expected to outbid residents in the core who face a trade-off between commuting and space consumption. Overall, residents value centrality not as much as firms do, use land less productively and are thus only able and willing to pay a lower price. Therefore the relationship between land rents and transport development indicates how intense land is actually used. This argument could also work in a reverse way (simultaneity).

Panel VAR estimations require stationary variables since non-stationary data lead to inconsistent estimates. There are a number of unit root tests, like Levin et al. (2002); Im et al. (2003) etc., which test for stationarity. However, most of these tests are designed for long macroeconomic panels whereas this dataset is a rather short microeconomic one (N large, T small). I therefore apply a modified Fisher type test which combines the test results of testing each panel individually for a unit-root based on a Philipps-Perron test. As suggested by Choi (2001) for a large number of cross-sectional observations, I use the modified version of the inverse χ^2 transformation in order to test the null hypothesis of all panels having a unit root. All aforementioned tests are first generation tests which assume cross-sectional independence, i.e. being independently and identically distributed across individuals. But particularly when working with spatial data this assumption is likely to be violated. Ignoring the problem of cross-sectional dependence could result in biased test statistics. I therefore additionally apply the second generation Pesaran (2007) test, which allows for some form of cross-sectional dependence.

Unit root test as well as regression results are reported and discussed in the next section.

5.5. Empirical results

5.5.1. Main results

The analysis begins with the unit root tests. The test results are given by Table 5.6 on page 100. I run the individual unit root tests of the modified Fisher type test as Philipps-Perron test (Phillips & Perron, 1988). The inverse χ^2 transformed test statistic (Choi, 2001) rejects the null hypothesis of all panels being non-stationary at a significance level of 1%. Hence, assuming cross-section independence the tests yield stationary series.

The Pesaran (2007) unit root test, to control for potential cross-section dependence, rejects the null hypothesis of the series being non-stationary, $I(1)$, for population and rail station density of the municipality-level dataset at a 1% level. Unfortunately, the land value samples are too short to perform a Pesaran unit root test. However, according to Sarafidis & Robertson (2009) the bias caused by potential cross-section dependence can be reduced when the series are time-demeaned prior to the estimation. I therefore expect the variables of the city block and grid datasets to sufficiently fulfil the stationarity requirements, too, as the PVAR is estimated using time-demeaned and forward-mean-differed series.

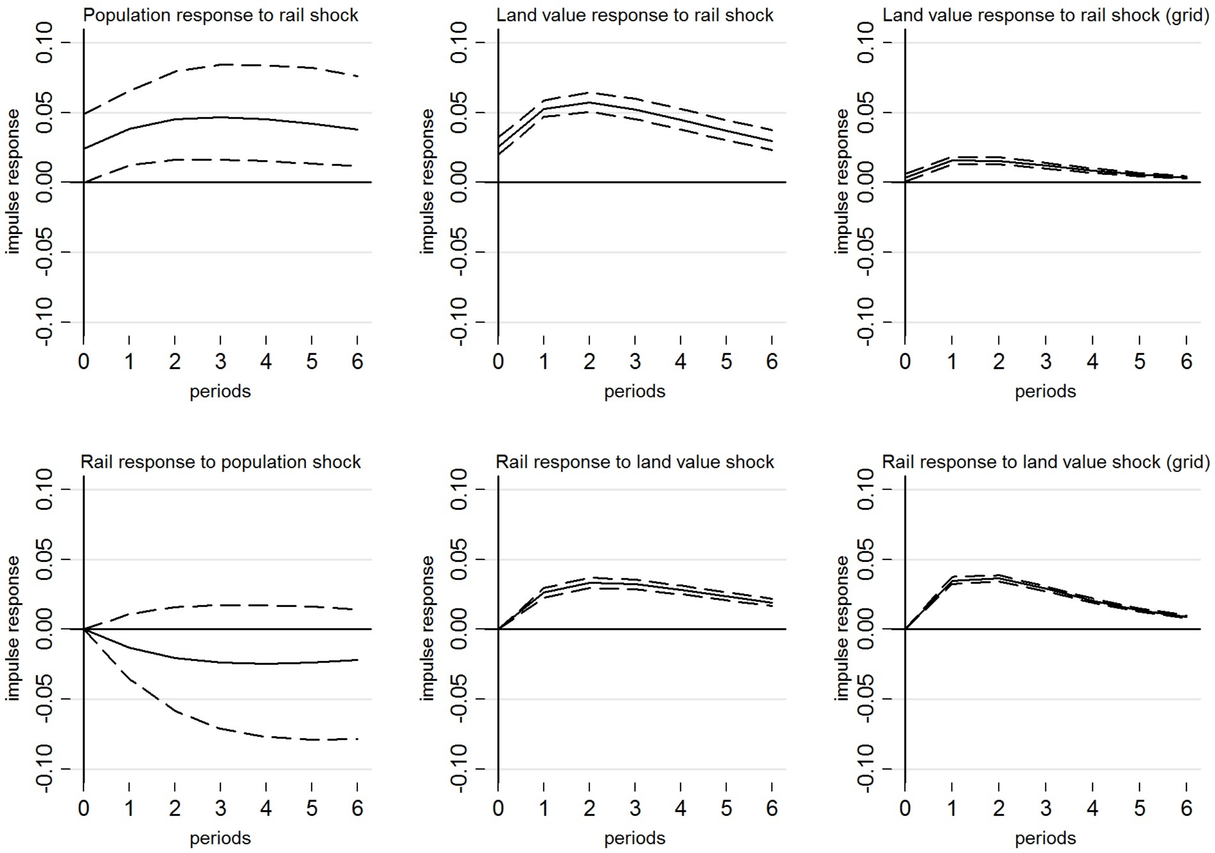
Table 5.6.: Panel unit root tests (separate samples)

Panel	Variable		Phillips-Perron (Choi 2001)	Pesaran (2007)
1	pop	test statistic	3.099***	-4.552***
		p-value	0.001	0.000
	railDens	test statistic	87.988***	-4.646***
		p-value	0.000	0.000
2	LV	test statistic	197.555***	–
		p-value	0.000	–
	railDens	test statistic	183.268***	–
		p-value	0.000	–
3	LV	test statistic	–	–
		p-value	–	–
	railDens	test statistic	114.779***	–
		p-value	0.000	–

Notes: (1) Variables shown are logarithmised, time-demeaned and Helmert transformed, (2) standard error in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Interpretation is restricted to the identified impulse response functions. The reduced form estimates are reported in the appendix (Section C.1). The IRFs are illustrated jointly with 5% confidence bands generated by Monte Carlo simulations in Figure 5.6 on page 101 for the total sample, Figure 5.7 on page 102 for the core, and Figure 5.8 on page 103 for the periphery subsample respectively. The top three graphs in each figure show the often investigated response of economic activity to transport shocks which represents the rather conventional supply-side driven perspective. Each column stands for a different panel, the municipality panel on the left hand side, followed by the block and grid panel (on the right hand side). The bottom graphs illustrate the reverse transport response.

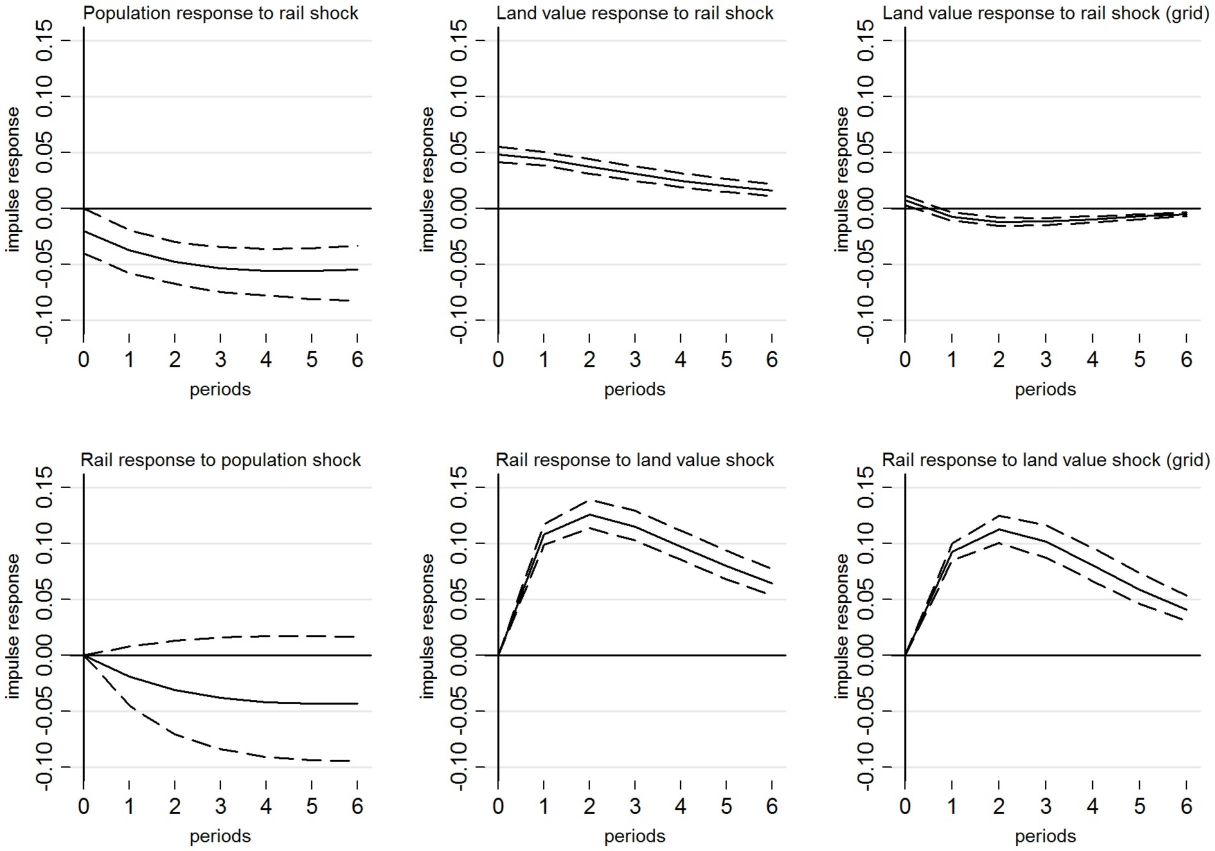
Figure 5.6.: Impulse responses for 2-PVAR model and total sample



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

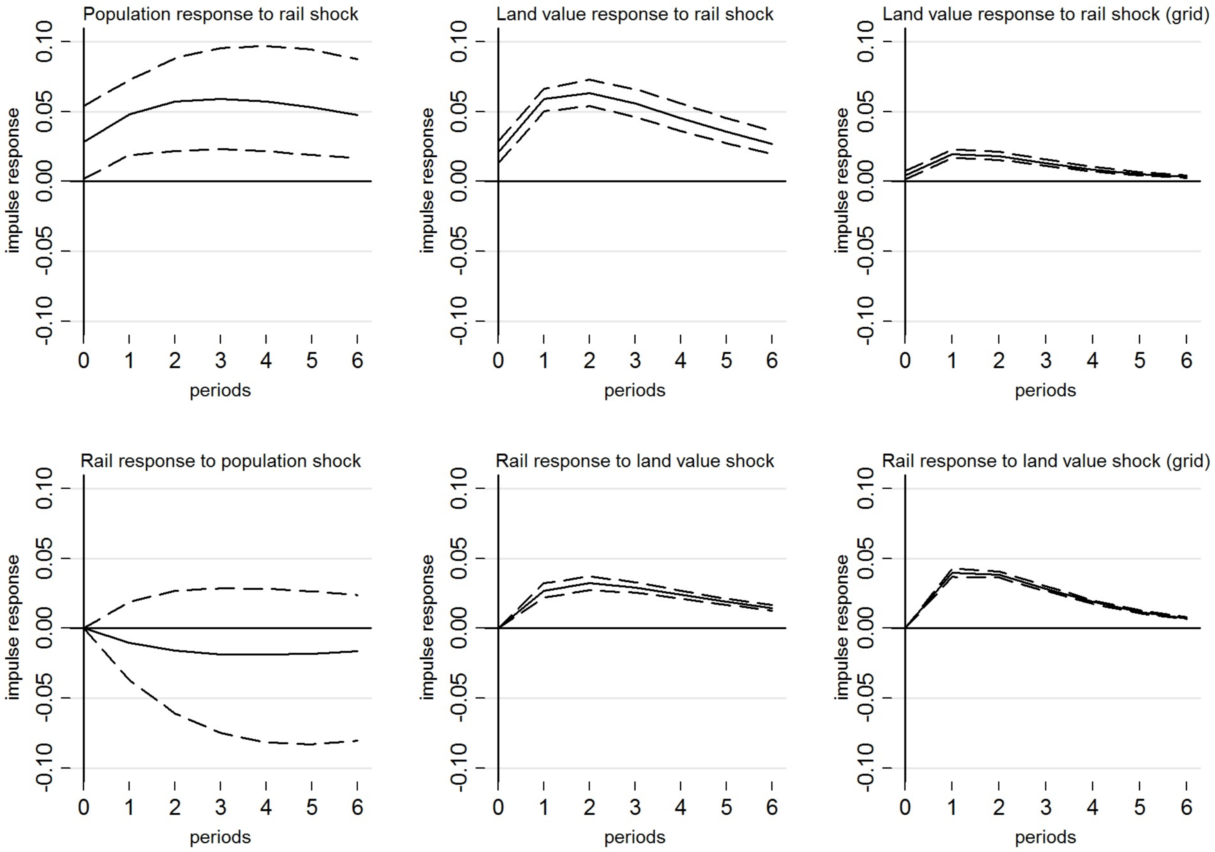
Beginning the analysis with the reaction to transport innovations for the total sample (Figure 5.6), one can observe a positive population response (top left), as well as a positive land value response to transport shocks (for both, the block and the grid sample). A one standard deviation increase in rail station density raises population relatively constantly over time by about 0.04 standard deviations in each period, and land values by about 0.05 respectively in the initial periods. The cumulative response over six periods in standard deviations is about 0.3 for the population response and between 0.08 and 0.3 for the land value response (see also the cumulative IRF which are presented in Section C.3 in the appendix). The land value response declines more quickly than the population one and is generally stronger for the block sample compared to the grid sample (top right). This might be due to the fact that the historic block samples only covers an inner sample of the city with a higher share of central areas and is thus slightly biased by centrality. Generally, the response to transport improvement is very much in line with expectations and the literature. Moving on to the reverse relation and the question whether transport also follows economic activity, the bottom left IRF in Figure 5.6

Figure 5.7.: Impulse responses for 2-PVAR model and core sample



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Figure 5.8.: Impulse responses for 2-PVAR model and periphery sample



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

indicates a negative but insignificant reaction. Conversely, new infrastructure is attracted by high-valued areas as transport reacts positively to land values (bottom middle and left). The estimates for the block and the grid sample are, moreover, very similar. Assuming that land is used efficiently, a positive response to land value and a negative response to population can be interpreted as (an unconditional) positive transport reaction to commercial activity or firms. Based on this interpretation, planners therefore do not follow residents but only firms.

To further exploit this this idea, the Berlin sample is split into two geographic subsamples for the core and the periphery. Population now responds negatively in the core (top left in Figure 5.7) but positively in the periphery (Figure 5.8). This can be interpreted as the outbidding of residents in central areas by firms as a results of a land use intensification triggered by the transport shock. Land value response is positive for the block sample, with the effect being immediate in the core and slightly lagged in the periphery. Land value response is also positive for the grid sample estimates but only in the periphery. In the core, transport has a negative effect on land values which is significant from period one onwards but small in magnitudes. As for the total sample, transport planners do not follow residents, neither in the core nor in the periphery. They are, however, attracted by economic activity expressed by high land values. This effect is particularly strong in the core, where a one stand deviation increase in land values raises station density by up to 0.12 stand deviations (second period) which accumulates up to 0.5 standard deviations over all six periods.

Overall, there is indeed evidence for a simultaneously determined relationship between urban development and transport. In fact, transport planners only follow businesses but they do not follow residents. An isolated focus on the supply-side driven relation of how transport affects urban development hence ignores important effects. Moreover, once can observe an outbidding of residents by businesses in response to transport shocks which is in line with traditional models.

In a next step, the variance decompositions (Table 5.7 on page 105) try to explain what shocks are most important in explaining a variable through time. Population hardly explains variation in rail station density two periods (about 10 years) ahead in the total sample, whereas transport explains about 0.6% of the population variation. Looking at the subsamples, explanation power is much higher in the core for both models. Here, station density explains about 5.3% of the population variation and about 0.8% in the periphery. The variance decomposition yields significantly smaller explanation power for the reverse relation. Overall the relationship between transport development and population is much stronger in the CBD than in the periphery.

Table 5.7.: Variance decompositions

Panel	Municipality			Block			Grid	
		pop	railDens		LV	railDens	LV	railDens
Total	pop	0.994	0.006	LV	0.976	0.024	0.997	0.003
	railDens	0.000	1.000	railDens	0.007	0.993	0.008	0.992
Core	pop	0.947	0.053	LV	0.914	0.086	0.997	0.003
	railDens	0.004	0.996	railDens	0.106	0.894	0.025	0.975

Periphery	pop	0.992	0.008	LV	0.978	0.022	0.996	0.004
	railDens	0.000	1.000	railDens	0.008	0.992	0.012	0.988

Notes: Percent of variation in the row variable explained by column variable (2 periods ahead).

The variance decomposition of the city block sample generally yields higher variations in terms of explanation power (also the number of observations is greater). Transport explains about 2.4% of variation in land values two periods ahead whereas it is only 0.7% for the reverse relation. Explanation power is significantly higher in the core again. In central areas almost 10.6% of transport variation is explained, conversely the explanation power is 8.6%. Interestingly, by explanation power, the relation is rather demand-side driven in the core and rather supply-side driven in the periphery. Finally, looking at the variance decomposition of the grid level data, variation is smaller in magnitudes but generally similar to the block sample. The explanation power is also higher in the core again. Overall, the variance decompositions reflect the simultaneity between transport and development.

5.5.2. Robustness

As stated earlier, one might be worried about price anticipation effects. Information might enter the market in advance and prices adjust directly after the construction of a new line has been announced. Investors anticipate a future rise in land values and the values increase even before the station has been officially inaugurated. To control for anticipation effects, I re-estimate the land value PVAR model using announced station density instead of actual station density.

The anticipation approach requires a careful collection of all announcement dates for each station constructed during the observation period. I use the first construction year for stations for which I do not find any information on the announcement date.[6]

I re-run the unit root tests for the updated land value sample. Like the other variables, announced station density, *AnRailDens*, is generally stationary when being time-demeaned and Helmert transformed (see Table 5.8 on page 106). However, only the Phillips-Perron test consistently rejects the null hypothesis of having a unit root for all panels, but not the Pesaran test. The following analysis, however, ignores any potential problems resulting from cross-dependence and treats the series as being stationary, relying only on the Phillips-Perron test.

Table 5.8.: Panel unit root tests (anticiptation model)

Panel	Variable		Phillips-Perron (Choi 2001)	Pesaran (2007)
1	AnRailDens	test statistic	41.515***	-4.104***
		p-value	0.000	1.000
2	AnRailDens	test statistic	183.278***	–
		p-value	0.000	–

---

[6] Research on the announcement dates of new stations is based on: Dudczak & Dudczak (2012); Kurpjuweit & Meyer-Kronthaler (2009); Mauruszat (2010); Loop (2012); Luisenstädtischer Bildungsverein e.V. (2012); Senatsverwaltung für Stadtentwicklung und Umwelt (2012); Straschewski (2011).

3	AnRailDens	test statistic	354.749***	–
		p-value	0.000	–

Notes: (1) Variables shown are logarithmised, time-demeaned and Helmert transformed, (2) standard error in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The impulse response functions of the anticipation PVAR models are shown in Figure 5.9 on page 107 for the total sample, Figure 5.10 on page 108 for the core and Figure 5.11 on page 109 for the periphery, the reduced VAR estimates are reported in the appendix. Using announced station density instead of the actual construction year does not severely change the overall findings. Most IRFs are in line with the ones reported earlier. The following interpretation therefore only highlights major differences.

Residential response to shock is only significant in the spatial subsamples (top left graphs). Residents negatively respond to the announcement of rail stations in the core and positively in the periphery. The IRF of the periphery sample moreover suggest that residents do not anticipate transport improvements. In contrast to the main models, announced station density, i.e. forward-looking planners, do now significantly respond to shocks in population. Transport reacts positively in the total as well as in the peripheral sample and negatively in the core one (bottom left graphs). This supports the tendencies presumed in the main models that planners follow economic activity in the core and, based on the magnitude of the shocks, less strongly follow residents in the periphery.

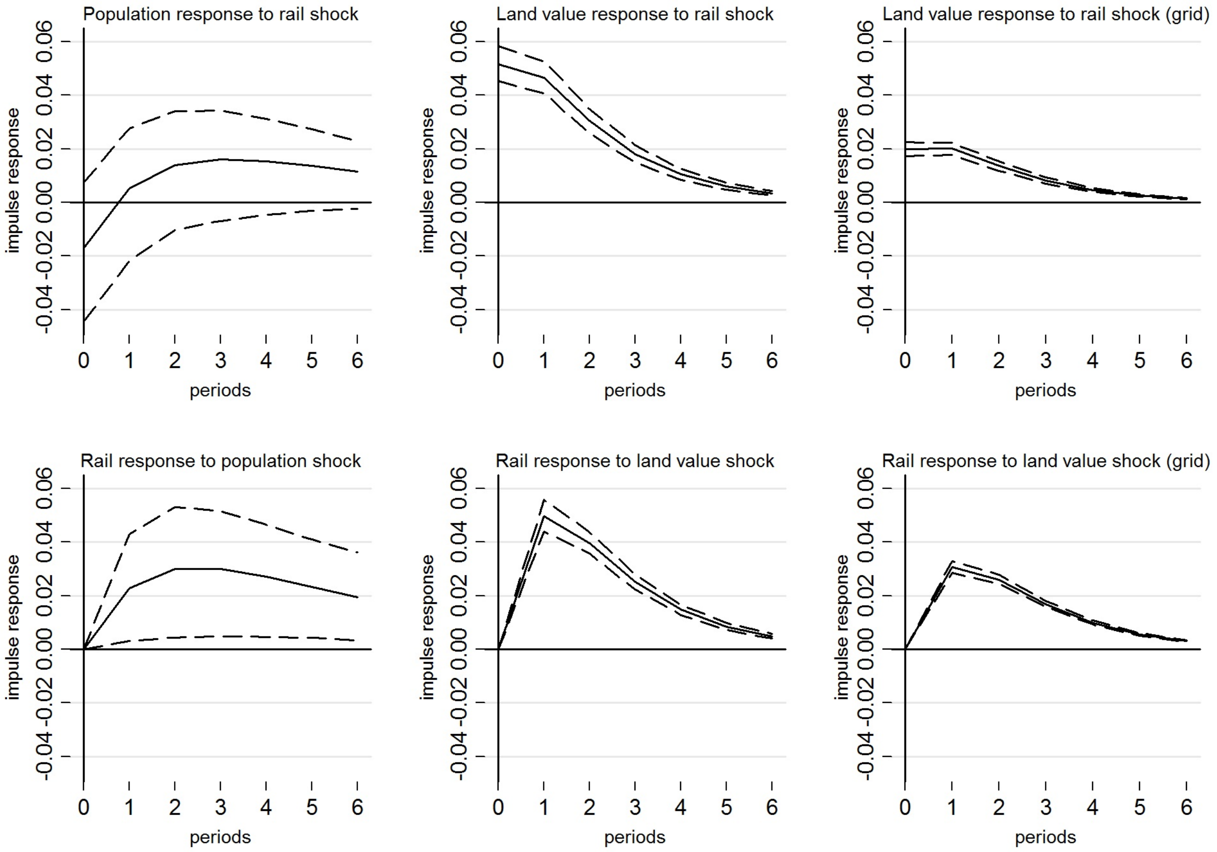
Block samples estimates of the anticipation model are very similar to the main model estimates except in core areas where announced station density's response to land values is insignificant for the anticipation model (bottom middle graph in Figure 5.10 on page 108). Estimating the anticipation model for the grid sample yields comparable results, too. In contrast to the main grid model, announced station density responds, however, negatively to land value shocks (bottom right graph in Figure 5.10 on page 108).

It is important to note that the idea of this robustness exercise is to control for price anticipation effects and thus motivated from a supply side driven perspective. Is there any price or population reaction due to the announcement of constructing new stations? The reverse relation of how station announcement responds to changes in land values/population is, however, less intuitive and mainly reported for completeness.

Overall, the anticipation models are mostly in line with our main models, rejecting any worries that the results are driven by anticipation effects.

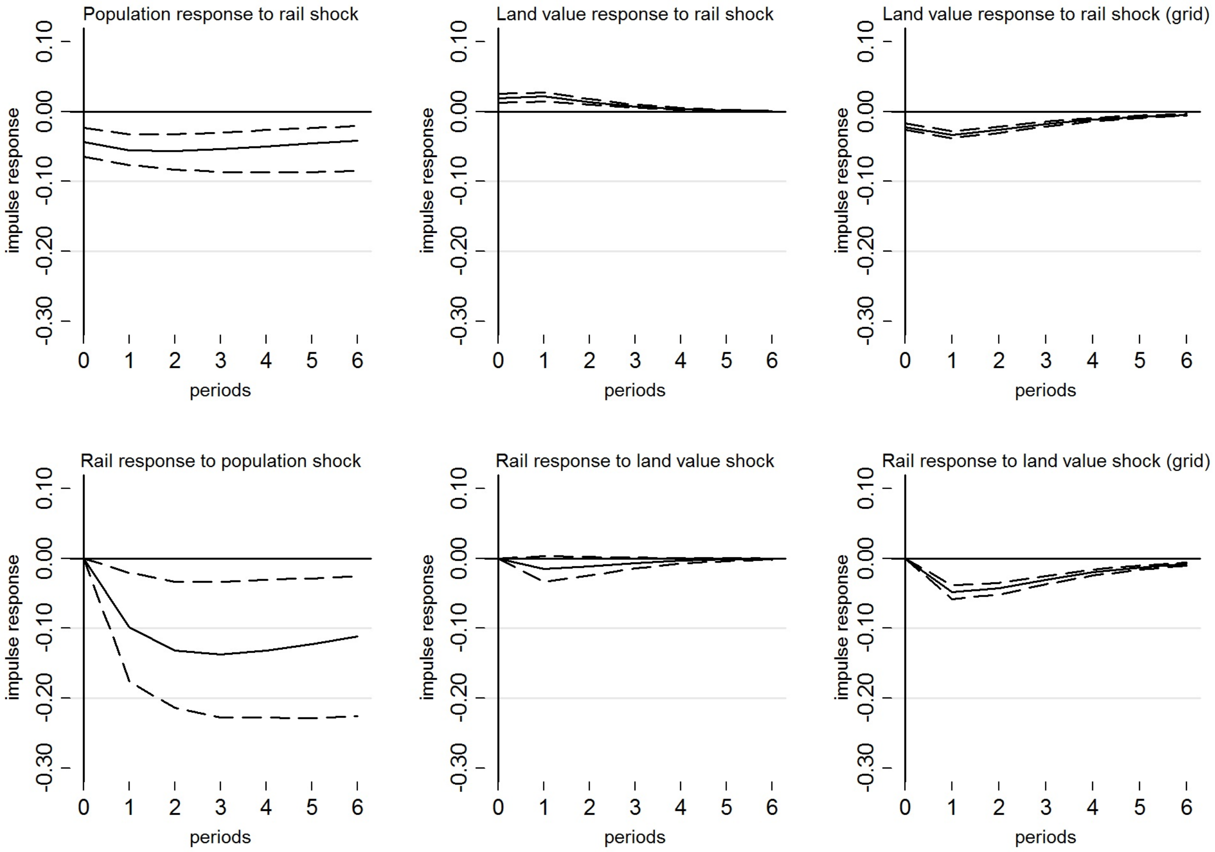
The variance decomposition of the price anticipation models is also mainly in line with the standard model (Table 5.9 on page 110). However, transport explains population in the core, as well as land values in all spatial grid samples comparably well two periods ahead, indicating that anticipation is at work. Nonetheless anticipation effects are not strong enough to drive our results.

Figure 5.9.: Impulse responses for 2-PVAR anticipation model and total sample



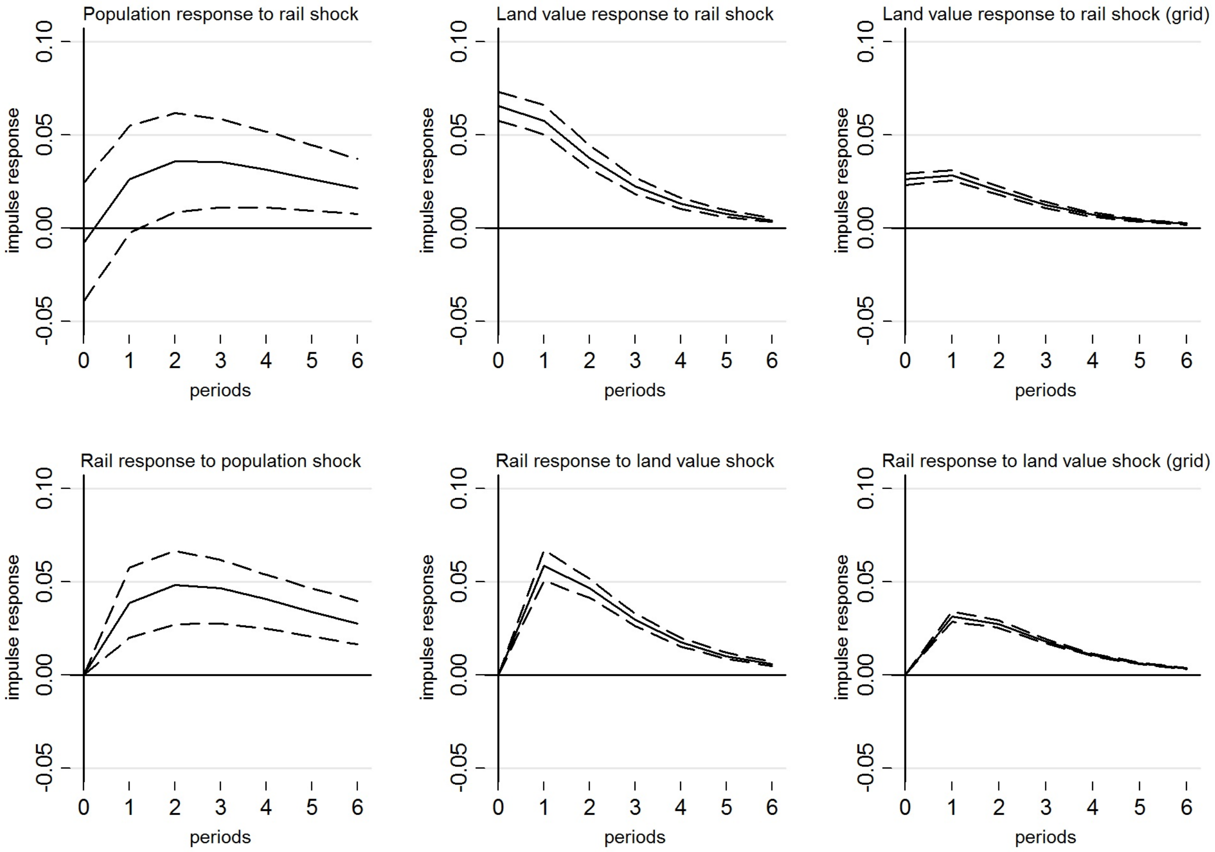
Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Figure 5.10.: Impulse responses for 2-PVAR anticipation model and core sample



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Figure 5.11.: Impulse responses for 2-PVAR anticipation model and periphery sample



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

Table 5.9.: Variance decompositions (anticipation model)

Panel		Municipality			Block		Grid	
		pop	railDens		LV	railDens	LV	railDens
Total	pop	0.999	0.001	LV	0.966	0.034	0.992	0.008
	railDens	0.002	0.998	railDens	0.009	0.991	0.006	0.994
Core	pop	0.865	0.135	LV	0.982	0.018	0.952	0.048
	railDens	0.019	0.981	railDens	0.001	0.999	0.009	0.991
Periphery	pop	0.998	0.002	LV	0.957	0.043	0.985	0.015
	railDens	0.009	0.991	railDens	0.012	0.098	0.007	0.993

Notes: Percent of variation in the row variable explained by column variable (2 periods ahead).

To address any multiple areal unit problem (MAUP) concerns, the raw land value and transport data have alternatively been aggregated to grid squares with a length of 300 m instead of 150 m. The reduced form estimates and the IRF are shown in the appendix (Section C.2) and reject any MAUP concerns.

I finally control for different land use types to make up for the lack of firm data in interpreting adjustments in the land use pattern due to transport innovations. Land use data and results are explained and interpreted in greater detail in Section C.4. The distinction by land use is in line with the outbidding of residents.

5.6. Summary

This chapter was intended to provide a new and purely empirical perspective on the chicken-and-egg-problem of transport economics. It was argued that despite conventional models' power in establishing causality, they lack of a demand side driven and hence a more complete, simultaneous perspective. Most analyses focus only on the uni-directional supply side driven effect of transport on an adjustment of prices and people. The reverse relation from urban development/economic activity to the attraction of new transport links is, however, highly understudied. Therefore, a method which is well suited to explore the structure of mutually related endogenous variables was proposed. In particular, I run a panel VAR analysis using municipality, block and grid level data for land values, population and rail stations. The models were estimated for historic Berlin between 1881 and 1935, a dynamic period characterised by a battery of shocks.

I find empirical evidence that the relation between urban development and transport is not uni-directional but highly simultaneous. In fact, transport planners follow economic activity in core areas but tend to ignore residential development. In line with traditional models, the estimates suggest an outbidding of residents by firms due to transport improvements.

6. Hundred years of transport in Chicago – a Panel VAR analysis

6.1. Motivation

The previous chapter highlighted the important role of transport amenities for urban development. Moreover, it was said that there is a lack of studies which incorporate the simultaneous determination of transport facilities and economic activity. In this chapter I extend the previously carried out empirical analysis of the Berlin rail sector by applying a similar methodology for the city of Chicago, Illinois. The motivation for this extension is two-fold: First of all, it can be interpreted as an additional robustness test. Similar results for Chicago would back-up the Berlin findings and could further support the applicability of the earlier Panel VAR approach. Secondly, extending the empirical investigation to Chicago results in a unified comparative analysis of this question which provides significant benefits over individual case studies. Previous studies into the question have all employed different methodologies and the results may differ for this reason and not due to the fundamental processes at hand. A consistent approach would ensure that differences in results reflect real differences in situations.

In particular, I look into the interaction between Chicago's 'L' (=Elevated) train and population for a period of 100 years, ranging from 1910 to 2010. Applying a similar PVAR approach as in the previous chapter, also distinguishing between a core and a periphery, yields results similar to those found for Berlin. Again, rail developers do not follow people but presumingly economic activity. Moreover residents are driven out of core areas and move to the periphery (outbidding).

The chapter proceeds as follows: Section 6.2 provides a brief introduction to the history of Chicago's famous 'L'. Section 6.3 introduces the data which will be used for the analysis carried out in Section 6.4. Finally, Section 6.5 sums up the insights gained from the analysis and compares them to the results described in the previous chapter.

6.2. Historic background

Chicago was officially founded in 1833.[1] Only four years later it was granted city rights and the City of Chicago was incorporated. The former village near Fort Dearborn quickly became

---

[1] For a more comprehensive summary of the history of Chicago see Pierce (1937, 1940); Andreas (1975); Grossman et al. (2004); Emmett (2012) and Cudahy (1982); Young (1998); Borzo (2007); Chicago"L".org (2013a) for the evolution of the 'L' in particular.

the fastest growing city in the world. Population grew from about 4,000 in the 1830's up to almost 1.1 million people in 1890 (Borzo, 2007). This massive growth was accompanied by a rapid expansion of the city border, stressing the need for mass transportation. Horse-drawn omnibuses were introduced in 1850, horse-drawn streetcars followed in 1859. Many attempts were made to introduce a (heavy) rapid transit line, with the first attempt failing in 1869. It took an additional 23 years until the first rapid transit line opened in 1892 (Chicago"L".org, 2013a). This was the birth of Chicago's 'L', which got its nickname from the all-elevated beginnings of the transport system (Chicago Transit Authority, 2013).

The first line (1892) was constructed and operated by the Chicago and South Side Rapid Transit Railroad Company. The so-called South Side 'L' went from a terminal at Congress Street to 39th Street. The route followed a straight line of 3.6 miles and ran through a completely city-owned alley. The Chicago citizens quickly adopted to the new means of intra-urban travel and public demand required an extension of the new line. The tracks were extended to Jackson Park, where the World's Colombian Exposition was held during the summer of 1893. Additional branches were opened subsequently: The Englewood Elevated Railroad Company made the Englewood neighboorhood accessible in 1905, the Normal Park Branch was built to serve a growing real estate development (1907), similar to the Kenwood Branch (1907) which was characterised by a massive residential development between 1905 and 1915, when the old housing stock was replaced by new apartment buildings. Anecdotal evidence suggests that early lines were constructed to respond to residential development, often in conjunction with great public events like the World's Colombian Exposition. Commuter demand is said to be responsible for the Stock Yard Branch which opened in 1908. The line went to "Packingtown" to transport "the vast quantity of workers to and from their south side homes" (Chicago"L".org, 2013a). The aforementioned branches jointly made up the South Side Elevated.

The Lake Street Elevated opened for service in 1893. It initially went from 52nd Avenue to Market & Madison. The franchise was owned by Michael C. McDonald ("King Mike"), whose main motive for developing new branches was to make a fortune via land speculations. His 'L' line served many factories along the track and became an important transport mode for workers to commute to their jobs. The Metropolitan Westside Elevated (1895) was the first branch serving a steadily growing population in the west. It was moreover the first 'L' to use electric traction technology. The franchise initially connected Franklin Street Terminal with Garfield Park, Douglas Park and Logan Square, splitting into three lines at Marshfield Avenue. The line was extended to Oak Park and to Garfield Park in 1913.

The early 'L' lines had one drawback: They did not share a joint terminal in the CBD but all ended outside the city centre.[2] The Union Loop (or just Loop) was supposed to solve this problem by connecting the early three 'L' lines. The first part of the Loop opened in 1895 (Lake Street 'L') and the first full circuit journey around the CBD could be made in 1897. It took about 25 minutes to circle through the Union Loop's twelve stations. Initially, companies along the planned route needed to be persuaded since they were worried that rail

---

[2] This is similar to the origins of Berlin's rapid transit system, which was a result of terminus stations connecting Berlin with other major cities.

disamenities could have a negative effect on their business. They were quickly proven wrong and ridership numbers rose rapidly together with their market potential. The Loop provided new opportunities to residents; the CBD was suddenly made accessible for commuting or shopping trips.

The fourth ‘L’, the North Western Elevated, started operations in 1900. After initial problems the line connected Wilson Avenue with the Loop. By 1907 the growing Ravenswood neighbourhood was linked to the system by a new branch line. Ridership numbers went up to 10,000 per day already within the first two months. The original four ‘L’s radiated from the city centre via the Loop into surrounding residential neighbourhoods, industrial as well as farm land being the backbone of Chicago’s ‘L’. But the system got further extended and in 1919 it was possible to go from downtown Chicago to Milwaukee.

Different attempts were made over the years to unify the various ‘L’ companies. The consortium Chicago Surface Lines (1913) was replaced by the Chicago Rapid Transit Company (CRT) in 1924, which completely consolidated all ‘L’ lines. The CRT experienced the ‘L”s historic peak in terms of rail service in the early 1930s, with 5,306 scheduled trains serving 627,157 passengers per day. The network consisted of 227 stations and a total track length of 227.49 miles (363.98 km). However, as far back as the 1940’s Chicago’s ‘L’ faced numerous difficulties. The CRT was no longer profitable as a private company and became public in 1945. The newly formed Chicago Transit Authority (CTA) took over business. The Elevates Lines’ biggest problem was the increasing competition from the automobile. The CRT failed to integrate the former private (sometimes parallelly running) lines, to close down old traffic routes or to convert them to bus operations, to adjust routes with respect to traffic-usage and to set fares which would allow the generation of funds for reinvestments and necessary modifications. The new CTA’s focus was on the repayment of debts and interests, which made important re-investments difficult. However, the CTA was facing the problem that 50% of its revenue had to originate from fare boxes. Another problem rose from the construction of the Congress Expressway (later Eisenhower Expressway) which was the first expressway within city borders and which provided a new and fast way of commuting into the city centre by car. The new organisation mainly responded to the rising problems by closing down stations and lines. By 1960, about one fourth of the rapid transit system had been abandoned.

In 1964, a new project, the Skokie Swift, was supposed to revitalise the ‘L’ and deal with the increasing car competition. The demonstration project was jointly carried out by the federal government and the mass transit agency. The Skokie Swift’s main intention was to complement car usage. The new line was built in proximity to the Edens Expressway, space was reserved for Park’n’Ride as well as Kiss’n’Ride and rail stations were connected by buses. The Skokie line was, moreover the fastest rapid transit in the world by the time. The new line was a success since it responded to the demand for good rapid transit service in the suburbs. This “prototype” was emulated by the Dan Ryan Line (1969), built in the centre of the new Dan Ryan Expressway, connecting the South Side with the West Side of the city as well as the suburbs of Oak Park. The line started on the 95th and joined the Loop at 18th Street. It facilitated the commuting of low-income households from the city centre to the suburbs (reverse commuting). The Kennedy Expressway (1970) was an extension of the Milwaukee Line to Jefferson Park. In 1984, the lines got further extended to O'Hare Airport. This was the first extension into new territory for more than fifty years. Against the planners' expectations not only airport workers but also airline passengers were using the new service. The new rapid transit lead urban development and new office parks and retail outlets located in proximity to the line.[3] Despite the adjustments, ridership significantly declined over the 1970s and CTA failed to cover operating costs. The number of commuters strongly decreased due to the outmigration of people and jobs. In 1973, the Regional Transit Authority was created as overseer and included representatives of the surrounding counties. However, the CTA continued its downward trend during the 1980s. Several raises in fares which were intended to cover the 'L"s operating costs drove away more riders. In the 1990s the CTA tried to reinvent itself. Changes like the reconfiguration of the North-South and West-South Routes (1993) as well as the introduction of color instead of name codes were supposed to make the service more attractive to customers. The Orange line was built between the Loop and Midway Airport based on the good O'Hare experience. The new line was inaugurated in 1993. In 1997, a private consulting company proposed further cuts via changes in services. After its implementation, CTA was eventually able to cover its costs again.

Ridership numbers eventually increased again over the last years. In the fourth quarter of 2012, an average number of 728,800 riders were taking the 'L'. The CTA manages America's third biggest heavy rail systems in terms of ridership. Only the Washington Metro Area TA (901,300) and the MTA New York City Transit counted more passengers (8,373,100) (Dickens, 2013). Today's 'L' system consists of 144 stations which are distributed over a network of about 242.2 miles (387.52 km) of track. There are eight rapid transit routes which run elevated above ground, in tubes and subway tunnels as well as in expressway medians (Chicago Transit Authority, 2013).

Overall, anecdotal evidence suggests that Chicago rail developers had similar motives to those of their counterparts in Berlin. New rapid transit lines were constructed to connect residential areas, to provide commuting opportunities to industrial areas and for reasons of financial speculation. In contrast to the period investigated in Berlin, this analysis covers long periods of decline where the 'L' was facing strong competition with the automobile and its expressways. However, as with Berlin, the historic review does not uncover the link between transport and urban development and further empirical investigations will therefore be carried out in the subsequent sections.

6.3. Data

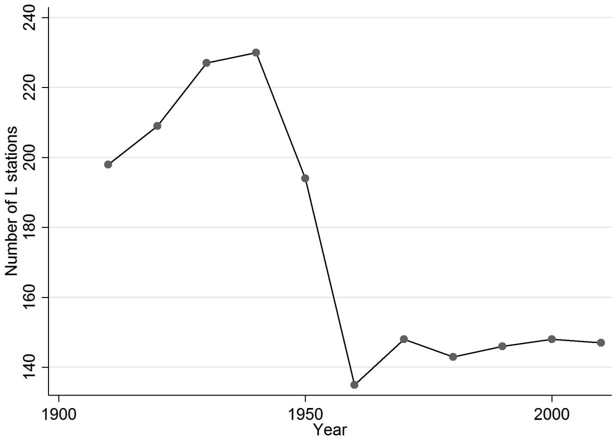
The empirical analysis of the interaction between transport and development is based on US census tract data for Chicago. The tract definition of 1990 is used as the baseline geography. It has the highest number of individual census tracts (i.e. 1,352) and covers the largest geographical

---

[3] O'Hare Airport's role as a new and important employment centre for Chicago has also been investigated and empirically confirmed by McMillen (1996).

area facilitating the (dis-)aggregation of the data for the other years. Overall, the panel covers eleven time periods (every ten years) between 1910 and 2010. Transport data for the ‘L’ network are provided by the City of Chicago (Chicago Transit Authority, 2012) and by Chicago“L”.org (2013b). The evolution of the number of stations is illustrated by Figure 6.1 on page 115. One can clearly see the rapid transit system’s peak in 1930/1940, as mentioned in the history review. The number of stations drastically declines from 230 (1940) down to 135 in 1960. During the last fifty years the number of ‘L’ stations has been roughly stable at around 150 stops.

Figure 6.1.: Number ‘L’ stations



As in the transport analysis of Berlin (Chapter 5), I compute station densities for the ‘L’ stations using a kernel with a radius of 2 km (Silverman, 1986) reflecting the maximal walking distance (Gibbons & Machin, 2005). Figure 6.2 on page 120 provides an overview of the development of the rail network during the observation period. The four years are chosen to reflect the beginning and the end of the period as well as it’s bust and boom periods in terms of the number of stations. One can clearly observe the network’s extension until 1940 followed by it’s decline and stabilisation process.

Population data come from the National Historical Geographic Information System database (Minnesota Population Center, 2011, NHGIS) and are depicted for a selection of years in Figure 6.3 on page 121. Between 1910 and 2010 the number of inhabitants of the study area rises from 2.2 to almost 5.2 million. The figure clearly describes a strong decentralisation pattern, where the historic city centre becomes depopulated over the years. Population density declines in central and rises in peripheral areas. Moreover, there seems to be a tendency of residents to move north.

In the analysis I distinguish between the CBD as defined by City of Chicago (2010) and the periphery to capture the presumably distinct behaviour of economic agents in the city.[4] An overview of the study area as well as the definition of the CBD is given in the appendix (Figure D.1 on page 184). Summary statistics of the panel data are shown in Table 6.1 on page 116. The next section briefly reviews the empirical approach and discusses its results.

---

[4] As in the previous chapter, core and CBD are used synonymously.

Table 6.1.: Census tract summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
L Dens	overall	0.321	0.777	0	10.676	N=14872
	between		0.718	0	8.791	n=1352
	within		0.298	-1.883	3.991	T=11
population	overall	3233.668	2708.459	0	21726.000	N=14872
	between		1665.356	27.450	12567.520	n=1352
	within		2136.400	-6366.936	20861.150	T=11

Notes: *LDens* denotes the density of “L” stations.

6.4. Empirical analysis

6.4.1. Empirical approach

The subsequent analysis follows the empirical strategy as outlined in Section 5.4 and builds on the estimation of a Panel VAR as defined by equation (5.1). Translating the reduced form VAR into a system of single equations with a lag length *a priori* determined to be one yields:

$$L_{i,t} = \delta_1 + \beta_{11} Pop_{i,t-1} + \beta_{12} L_{i,t-1} + \mu_i + v_t + \epsilon_{1i,t} \tag{6.1}$$

$$Pop_{i,t} = \delta_2 + \beta_{21} Pop_{i,t-1} + \beta_{22} L_{i,t-1} + \mu_i + v_t + \epsilon_{2i,t} \tag{6.2}$$

with $L_{i,t}$ denoting Elevated line station density in census tract i at time t and Pop denoting population.

I control for individual fixed effects (μ_i) via Helmert transforming the data as well as for period effects (v_t), which are removed by time demeaning. Since all variables are in logs and identification comes from variation over time a transformation into density becomes empirically obsolete. The estimation of the reduced form VAR is followed by a Wold decomposition to derive the impulse response functions (IRF). For a structural interpretations a recursive correlation scheme (Cholesky decomposition) needs to be assumed. Identification builds on the assumption that transport is not affected by a contemporaneous population shock due to physical constraints (time-to-build) but only responds with a one period lag. Population, however, is subject to contemporaneous rail shocks and allowed to respond immediately.

The analysis is initially carried our for the total sample and then separately for the CBD as well as for the periphery. The idea is to disentangle residential and firm behaviour in the absence of explicit firm data. This approach has also been followed by Levinson (2008); Xie & Levinson (2010) as well as in Chapter 5. Because the estimations require the series to be stationary, a modified Fisher type unit root test based on a Philipps-Perron test (Choi, 2001; Phillips & Perron, 1988) precedes the PVAR estimations. Test as well as estimation results are presented in the next section.

6.4.2. Results

The results of the unit root test are shown in Table 6.2 on page 117. Population data fulfil the stationary requirements and the inverse χ^2 transformed test statistic rejects the null hypothesis of all panels being non-stationary at a 1% significance level. Unfortunately, the 'L' system covers only a smaller part of the city (see again Figure D.1 on page 184) such that the majority of the census tracts is characterised by an excess of zeros. This makes a unit root testing of the 'L' density variable virtually impossible. However, to get a rough indication for the presence of stationarity, I restrict the unit root test for the transport variable for a sample which has a density of at least 0.14, i.e. 25% of the density distribution. This subsample estimate rejects the null hypothesis of all panels containing a unit root at a 1% level. In contrast to the analysis of Berlin, I cannot control for cross-sectional dependence. I therefore need to rely on the assumption that the time-demeaning and forward-mean-differentiation sufficiently eliminates any unit root.

Table 6.2.: Panel unit root tests

Variable		Phillips-Perron (Choi 2001)
pop	test statistic	286.336***
	p-value	0.000
L Dens	test statistic	9.8762***
	p-value	0.000

Notes: (1) Variables shown are logarithmised, time-demeaned and Helmert transformed, (2) *LDens* testing is reduced to a subsample which at least has a density value that is equivalent to 25% of the density distribution due to an excess of zeros, (3) standard error in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

The interpretation of the empirical results is restricted to the impulse response functions and the variance decomposition. Reduced form estimates are moved to the appendix (Table D.1 on page 183). Figure 6.4 on page 122 illustrates the IRFs with 5% confidence bands generated by Monte Carlo simulations.[5] The impulse responses for the total sample are given by the graphs on the left-hand side. Population positively follows transport shocks over time (top left) whilst transport planners respond negatively to population shocks (bottom left). However, the estimated rail IRFs are statistically insignificant as shown by the confidence intervals. These initial results suggest a clearly supply-side driven relation between transport and development, where only population responds to transport improvements. This view is strengthened by the cumulated impulse response which is about 0.23 standard deviations for the population response and only -0.026 for the transport sector (cumulative IRFs are presented in Section D.3 in the appendix).

---

[5] The IRFs shown in the main text share a common Y-axis to make the responses comparable. Alternatively see Figure D.2 on page 185 where the IRFs have individual Y-axes.

Moving on to the geographic subsamples, impulse responses of ‘L’ station density to population shocks in the core (bottom middle) as well as in the periphery (bottom right) turn out to be insignificant. There are, however, significant population reactions. Residents seem to get driven out of the CBD due to transport improvements. A one standard deviation increase in station density leads to a population decline of about 0.1 standard deviations in the initial periods. The effect declines over time (note that a period denotes one decade, i.e. 10 years) and seems to converge back to zero (top middle). Cumulated over all six periods, the population response is -0.6 standard deviations (see Figure D.4 on page 186). In the periphery, population responds positively to shocks in the ‘L’ system. A one standard deviation increase in rail station density raises population by 0.06 standard deviations in the second period (top right, accumulated response of 0.3 standard deviations). Overall, people seem to follow transport but transport does not responds to residents. Moreover, the negative population response in the CBD and positive one in the periphery suggest an outbidding of residents by firms. The variance decomposition in Table 6.3 on page 118 confirms the earlier presumed weak demand-side driven link between transport and population: Population does not explain ‘L’ station density variation two periods ahead for any of the samples. Conversely, transport has the strongest explanation power in the periphery, explaining about 2.5% of the future variation while it is only 0.2% in the core. Overall, the variance decomposition yields relatively weak links between the location of residents and of the ‘L’ network. This might be due to the fact that the ‘L’ system only covers a smaller part of Chicago, being the potential transport mode of only a minority of residents.

Table 6.3.: Variance decompositions

		pop	L Dens
Total	pop	0.992	0.008
	L Dens	0.000	1.000
Core	pop	0.997	0.003
	L Dens	0.000	1.000
Periphery	pop	0.975	0.025
	L Dens	0.000	1.000

Notes: Percent of variation in the row variable explained by column variable (2 periods ahead).

To conclude, the interaction between the Elevated network and population development in Chicago is not simultaneous but supply side driven. People follow transport but planners extend/reduce the rapid transit network independent of people. In line with urban economic literature transport innovations lead to a decentralisation of residential activity, most probably because residents are outbid by firms in the CBD (land use densification).

6.5. Summary

The second chapter studying transport was motivated by the idea of extending the Panel VAR analysis previously carried out on Berlin to a second city. Applying the same methodology, but for the city of Chicago, Illinois, the investigation not only provides an additional robustness test but also allows for a comparison of the two cities.

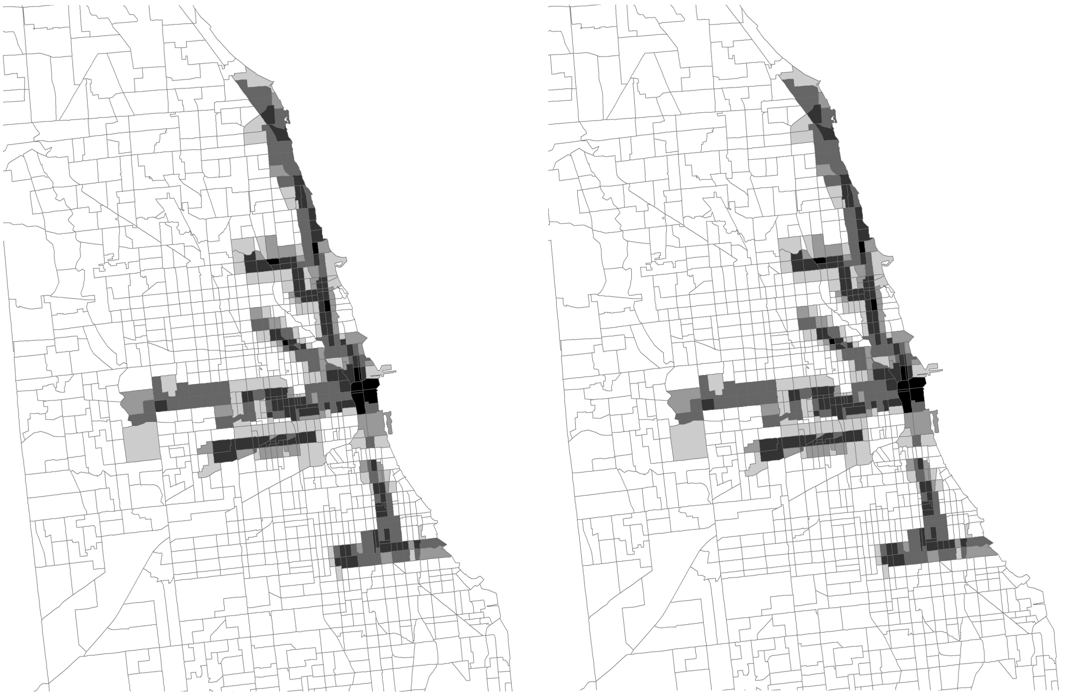
First of all, the estimates for Chicago are mainly in line with the results for Berlin. Both cities are characterised by a presumed outbidding of residents by firms due to transport improvements: Population declines by 0.1 standard deviations in Chicago and by about 0.06 in Berlin following a one standard deviation increase in rail station density in the core. In the periphery, the transport shocks lead to increases in population which are comparable between the two cities in terms of their magnitudes (both between 0.05 and 0.06 standard deviations). The similar reactions are quite astonishing when taking into account the different historical development of the (public) transport system in the two cities. While the rapid transit as well as the underground system have always played an important role for residents in Berlin and moreover been crucial for the development of the city, Chicago's 'L' faced severe competition from intra-urban expressways and the automobile resulting in a decline of the system. Moreover the study periods are quite different; the investigation of Berlin is restricted to the early years of massive transit until 1935 while the analysis of the US city makes use of a temporally larger sample. Despite all these facts, the long-run relation between population and transport describes comparable patterns.

Only restricting the analysis to population, transport planners in Berlin as well as in Chicago do not responds to urban development and do not follow residents. The provision of transport is rather supply-side driven with people responding to new commuting opportunities. The inclusion of land values in the study of Berlin, however, yields positive transport responses. Chapter 5 therefore concludes that planners only follow economic activity in core areas but tend to ignore residential development. Based on the comparable interaction of transport and population in the two cities, which moreover originates from a consistent investigation approach, one might infer that transport planners in Chicago also only follow firms but ignore residential needs.

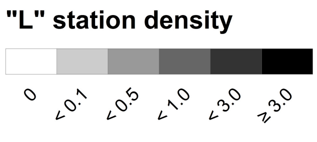
The two transport studies carried out in this work both provide comparable results for the spatial allocation of transport, households and indirectly firms/commercial activity. By applying a consistent empirical investigation approach to another study area, the work leaves the domain of individual case studies and moves towards a unified comparative analysis. Implementing a consistent method across a large number of cities would eventually help to provide a categorical answer to transport questions and would facilitate the derivation of more general policy implications.

Figure 6.2.: 'L' station density on census tracts

(a) 1910 (b) 1940



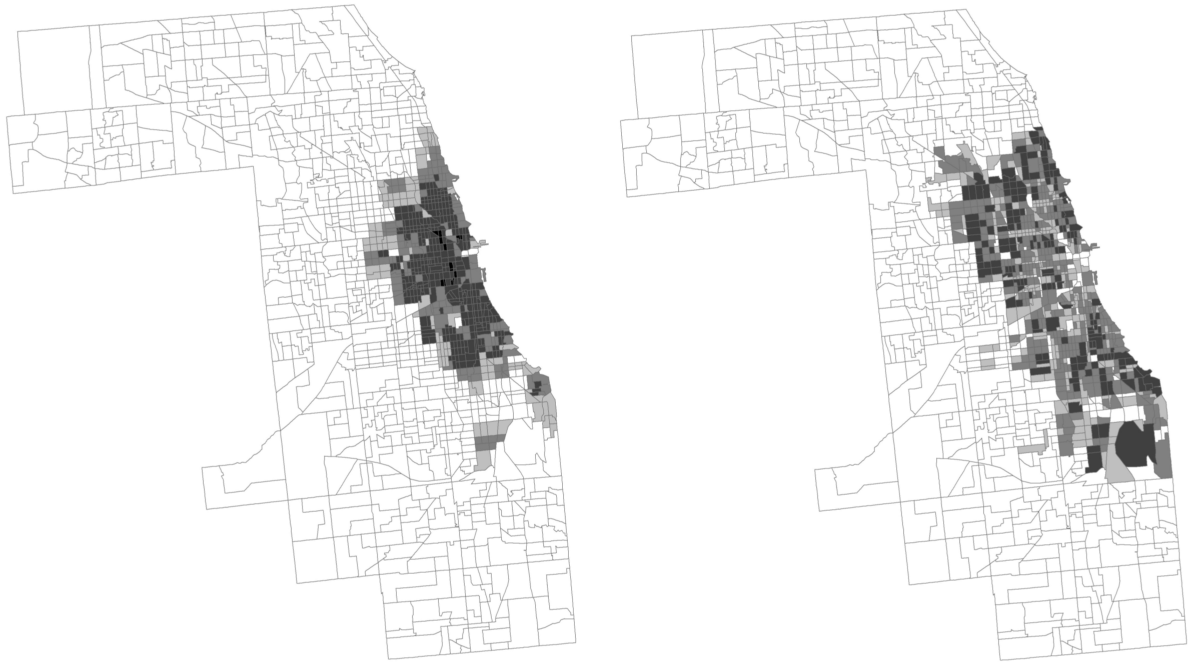
(c) 1960 (d) 2010



(e) Legend

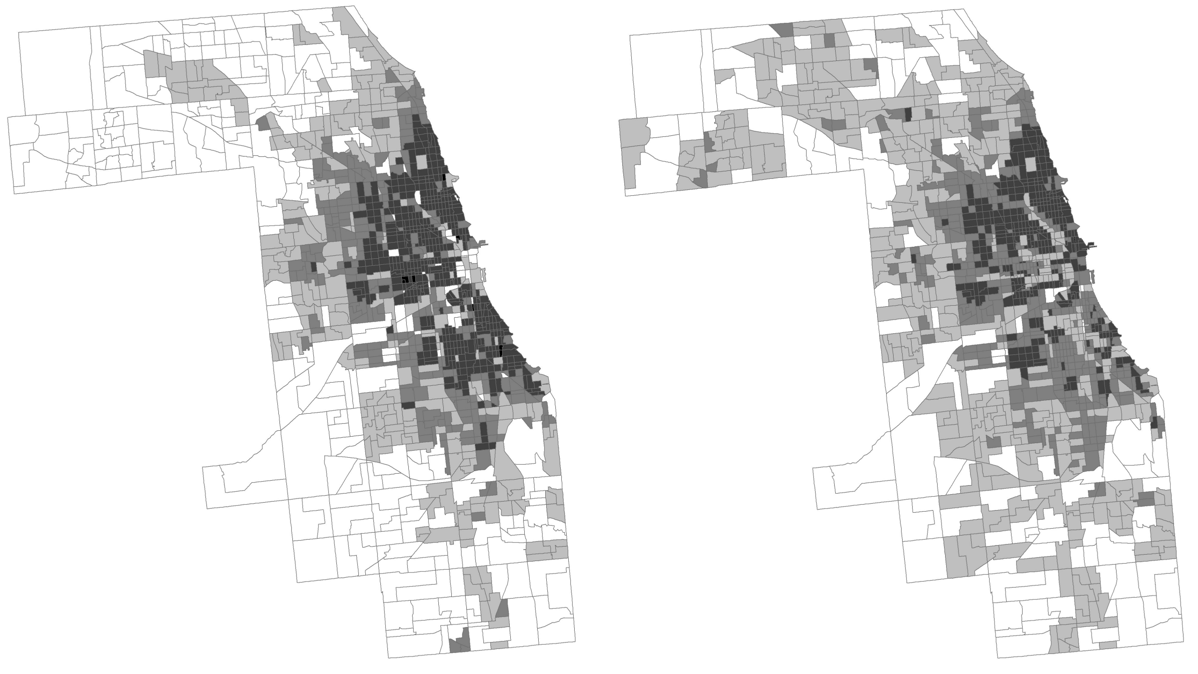
Notes: Transport data come from Chicago Transit Authority (2012) and Chicago"L".org (2013b).

Figure 6.3.: Population density on census tracts



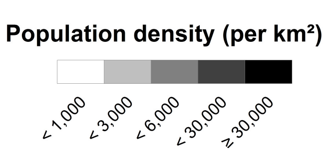
(a) 1910

(b) 1940



(c) 1960

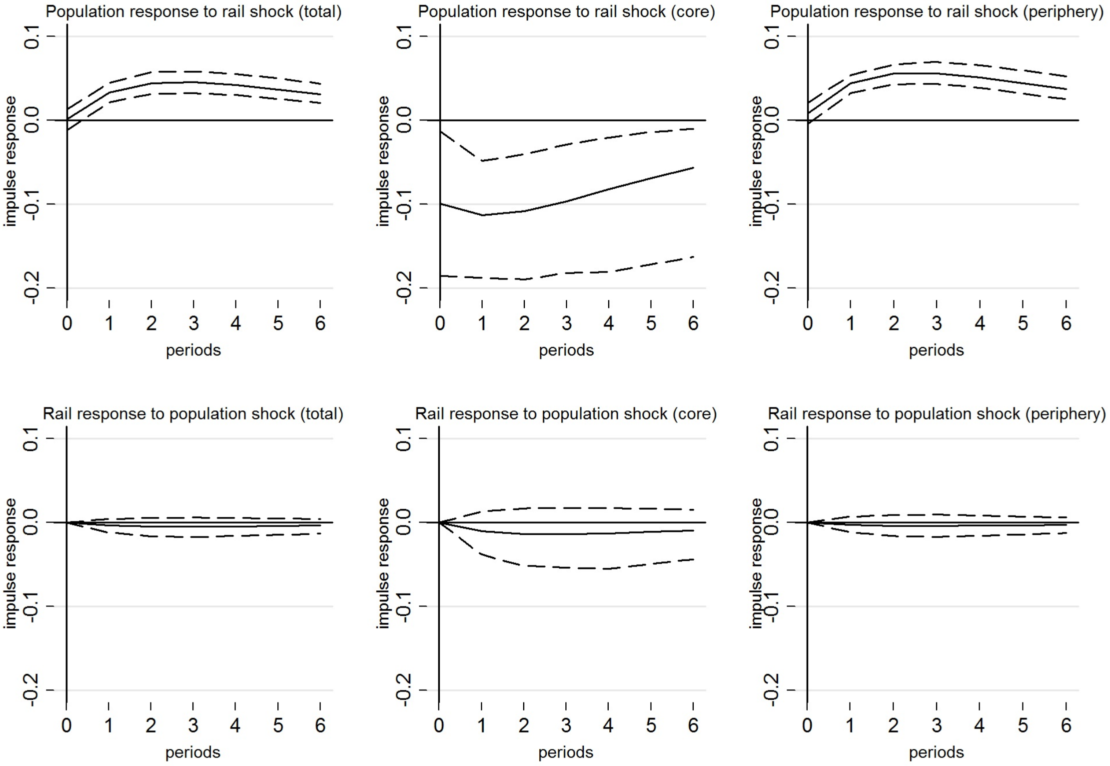
(d) 2010



(e) Legend

Notes: Population are extracted from Minnesota Population Center (2011).

Figure 6.4.: Impulse responses for 2-PVAR model



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

7. Conclusion

This dissertation's goal is to elaborate on the consumer city idea by investigating different aspects of the role of urban amenities in the location decision of households and firms. In the following, the main findings of the empirical analyses are summarised. I also highlight key contributions and limitations which leave room for future research.

The first analysis, which is carried out in Chapter 3, explains firm location by an area's endowment with localised consumption amenities. Berlin internet start-ups are found to be attracted by cultural amenities: An increase in the amenity density by 1% causes an increase in the probability of a young web firm to locate in a block by 1.2%. The positive impact remains significant across a number of robustness checks, testing for alternative explanations such as centrality and for the validity of the applied instrumental variables. Moreover, according to robustness specifications using placebo firms, it is indeed creative firms like agencies (probability increase of 0.84% as a response to a 1% rise in amenity density) and consultancies (with an increased probability of 0.57%) which are attracted by cultural amenities, whilst traditional service industries such as financial advisories (-0.55%) respond negatively. I contribute to an understudied field of research by being the first to explain firm location with highly endogenous urban amenities. It is shown that cultural amenities, a city's diversity and tolerance play an important role in attracting start-ups. It is therefore not very likely that young innovative firms will (re-)locate at (to) artificially created science parks in the periphery but will stay in more central and amenity-rich areas. Due to data availability this analysis is limited to cross-sectional estimation techniques. Time-variation would allow for estimating the start-up model in differences differentiating out time invariant unobservables and also for investigating the dynamics of the amenity effect. This could also enable us to disentangle potential agglomeration effects from the amenity effects. Moreover, it would be interesting not just to distinguish between service sectors but also between types of amenities. This is left for further research.

The analysis in Chapter 4 is centered around heritage as an example for aesthetic amenities. The key contribution is to provide insights into the political economy of conservation area designation. It is shown that the outcome is Pareto-efficient for local home-owners based on the following findings: An increase in preferences for historic character, proxied by the local share of residents holding a university degree, leads to an increase in the designated share of land area. Moreover, the property price effect inside newly designated conservation areas is not statistically distinguishable from zero, whereas the effect just outside the conservation area is positive and significant. Conservation area designation is therefore interpreted as solving a coordination problem among homeowners (and landlords) within a neighbourhood, i.e. the policy is Pareto-optimal at the local level. However, the analysis does not allow for conclusions regarding any global welfare effects. Future research must therefore depart from the localised mobility assumption and needs to consider potential supply restrictions due to the heritage preservation.

The idea of the two transport analyses carried out in Chapter 5 for Berlin and Chapter 6 for Chicago is to propose and apply a new methodology borrowed from macroeconomics which allows for the simultaneous estimation of transport and urban development. The intention is not to reject existing uni-directional causal inference but to complement and provide an alternative view on the link between transport and economic activity. The Panel VAR estimates indeed suggest that the above stated relation is not uni-directional but simultaneous. In fact, concluding from the land value analysis of Berlin, transport planners follow economic activity in core areas but tend to ignore residential development. In line with traditional models, the estimates for both Berlin and Chicago suggest an outbidding of residents in the CBD by firms due to transport improvements. Future research could follow up on the comparative study idea and extend the analysis to additional cities. Moreover, an extension to a multi variable VAR would allow for the estimation of conditional effects over time.

To conclude, the elaboration of the consumer city idea and the analysis of various different aspects makes clear that amenities are important when explaining how households and firms act in space, where they locate and why. The spatial allocation of people and firms can therefore be explained by more than just economic activity. Potential determinants are as manifold as the aspects one can study. Novel and improved empirical methods, an increase in the availability of spatial data (such as open data triggered by the availability of GPS devices), as well as improved data analysis and management software (like GIS) provide valuable research opportunities which help us to understand where we live and why we live there.

A. Appendix to Chapter 3

This appendix complements Chapter 3. Section A.1 provides additional information on the data used. The full tables of the empirical analyses providing all coefficient estimates are presented in Section A.2.

A.1. Data

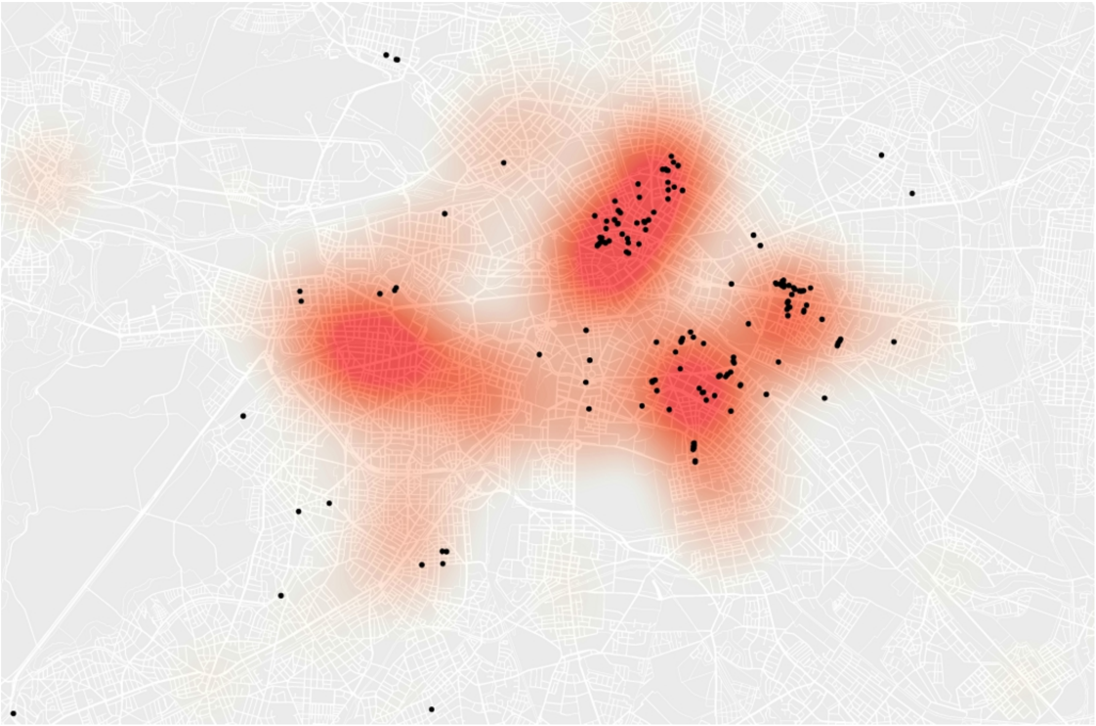
To motivate the first stage of the secondary instrument used in the analysis in the main text, squat density, Figure A.1 jointly maps the location of squatted houses, represented by the black dots, and the kernel density measure of cultural amenities. As reviewed in Section 3.2.1, a lot of abandoned buildings in proximity to the Berlin Wall and above all in the former Eastern part were squatted during the 1990s. There is a high number of squats (black dots) in the districts of Mitte, Prenzlauer Berg, Friedrichshain as well as Kreuzberg, whereas there are very few squatted houses in the proximity to Kurfürstendamm. The squats in the east are visually correlated with the red amenity clouds, strengthening the applicability of squat density as an instrument for cultural amenities.

Figure A.2 on page 126 shows the location of agencies (white dots) as well as financial advisories (black dots) jointly mapped with amenity density. Generally, the two service firms, which are used as placebos to test the robustness of the footloose start-up model, are distributed more equally over the city compared to internet start-ups (Figure 3.3 on page 37). Both exemplary sectors seem to be slightly clustered around Kurfürstendamm in the South West. Additionally, agencies are clustered in the northern red amenity cloud around Mitte/Prenzlauer Berg. Overall, visual inference suggests that agencies are more likely be affected by the cultural amenity distribution than financial advisories.

A.2. Empirical results

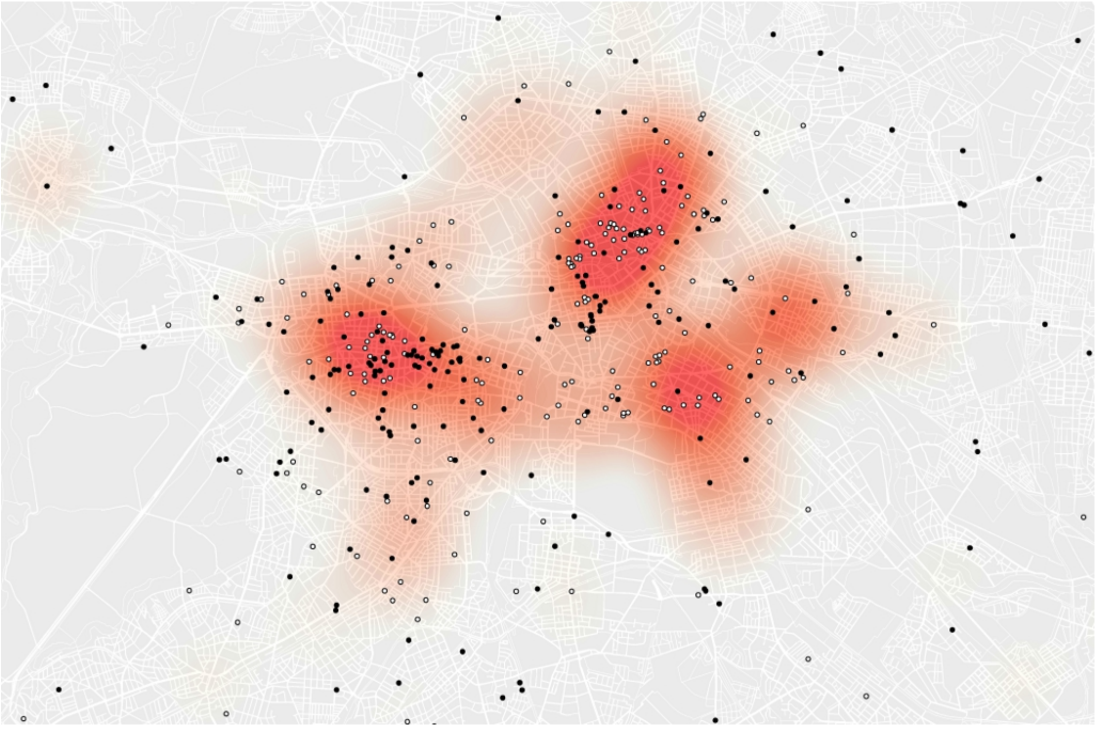
This section complements the estimation results from Section 3.6 by showing the full regression tables, i.e. including all control variables. Table A.1 on page 128 reports the estimates of the main models, starting with an uninstrumented baseline specification in column (1) and ending with the benchmark specification in column (5). The respective first stage regressions for the benchmark as well as for the remaining IV robustness models are given by Table A.2 on page 130. To address the concerns that proximity to Wall only reflects centrality or is correlated with unobservables driving the results, Table A.3 on page 132 presents the results of a first set of

Figure A.1.: Distribution of squatted houses and cultural amenities.



Notes: Black dots denote the location of squatted houses extracted from Hausbesetzungs Geschichte Berlin (2010). Red amenity clouds represent the amenity density measure with a radius of 2 km (Silverman, 1986), with dark red indicating a high amenity density.

Figure A.2.: Distribution of placebo firms and cultural amenities.



Notes: White dots denote the location of agencies and black dots of financial advisories respectively, both extracted from Gelbe Seiten Deutschland (2012). Red amenity clouds represent the amenity density measure with a radius of 2 km (Silverman, 1986), with dark red indicating a high amenity density.

robustness exercises. In a second set of robustness tests, I verify the validity of the measures and indicators used in the analysis and experiment with alternative instrumental variables. Results are shown in Table A.4 on page 135. The estimates using the placebo firms are eventually reported in Table A.5 on page 137 with their respective first stage results in Table A.6 on page 139.

Table A.1.: Estimation results: Footloose start-up model

	(1) # start-up	(2) # start-up	(3) # start-up	(4) # start-up	(5) # start-up
log cult dens	1.547***	1.640***	0.885***	0.671***	1.195***
	(0.098)	(0.118)	(0.190)	(0.157)	(0.307)
employment		4E-4***	4E-4***	4E-4***	4E-4***
		(0.000)	(0.000)	(0.000)	(0.000)
rent		-5E-5**	-3E-5	-3E-5	-3E-5
		(0.000)	(0.000)	(0.000)	(0.000)
migrants			0.004***	0.004***	0.004***
			(0.001)	(0.001)	(0.001)
dist research inst			0.024	0.166*	0.269**
			(0.115)	(0.101)	(0.115)
dist university			0.209**	0.176**	0.203**
			(0.097)	(0.084)	(0.095)
dist VC			-0.284***	-0.291***	-0.215***
			(0.079)	(0.074)	(0.080)
dist to water			-0.512***	-0.392***	-0.384***
			(0.119)	(0.109)	(0.110)
dist to green space			0.593	0.655	0.654*
			(0.458)	(0.425)	(0.359)
sport			0.209***	0.191***	0.208***
			(0.073)	(0.064)	(0.061)
bus dens			0.100**	-0.006	-0.035
			(0.042)	(0.040)	(0.045)
light rail dens			0.316	0.721*	0.790*
			(0.509)	(0.432)	(0.427)
undergr. dens			-0.350	0.216	-0.086
			(0.223)	(0.264)	(0.283)
tram dens			0.010	-0.034	-0.128
			(0.073)	(0.068)	(0.087)
U/tram noise			0.015***	0.010**	0.010**
			(0.005)	(0.005)	(0.005)
train noise			-0.046***	-0.034***	-0.035***
			(0.012)	(0.010)	(0.011)
street noise			-0.021**	-0.021**	-0.019**
			(0.009)	(0.009)	(0.009)

East Berlin			1.457***	0.896***	1.114***
			(0.402)	(0.302)	(0.386)
historic CBD			0.076	0.019	0.081
			(0.284)	(0.283)	(0.247)
urban renewal			0.314	0.370*	0.296
			(0.234)	(0.204)	(0.211)
x coord			-0.013	0.047	0.021
			(0.033)	(0.036)	(0.036)
y coord			-0.017	-0.037	-0.036
			(0.031)	(0.028)	(0.031)
Constant	-8.026***	-8.315***	-3.143**	-3.856***	-5.024***
	(0.370)	(0.427)	(1.311)	(1.261)	(1.492)
Controls	No	No	Yes	Yes	Yes
FE	No	No	No	Yes	Yes
IV	No	No	No	No	Yes
N	15850	15850	15850	15850	15850
OVERID					1.407
OVERIDP					0.236

Notes: Clustered standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01, Instruments: distance to Wall, squat density, OVERID (OVERIDP) denotes Hansen's J statistic of the overidentification test (and its p-value).

Table A.2.: Estimation results: Footloose start-up model - first stage

	(1) log cult dens	(2) log cult dens	(3) cult dens	(4) cult pot	(5) cult pot	(6) cult pot	(7) log cult dens
dist wall	-0.022***	-0.023***	-0.062***	-0.163***	0.239***		
	(0.002)	(0.002)	(0.021)	(0.032)	(0.039)		
pot squats					1.317***	1.195***	
					(0.029)	(0.029)	
pot wall						0.460***	
						(0.059)	
cult1936							-0.523***
							(0.031)
squat density	0.115***		8.025***	11.759***			
	(0.007)		(0.083)	(0.129)			
employment	-3E-5*	-3E-5**	-0.001***	-0.001**	-0.001**	-0.001**	-2E-5
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
rent	3E-5***	3E-5***	0.003***	0.005***	0.005***	0.005***	9E-5***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
migrants	0.001***	0.001***	0.016***	0.031***	0.035***	0.034***	0.001***
	(0.000)	(0.000)	(0.001)	(0.002)	(0.002)	(0.002)	(0.000)
dist research inst	-0.054***	-0.057***	0.105**	-0.025	-0.331***	-0.267***	-0.052***
	(0.004)	(0.004)	(0.046)	(0.071)	(0.082)	(0.082)	(0.004)
dist university	0.011***	0.024***	0.390***	0.562***	0.788***	0.802***	0.030***
	(0.003)	(0.003)	(0.042)	(0.066)	(0.078)	(0.078)	(0.003)
dist VC	-0.060***	-0.068***	-0.436***	-0.723***	-0.144**	-0.166***	-0.073***
	(0.003)	(0.003)	(0.032)	(0.050)	(0.064)	(0.064)	(0.003)
dist to water	-0.101***	-0.109***	-0.289***	-0.414***	-0.513***	-0.469***	-0.117***
	(0.005)	(0.005)	(0.056)	(0.087)	(0.101)	(0.101)	(0.005)
dist to green space	-0.157***	-0.170***	1.780***	3.011***	1.839***	1.593***	-0.146***
	(0.018)	(0.018)	(0.219)	(0.341)	(0.395)	(0.393)	(0.018)
sport	0.021**	0.018*	-0.275**	-0.401**	-0.365*	-0.329	0.014
	(0.009)	(0.010)	(0.117)	(0.182)	(0.211)	(0.211)	(0.010)

bus dens	0.096***	0.099***	0.097***	0.039	0.206***	0.040	0.109***
	(0.002)	(0.002)	(0.028)	(0.043)	(0.050)	(0.049)	(0.002)
light rail dens	1.268***	1.242***	3.610***	5.578***	-0.125	0.436	1.157***
	(0.036)	(0.037)	(0.449)	(0.698)	(0.811)	(0.810)	(0.037)
undergr dens	1.070***	1.134***	21.365***	30.483***	28.238***	28.907***	1.245***
	(0.020)	(0.020)	(0.249)	(0.387)	(0.481)	(0.486)	(0.021)
tram dens	0.276***	0.311***	0.929***	0.526***	2.118***	2.112***	0.321***
	(0.006)	(0.005)	(0.069)	(0.107)	(0.123)	(0.123)	(0.005)
U/tram noise	0.001**	0.000	-0.018***	0.004	-0.053***	-0.048***	0.000
	(0.000)	(0.000)	(0.005)	(0.007)	(0.008)	(0.008)	(0.000)
train noise	0.003***	0.003***	0.049***	0.075***	0.091***	0.087***	0.004***
	(0.001)	(0.001)	(0.007)	(0.011)	(0.013)	(0.013)	(0.001)
street noise	-0.002**	-0.002***	-0.112***	-0.145***	-0.256***	-0.251***	-0.002**
	(0.001)	(0.001)	(0.009)	(0.013)	(0.016)	(0.016)	(0.001)
East Berlin	-0.598***	-0.602***	2.710***	4.118***	4.668***	5.201***	-0.621***
	(0.025)	(0.026)	(0.310)	(0.482)	(0.560)	(0.548)	(0.025)
historic CBD	-0.326***	-0.212***	-6.896***	-9.114***	-7.478***	-8.853***	-0.072*
	(0.037)	(0.037)	(0.459)	(0.713)	(0.841)	(0.861)	(0.037)
urban renewal	0.047	0.159***	3.031***	7.596***	14.857***	15.235***	0.157***
	(0.034)	(0.034)	(0.420)	(0.653)	(0.749)	(0.751)	(0.034)
x coord	0.011***	0.012***	0.010	0.038	-0.209***	-0.267***	0.020***
	(0.001)	(0.001)	(0.018)	(0.028)	(0.034)	(0.032)	(0.001)
y coord	-0.011***	-0.012***	-0.051***	-0.057***	-0.163***	-0.221***	-0.007***
	(0.001)	(0.001)	(0.012)	(0.019)	(0.022)	(0.021)	(0.001)
Constant	1.008***	0.994***	2.125***	3.708***	3.803**	8.637***	0.397***
	(0.067)	(0.067)	(0.822)	(1.270)	(1.478)	(1.244)	(0.056)
F	1775.682	1772.545	2617.249	2440.989	1735.954	1738.901	1786.564
R^2	0.851	0.849	0.894	0.887	0.848	0.848	0.850
N	15850	15850	15937	15937	15937	15937	15850

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A.3.: Estimation results: Robustness exercises (1)

	(1) # start-ups	(2) # start-ups	(3) # start-ups	(4) # start-ups	(5) # start-ups
log cult dens	0.673***	0.573	-0.644	1.121***	1.507***
	(0.160)	(1.035)	(1.446)	(0.405)	(0.459)
dist to CBD	0.007				
	(0.078)				
employment	4E-4***	4E-4***	4E-4***	5E-4***	4E-4***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
rent	-4E-5	-4E-5	-3E-5	-1E-5	-3E-5
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
migrants	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
dist research inst	0.164	0.144	-0.191	0.058	0.095
	(0.105)	(0.255)	(0.503)	(0.130)	(0.143)
dist university	0.177**	0.172*	0.112	0.265**	0.228
	(0.084)	(0.094)	(0.146)	(0.112)	(0.139)
dist VC	-0.297***	-0.308	-0.559	-0.110	-0.092
	(0.093)	(0.194)	(0.344)	(0.097)	(0.098)
dist to water	-0.393***	-0.391***	-0.328*	-0.495***	-0.628***
	(0.111)	(0.105)	(0.172)	(0.141)	(0.154)
dist to green space	0.648	0.655*	0.352	0.567	0.665
	(0.437)	(0.375)	(1.449)	(0.469)	(0.460)
sport	0.190***	0.189***	0.172***	0.217***	0.225***
	(0.063)	(0.065)	(0.060)	(0.076)	(0.077)
bus dens	-0.006	-0.001	0.082	0.114**	0.076
	(0.040)	(0.067)	(0.161)	(0.057)	(0.063)
light rail dens	0.727*	0.709	0.566	0.238	0.483
	(0.440)	(0.457)	(0.538)	(0.626)	(0.464)
undergr dens	0.221	0.268	0.873	-0.619*	-0.550*
	(0.278)	(0.610)	(0.622)	(0.338)	(0.316)
tram dens	-0.035	-0.018	0.182	-0.019	-0.105
	(0.068)	(0.184)	(0.279)	(0.092)	(0.093)
U/tram noise	0.010**	0.010**	0.012*	0.009**	0.016***
	(0.005)	(0.005)	(0.007)	(0.005)	(0.005)
train noise	-0.035***	-0.034***	-0.032***	-0.048***	-0.055***
	(0.011)	(0.011)	(0.011)	(0.013)	(0.014)
street noise	-0.020**	-0.021**	-0.027**	-0.018**	-0.021**
	(0.009)	(0.010)	(0.013)	(0.009)	(0.009)
East Berlin	0.890***	0.864*	0.495	1.791***	0.816*
	(0.308)	(0.513)	(0.544)	(0.580)	(0.428)
historic CBD	0.019	0.003	-0.263	0.001	0.188
	(0.283)	(0.294)	(0.466)	(0.319)	(0.387)
urban renewal	0.372*	0.387	0.620	0.415	0.199
	(0.203)	(0.264)	(0.381)	(0.269)	(0.221)
x coord	0.050	0.053	0.137	0.021	-0.012
	(0.045)	(0.068)	(0.130)	(0.050)	(0.042)
y coord	-0.035	-0.037	-0.055	0.044	-0.053
	(0.032)	(0.027)	(0.047)	(0.051)	(0.047)

Constant	-3.981**	-3.635	-1.196	-8.042***	-4.372**
	(1.755)	(2.578)	(2.829)	(2.232)	(1.707)
Controls	Yes	Yes	Yes	Yes	Yes
FE	voting	voting	voting	district	municipality
IV1		dist to CBD	placebo Wall	dist Wall	dist Wall
IV2				d. squat	p. squat
N	15850	15850	15850	15850	15850
OVERID		0.000	0.000	1.360	0.002
OVERIDP				0.244	0.968

Notes: Clustered standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$. OVERID (OVERIDP) denotes Hansen's J statistic of the over-identification test (p-value).

Table A.4.: Estimation results: Robustness exercises (2)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	start-ups	start-ups	start-ups	start-ups	start-ups	start-ups	start-ups	start-ups	start-ups	start-ups
cult amenity	1.570*	0.021***	0.014***	0.032**	0.031***	0.019	0.000	1.600	0.987**	1.590***
	(0.844)	(0.006)	(0.004)	(0.013)	(0.011)	(0.015)	(0.006)	(1.269)	(0.377)	(0.431)
ring 250m						0.019				
						(0.015)				
ring 500m						-0.007	2E-4			
						(0.011)	(0.006)			
ring 1000m						0.012***	0.012**			
						(0.004)	(0.005)			
ring 1500m						-0.006**	-0.007***			
						(0.002)	(0.003)			
ring 2000m						0.005**	0.006***			
						(0.002)	(0.002)			
employment	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	5E-4***	4E-4***	5E-4
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
land value 1992	-3E-5	-3E-5	-2E-5	-1E-5	-3E-5	-3E-5	-2E-5	-2E-5	-3E-5	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
rent 2010										0.056
										(0.059)
migrants	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.001)
dist research institute	0.162	-0.162	-0.148	0.014	-0.019	-0.113	-0.090	0.167	0.236*	0.155
	(0.166)	(0.106)	(0.107)	(0.163)	(0.145)	(0.105)	(0.105)	(0.223)	(0.123)	(0.123)
dist university	0.269*	0.201**	0.195*	0.270**	0.275**	0.171*	0.156	0.271	0.197**	0.284***
	(0.143)	(0.100)	(0.101)	(0.127)	(0.119)	(0.103)	(0.100)	(0.171)	(0.091)	(0.110)
dist VC	-0.170	-0.426***	-0.435***	-0.373***	-0.370***	-0.359***	-0.335***	-0.166	-0.244**	-0.155*
	(0.134)	(0.071)	(0.073)	(0.084)	(0.080)	(0.097)	(0.095)	(0.184)	(0.091)	(0.093)
dist to water	-0.522***	-0.395***	-0.371***	-0.267**	-0.295**	-0.369***	-0.372***	-0.522***	-0.390***	-0.487***
	(0.128)	(0.119)	(0.121)	(0.119)	(0.116)	(0.124)	(0.124)	(0.129)	(0.108)	(0.128)

dist to green	0.704	0.603	0.534	0.233	0.165	0.563	0.652	0.709	0.649*	0.516
space	(0.487)	(0.459)	(0.476)	(0.474)	(0.476)	(0.437)	(0.430)	(0.442)	(0.364)	(0.449)
sport	0.228***	0.197***	0.189**	0.194**	0.196**	0.131	0.146*	0.229**	0.203***	0.232***
	(0.077)	(0.073)	(0.075)	(0.088)	(0.084)	(0.094)	(0.088)	(0.089)	(0.062)	(0.071)
bus dens	0.063	0.080*	0.067	0.020	0.021	0.079*	0.086**	0.061	-0.020	0.043
	(0.060)	(0.044)	(0.044)	(0.042)	(0.040)	(0.043)	(0.043)	(0.085)	(0.049)	(0.045)
light rail dens	0.642	0.665	0.420	1.132	1.123	0.504	0.583	0.657	0.755*	0.466
	(0.695)	(0.532)	(0.552)	(0.967)	(0.867)	(0.584)	(0.577)	(0.679)	(0.432)	(0.465)
undergr dens	-0.629	-0.449*	-0.435	-1.116**	-0.986**	-0.542*	-0.574*	-0.641	0.039	-0.731***
	(0.408)	(0.265)	(0.265)	(0.524)	(0.441)	(0.280)	(0.297)	(0.562)	(0.335)	(0.263)
tram dens	-0.088	-0.035	-0.000	-0.172	-0.181	-0.035	-0.036	-0.093	-0.089	-0.114
	(0.142)	(0.085)	(0.081)	(0.138)	(0.124)	(0.081)	(0.082)	(0.198)	(0.103)	(0.087)
U/tram noise	0.015***	0.010**	0.008	0.000	0.004	0.014**	0.016***	0.015***	0.010**	0.015***
	(0.005)	(0.005)	(0.005)	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
train noise	-0.051***	-0.045***	-0.044***	-0.059***	-0.057***	-0.046***	-0.048***	-0.051***	-0.034***	-0.046***
	(0.015)	(0.013)	(0.012)	(0.018)	(0.016)	(0.013)	(0.014)	(0.014)	(0.011)	(0.013)
street noise	-0.017	-0.013	-0.013	0.004	0.001	-0.018*	-0.016	-0.017	-0.019**	-0.019**
	(0.011)	(0.010)	(0.010)	(0.013)	(0.012)	(0.010)	(0.010)	(0.012)	(0.009)	(0.009)
East Berlin	1.748***	1.269***	1.194***	1.957***	1.983***	1.488***	1.434***	1.761**	1.043***	1.704***
	(0.568)	(0.429)	(0.452)	(0.751)	(0.664)	(0.439)	(0.426)	(0.717)	(0.390)	(0.437)
historic	0.277	-0.097	-0.047	0.084	0.023	-0.240	-0.260	0.285	0.056	0.292
CBD	(0.320)	(0.288)	(0.308)	(0.281)	(0.271)	(0.307)	(0.304)	(0.432)	(0.248)	(0.272)
urban	0.202	0.265	0.234	-0.156	-0.134	-0.056	0.052	0.198	0.344	0.222
renewal	(0.280)	(0.254)	(0.256)	(0.395)	(0.359)	(0.280)	(0.309)	(0.265)	(0.220)	(0.232)
x coord	-0.050	0.064**	0.061*	0.052	0.038	0.019	0.014	-0.052	0.301	-0.040
	(0.055)	(0.032)	(0.032)	(0.037)	(0.033)	(0.038)	(0.038)	(0.068)	(0.039)	(0.037)
y coord	-0.022	-0.017	-0.016	-0.029	-0.022	-0.019	-0.019	-0.022	-0.344	-0.027
	(0.042)	(0.028)	(0.028)	(0.030)	(0.028)	(0.027)	(0.027)	(0.046)	(0.029)	(0.039)
Constant	-4.812*	-2.529**	-2.402**	-3.394**	-3.057**	-1.669	-1.862	-4.883	-4.657***	-5.380***
	(2.476)	(1.232)	(1.215)	(1.641)	(1.480)	(1.182)	(1.195)	(3.418)	(1.558)	(1.707)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Rent 2010
FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Cult (X)	ln cult dens	cult dens	cult pot	cult pot	cult pot	rings	rings	ln cult dens	ln cult dens	ln cult dens
IV1	dist Wall	dist Wall	dist Wall	dist Wall	p. Wall	hist rings	hist rings	cult 1936	cult 1998	dist Wall
IV2		d. squat	d. squat	p. squat	p. squat				dist Wall	d. squat
N	15850	15937	15937	15937	15937	15937	15937	15850	15850	15850
OVERID	0.000	5.526	6.942	7.049	3.219	0.000	0.000	0.000	1.043	0.012
OVERIDP		0.019	0.008	0.008	0.073				0.307	0.911

Notes: Standard errors in parentheses, * p<0.1, ** p<0.05, *** p<0.01, OVERID (OVERIDP) denotes Hansen's J statistic of the overidentification test (and its p-value). IVs: dens denotes a density measure, dist a Euclidean distance measure, pot a potentiality measure, rings the amenity rings. Cult stands for the the cultural amenity to be instrumented. Column (10) includes rents from 2010 (Immobilien Scout, 2012) instead of 1992 land values.

Table A.5.: Estimation results: Placebo firms

	(1) architects	(2) consultancies	(3) engineering	(4) insurance	(5) law	(6) publisher	(7) finance	(8) agencies
log cult dens	0.110	0.567***	-0.187	-0.221	-0.147	0.234	-0.551*	0.836**
	(0.196)	(0.216)	(0.336)	(0.465)	(0.176)	(0.376)	(0.320)	(0.333)
employment	3E-4***	3E-4***	3E-4***	3E-4***	3E-4***	3E-4***	3E-4***	4E-4***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
rent	-5E-6	1E-4***	5E-5*	8E-5**	8E-5***	6E-5*	1E-4***	4E-5
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
migrants	0.003***	0.003***	0.003***	0.003***	0.003***	0.004***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
dist research inst	0.039	0.071	-0.004	-0.022	0.033	0.162	0.011	0.012
	(0.048)	(0.047)	(0.063)	(0.082)	(0.040)	(0.132)	(0.062)	(0.108)
dist university	-0.042	-0.054	-0.002	0.039	-0.085***	0.041	-0.007	-0.010
	(0.030)	(0.038)	(0.046)	(0.062)	(0.025)	(0.083)	(0.059)	(0.080)
dist VC	-0.089**	0.021	-0.099**	-0.087	-0.051	-0.128	-0.147***	-0.025
	(0.035)	(0.043)	(0.049)	(0.084)	(0.036)	(0.106)	(0.055)	(0.086)
dist to water	0.005	0.081	-0.082	-0.055	-0.027	-0.017	0.005	0.052
	(0.041)	(0.049)	(0.070)	(0.097)	(0.037)	(0.098)	(0.074)	(0.095)
dist to green space	0.197	0.180	-0.001	-0.230	-0.138	-0.237	-0.260	-0.235
	(0.169)	(0.207)	(0.226)	(0.392)	(0.156)	(0.490)	(0.492)	(0.457)
sport	0.133***	0.125***	0.062	0.242***	0.124***	0.199**	0.194***	0.124*
	(0.045)	(0.047)	(0.068)	(0.057)	(0.037)	(0.091)	(0.061)	(0.067)
bus dens	-0.004	-0.032	0.021	0.019	0.052**	0.030	0.080	0.043
	(0.024)	(0.028)	(0.043)	(0.054)	(0.023)	(0.052)	(0.058)	(0.048)
light rail dens	1.429***	1.260***	1.123**	1.481**	1.576***	1.574***	1.837***	0.760
	(0.269)	(0.285)	(0.488)	(0.609)	(0.245)	(0.521)	(0.601)	(0.474)
undergr dens	0.362	-0.292	0.343	0.570	0.717***	0.619*	0.830**	-0.695**
	(0.220)	(0.240)	(0.402)	(0.538)	(0.180)	(0.321)	(0.370)	(0.307)
tram dens	-0.126**	-0.210***	-0.129	-0.028	0.007	-0.018	0.005	-0.219**
	(0.059)	(0.059)	(0.108)	(0.123)	(0.050)	(0.091)	(0.116)	(0.086)

U/tram noise	0.009***	0.008**	0.012***	0.010**	0.009***	0.007	0.002	0.010**
	(0.003)	(0.003)	(0.004)	(0.005)	(0.002)	(0.006)	(0.004)	(0.005)
train noise	-0.025***	-0.034***	-0.015**	-0.033***	-0.046***	-0.053***	-0.039***	-0.044***
	(0.005)	(0.006)	(0.007)	(0.011)	(0.005)	(0.016)	(0.009)	(0.013)
street noise	-0.055***	-0.028***	-0.016**	-0.013	-0.023***	-0.053***	0.002	-0.025**
	(0.005)	(0.006)	(0.007)	(0.011)	(0.004)	(0.013)	(0.008)	(0.011)
East Berlin	0.747***	1.012***	0.709*	0.395	0.468***	0.565	0.649*	1.257***
	(0.220)	(0.244)	(0.374)	(0.449)	(0.175)	(0.433)	(0.382)	(0.410)
historic CBD	-0.035	0.092	-0.895**	-1.128**	-0.779***	-0.294	-0.916**	0.687**
	(0.191)	(0.211)	(0.394)	(0.546)	(0.183)	(0.346)	(0.427)	(0.288)
urban renewal	0.701***	0.563***	0.782***	0.702**	1.027***	0.378	0.681**	1.001***
	(0.144)	(0.163)	(0.251)	(0.345)	(0.149)	(0.379)	(0.318)	(0.254)
x coord	-0.056***	-0.067***	-0.011	0.004	-0.026*	-0.015	-0.001	-0.072**
	(0.015)	(0.017)	(0.022)	(0.032)	(0.014)	(0.041)	(0.024)	(0.034)
y coord	-0.032***	-0.017	0.004	-0.008	-0.027***	-0.045	-0.008	-0.022
	(0.010)	(0.012)	(0.013)	(0.016)	(0.008)	(0.028)	(0.018)	(0.025)
Constant	2.350***	-0.437	-2.143***	-2.594***	1.352***	-0.433	-2.539***	-1.895
	(0.436)	(0.537)	(0.692)	(0.918)	(0.366)	(1.136)	(0.834)	(1.180)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	15850	15850	15850	15850	15850	15850	15850	15850
OVERID	5.102	0.677	0.001	2.167	2.510	3.117	0.014	0.241
OVERIDP	0.024	0.410	0.974	0.141	0.113	0.077	0.906	0.624

Notes: Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Instruments: distance to Wall, squat density, OVERID (OVERIDP) denotes Hansen's J statistic of the overidentification test (and its p-value). There are at least ten law firms per voting precincts (instead of five).

Table A.6.: Estimation results: Placebo firms - first stage

	(1) architects log cult dens	(2) consultancies log cult dens	(3) engineering log cult dens	(4) insurance log cult dens	(5) law log cult dens	(6) publisher log cult dens	(7) finance log cult dens	(8) agencies log cult dens
dist wall	-0.021***	-0.023***	-0.023***	-0.023***	-0.024***	-0.023***	-0.023***	-0.023***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
squat density	0.095***	0.104***	0.107***	0.104***	0.108***	0.105***	0.102***	0.112***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
employment	-3E-5*	-3E-5**	-3E-5**	-3E-5*	-1E-5	-3E-5**	-3E-5**	3E-5**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
rent	5E-5***	5E-5***	2E-6	5E-6	5E-5***	2E-5***	4E-5***	3E-5***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
migrants	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
dist research inst	-0.053***	-0.055***	-0.056***	-0.056***	-0.053***	-0.055***	-0.055***	-0.056***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
dist university	0.011***	0.013***	0.015***	0.015***	0.014***	0.013***	0.012***	0.012***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
dist VC	-0.060***	-0.064***	-0.071***	-0.071***	-0.064***	-0.066***	-0.066***	-0.064***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
dist to water	-0.099***	-0.099***	-0.100***	-0.101***	-0.100***	-0.102***	-0.101***	-0.101***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
dist to green space	-0.152***	-0.157***	-0.141***	-0.144***	-0.149***	-0.151***	-0.149***	-0.147***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
sport	0.020**	0.022**	0.023**	0.022**	0.021**	0.022**	0.023**	0.022**
	(0.009)	(0.010)	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)
bus dens	0.095***	0.096***	0.095***	0.095***	0.096***	0.095***	0.095***	0.096***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
light rail dens	1.256***	1.172***	1.138***	1.125***	1.159***	1.170***	1.170***	1.215***
	(0.038)	(0.037)	(0.037)	(0.037)	(0.038)	(0.037)	(0.037)	(0.037)
undergr dens	1.053***	1.044***	1.097***	1.097***	1.051***	1.065***	1.049***	1.045***
	(0.022)	(0.021)	(0.020)	(0.020)	(0.022)	(0.020)	(0.020)	(0.020)

tram dens	0.282***	0.267***	0.252***	0.253***	0.268***	0.261***	0.261***	0.262***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
U/ram noise	0.001**	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
train noise	0.002***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
street noise	-0.002**	-0.002**	-0.001	-0.001	-0.002**	-0.001*	-0.001**	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
East Berlin	-0.633***	-0.607***	-0.655***	-0.652***	-0.612***	-0.649***	-0.636***	-0.643***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
historic CBD	-0.370***	-0.473***	-0.727***	-0.713***	-0.428***	-0.584***	-0.565***	-0.575***
	(0.039)	(0.036)	(0.033)	(0.033)	(0.037)	(0.034)	(0.035)	(0.034)
urban renewal	0.093***	0.014	0.056*	0.058*	0.048	0.034	0.036	0.065*
	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)	(0.034)
x coord	0.012***	0.013***	0.016***	0.016***	0.014***	0.015***	0.015***	0.014***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
y coord	-0.011***	-0.011***	-0.010***	-0.010***	-0.011***	-0.010***	-0.010***	-0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.968***	0.964***	0.977***	0.971***	0.939***	0.967***	0.974***	0.981***
	(0.067)	(0.068)	(0.069)	(0.069)	(0.068)	(0.068)	(0.068)	(0.068)
F	856.351	1856.923	3059.642	3419.338	1093.501	3356.475	3017.632	3036.796
R^2	0.855	0.849	0.844	0.844	0.852	0.847	0.847	0.848
N	15850	15850	15850	15850	15850	15850	15850	15850

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

B. Appendix to Chapter 4

B.1. Introduction

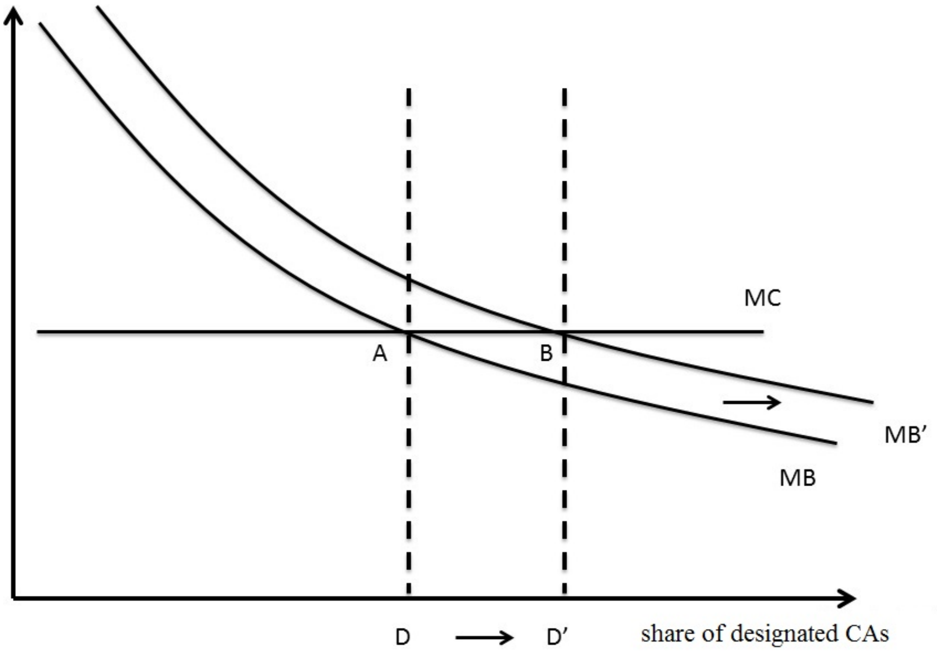
This appendix complements Chapter 4. Section B.2 provides an illustration of how a planner determines the Pareto-efficient designation share and adds to the theory section of the main chapter. Section B.3 complements the empirical strategy section of the main text by providing a more detailed discussion of the control variables in tobit designation process models. The section also links the reduced form difference-in-differences parameters to the marginal policy effect in the theoretical model. Section B.4 provides a detailed overview of the data we use, its sources, and how they are processed. Finally, Section B.5 complements the empirical results section of the main text by showing the results of a variety of robustness tests and model alterations not reported in the main chapter.

B.2. Theory and context

This section briefly illustrates how a planner determines the Pareto-efficient designation share. The equilibrium between the social marginal benefits (MB) of designation (equation (4.9) in the main text) and the marginal costs (MC) (equation (4.10)) is depicted by Figure B.1 on page 142. At point A the designation share D is Pareto-efficient. Social marginal benefits equal the private marginal costs associated with designation. A further extension would benefit all owners to the left of A as they would profit from increasing the expected heritage in the neighbourhood without experiencing a change in marginal cost. To the right of A, however, the social marginal benefit would also increase, but the increase would not compensate for the private marginal costs associated with a change in the designation status from undesignated to designated. The expansion would not be Pareto-optimal.

If there is, for instance, a change in preferences and residents develop a greater taste for external heritage γ their marginal benefits curve shifts to the right. A Pareto-optimal planner adapts to this situation and raises the designation share to set marginal benefits equal to marginal costs again. This new Pareto-optimal equilibrium is illustrated by point B where the designation share increases to $D^{'}$.

Figure B.1.: Designation equilibrium



B.3. Empirical strategy

B.3.1. Designation process - control variables

This section provides a detailed description and motivation of the control variables we use to account for the determinants of conservation area designation that are unrelated to the mechanisms modelled in our theory. In particular we try to control for composition effects, neighbourhood sorting, heterogeneity in terms of homeownership, and whether the heritage in a neighbourhood is at particular risk.

We add the initial period (1991) degree share for two reasons. First, we assume that the highly educated derive higher (net-)benefits from neighbourhood heritage. To the extent that this group is capable of more efficiently articulating their will in a political bargaining a higher degree share will make the designation more likely. It is important to control for the initial degree share since levels and changes may be correlated in either direction. On the one hand there may be catch-up growth in the degree share of less educated regions, i.e. mean reversion. On the other hand, people with degrees may be more likely to move to areas with an already high share of people with degrees, which would imply a self-reinforcing process leading to spatial segregation.

We also include a control for the extent of designation in the initial period (1991). The share of designated land area in the total ward area would be (positively) correlated with the change in the designation share if designations spark further designations as in a contagion model. Initial designation also helps to control for the possibility that the skilled may be attracted to areas with a lot of designated land. Another set of controls is driven by the interest in homeowners within the designation process. Homeowners experience extra benefits/costs from designation since, unlike renters, they are not compensated for changes in neighbourhood quality by increases in degrees or rents. Homeowners, thus have additional incentives to engage in political bargaining. Similar to the other controls, homeownership status enters in lagged levels and differences. In a final specification we also add an interaction of the logged change in degree with homeownership (rescaled to a zero mean to make coefficients comparable). We use average household size (both in differences and lagged levels) to control for the presumption that larger households are more likely to lobby against designation and the resulting constraint on available floor space.

We add a measure of property price appreciation, which we obtain from ward-level regressions of log property prices on a time trend (and property controls).

A larger risk aversion increases the benefit from a policy that increases certainty regarding the future of the neighbourhood and, thus, potentially increases the optimal designation share. To control for a potentially positive correlation between owners' risk aversion and the value of their properties - typically their largest assets - we add a measure of neighbourhood appreciation. We generate ward-level property price trends in n separate auxiliary regressions of the following type:

$$\log\left(P_{itn}\right) = a_n + X_{in}b_n + \beta_n T_t + \varepsilon_{itn} \tag{B.1}$$

where X is a vector of property and neighbourhood characteristics and T is a linear time trend. To avoid a reverse effect of designation on the property price trend we only consider transactions that occur outside conservation areas.

A second set of controls deals with potential development risk. Areas that experience development pressure or are in poor and/or declining condition may be more likely to be designated in order to protect against the threats to the heritage character of the neighbourhood. We use the vacancy rate, a density measure of listed buildings as well as score measures for a conservation area's condition, vulnerability and trajectory provided by English Heritage to capture development pressure. We expect that neighbourhoods with few vacancies will be put under higher development pressure. Vacancies enter the specification both in differences and lagged levels. The reason for the differenced term is that a change in development pressure is likely to lead to a change in designation status as a result. We argue that the lagged level may also capture changes (not just levels) in development pressure. This is because of external factors and conditions (i.e. population growth) that effect areas unevenly depending on their level in certain attributes (e.g. vacant housing). It seems likely that general population growth would put greater development pressure on neighbourhoods with lower vacancy rates. The score measures reflect the development risk inside a conservation area and come from a survey provided by English Heritage. The higher the condition score, the worse the heritage conditions. A higher vulnerability as well as a higher trajectory are also indicated by higher scores. Except for the score variables, all control variables enter our empirical specification in logs.

While taking first-differences of the empirical specification will remove all time-invariant ward-specific effects that might impact on the level of designation (including the heritage itself), it will not help if there are location-specific effects that impact on the changes in designation status. For example, if there is heterogeneity across Local Authorities (LAs) about how difficult or easy it is to designate arising from different bureaucratic practices then this would affect changes in designation for all wards within a particular LA. We therefore estimate a fixed effects specification for the 166 English Travel To Work Areas (TTWAs). The TTWAs are designed to approximate city regions which can be described as somehow self-contained economic areas from a job market perspective. By applying a TTWA fixed effect model we are therefore able to control for socio-economic heterogeneity across TTWAs.

B.3.2. Difference-in-differences

This section motivates the difference-in-differences approach for the estimation of the marginal policy effect. Firstly, we illustrate how the policy and heritage effects are difficult to disentangle in a simple cross-sectional hedonic estimation. Secondly, we lay out how the difference-in-differences treatment effect is used to estimate the marginal policy effect laid out in terms of the structural parameters of our model.

Cross-sectional hedonics

Taking logs of the spatial equilibrium price equation (4.17) from the main text gives:[1]

$$\log\theta(x) = \tau + \frac{1}{1-\delta}\log\alpha(x) + \frac{\varphi h(x)}{1-\delta} + \frac{\gamma E[H \mid D]}{1-\delta} - \frac{c\tilde{D}(x)}{1-\delta} \tag{B.2}$$

The following heritage and policy effects determine the bid rent:

$$\text{Policy cost} = \frac{c\tilde{D}(x)}{1-\delta} \tag{B.3}$$

$$\text{External heritage effect (conditional on designation)} = \frac{\gamma E[H \mid D]}{1-\delta} \tag{B.4}$$

$$\text{Internal heritage effect} = \frac{\varphi h(x)}{1-\delta} \tag{B.5}$$

Consider the cross-sectional reduced form equation:

$$p_{it} = \aleph I_i + X_i^{'}\mu + f_n + Y_t + \epsilon_{it} \tag{B.6}$$

where p_{it} is the natural logarithm of the transaction price for property i in time period t, I_i is a dummy variable equal to one if the observation is internal to a treated conservation area, X_i is a vector of controls for property, neighbourhood, and environmental characteristics, f_n is a set of n location fixed effects and Y_t are year effects. The coefficient $\aleph$ on the CA_i dummy identifies the policy cost associated with the location of a property inside a conservation area $\tilde{D}(x) = 1$. The policy cost should have a negative effect on logged house prices. The coefficient also partly identifies the internal heritage effect. Specifically, it identifies the value

---

[1] Where τ is a constant and equal to: $\log(1-\delta) + \frac{\delta}{1-\delta}\log\delta + \frac{1}{1-\delta}\log W$.

of the difference between the mean internal heritage inside conservation areas and the mean internal heritage outside conservation areas (i.e. $\varphi/(1-\delta)(\overline{h_{CA_i=1}} - \overline{h_{CA_i=0}})$. This should be positive because the policymaker would normally designate areas that have the most heritage. Finally, under the existence of some spatial decay in externalities, it will also identify the value of the difference inside and outside conservation areas in the external heritage effect (i.e. $\gamma(1-\delta)(\overline{E[H \mid D]_{CA_i=1}} - \overline{E[H \mid D]_{CA_i=1}})$. This is a function of internal heritage and will therefore also be positive.

The coefficient $\aleph$ thus reflects a composite effect of policy costs, policy benefits, and correlated internal heritage effect. Furthermore, in reality the actual distribution of internal heritage is unknown and there is likely a spatial decay to externalities, further complicating the estimate.[2] In practice, $\aleph$ will also be affected by unobserved neighbourhood characteristics that are correlated with the distance to the conservation area. A positive $\aleph$ parameter, at best, tells us only that the overall higher levels of heritage (internal and external) combined with the policy benefits of conservation outweigh the policy costs. This does not provide a comprehensive evaluation of the policy effect itself. To try and disentangle these effects we implement a different empirical approach.

Difference-in-differences

Using the difference-in-differences (DD) approach to estimate the marginal effect of a change in designation status offers an improved identification.

Our empirical difference-in-differences specification is equation (4.26) from the main paper:

$$p_{it} = \beta^I I_i + \beta^E E_i + \beta^{IPost}(I_i \times Post_{it}) + \beta^{EPost}(E_i \times Post_{it}) + X_i^{'}\mu + f_n + Y_t + \epsilon_{it} \tag{B.7}$$

Table B.1 on page 146 illustrates the conditional mean prices (after controlling for time effects) for the treatment and control group in the pre- and post-treatment periods. It is important to note that the year fixed effects Y_t capture the general development of price over time. Without this feature it would be necessary to control for the overall growth in price between the pre- and post-treatment periods via the inclusion of a non-interacted version of $Post_{it}$.

Our treatment coefficient β^{IPost} essentially differentiates across the treatment and control groups before and after designation and is, thus defined as follows:

$$\beta^{IPost} = (\bar{p}^{Treat}_{Post} - \bar{p}^{Treat}_{Pre}) - (\bar{p}^{Con}_{Post} - \bar{p}^{Con}_{Pre}) \tag{B.8}$$

Let us assume that the relationship between the observed conditional mean and the theoretical bid rent is given by:

$$\bar{p}^{Treat}_{Post} = \theta^{Treat}_{Post} + u^{Treat}_{Post} \tag{B.9}$$

where u^{Treat}_{Post} are partially unobservable factors specific to properties in the Treated-Post cell. The same relationship applies for the other cells (Treated-Pre, Control-Post and Control-Pre).

---

[2] In a general case the estimate would be equal to:
$\aleph = \frac{\varphi}{1-\delta}(\overline{h_{CA_i=1}} - \overline{h_{CA_i=0}}) + \frac{\gamma}{1-\delta}(\overline{E[H \mid D]_{CA_i=1}} - \overline{E[H \mid D]_{CA_i=1}})$

At the heart of our identification strategy we assume that the price trends unrelated to the policy are the same within the treatment and the control group. The typical identifying assumption on which the difference-in-differences identification strategy relies can be expressed as follows:

$$(u^{Treat}_{Post} - u^{Treat}_{Pre}) = (u^{Con}_{Post} - u^{Con}_{Pre}) \tag{B.10}$$

The credibility of the counterfactual rests on the likelihood that the treatment group, in the absence of the intervention, would have followed a trend that is similar to that of the control group. An appropriate definition of the control group is therefore a critical element of the identification strategy. We therefore consider a number of different control groups in which we try to reduce the potential heterogeneity between properties in the treatment and control group.

Table B.1.: Conditional mean prices

Conditional mean of prices	Pre	Post
Treated (internal)	$\bar{p}^{Treat}_{Pre} = \beta^{I}$	$\bar{p}^{Treat}_{Post} = \beta^{I} + \beta^{IPost}$
Control	$\bar{p}^{Con}_{Pre} = 0$	$\bar{p}^{Con}_{Post} = 0$
Treatment effect $= (\bar{p}^{Treat}_{Post} - \bar{p}^{Treat}_{Pre}) - (\bar{p}^{Con}_{Post} - \bar{p}^{Con}_{Pre})$		
Treatment effect $= ([\beta^{I} + \beta^{IPost}] - [\beta^{I}]) - ([0] - [0])$		
Treatment effect $= \beta^{Post}$		

Notes: The conditional mean of prices in the treatment group in the pre-period is denoted $\bar{p}^{Treat}_{Pre}$. This represents the log of prices conditional on fixed and year effects $(f_n + Y_t)$ and controls X_i. The same notation is used for the other groups.

The first treatment group is a spatial match where we choose the observations that fall within a 2 km buffer surrounding conservation areas that changed designation status during the observation period (1995–2010). As an alternative, we consider a number of matching procedures that rest on the idea that properties inside conservation areas generally share similarities. Properties in conservation areas that did not change designation status therefore potentially qualify as a control group. To make the areas in the treatment and control group more similar, we select conservation areas based on similarities with those in our treatment group (Rosenbaum & Rubin, 1983). For the matching procedure we only make use of variables that turn out to have significant impact in the auxiliary propensity score matching regression.[3] We use a nearest neighbour matching procedure, which produces a broader and a narrower group.

Under the assumptions made it is straightforward to demonstrate that the DD treatment coefficient gives the pure policy effect we are interested in. Combining the theoretical bid rent of equation (4.17) from the main paper with the definition of $\bar{p}^{Treat}_{Post}$ in equation (B.9) gives

---

[3] A list of significant controls in propensity score matching regressions is included in the next subsection.

the conditional mean price of (treated) properties inside newly designated conservation areas before (pre) and after (post) designation can be expressed as follows:[4]

$$\bar{p}_{Pre}^{Treat} = \tau + \frac{1}{1-\delta}\log\alpha_i + \frac{\varphi h_i}{1-\delta} + \frac{\gamma E[H \mid D]}{1-\delta} + u_{Pre}^{Treat} \tag{B.11}$$

$$\bar{p}_{Post}^{Treat} = \tau + \frac{1}{1-\delta}\log\alpha_i + \frac{\varphi h_i}{1-\delta} + \frac{\gamma}{1-\delta}\left(E[H \mid D] + \frac{\mathrm{d}E[H \mid D]}{\mathrm{d}D}\right) - \frac{c\tilde{D}_i}{1-\delta} + u_{Post}^{Treat} \tag{B.12}$$

where a new designation is represented as an increase in designation share D. For a control group sufficiently far away to not be exposed to the heritage externality we similarly get:

$$\bar{p}_{Pre}^{Con} = \tau + \frac{1}{1-\delta}\log\alpha_i + \frac{\gamma E[H \mid D]}{1-\delta} + u_{Pre}^{Con} \tag{B.13}$$

$$\bar{p}_{Post}^{Con} = \tau + \frac{1}{1-\delta}\log\alpha_i + \frac{\gamma E[H \mid D]}{1-\delta} + u_{Post}^{Con} \tag{B.14}$$

where there is (by definition) no new designation. Given the common trend assumption of equation (B.10), β^{IPost} identifies the pure net policy effect of designation:

$$\beta^{IPost} = \frac{\gamma}{1-\delta}\frac{\mathrm{d}E[H \mid D]}{\mathrm{d}D} - \frac{c\tilde{D}_i}{1-\delta} \tag{B.15}$$

In the empirical implementation of the DD strategy we also consider alternative treatment groups that consist of properties just outside conservation areas, which are potentially exposed to spillovers, but not to the cost of designation. The interpretation of the external treatment co-efficient can be derived analogically where designation leads to benefits but without the associated costs:

$$\bar{p}_{Pre}^{Treat} = \tau + \frac{1}{1-\delta}\log\alpha_i + \frac{\gamma E[H \mid D]}{1-\delta} + u_{Pre}^{Treat} \tag{B.16}$$

$$\bar{p}_{Post}^{Treat} = \tau + \frac{1}{1-\delta}\log\alpha_i + \frac{\gamma}{1-\delta}\left(E[H \mid D] + \frac{\mathrm{d}E[H \mid D]}{\mathrm{d}D}\right) + u_{Post}^{Treat} \tag{B.17}$$

Under the common trends assumption the treatment coefficient reflects the pure policy benefit associated with the reduction in uncertainty as predicted by the stylized theory:

$$\beta^{EPost} = \frac{\gamma}{1-\delta} + \frac{\mathrm{d}E[H \mid D]}{\mathrm{d}D} \tag{B.18}$$

Propensity score matching regression

In order to determine the control group for the difference-in-differences specification a propensity score matching approach was employed. We used a stepwise elimination approach in order to determine which variables have a significant impact on propensity score. With a significance

---

[4] Where the theoretical locations x have been replaced by observed housing transactions i.

level criterion of 10% the following variables remained in the final CA propensity score estimation:

CA characteristics: Urban, Commercial, Residential, Industrial, World Heritage Site, At Risk and Article 4 Status.

Environmental characteristics: Land Cover Type 9 (Inland bare ground), Land Cover Type 3 (Mountains, moors and heathland), distance to nearest National Nature Reserve, distance to nearest National Park, National Park (kernel density) and Area of Outstanding Natural Beauty (kernel density).

Neighbourhood characteristics: Median Income and Ethnicity Herfindahl index.

Amenities: Distance to nearest Bar, distance to nearest Underground Station, distance to nearest Hospital, distance to nearest Motorway and distance to nearest TTWA centroid.

Semi-parametric temporal and spatial estimations of treatment effects

We estimate a semi-parametric version of (4.27) that replaces the YD_{it} variables with a full set of years-since-designation bins. We group transactions into bins depending on the number of years that have passed since the conservation area they fall into or are near to had been designated. Negative values indicate years prior to designation. These bins (b) are captured by a set of dummy variables PT_b:

$$p_{it} = \sum_b \beta_b^I (PT_i^b \times I_i) + \sum_b \beta_b^E (PT_i^b \times E_i) + \sum_b \beta_b PT_i^b + X_i^{'}\mu + f_n + Y_t + \epsilon_{it} \tag{B.19}$$

The parameters β_b^I and β_b^E give the difference in prices between treatment and control groups in each years-since-designation bin b. The results of this semi-parametric estimation are plotted in Figure B.2 on page 169 in Section B.5.2. In order to allow for a casual inspection of the fit of the parametric models the semi-parametric point-estimates are also plotted in Figure 4.2 on page 77 (internal) and Figure 4.3 on page 78 (external) of the main text.

As with the temporal models, we relax the parametric constraints of the spatial estimations by replacing the distance variable in equation with distance bins:

$$p_{it} = \sum_d \beta_d (DB_i^d \times T_i) + \sum_d \beta_d^{Post} (DB_i^d \times T_i \times Post_{it}) + X_i^{'}\mu + f_n + Y_t + \epsilon_{it} \tag{B.20}$$

where DB_i^d are positive (external) and negative (internal) distance bins from the designation area boundary and β_d^{Post} are d treatment effect parameters at different distances inside and outside the conservation area. If the planner designates in a Pareto-optimal manner then the bin that corresponds to the locations just inside the treated conservation area should indicate a zero treatment effect. This may or may not be associated with a positive effect for the bins deepest inside the conservation area. Furthermore, if there are significant externalities associated with the designation (and heritage in general) then the bins just outside the boundary should indicate a positive effect. A lower effect for further out bins would indicate a spatial decay to this externality. The results from this specification are presented Figure B.3 on page 170 in Section B.5.2 and in Figure 4.4 on page 81 of the main text.

B.4. Data

B.4.1. Data sources

Housing transactions

The transactions data relates to mortgages for properties granted by the Nationwide Building Society (NBS) between 1995 and 2010. The data for England comprise 1,088,446 observations and include the price paid for individual housing units along with detailed property characteristics. These characteristics include floor space (m^2), the type of property (detached, semi-detached, flat, bungalow or terraced), the date of construction, the number of bedrooms and bathrooms, garage or parking facilities and the type of heating. There is also some buyer information including the type of mortgage (freehold or leasehold) and whether they are a first-time buyer.

Importantly, the transaction data includes the full UK postcode of the property sold allowing it to be assigned to grid-reference coordinates. With this information it is possible within a Geographical Information System environment to calculate distances to conservation area borders and to determine whether the property lies inside or outside these borders. Furthermore it is possible to calculate distances and other spatial measures (e.g. densities) for the amenities and environmental characteristics that will be used as control variables. Since the data set refers to postcodes rather than individual properties, it is not possible, however, to analyze repeated sales of the same property. This is a limitation shared with most property transaction data sets available in England, including the land registry data.

Neighbourhood characteristics

The main variables used for estimating capitalization effects of neighbourhood characteristics are median income and ethnic composition. The income data is a model-based estimate of median household income produced by Experian for Super Output Areas of the lower level (LSOA). This is assigned to the transaction data based on postcode. The data on ethnicity was made available by the 2001 UK Census at the level of Output Area (OA). Shares of each of the 16 ethnic groups and a Herfindahl index were computed to capture the ethnic composition of neighbourhoods. The Herfindahl index (HI) is calculated according to the following relation: $HI = \sum_{i=1}^{N} s_i^2$, where s_i is the share of ethnicity i in the LSOA, and N is the total number of ethnicities.

Environmental variables

The environmental variables capture the amenity value of environmental designations, features of the natural environment, different types of land cover and different types of land use.

Geographical data (in the form of ESRI shapefiles) for UK National Parks, Areas of Outstanding Natural Beauty, and National Nature Reserves are available from Natural England. National Parks and Areas of Outstanding Natural Beauty are protected areas of countryside designated because of their significant landscape value. National Nature Reserves are "established to protect sensitive features and to provide 'outdoor laboratories' for research" (National England website). Straight line distances to these designations were computed for the housing units as geographically located by their postcodes. Furthermore, density measures that take into account both the distance to and the size of the features were created. We apply a kernel density measure (Silverman, 1986) with a radius of 2 km which is considered to be the maximum distance people are willing to walk (Gibbons & Machin, 2005).

The location of lakes, rivers and coastline are available from the GB Ordinance Survey. The distance to these features is also computed for the housing units from the transaction data. The UK Land Cover Map produced by the Centre for Ecology and Hydrology describes land coverage by 26 categories as identified by satellite images. We follow Mourato et al. (2010) who construct nine broad land cover types from the 26 categories. Shares of each of these nine categories in 1 km grid squares are calculated and the housing units take on the value of the grid square in which they reside.

The generalized Land Use Database (GLUD) available from the Department for Communities and Local Government gives area shares of nine different types of land use within Super Output Areas, lower level (LSOA). These nine land use types are domestic buildings, non-domestic buildings, roads, paths, rail, domestic gardens, green space, water, and other land use. These shares are assigned to the housing units based on the LSOA in which they are located.

Amenities

The locational amenities variables capture the benefits a location offers in terms of accessibility, employment opportunities, schools quality, and the proximity of cultural and entertainment establishments.

Employment accessibility is captured both by the distance to Travel to Work Area (TTWA) centroid and a measure of employment potentiality. TTWAs are defined such that 75 per cent of employees who work in the area also live within that area. Thus they represent independent employment zones and the distance to the centre of these zones is a proxy for accessibility to employment locations. A more complex measure of accessibility is the employment potentiality index (Ahlfeldt, 2011b).[5] This is computed at the Super Output Area, lower level (LSOA) and represents an average of employment in neighbouring LSOAs weighted by their distance.

Key Stage 2 (ages 7–11) assessment scores are available from the Department for Education at the Super Output Area, middle layer (MSOA). School quality is thus captured at the housing unit level by computing a distance-weighted average of the KS2 scores of nearby MSOA centroids.[6]

---

[5] Further detail on the construction of the employment potentiality measure is provided in Section B.4.2.

[6] This is calculated as an Inverse Distance Weighting (IDW) with a threshold distance of 5 km and a power of two.

Geographical data on the locations of motorways, roads, airports, rail stations and rail tracks are available from the GB Ordinance Survey. Distances were computed from housing units to motorways, A-roads, B-roads and rail stations to capture accessibility. Buffer zones were created around the motorways and roads along with distance calculations to rail tracks and airports in order to capture the disamenity noise effects of transport infrastructure.

Further data on local amenities were taken from the Ordinance Survey (police stations, places of worship, hospitals, leisure/sports centers) and OpenStreetMap (cafés, restaurants/fast food outlets, museums, nightclubs, bars/pubs, theatres/cinemas, kindergartens and monuments, memorials, monuments, castles, attractions, artwork). The number of listed buildings was provided by English Heritage. Kernel densities for these amenities were computed for housing units using a kernel radius of 2 km and a quadratic kernel function (Silverman, 1986). The radius of 2 km is consistent with amenities having a significant effect on property prices only when they are within walking distance.

Table B.2.: Variable description

Variable	Description
Dependent variable	
Price	Per square meter transaction price in Euro of the corresponding plot of land (expressed as natural logarithm). Transaction data from the Nationwide Building Society (NBS).
Independent variables	
CA Effects	Dummy variables denoting property transactions taking place within the boundaries of an currently existing conservation area, in a conservation area at the time when designated or where the designation date is unknown as well as various buffer areas surrounding current or treated conservation areas.
Fixed Effect Control	Travel to Work Areas, nearest conservation area catchment areas and interactives with year effects.
Housing information	Set of property variables from the NBS including: Number of bedrooms, number of bathrooms, floor size (in square meter), new property (dummy), building age (years), tenure (leasehold/freehold), central heating (full: gas, electric, oil, solid fuel), central heating (partial: gas, electric, oil, solid fuel), garage (single or double), parking space, property type (detached, semi-detached, terraced, bungalow, flat-maisonette).

Neighbourhood information	Set of neighbourhood variables including: media income (2005, LSOA level), share of white population at total population (2001 census, output area level), share of mixed population at total population (2001 census, output area level), share of black population at total population (2001 census, output area level), share of Asian population at total population (2001 census, output area level), share of Chinese population at total population (2001 census, output area level), Herfindahl of ethnic segregation (including population shares of White British, White Irish, White others, Mixed Caribbean, Mixed Asian, Mixed Black, Mixed other, Asian Indian, Asian Pakistani, Asian others, Black Caribbean, Black African, Black other, Chinese, Chinese other population, 2001 census output area).
Conservation area Characteristics	Set of characteristic variables for conservation areas from English Heritage including: Conservation area land use (dummy variables for residential, commercial, industrial or mixed land use), conservation area type (dummy variable for urban, suburban or rural type), conservation area size (dummy for areas larger than mean of 128,432.04 square meters), conservation area (square meter), conservation area has an Article 4 Direction implemented (dummy), oldness of conservation area (dummy for areas older than mean of 1981), conservation area at risk (dummy), conservation area with community support (dummy), conservation area is World Heritage Site (dummy), score measures on the condition, vulnerability and trajectory.
Environment Characteristics and Amenities	Set of locational variables processed in GIS including: National Parks (distance to, density), Areas of Outstanding Beauty (distance to, density), Natural Nature Reserves (distance to, density), distance to nearest lake, distance to nearest river, distance to nearest coastline, land cover in 1 km square: Marine and coastal margins; freshwater, wetland and flood plains; mountains, moors and heathland; semi-natural grassland; enclosed farmland; coniferous woodland; broadleaved/mixed woodland; urban; inland bare ground.

Other amenities	Set of locational variables created in GIS including: Average key stage 2 test score (MSOA averages as well as interpolated in GIS), distance to electricity transmission lines, A-Roads (distance to, buffer dummy variables within 170m), B-Roads (distance to, buffer dummy variable within 85m), motorway (distance to, buffer dummy variable within 315m; buffer distances refer to the distance were noise of maximum speed drops drown to 50 decibel), distance to all railway stations, distance to London Underground stations, distance to railway tracks, distance to bus stations, distance to airports, densities of cafés, restaurants/fast food places, museums, nightclubs, bars/pubs, theatres/cinemas, kindergartens, monuments (memorial, monument, castles, attraction, artwork), hospitals, sports/leisure centres, police stations and worship locations, distance to Travel to Work Areas, employment potentiality (based on Travel to Work Areas with an time decay parameter of 0.073).
Neighbourhood Distance Controls	Set of neighbourhood distance dummy variables created in GIS including: Distances outside conservation area border (up to 50 m, 100 m, 150 m, 200 m, 250 m, 300 m, 350 m, 400 m, 1 km, 2 km and 3 km), distances inside conservation area border (up to 50 m, 100 m, 150 m, 200 m).

B.4.2. Further notes on data methods

Employment potentiality

The employment potentiality index is computed at the Super Output Area, lower level (LSOA) and represents an average of employment in neighbouring LSOAs weighted by their distances. Employment potentiality is calculated for each Lower Layer Super Output Area i (LSOA) based on employment in all other LSOAs j using the following equation:

$$EP_i = \sum_j E_j e^{-ad_{ij}}, \text{ with } i \neq j \tag{B.21}$$

where d measures the straight line distance converted into travel time assuming an overall average speed of 25km/h (Department for Transport, 2009) and Employment the absolute number of workers in the respective LSOA. The indicator is weighted by a decay parameter of $a = -0.073$ estimated by Ahlfeldt (in press). Internal distances are calculated as:

$$d_{ii} = \frac{1}{3}\sqrt{\frac{Area_i}{\pi}} \tag{B.22}$$

Kernel densities for National Parks, Areas of Outstanding Natural Beauty and National Nature Reserves

The kernel density is a measure that takes into account both the proximity and the size of National Parks (NPs), Areas of Outstanding Natural Beauty (AONBs) and National Nature Reserves (NNRs). Every 100x100 m piece of designated area is assigned a point and the density of these resulting points calculated for 10 km kernels and a quadratic kernel function (Silverman, 1986, p. 76, equation 4.5) around each housing unit using a kernel density method. The result is similar to calculating a share of NP area within a circle, the one difference being that the points are additionally weighted by distance to the housing units according to a normal distribution.

Buffers for motorways and roads

The buffer sizes for the different roads are as follows: B-Road (85 m), A-Road (170 m) and Motorway (315 m). These distances are calculated based on how far it is expected that the noise from traffic travelling at the speed limit of the respective roads (Steven, 2005) would decline to an assumed disamenity threshold level of noise of 50 db (J. P. Nelson, 2008).

Land cover map Broad Categories

Table B.3.: Land Cover Broad categories

1	Marine and coastal margins
2	Freshwater, wetlands, and flood plains
3	Mountains, moors, and heathland
4	Semi-natural grasslands
5	Enclosed farmland
6	Coniferous woodland
7	Broad-leaved/mixed woodland
8	Urban
9	Inland bare ground

Broad categories as defined by Mourato et al. (2010).

B.5. Estimation results

B.5.1. Designation process

In order to test our theoretical implication that changes in heritage preferences lead to changes in designation we estimate the regression model as outlined in Section 4.3.1. The prediction of the model is that positive changes in heritage preferences should lead to negative changes in the share of non-designated land in a neighbourhood. OLS regression results are reported in Table B.4 on page 155. We drop all zeros and identify the effect based on the sample of observations with observable changes in conservation area shares. The standard OLS estimates without (1) and with a basic set of composition controls (2) are insignificant. Due to the potential sources of bias in OLS discussed in the main part we re-estimate the two models using our instrumental variables. The 2SLS estimates (3) and (4) are in line with the tobit results reported in the main paper and support the theory that a positive change in degree share leads to higher designation.

Table B.4.: Designation process

	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
	Δ log non designation share$_t$			
Δ log degree share$_t(\vartheta)$	-0.009	-0.017	-0.674***	-0.279***
	(0.012)	(0.015)	(0.113)	(0.085)
log degree share$_{t-1}$		-0.009		-0.106***
		(0.013)		(0.033)
log designation share$_{t-1}$		0.169***		0.147***
		(0.010)		(0.023)
Δ log homeownership$_t$		0.120***		0.148***
		(0.029)		(0.035)
log homeownership$_{t-1}$		0.025		-0.030
		(0.022)		(0.026)
Δ log aver. household size$_t$		-0.015		-0.067
		(0.064)		(0.052)
log aver. household size$_{t-1}$		0.005		-0.041
		(0.030)		(0.064)
Constant	-0.044***	-0.033	0.430***	-0.017
	(0.012)	(0.031)	(0.081)	(0.062)
IV	No	No	Yes	Yes
Controls	No	Yes	No	Yes
R^2	0.000	0.075	-1.023	0.015
F	0.604	139.420	35.637	19.008
AIC	-931.737	-1045.816	210.758	-944.019
Overid			0.966	2.372
OveridP			0.326	0.124
Observations	1621	1621	1621	1621

Notes: See the data section for a description of control variables. IVs are station density and employment potential. Standard errors in parentheses and clustered on fixed effects in (3). *p< 0.05, **p< 0.01, ***p< 0.001.

Table B.5 on page 157 reports the first stage results to the second-stage results reported in Table 4.1 on page 70 in the main text. Both IVs are (conditionally) positively correlated with the change in degree share, as expected.

Table B.5.: Standard IV models – First stage regressions

	(1)	(2)	(3)	(4)	(5)	(6)
			Δ log degree share$_t$			Δ log degree share$_t \times$ homeownership$_{t-1}$
rail station density	0.124***	0.112***	0.124***	0.103***	0.128***	-0.007***
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.002)
employment potentiality	3.2E-08***	2.5E-08***	3.2E-08***	3.8E-08***	3.2E-08***	9.5E-10***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
predicted Δ log degree share$_t \times$ homeownership$_{t-1}$					-0.403***	1.118***
					(0.048)	(0.009)
log degree share$_{t-1}$	-0.409***	-0.454***	-0.409***	-0.429***	-0.400***	0.002***
	(0.013)	(0.009)	(0.004)	(0.004)	(0.004)	(0.001)
log designation share$_{t-1}$	-0.021	-0.20	-0.021***	-0.001	-0.025***	-0.005***
	(0.017)	(0.015)	(0.008)	(0.008)	(0.008)	(0.001)
Δ log homeownership$_t$	0.293***	0.339***	0.294***	0.376***	0.275***	0.012***
	(0.060)	(0.065)	(0.016)	(0.017)	(0.016)	(0.003)
log homeownership$_{t-1}$	0.016	0.054	0.016*	0.073***	0.210***	-0.060***
	(0.021)	(0.031)	(0.011)	(0.011)	(0.025)	(0.005)
Δ log aver. household size$_t$	-0.075	-0.140*	-0.075***	-0.032*	-0.052***	0.020***
	(0.084)	(0.055)	(0.024)	(0.024)	(0.024)	(0.004)
log aver. household size$_{t-1}$	-0.170	-0.315***	-0.169***	-0.070***	0.171***	0.018***
	(0.087)	(0.090)	(0.028)	(0.005)	(0.028)	(0.005)
log price trend			0.004			
			(0.012)			
Δ log vacancy rate$_t$				0.024***		
				(0.004)		
log vacancy rate$_{t-1}$				0.070***		
				(0.005)		
log listed building density				0.017***		
				(0.002)		

aver. condition score (1 best, 4 worst)				0.011		
				(0.014)		
aver. vulnerability score (1 low, 8 high)				-0.013		
				(0.013)		
aver. trajectory score (-2 improving, +2 deteriorating)				0.002		
				(0.027)		
Constant	0.017	0.136	0.025	0.316***	0.106***	-0.032***
	(0.098)	(0.107)	(0.039)	(0.034)	(0.030)	(0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	No	Yes	No	No	No	No
Price trend	No	No	Yes	No	No	No
Housing Conditions	No	No	No	Yes	No	No
F	420.662	123.00	1756.16	1320.37	2093.28	18708.76
R^2	0.688	0.732	0.688	0.699	0.703	0.955
Observations	7965	7965	7965	7965	7965	7965

Notes: See the data section for a description of control variables. IVs are station density and employment potential in all models. Model (4) includes a dummy variable indicating 60 wards for which no price trend could be computed due to insufficient transactions. We derive the instrument (predicted $\Delta \log \text{degree share}_t \times \text{homeownership}_{t-1}$) for the interaction term in model (5) by interacting $\text{homeownership}_{t-1}$ with the predicted values of an auxiliary regression where we regress $\Delta \log$ degree share on the exogenous variables, i.e. on the standard IVs and controls. Standard errors in parentheses and clustered on fixed effects, *$p< 0.05$, **$p< 0.01$, ***$p< 0.001$.

We have tried four alternative IV models which are based on the benchmark model, i.e. including the set of controls (Table 4.1, column 2 in the main text). The coefficient estimates reported in Table B.6 on page 158 remain qualitatively similar and quantitatively close to the main model. First stage results are reported in Table B.7 on page 159. The alternative instruments, again, pass the validity tests. Only the overidentification test is failed by specification (1) using employment potentiality and museum density as instruments.

Table B.6.: Alternative IV models

	(1)	(2)	(3)	(4)
	Δ log non designation share$_t$			
Δ log degree share$_t(\vartheta)$	-0.488***	-0.512***	-0.502***	-0.557***
	(0.063)	(0.065)	(0.063)	(0.070)
log degree share$_{t-1}$	-0.265***	-0.274***	-0.270***	-0.291***
	(0.025)	(0.025)	(0.025)	(0.027)
log designation share$_{t-1}$	-0.020	-0.022*	-0.022*	-0.021
	(0.013)	(0.013)	(0.013)	(0.013)
Δ log homeownership$_t$	0.259***	0.263***	0.262***	0.251***
	(0.029)	(0.029)	(0.029)	(0.028)
log homeownership$_{t-1}$	0.065***	0.061***	0.061***	0.046**
	(0.018)	(0.018)	(0.018)	(0.019)
Δ log aver. household size$_t$	0.033	0.033	0.032	0.013
	(0.040)	(0.040)	(0.040)	(0.039)
log aver. household size$_{t-1}$	-0.009	-0.014	-0.012	-0.010
	(0.049)	(0.050)	(0.049)	(0.050)
Constant	0.049	0.052	0.051	0.040
	(0.048)	(0.048)	(0.048)	(0.048)
IV	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Chi2	340.356	341.226	342.655	331.908
EXOGP	0.000	0.000	0.000	0.000
Overid	3.544	0.078	0.752	0.201
OveridP	0.060	0.780	0.386	0.654
Observations	7965	7965	7965	7968
Instruments (as densities except employment potentiality)	Employment potentiality Museum	Employment potentiality Coffee place	Employment potentiality Bar	Rail station Coffee place

Notes: See the data section for a description of control variables. Standard errors in parentheses and clustered on fixed effects, *p< 0.05, **p< 0.01, ***p< 0.001.

Table B.7.: Alternative IV models - First stage regressions

	(1)	(2)	(3)	(4)
	Δ log degree share$_t$			
employment potentiality	4.4E-08***	4.3E-08***	4.1E-08***	
	(0.000)	(0.000)	(0.000)	

museum density	0.142***			
	(0.034)			
coffee place density		0.004***		-0.011***
		(0.001)		(0.003)
bar density			0.006**	
			(0.002)	
rail station density				0.282***
				(0.022)
log degree share$_{t-1}$	-0.406***	-0.406***	-0.408***	-0.399***
	(0.012)	(0.012)	(0.012)	(0.010)
log designation share$_{t-1}$	-0.019	-0.022	-0.018	-0.031*
	(0.015)	(0.016)	(0.016)	(0.012)
Δ log homeownership$_t$	0.285***	0.274***	0.285***	0.260***
	(0.065)	(0.063)	(0.068)	(0.061)
log homeownership$_{t-1}$	0.007	-0.002	0.011	-0.038
	(0.024)	(0.024)	(0.025)	(0.031)
Δ log aver. household size$_t$	-0.101	-0.098	-0.088	-0.040
	(0.078)	(0.077)	(0.079)	(0.108)
log aver. household size$_{t-1}$	-0.192*	-0.204*	-0.188*	-0.127
	(0.083)	(0.084)	(0.082)	(0.086)
Constant	0.039	0.051	0.035	-0.015
	(0.092)	(0.094)	(0.091)	(0.091)
Controls	Yes	Yes	Yes	Yes
F	396.188	517.118	441.850	552.553
R^2	0.686	0.685	0.686	0.681
Observations	7965	7965	7965	7968

Notes: See the data section for a description of control variables. Standard errors in parentheses and clustered on fixed effects, *p< 0.05, **p< 0.01, ***p< 0.001.

Furthermore, we have split the long difference between 1991 and 2011 into two shorter differences of 1991 to 2001 and 2001 to 2011 (Table B.8 on page 160). For the latter short difference we moreover used the change in income instead of change in degree as a proxy for heritage preferences (Table B.10 on page 162). The coefficient estimates remain qualitatively similar to the main model. Their first stages are reported in Table B.9 on page 161. The coefficient of the key variable is slightly smaller in the benchmark specification of the short different between 1991 and 2001 (Table B.8, column 4) and considerably larger for the period between 2001 and 2011 (column 8). This could be explained with an increase in gentrification over time. In Table B.10 we use income as a proxy of heritage preference. Focusing on the benchmark specification in the final column, doubling income more than quadruples the designation share. The respective instruments are valid and sufficiently strong. Overall, the results are in line with our theory; increases in heritage preferences, proxied by change in degree or change in income, lead to increases in designation shares.

Table B.8.: Short differences models

	1991-2001				2001-2011			
	Δ log non designation share$_t$				Δ log non designation share$_t$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ log degree share$_t(\vartheta)$	-0.013	-0.229***	-0.064***	-0.193***	0.464***	1.618***	-0.066	-2.790***
	(0.008)	(0.022)	(0.011)	(0.031)	(0.051)	(0.124)	(0.077)	(0.910)
log degree share$_{t-1}$			-0.063***	-0.291***			-0.140***	-0.689***
			(0.005)	(0.027)			(0.022)	(0.036)
log designation share$_{t-1}$			-0.034***	-0.021			0.022	0.036
			(0.011)	(0.013)			(0.018)	(0.022)
Δ log homeownership$_t$			0.109***	0.251***			0.225***	0.895***
			(0.024)	(0.028)			(0.084)	(0.236)
log homeownership$_{t-1}$			0.061***	0.046**			0.185***	0.390***
			(0.010)	(0.019)			(0.025)	(0.074)
Δ log aver. household size$_t$			0.039	0.013			-0.220*	-0.813***
			(0.034)	(0.039)			(0.132)	(0.257)
log aver. household size$_{t-1}$			0.090***	-0.010			0.108	-0.140
			(0.031)	(0.050)			(0.068)	(0.107)
Constant	0.153***	0.231***	-0.011	0.040	0.311***	-0.125***	0.246***	0.615***
	(0.005)	(0.010)	(0.031)	(0.048)	(0.021)	(0.042)	(0.066)	(0.140)
IV	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	No	Yes	Yes	No	No	Yes	Yes
Chi2		106.812		215.197		171.695		169.534
EXOGP		0.000		0.000		0.000		0.000
Overid		2.165		1.118		1.485		15.948
OveridP		0.141		0.276		0.223		0.000
Observations	7968	7965	7968	7965	7969	7966	7969	7966

Notes: See the data section for a description of control variables. Standard errors in parentheses and clustered on fixed effects, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table B.9.: Short differences and income models - first stage regressions

	1991-2001		2001-2011		2001-2011	
	Δ log degree share$_t$		Δ log degree share$_t$		Δ log income$_t$	
	(1)	(2)	(3)	(4)	(5)	(6)
rail station density	0.117***	0.102***	-0.061***	0.038***	-0.012	0.018
	(0.032)	(0.012)	(0.010)	(0.007)	(0.037)	(0.029)
employment potentiality	5.14E-8***	4.99E-8***		-1.87E-9	5.69E-9*	5.44E-9
	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
log degree share$_{t-1}$		-0.262***		-0.209***		
		(0.011)		(0.009)		
log income$_{t-1}$						-0.095***
						(0.020)
log designation share$_{t-1}$		-0.040*		0.006		-0.011*
		(0.018)		(0.006)		(0.005)
Δ log homeownership$_t$		0.411***		0.253***		-0.017
		(0.083)		(0.029)		(0.064)
log homeownership$_{t-1}$		-0.038		0.092***		0.040*
		(0.022)		(0.009)		(0.017)
Δ log aver. household size$_t$		-0.145*		-0.217***		0.220***
		(0.064)		(0.065)		(0.037)
log aver. household size$_{t-1}$		-0.236**		-0.069*		0.130**
		(0.077)		(0.030)		(0.044)
Constant	0.327***	-0.036	0.390***	0.112***	0.255***	0.741***
	(0.008)	(0.083)	(0.005)	(0.004)	(0.004)	(0.113)
Controls	No	Yes	No	Yes	No	Yes
F	34.876	443.629	74.997	544.976	8.308	12.770
R^2	0.103	0.504	0.095	0.602	0.004	0.068
Observations	7965	7965	7966	7966	7966	7966

Notes: See the data section for a description of control variables. Standard errors in parentheses and clustered on fixed effects, *$p < 0.05$,**$p < 0.01$, ***$p < 0.001$.

Table B.10.: Income models

	2001-2011			
	Δ log non designation share$_t$			
	(1)	(2)	(3)	(4)
Δ log income$_t$	-0.204***	-9.152***	-0.210***	-4.357***
	(0.068)	(1.981)	(0.067)	(0.959)
log income$_{t-1}$			-0.201***	-0.498***
			(0.026)	(0.078)
log designation share$_{t-1}$			0.032*	-0.019
			(0.018)	(0.028)
Δ log homeownership$_t$			0.278***	0.134
			(0.082)	(0.113)
log homeownership$_{t-1}$			0.225***	0.244**
			(0.026)	(0.036)
Δ log aver. household size$_t$			-0.227*	0.774***
			(0.130)	(0.276)
log aver. household size$_{t-1}$			0.198***	0.671***
			(0.070)	(0.145)
Constant	0.535***	2.825***	1.702***	4.203***
	(0.026)	(0.513)	(0.179)	(0.630)
IV	No	Yes	No	Yes
Controls	No	No	Yes	Yes
Chi2			21.347	122.739
EXOGP		0.000		0.000
Overid		13.591		0.061
OveridP		0.000		0.000
Observations	7969	7966	7969	7966

Notes: See the data section for a description of control variables. Standard errors in parentheses and clustered on fixed effects, *p< 0.05, **p< 0.01, ***p< 0.001.

B.5.2. Pareto optimality

Table B.11 on page 167 reports the conservation area effects as well as the full set of hedonic controls, housing characteristics in particular, for the difference-in-differences estimation given by equation (4.26) in the main text. Column (7) shows that housing units with more bathrooms and bedrooms fetch higher prices, as do detached, semi-detached, and bungalows (over the omitted category flats/maisonettes). The sales price of terraced housing is insignificantly different from flats/maisonettes. Larger floor spaces are associated with higher price but with significant diminishing effects. There is a premium for new properties. Leased properties are of less value than those owned. Properties with parking spaces, single garages and double garages sell for higher prices than those without any parking facilities. There is a house price premium for properties with central heating over other types of heating. In order to control for a potentially non-linear relationship between housing age and house prices we included a series of house age bins. In order to separate the effects of pure building age (which may be associated with deterioration) from the build date (which may strongly determine the architectural style) we allow for age cohort and building data cohort effects. Since the ‘New property’ variable identifies all properties where the build age is zero years, the omitted category from the age variables is 1–9 years. All of the bins for properties older than this indicate significant negative premiums. The negative premium increases with age, mostly quickly over the first few categories and then more slowly until the penultimate category and finally decreases for buildings over 100 years. The effect of the build date is also non-linear. The general tendency is for buildings built in earlier periods to have higher prices than buildings built in the omitted period 2000–2010. However, this effect becomes insignificant in the 60s and 70s; periods associated with the architectural styles of the post-ward reconstruction phase that are today less appreciated than other styles. The greatest premium is attached to houses built pre-1900, the earliest category.

Table B.11.: Conservation area premium – designation effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				log property transaction price			
Inside treated CA ×	0.028***	0.014	0.014	0.003	-0.024	-0.077	-0.003
Post designation	(0.009)	(0.009)	(0.010)	(0.012)	(0.070)	(0.111)	(0.015)
Within 500m buffer of	0.023***	0.013***	0.012***	0.004	0.012	-0.005	-0.005
treated CA × Post des.	(0.004)	(0.004)	(0.005)	(0.006)	(0.027)	(0.022)	(0.010)
Inside treated CA	-0.043***	-0.038***	-0.048***	-0.037***	-0.062	0.029	-0.024
	(0.009)	(0.009)	(0.010)	(0.012)	(0.057)	(0.108)	(0.021)
Within 500m buffer	-0.10**	-0.004	-0.011**	-0.005	0.003	0.006	-0.002
of treated CA	(0.004)	(0.004)	(0.005)	(0.005)	(0.030)	(0.023)	(0.013)
No. of bathrooms	0.007***	0.007***	0.006***	0.013***	0.057***	0.059***	0.014***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.008)	(0.006)	(0.002)
No. of bedrooms	0.166***	0.172***	0.169***	0.165***	0.170***	0.179***	0.158***
	(0.002)	(0.004)	(0.005)	(0.005)	(0.014)	(0.011)	(0.006)
No. of bedrooms^2	-0.019***	-0.020***	-0.020***	-0.019***	-0.019***	-0.019***	-0.018***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Detached house	0.254***	0.222***	0.211***	0.194***	0.235***	0.216***	0.193***
	(0.003)	(0.005)	(0.008)	(0.007)	(0.015)	(0.014)	(0.007)
Semi-detached house	0.119***	0.097***	0.088***	0.070***	0.082***	0.066***	0.073***
	(0.003)	(0.004)	(0.007)	(0.006)	(0.014)	(0.012)	(0.006)
Terraced house /	0.040***	0.026***	0.015**	0.001	0.002	-0.013	-0.000
country cottage	(0.003)	(0.004)	(0.006)	(0.006)	(0.013)	(0.012)	(0.006)
Bungalow	0.311***	0.285***	0.281***	0.257***	0.292***	0.269***	0.257***
	(0.003)	(0.006)	(0.008)	(0.009)	(0.019)	(0.016)	(0.009)
Floor size (m^2)	0.006***	0.006***	0.007***	0.007***	0.008***	0.007***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Floorsize^2	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
New property	0.084***	0.087***	0.088***	0.088***	0.047***	0.076***	0.077***
	(0.002)	(0.004)	(0.005)	(0.006)	(0.024)	(0.017)	(0.006)

Leasehold	-0.054***	-0.067***	-0.065***	-0.073***	-0.100***	-0.104***	-0.070***
	(0.003)	(0.004)	(0.006)	(0.006)	(0.014)	(0.012)	(0.006)
Single garage	0.112***	0.097***	0.100***	0.097***	0.096***	0.097***	0.098***
	(0.001)	(0.002)	(0.003)	(0.003)	(0.007)	(0.005)	(0.003)
Double garage	0.190***	0.162***	0.161***	0.159***	0.160***	0.156***	0.158***
	(0.002)	(0.003)	(0.005)	(0.005)	(0.015)	(0.010)	(0.005)
Parking space	0.076***	0.063***	0.065***	0.061***	0.052***	0.049***	0.063***
	(0.001)	(0.002)	(0.003)	(0.003)	(0.007)	(0.005)	(0.003)
Central heating	0.089***	0.094***	0.098***	0.100***	0.085***	0.094***	0.095***
	(0.001)	(0.002)	(0.003)	(0.003)	(0.007)	(0.007)	(0.003)
Building age: 10-19 years	-0.047***	-0.063***	-0.062***	-0.075***	-0.071***	-0.068***	-0.069***
	(0.002)	(0.003)	(0.004)	(0.005)	(0.016)	(0.015)	(0.005)
Building age: 20-29 years	-0.079***	-0.106***	-0.104***	-0.125***	-0.133***	-0.126***	-0.113***
	(0.002)	(0.005)	(0.007)	(0.008)	(0.026)	(0.021)	(0.007)
Building age: 30-39 years	-0.092***	-0.127***	-0.123***	-0.150***	-0.169***	-0.141***	-0.133***
	(0.003)	(0.006)	(0.010)	(0.011)	(0.032)	(0.027)	(0.009)
Building age: 40-49 years	-0.104***	-0.148***	-0.142***	-0.180***	-0.199***	-0.165***	-0.158***
	(0.004)	(0.008)	(0.012)	(0.013)	(0.036)	(0.031)	(0.011)
Building age: 50-59 years	-0.121***	-0.171***	-0.167***	-0.207***	-0.232***	-0.204***	-0.175***
	(0.004)	(0.009)	(0.015)	(0.016)	(0.044)	(0.038)	(0.014)
Building age: 60-69 years	-0.135***	-0.198***	-0.194***	-0.238***	-0.320***	-0.265***	-0.215***
	(0.005)	(0.011)	(0.019)	(0.020)	(0.051)	(0.042)	(0.018)
Building age: 70-79 years	-0.136***	-0.213***	-0.207***	-0.263***	-0.326***	-0.273***	-0.234***
	(0.006)	(0.013)	(0.021)	(0.022)	(0.053)	(0.046)	(0.019)
Building age: 80-89 years	-0.132***	-0.218***	-0.213***	-0.277***	-0.339***	-0.313***	-0.243***
	(0.007)	(0.014)	(0.023)	(0.024)	(0.062)	(0.054)	(0.021)
Building age: 90-99 years	-0.111***	-0.208***	-0.204***	-0.280***	-0.360***	-0.304***	-0.248***
	(0.008)	(0.016)	(0.025)	(0.027)	(0.068)	(0.063)	(0.023)
Building age: over 100 years	-0.083***	-0.176***	-0.176***	-0.261***	-0.348***	-0.284***	-0.227***
	(0.009)	(0.017)	(0.027)	(0.030)	(0.074)	(0.065)	(0.025)
Build date: 1900-1909	0.040***	0.121***	0.128***	0.208***	0.256***	0.222***	0.173***
	(0.009)	(0.018)	(0.028)	(0.031)	(0.077)	(0.067)	(0.025)

Build date: 1910-1919	0.074***	0.153***	0.158***	0.226***	0.262***	0.256***	0.196***
	(0.008)	(0.016)	(0.027)	(0.028)	(0.071)	(0.059)	(0.024)
Build date: 1920-1929	0.093***	0.157***	0.162***	0.215***	0.225***	0.189***	0.190***
	(0.007)	(0.014)	(0.024)	(0.025)	(0.062)	(0.050)	(0.021)
Build date: 1930-1939	0.082***	0.128***	0.130***	0.168***	0.187***	0.163***	0.151***
	(0.006)	(0.013)	(0.021)	(0.023)	(0.058)	(0.045)	(0.020)
Build date: 1940-1949	0.040***	0.078***	0.078***	0.111***	0.063***	0.053***	0.096***
	(0.005)	(0.012)	(0.018)	(0.021)	(0.058)	(0.048)	(0.018)
Build date: 1950-1959	0.017***	0.033***	0.041***	0.057***	0.017	-0.004	0.046***
	(0.004)	(0.010)	(0.016)	(0.018)	(0.047)	(0.039)	(0.015)
Build date: 1960-1969	0.001	0.007	0.018	0.023	-0.017	-0.012	0.011
	(0.004)	(0.009)	(0.013)	(0.015)	(0.044)	(0.037)	(0.013)
Build date: 1970-1979	-0.015***	-0.016***	-0.008	-0.004	-0.059	-0.046	-0.011
	(0.003)	(0.007)	(0.011)	(0.012)	(0.042)	(0.033)	(0.011)
Build date: 1980-1989	0.013***	0.017***	0.025***	0.029***	-0.023	-0.010	0.024***
	(0.003)	(0.006)	(0.008)	(0.010)	(0.038)	(0.029)	(0.008)
Build date: 1990-1999	0.022***	0.020***	0.022***	0.029***	-0.020	-0.008	0.017**
	(0.002)	(0.005)	(0.006)	(0.008)	(0.034)	(0.025)	(0.008)
Build date: pre 1900	0.098***	0.149***	0.162***	0.244***	0.312***	0.259***	0.216***
	(0.009)	(0.018)	(0.029)	(0.031)	(0.081)	(0.070)	(0.026)
Location controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Neighbourhood controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward effects	Yes	Yes	No	No	No	No	No
Nearest treated CA effects	No	No	Yes	Yes	Yes	Yes	No
Matched CA effects	No	No	No	No	No	No	Yes
Treatment group: CAs designated	1996-2010	1996-2010	1996-2010	1996-2010	1996-2002	1996-2002	1996-2010
Control group	Full England sample	Within 2km of treated CA	Within 2km of treated CA	Within 500m of CA designated before 1996 & within	Within 500m of CA designated before 1987-95 &	Within 500m of CA designated before 2003-10 &	Within 500m of pre-1996 CA matched on propen-

				2km of treated CA	within 2km of treated CA	within 2km of treated CA	sity score
R^2	0.921	0.922	0.915	0.915	0.861	0.864	0.909
AIC	-587375.2	-156426.4	-130469.1	-67046.3	-5408.8	-8475.7	-41184.2
Observation	1,088k	302k	302k	178k	21k	32k	133k

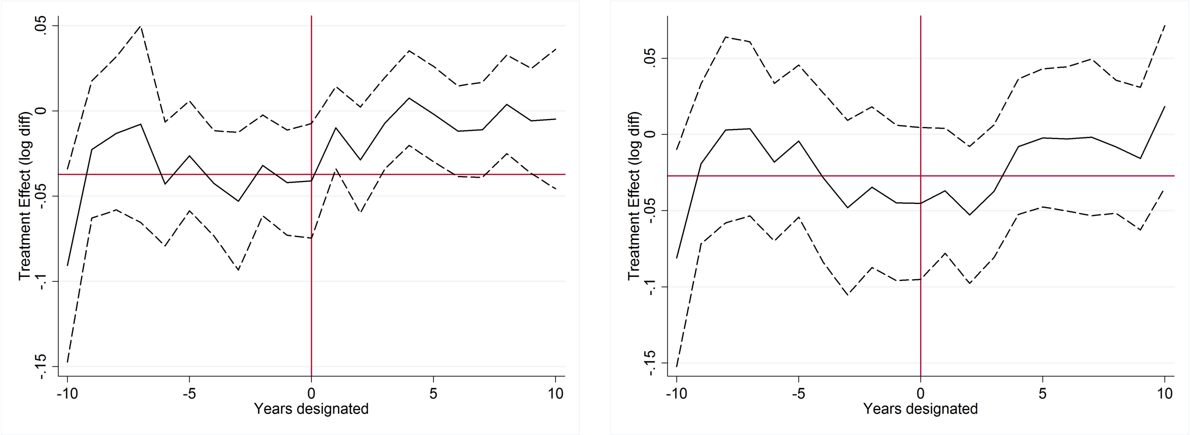
Notes: Standard errors in parentheses are clustered on location fixed effects. Conservation area control groups in columns (4)-(7) have separate fixed effects for the areas inside and outside a conservation area. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Semi-parametric temporal and spatial treatment effects

Figure B.2 on page 169 reports the results for the semi-parametric estimation of the temporal effects of designation using equation (B.19). Instead of simply presenting our two strongest specifications, as we do in the main text, here we present a different dimension to the results bin by comparing the bin estimates for the naïve DD in the left panels to the matched CA control group in the right panels. The left charts show that the post-period internal and external estimates deviate significantly from the pre-period mean (hence the significant DD estimates) but that this is driven by a general upward trends. This corroborates the results in Table 4.2 on page 73, column (1) of the main text where no significant discontinuity nor shift in trend for the naïve control group exists and hence the advantages of the RDD-DD over the standard DD method is high-lighted. The charts in the right panels also corroborate the evidence presented using the parametric trends equations in the main text. Specifically, they show that for the internal effects the post-treatment estimates tend not to deviate significantly from the pre-treatment effects but that there are upward shifts in the trend when compared to the pre-treatment trend. For the external effects there is a general upward trend in the less carefully matched control groups and a downward trend in the stronger control groups but no shift in the trend at the designation date.

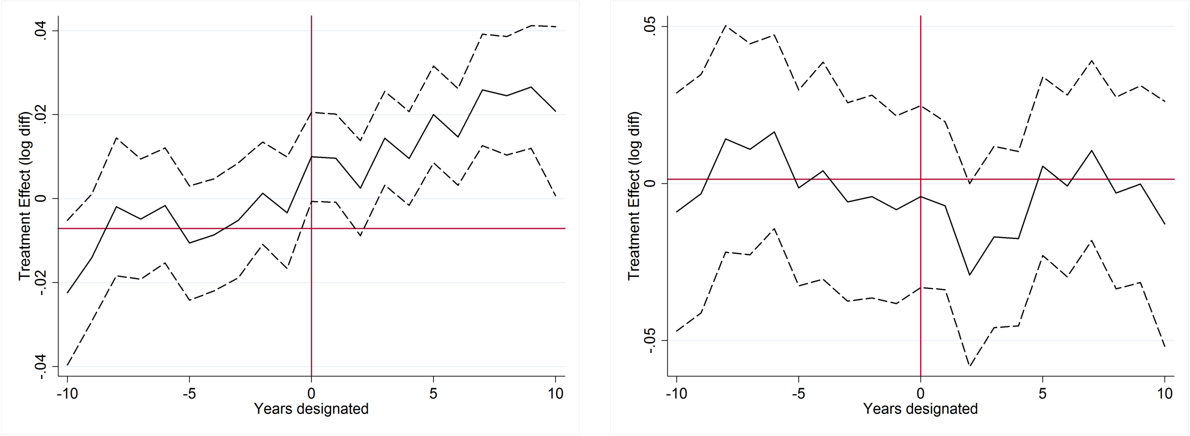
Figure B.3 on page 170 demonstrates the semi-parametric spatial effects using different bin sizes of 100 m and 200 m using appendix equation (B.20). These semi-parametric charts closely resemble their parametric counterparts. Notably, there is no significant and positive effect in the first bin outside the conservation area when using the preferred specification of column (7) from Table B.11 on page 167. This is consistent with the parametric findings and baseline DD findings that there is no significant external policy effect and that our second hypothesis cannot be accepted. There is, however, one significant bin inside the conservation area at 200-300 m. This provides some support for the idea that heritage externalities are stronger deeper within the conservation areas such that there may be a positive policy effect. This effect then declines to zero for the deepest bin of greater than 300 m.

Figure B.2.: Semi-parametric temporal bins estimates



(a) Internal effects: Full dataset
Tab. B.11, column (1)

(b) Internal effects: Matched CA
Tab. B.11, column (7)

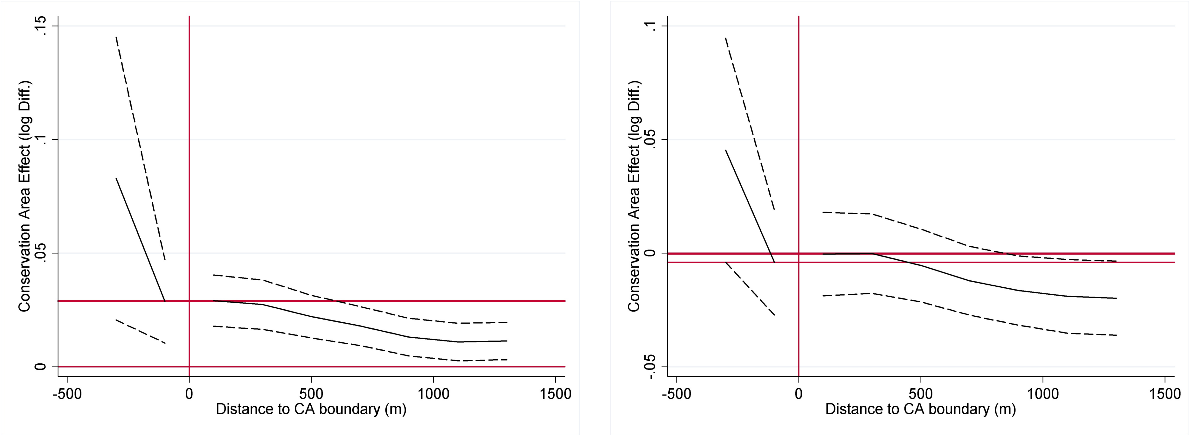


(c) External effects: Full dataset
Tab. B.11, column (1)

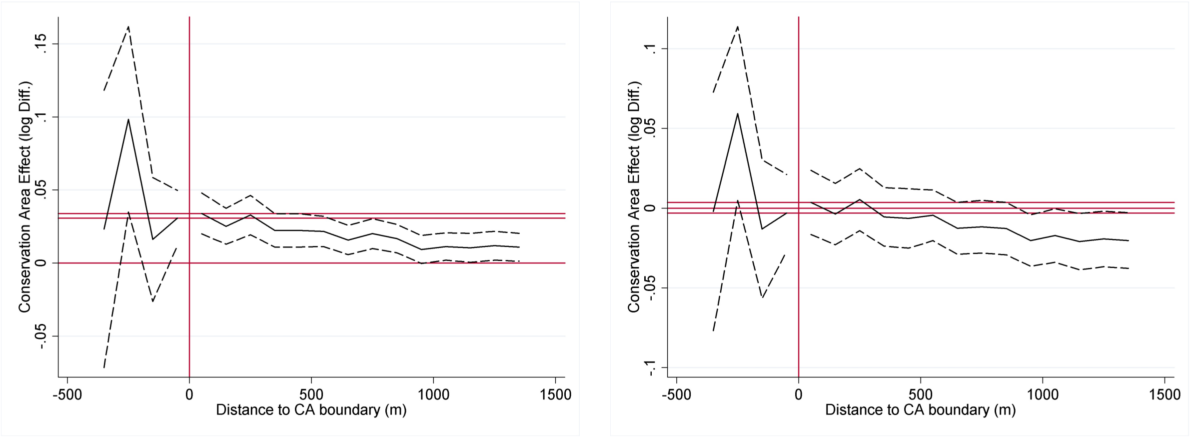
(d) External effects: Matched CA
Tab. B.11, column (7)

Note: The solid black line plots the estimated differences between treatment group and control group against year since designation date using equation (B.19). The dashed lines indicate the 5% confidence intervals. The left charts show results for the control group used in column (1) of appendix Table B.11 on page 167. The right charts show results for the control group used in column (7) of appendix Table B.11. The horizontal red line illustrates the mean of the pre-treatment estimates.

Figure B.3.: Semi-parametric spatial bins estimates



(a) 200 m bins: Full datasetTab. 4.4, column (1) (b) 200 m bins: Matched CATab. 4.4, column (5)



(c) 100 m bins: Full datasetTab. 4.4, column (1) (d) 100 m bins: Matched CATab. 4.4, column (5)

Note: The solid black line plots estimate the difference-in-differences treatment effect at different distances from the conservation area boundary using appendix equation (B.20). The dashed lines indicate the 5% confidence intervals. The left charts show results for the control group used Table 4.4 on page 80, column (1). The right charts show results for the control group used in Tab. 4.4, column (5). The horizontal red lines illustrate the mean of the pre-treatment estimates, the final pre-period bin and the first post-period bin.

C. Appendix to Chapter 5

The appendix to Chapter 5 is structured as follows: I begin by reporting the reduced form PVAR estimates in Section C.1 which complement the results from Section 5.5. As the reduced form results cannot be interpreted structurally, the interpretation is limited to the corresponding IRF in the main text. This is followed by a robustness test of modifiable areal unit problem (Section C.2), the presentation of the cumulative IRF (Section C.3) and finally the introduction and discussion of a land use location choice model which serves as an additional robustness test (Section C.4).

C.1. Reduced form PVAR estimates

Table C.1.: Results for reduced form 2-PVAR (municipality level)

	(1) total pop	(2) core pop	(3) periphery pop	(4) total railDens	(5) core railDens	(6) periphery railDens
pop_{t-1}	0.800***	0.818***	0.792***	-0.029	-0.0139	-0.022***
	(0.030)	(0.049)	(0.031)	(0.031)	(0.118)	(0.035)
railDens_{t-1}	0.029**	-0.086***	0.036**	0.745***	0.803***	0.739***
	(0.015)	(0.020)	(0.016)	(0.079)	(0.072)	(0.084)
Observations	1015	132	883	1015	132	883

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table C.2.: Results for reduced form 2-PVAR (block level)

	(1)	(2)	(3)	(4)	(5)	(6)
	total	core	periphery	total	core	periphery
	LV	LV	LV	railDens	railDens	railDens
LV$_{t-1}$	0.529***	0.524***	0.515***	0.079***	0.571***	0.073***
	(0.012)	(0.025)	(0.014)	(0.006)	(0.028)	(0.008)
railDens$_{t-1}$	0.157***	0.076***	0.200***	0.738***	0.641***	0.677***
	(0.013)	(0.015)	(0.019)	(0.024)	(0.022)	(0.030)
Observations	7759	1992	5767	7759	1955	5767

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table C.3.: Results for reduced form 2-PVAR (150m grid level)

	(1)	(2)	(3)	(4)	(5)	(6)
	total	core	periphery	total	core	periphery
	LV	LV	LV	railDens	railDens	railDens
LV$_{t-1}$	0.512***	0.593***	0.503***	0.128***	0.588***	0.054***
	(0.004)	(0.014)	(0.005)	(0.015)	(0.006)	(0.005)
railDens$_{t-1}$	0.039***	-0.024***	0.054***	0.537***	0.624***	0.470***
	(0.004)	(0.003)	(0.005)	(0.063)	(0.018)	(0.019)
Observations	26083	3192	22886	26083	3192	22886

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table C.4.: Results for reduced form 2-PVAR anticipation model (municipality level)

	(1)	(2)	(3)	(4)	(5)	(6)
	total	core	periphery	total	core	periphery
	pop	pop	pop	railDens	railDens	railDens
pop$_{t-1}$	0.796***	0.850***	0.772***	0.049**	-0.727***	0.077***
	(0.033)	(0.059)	(0.037)	(0.027)	(0.334)	(0.022)
railDens$_{t-1}$	0.045***	-0.029***	0.089**	0.517***	0.484***	0.481***
	(0.016)	(0.013)	(0.026)	(0.067)	(0.162)	(0.072)
Observations	993	123	810	933	123	810

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table C.5.: Results for reduced form 2-PVAR anticipation model (block level)

	(1) total LV	(2) core LV	(3) periphery LV	(4) total railDens	(5) core railDens	(6) periphery railDens
LV_{t-1}	0.547***	0.472***	0.546***	0.153***	-0.078*	0.163***
	(0.012)	(0.030)	(0.013)	(0.012)	(0.060)	(0.015)
railDens_{t-1}	0.037***	0.026***	0.043***	0.250***	0.282***	0.250***
	(0.006)	(0.001)	(0.008)	(0.024)	(0.032)	(0.029)
Observations	7590	1955	5635	7590	1955	5635

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table C.6.: Results for reduced form 2-PVAR anticipation model (150m grid level)

	(1) total LV	(2) core LV	(3) periphery LV	(4) total railDens	(5) core railDens	(6) periphery railDens
LV_{t-1}	0.514***	0.586***	0.503***	0.113***	-0.311***	0.110***
	(0.004)	(0.014)	(0.005)	(0.011)	(0.036)	(0.006)
railDens_{t-1}	0.027***	-0.043***	0.044***	0.334***	0.310***	0.367***
	(0.003)	(0.005)	(0.004)	(0.025)	(0.021)	(0.015)
Observations	25856	3091	22765	25856	3091	22765

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

C.2. IRF for the 300m grid model

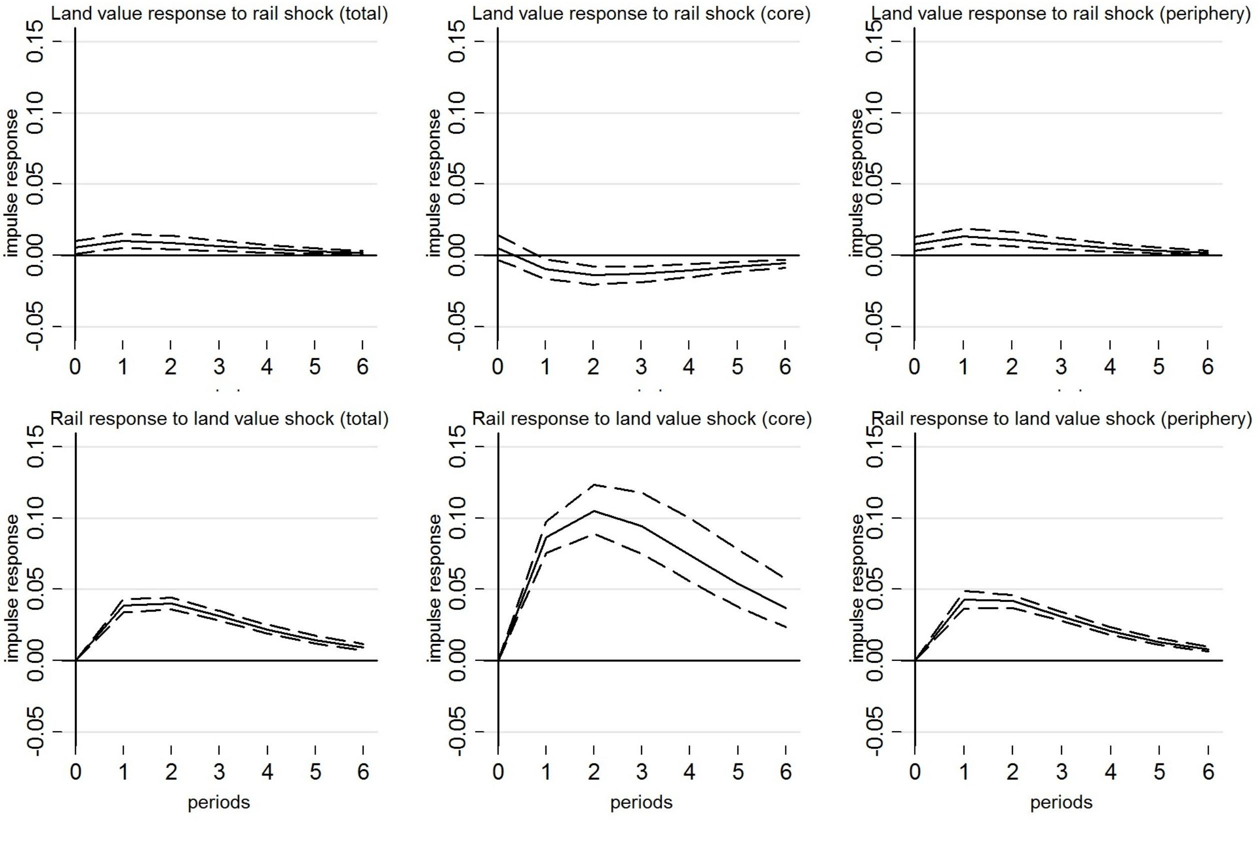
To address any multiple areal unit problem (MAUP) concerns, the raw land value and transport data have alternatively been aggregated to grid squares with a length of 300 m instead of 150 m. The reduced form estimates are given by Table C.7 on page 174 and the respective IRF by Figure C.1 on page 174. The IRF patterns are very similar to the 150m grid sample IRF derived in the main text, rejecting any MAUP concerns regarding the aggregation of the grid level data.

Table C.7.: Results for reduced form 2-PVAR (300m grid level)

	(1) total LV	(2) core LV	(3) periphery LV	(4) total railDens	(5) core railDens	(6) periphery railDens
LV_{t-1}	0.547***	0.604***	0.542***	0.147***	0.608***	0.158***
	(0.008)	(0.026)	(0.008)	(0.012)	(0.045)	(0.013)
$railDens_{t-1}$	0.019***	-0.028***	0.026***	0.490***	0.613***	0.437***
	(0.007)	(0.006)	(0.005)	(0.008)	(0.046)	(0.041)
Observations	7817	828	6989	7817	828	6989

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Figure C.1.: Impulse responses for 2-PVAR 300 m grid model



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

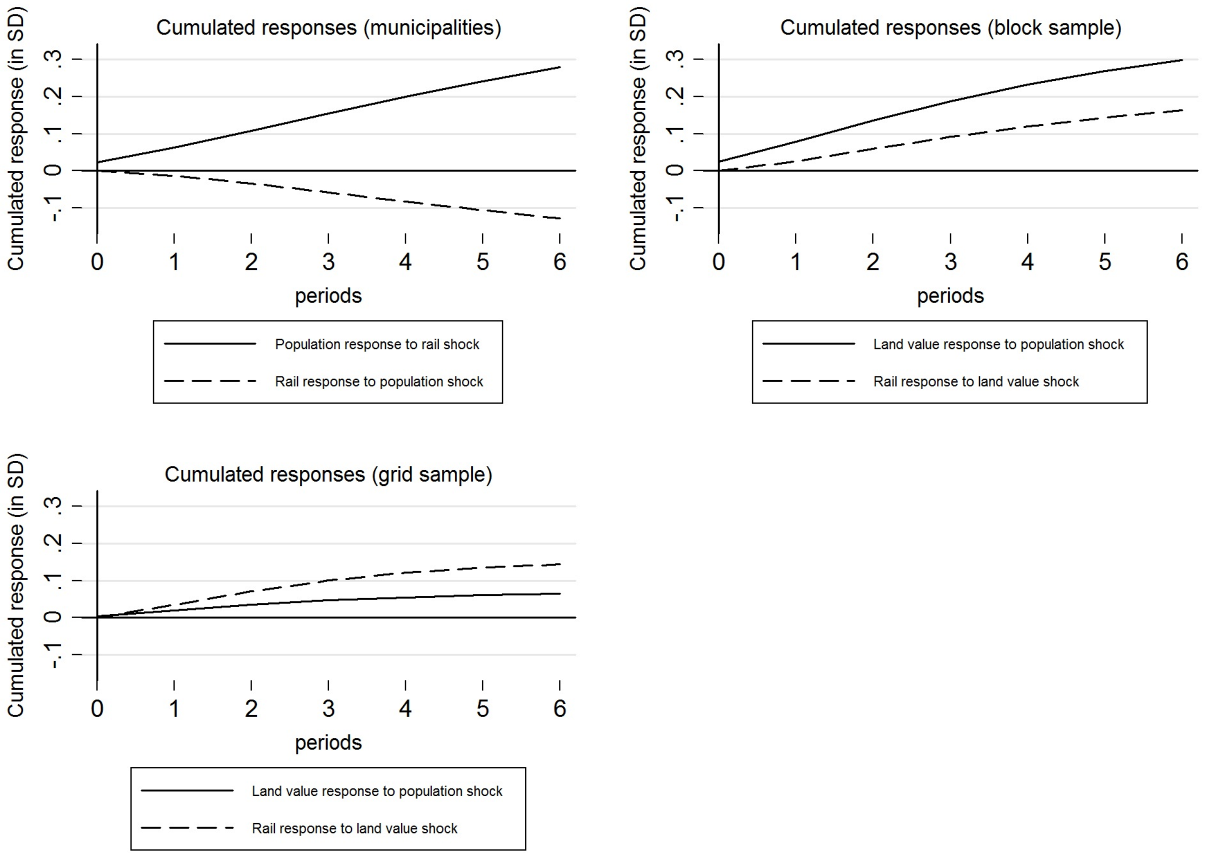
C.3. Cumulative impulse response functions

The cumulative impulse responses (CIRF) for the total sample are illustrated by Figure C.2 on page 175, for the core sample by Figure C.3 on page 176 and for the periphery sample by

Figure C.4 on page 177. The curves illustrate the response to a shock accumulated over all time periods.

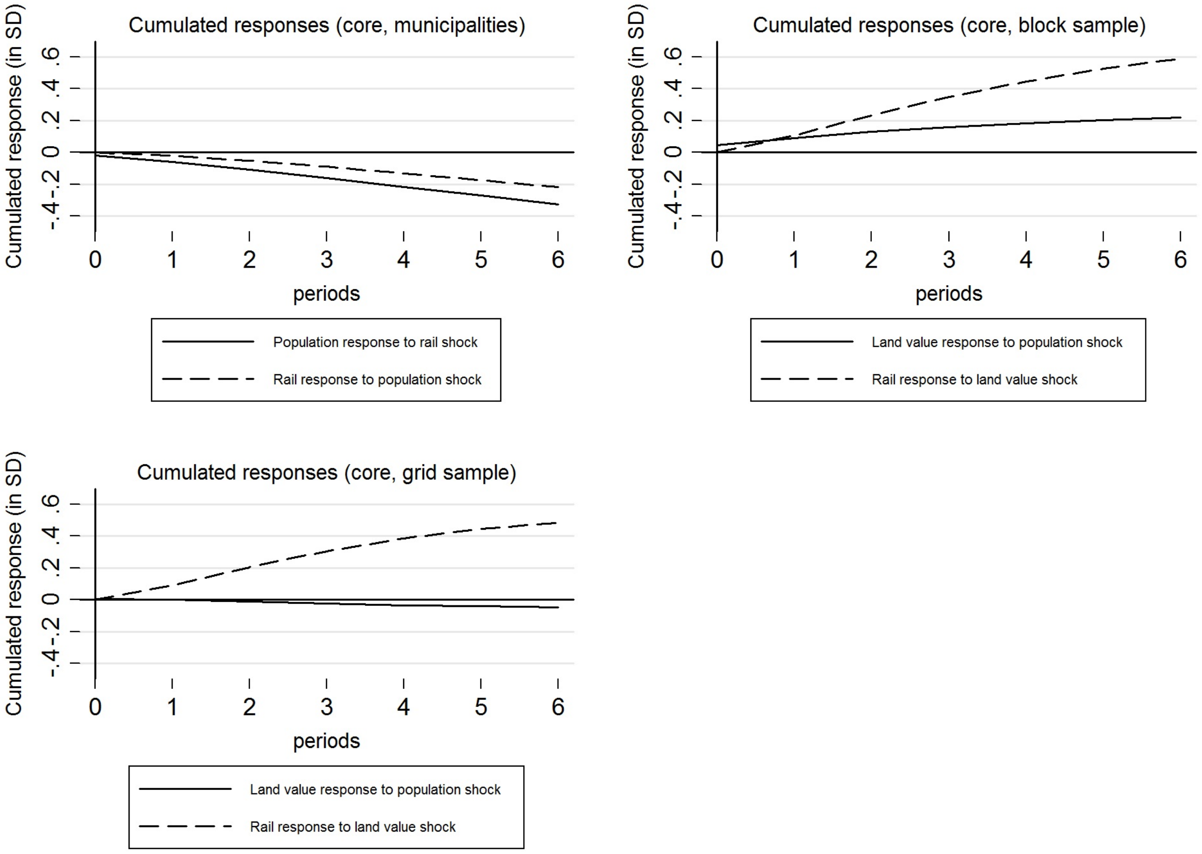
Looking at the CIRF based on the total sample (Figure C.2), population over time responds positively to transport shocks (almost up to 0.3 standard deviations after six periods) while the planner's response to population shocks is negative over time. The CIRF for the land value - transport relationship is positive in both ways. The CIRF for the block sample estimates indicate again that the inner block sample is slightly biased by centrality. In the core, population and transport respond negatively to each other (Figure C.3) which is in line with transport being more responsive to economic as opposed to residential activity. The CIRF for the land value samples are positive again where the land value response is significantly less pronounced than the transport response, which constantly increases over time. The CIRF derived from the periphery estimates (Figure C.4) further illustrate the positive population response due to the decentralisation partly driven by the outbidding of residents by firms. Land values and transport responses accumulated over time also stay positive in both directions in the periphery.

Figure C.2.: Cumulative impulse responses (total sample)



Notes: Cumulative IRFs illustrate accumulated effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation.

Figure C.3.: Cumulative impulse responses (core sample)



Notes: Cumulative IRFs illustrate accumulated effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation.

C.4. Land use location choice model

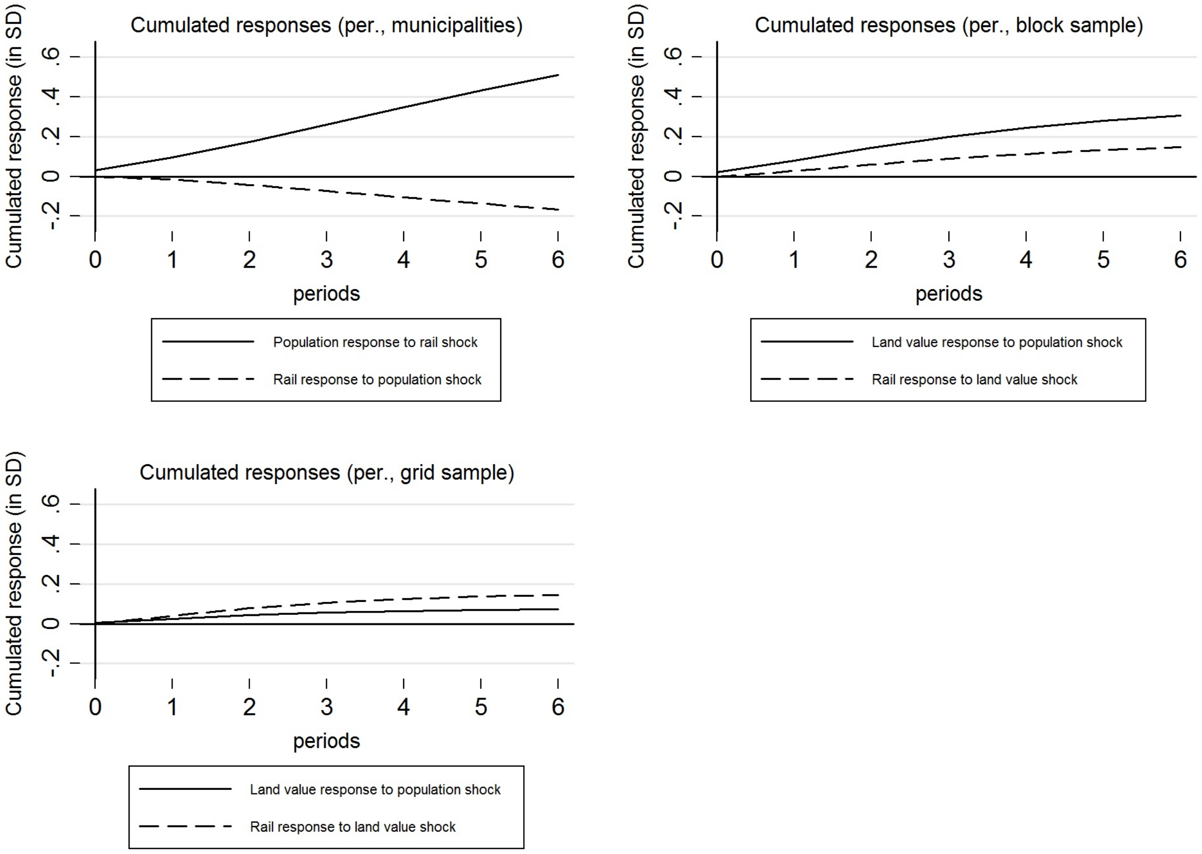
C.4.1. Motivation

As a robustness test and accompanying the PVAR models a multinomial choice model of different land use is developed in order to predict actual land use patterns. The discrete model's intention is to confirm the expectations about the distinct behaviour of firms and residents and the assumed outbidding process in particular. Owing to the nature of discrete models, I only test the supply-sided relationship of transport affecting land use. Moreover, it is important to note that only correlations are reported to shed some light into the outbidding process but it is not the author's intention to establish any causal relations.

C.4.2. Data

This section gives a brief overview of additional land use data I make use of to estimate a location choice model of different land use patterns.

Figure C.4.: Cumulative impulse responses (periphery sample)



Notes: Cumulative IRFs illustrate accumulated effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation.

The data were extracted from specific land use maps (Aust, 1986/1987) of Berlin for 1850, 1880, 1910 and 1940. These historical maps illustrate areas of different land use types, each devoted to either industrial, public, residential, business or mixed use. To control for natural amenities which could drive the results, water bodies and green spaces were also extracted from the Aust maps. Proximity to water, green space and overground tracks were computed in the GIS environment.

Over time, land use does not only change by designation but also by location and size. Therefore a unit of analysis is required which is able to capture these changes. Housing blocks as well as land value plots would not be able to describe the evolution appropriately. A grid square is able to take on a different land use type in every period; clustered grid squares then reflect spatial adjustments. Above all, the grid approach enables me to derive initially undeveloped land. I designate all grids as being undeveloped if they are covered by neither built-up structure (including streets), nor green space, nor water. Moreover, there are several technical advantages of using grid squares as discussed in Section 5.3. The selection of a reasonable grid size is based on the smallest land use areas in the data. The smallest areas range between 150 and 1,136 square meters. Based on this, grids with an area of 650 square meters were chosen, hence with a length of 25 meters. A grid square takes on a land use type if at least 50% of the square is covered with that respective land use. The grid panel is strongly balanced over 4 periods and consists of 215,600 observations per period, covering a total area of 140.14 square kilometres (entire Gross-Berlin: 891.85 square kilometres).

As for the PVAR models, I distinguish between a core and peripheral region within the analysis to capture the presumably distinct behaviour of economic agents in the city. As suggested (and tested) by Ahlfeldt & Wendland (2011), I define the underground station "Stadtmitte" (downtown) as centroid of the CBD. The station lies in the prestigious Friedrichstraße, surrounded by the boulevards Unter den Linden and Leipziger Straße.

Table C.8 on page 178 reports the number of grid squares for each land use type and time period. Public land use has been excluded due to its broad definition. The indicated public areas range from governmental buildings to military zones. It is important to note that there are two independent grid samples for distinct empirical models. To clarify, in one instance I use 150 m grid squares for estimating the relation between transport and land values in a PVAR model (300 m grid squares in a robustness test) and in another 25 m grids for the land use location choice model.

Table C.8.: Land use grid sample summary statistics

Land use	1850	1880	1910	1940
Business	197	453	2,869	4,789
Mixed	7,688	11,756	7,788	7,990
Residential	4,252	10,179	39,006	54,194
Industry	2,407	7,141	10,680	13,387
Undeveloped	78,088	66,187	27,481	17,050
Total	92,612	95,716	87,824	97,410

Notes: A grid square takes on a land use type if at least 50% of the square is covered with that respective land use. Grids are considered to be undeveloped if they are neither covered by built-up structure (including streets), nor green space, nor water.

C.4.3. Empirical strategy

Similar to the location choice model for footloose start-ups established in Section 3.4 the multinomial choice model is derived from utility theory. In random utility models, an urban economic agent (either resident or firm) chooses a location i over all other locations if location i's utility is at least as high as the utility of all other locations, i.e. if $u_i \geq u_j$, for all $j \neq i$. The probability of an agent choosing i over j can then be expressed by:

$$P_i = Prob\{u_i \geq u_j; \forall j \neq i\} \tag{C.1}$$

In this context, utility is assumed to be mainly defined by the local transport infrastructure. Hence, the choice model predicts the probability of a certain type of land use to locate at location i, conditional on the transport network at that particular location. I expect the increase in station density to have a larger effect on businesses than on residents (land use intensification). In fact, in central locations residents are expected to be driven out by high rail density.

The multinomial logit model describes a choice situation where the values have no natural order, $\gamma_m \in \{1, 2, \dots J\}$. The empirical specification takes the following form:

$$LU_i = \alpha_i + \beta railDens_i + \gamma \sum_i L_i + \delta \sum_i T_t + \varepsilon_i \tag{C.2}$$

where LU_i indicates the type of land use, $railDens_i$ is station density, L_i a vector of locational controls and T_t time fixed effects. The discrete land use variable LU_i is defined by:

$$LU_i = \left\{ \begin{matrix} 1 & \text{if business} \\ 2 & \text{if mixed} \\ 3 & \text{if residential} \\ 4 & \text{if industry} \\ 5 & \text{if undeveloped} \end{matrix} \right\} \tag{C.3}$$

The estimation results of the land use model as well as further robustness exercises are reported and discussed in the next part.

C.4.4. Results

The estimates of the multinomial logit model are reported in Table C.10 on page 181. The model is estimated separately for the core and the periphery sample.

Table C.9.: Results for multinomial land use model

	Core			Periphery		
	(1)	(2)	(3)	(4)	(5)	(6)
Business						
railDens	1.901***	1.139***	1.105***	3.659***	2.536***	2.128***
	(0.036)	(0.087)	(0.089)	(0.031)	(0.035)	(0.038)
Mixed						
railDens	0.193***	-0.930***	-1.062***	3.116***	2.549***	2.075***
	(0.032)	(0.083)	(0.086)	(0.019)	(0.022)	(0.024)
Residential						
railDens	0.808***	-2.209***	-2.092***	2.725***	1.443***	1.041***
	(0.031)	(0.083)	(0.085)	(0.014)	(0.016)	(0.017)
Industrial						
railDens	1.440***	-2.030***	-2.510***	1.936***	0.722***	0.092***
	(0.036)	(0.095)	(0.102)	(0.017)	(0.020)	(0.023)
Time fixed effects	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Observations	41253	41253	41253	239697	239697	239697
pseudo R^2	0.073	0.144	0.153	0.114	0.160	0.203

Notes: Base outcome 5 (undeveloped). Constants are significant in all models but not reported for brevity. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Starting with the core area, the probability that a location is of business land use increases with station density compared to undeveloped areas (column 1). Station density also increases the likelihood of being a mixed, a residential and an industrial area compared to the base outcome. Nevertheless, the effect is largest for business areas, followed by residential and industrial areas. When controlling for time fixed effects, see column (2), the probability of locations being mixed, residential or industrial decreases significantly and is even negative compared to undeveloped land. Only the likelihood of business land use increases with higher station density. The estimates stay relatively robust when additionally controlling for locational amenity effects (column 3). The amenity controls are distance to the nearest overground tracks (considered as disamenity due to negative noise/view effects (Ahlfeldt et al., 2011)), nearest green space and nearest water body (positive amenities). The individual effects are not reported in the table for reasons of brevity. Overall, the improved transport only increases the likelihood for areas to be used as business land (compared to the base outcome). This is not surprising since the choice model is only estimated for the core area.

In the baseline model of the sample covering only peripheral Berlin (same table, column 4), increased transport density raises the probability of being developed compared to undeveloped. The effect is still the largest for business areas but not as strong as in the core-area estimation. Controlling for time fixed effects and amenity controls slightly changes the picture. While the likelihood of being a business area stays high, its effect decreases relatively compared to mixed land use. When time dummies are added (column 5) the likelihood of being a location of mixed land use is even slightly higher than being a business location as a response to transport shocks. The probabilities of the two land use categories are fairly similar when adding the amenity controls. In contrast to the core-model, the likelihood of residential or industrial location increases with transport improvement. Comparing the estimates for the core and periphery one can observe the assumed outbidding of residents due to transport shocks. In the core, businesses are attracted by an increase in station density whereas all other land use types experience a negative effect. In the periphery, however, transport also attracts residential, industrial and especially mixed land use.

Table **??** on page **??** reports a rather flexible distinction between core and periphery by interacting each transport variable with distance to CBD (and distance to CBD^2). In all three models the probability of a business location is described by a positive binomial interaction between land use and distance to CBD, whereas mixed, residential and industrial land use is described by a negative binomial one. In all models the probability of a business location is highest in the CBD. This likelihood decreases with distance until the parabola is upward sloped again. The other land use categories' probabilities decline continuously with distance due to their negative slope.

Table C.10.: Results for multinomial land use model (total sample)

	(1)	(2)	(3)
Business			
railDens	4.931***	4.863***	4.039***
	(0.053)	(0.056)	(0.052)
railDens x distCBD	-0.875***	-1.107***	-1.104***
	(0.027)	(0.028)	(0.027)
railDens x distCBD^2	0.063***	0.068***	0.077***
	(0.004)	(0.004)	(0.004)
Mixed			
railDens	2.734***	3.272***	2.432***
	(0.051)	(0.054)	(0.050)
railDens x distCBD	0.172***	-0.105***	-0.087***
	(0.025)	(0.026)	(0.025)
railDens x distCBD^2	-0.098***	-0.054***	-0.048***
	(0.003)	(0.003)	(0.003)
Residential			
railDens	2.187***	2.007***	1.450***
	(0.047)	(0.049)	(0.045)
railDens x distCBD	0.246***	-0.019***	-0.044*
	(0.020)	(0.021)	(0.020)
railDens x distCBD^2	-0.098***	-0.054***	-0.048***
	(0.003)	(0.003)	(0.003)
Industrial			
railDens	1.559***	1.421***	0.169**
	(0.055)	(0.057)	(0.056)
railDens x distCBD	0.263***	-0.013	0.214***
	(0.025)	(0.026)	(0.026)
railDens x distCBD^2	-0.037***	-0.034***	-0.058***
	(0.003)	(0.003)	(0.003)
Time fixed effects		Yes	Yes
Controls			Yes
Observations	280950	280950	280950
pseudo R^2	0.073	0.144	0.153

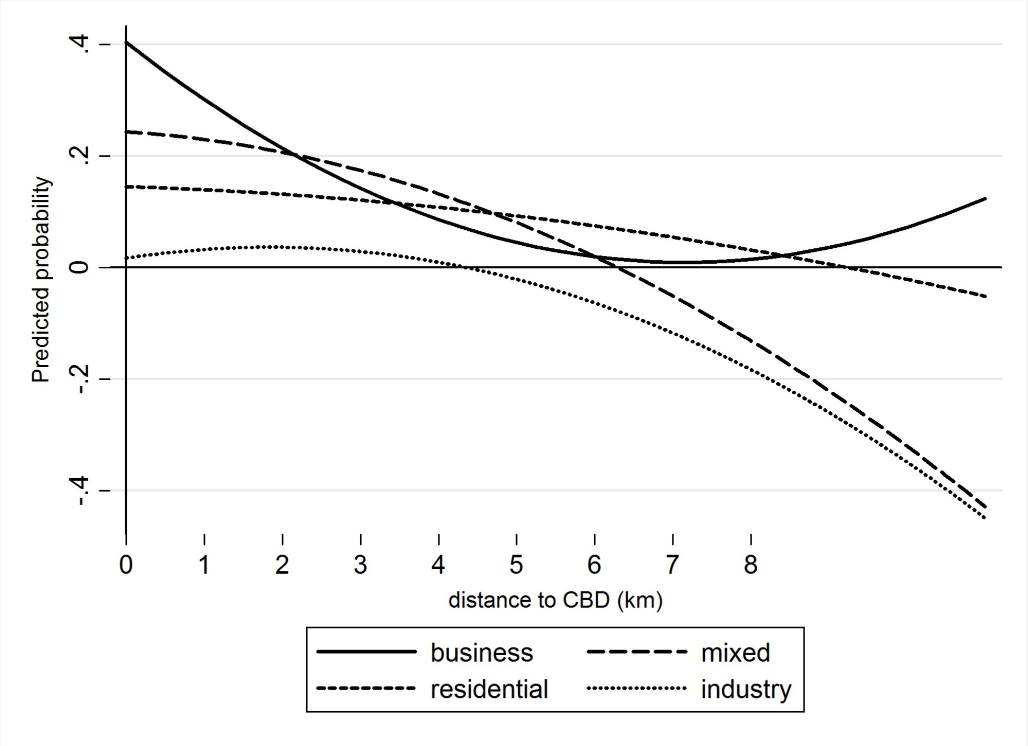
Notes: Base outcome 5 (undeveloped). Constants are significant in all models but not reported to save space. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Interpretation is restricted to the more demanding model, the estimates of which are reported in column (3). Plotting the estimated probability against distance to CBD (see Figure C.5 on page 182) results in a picture which shares similarities with the Alonso-Mills-Muth Model (AMM). Starting from the CBD, a location is most likely to be of business land use. The effect declines with distance. At approximately 2 km from the CBD, mixed land use takes over, i.e.

between 2 km and 4.5 km an increase in station density most likely results in an area becoming “mixed”. After that, improved transport has the strongest effect on residential land use, i.e. up to a distance of about 6.8 km from the CBD station density is most likely to attract residents. The estimated probability for industrial land use is relatively low and becomes negative at around 4.25 km. This means that improved transport does not attract any industry within the city. As the land use sample covers a rather central area, it contains only a few industrial areas, which most probably results in an underestimation of industrial land use probability.

Summing up, the estimates from the multinomial land use model provide further evidence for the outbidding process. In line with traditional models, a transport improvement leads to an outbidding of residents by firms in the core.

Figure C.5.: Predicted land use probabilities over distance to CBD (in km)



Notes: Land use probability estimates come from Table C.10.

D. Appendix to Chapter 6

The appendix complements Chapter 6 and is structured as follows: First of all, the study area of Chicago is illustrated on a map (Section D.1). Secondly, the appendix reports the reduced form PVAR estimates in Section D.2. And finally Section D.3 presents the cumulative IRF.

D.1. Data

The empirical analysis on the interaction between transport and development is based on US census tract data for Chicago. The tract definition of 1990 is used as the baseline geography (Minnesota Population Center, 2011). Figure D.1 on page 184 provides an overview of the study area. As one can see, the 'L' network does not cover the entire study area but only the area framed in black by Figure (a). In Figure (c) the Central Business District is indicated by a red border. I hereby follow the definition from City of Chicago (2010). The Union Loop, which circulates though downtown Chicago, is located in the middle of the CBD.

D.2. Empirical results

The following tables complement the estimation results from Section 6.4.2. As the reduced form results cannot be interpreted structurally, the interpretation is limited to the corresponding IRF in the main text. Figure D.2 on page 185 illustrated the IRF from the main text but with individual Y-axes.

Table D.1.: Results for reduced form 2-PVAR

	(1) total pop	(2) core pop	(3) periphery pop	(4) total L Dens	(5) core L Dens	(6) periphery L Dens
pop_{t-1}	0.567***	0.544***	0.562***	-0.008	-0.139	-0.006
	(0.260)	(0.078)	(0.026)	(0.011)	(0.118)	(0.012)
L Dens_{t-1}	0.072***	-0.134***	0.088***	0.795***	0.795***	0.795***
	(0.013)	(0.068)	(0.012)	(0.038)	(0.105)	(0.040)
Observations	3603	251	3352	3603	251	3352

Notes: (1) 1-lag VAR is estimated by GMM, (2) variables are estimated in logs, (3) variables are time-demeaned and Helmert transformed prior to estimation, (4) standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Figure D.1.: Study area of Chicago (2010)

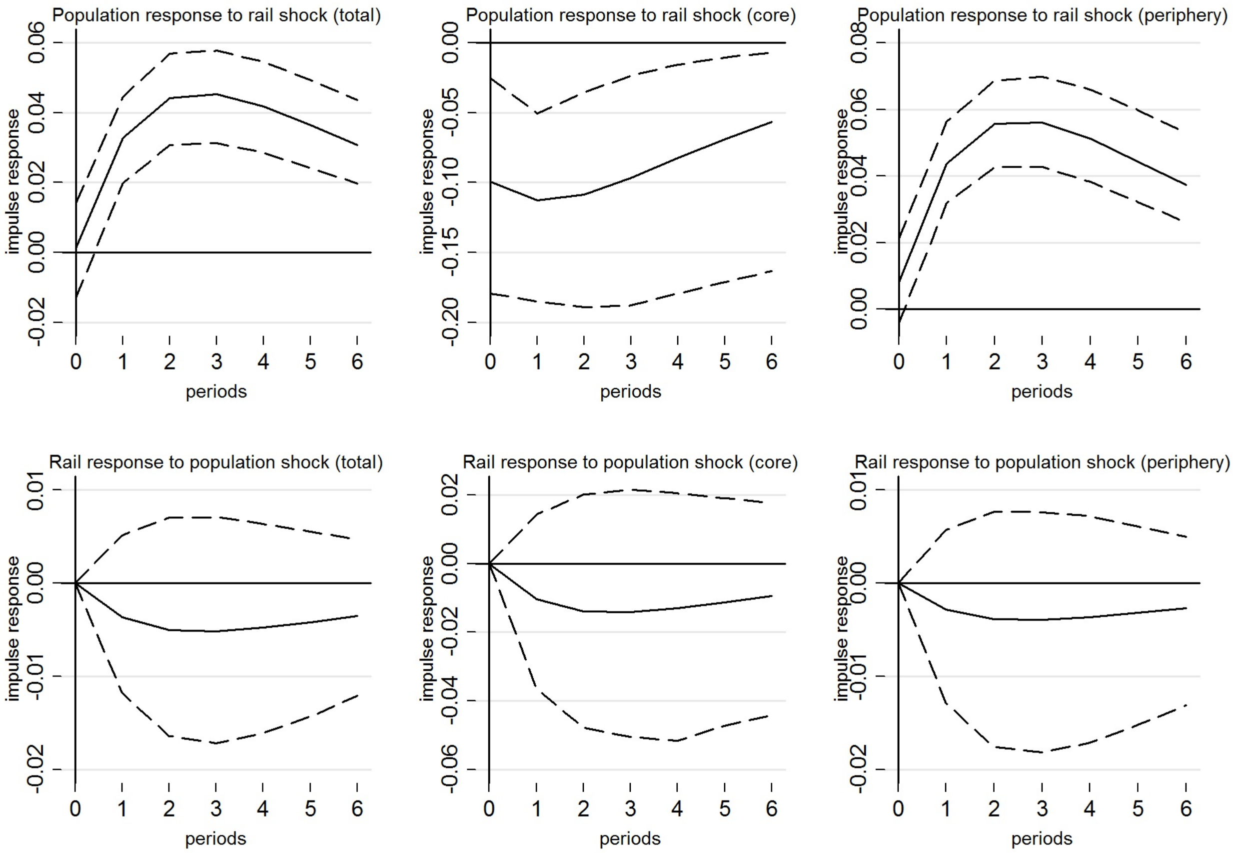
(a) Census tracts (b) The “L” network

(c) Central Business District (d) Legend

D.3. Cumulative impulse response functions

The cumulative impulse responses (CIRF) for the total sample are illustrated by Figure D.3 on page 186, for the core sample by Figure D.4 on page 186 and for the periphery sample by Figure D.5 on page 187. The curves illustrate the response to a shock accumulated over six time periods.

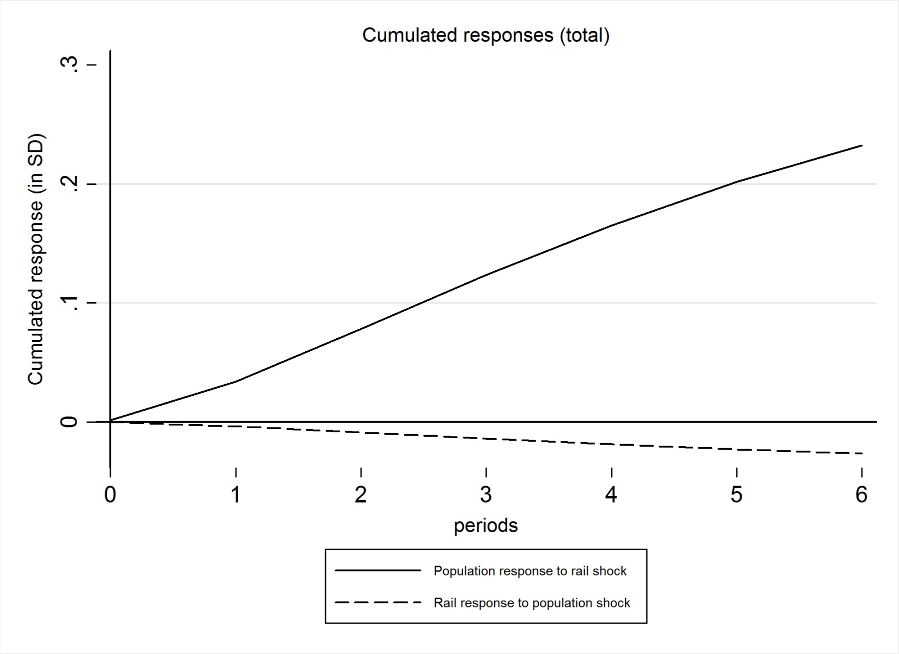
Figure D.2.: Impulse responses for 2-PVAR model (individual Y-axes)



Notes: Dotted lines indicate 5% error bands generated by Monte-Carlo with 500 repetitions.

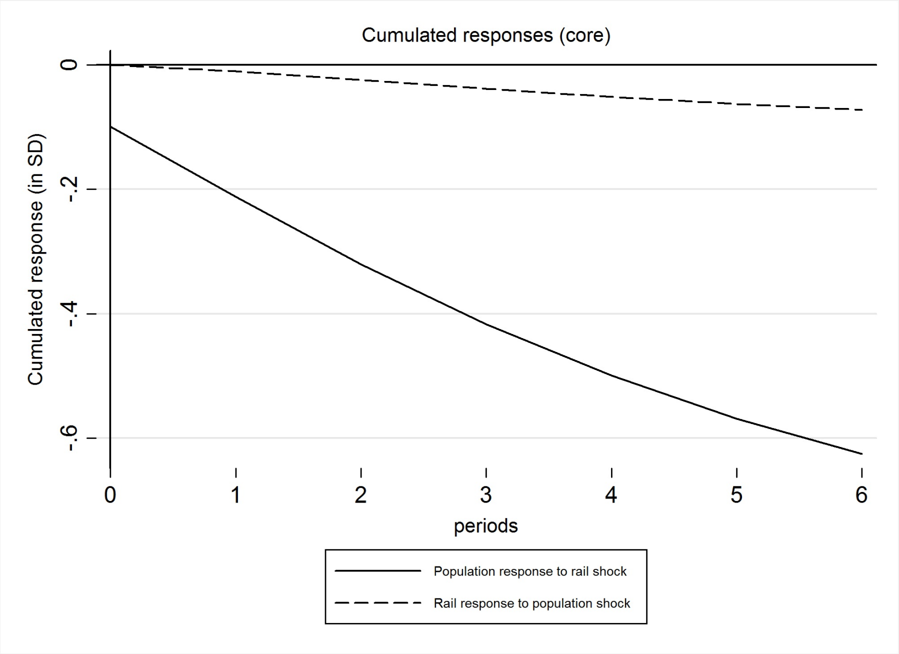
Looking at the CIRF based on the total sample (Figure D.3), population over time responds positively to transport shocks (about 0.23 standard deviations after six periods) while the planner's response to population shocks is negative over time. In the core, population and transport respond negatively to each other (Figure D.4) which is in line with transport being more responsive to economic instead of residential activity. The CIRF derived from the periphery estimates (Figure D.5) further illustrate the positive population response due to the decentralisation partly driven by the outbidding of residents by firms.

Figure D.3.: Cumulative impulse responses (total sample)



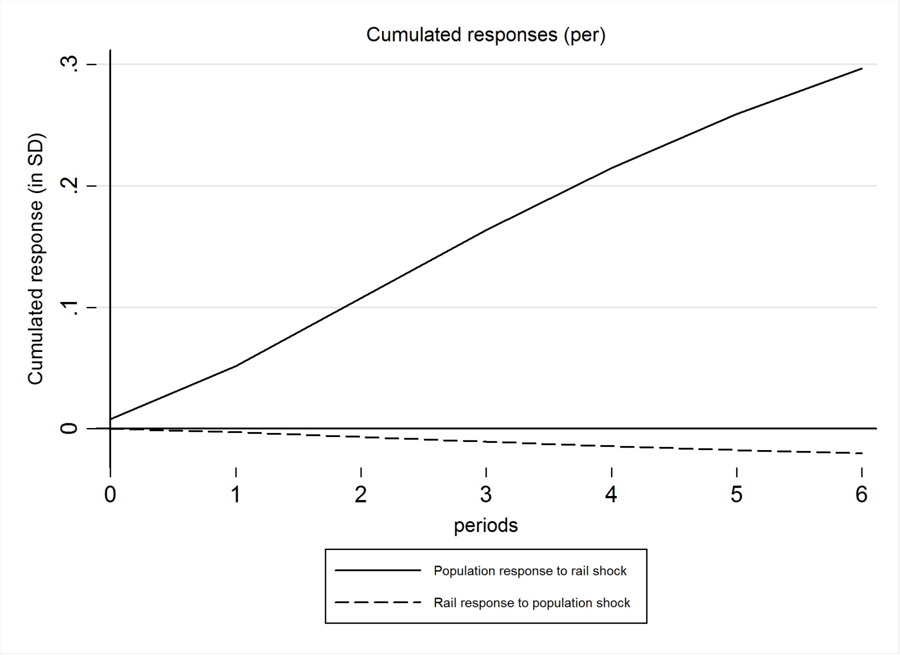
Notes: Cumulative IRFs illustrate accumulated effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation.

Figure D.4.: Cumulative impulse responses (core sample)



Notes: Cumulative IRFs illustrate accumulated effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation.

Figure D.5.: Cumulative impulse responses (periphery sample)



Notes: Cumulative IRFs illustrate accumulated effect of a one standard deviation shock (in logs) on the response variable (in logs) in units of standard deviation.

References

Ahlfeldt, G. (2011a, January). Blessing or curse? appreciation, amenities and resistance to urban renewal. *Regional Science and Urban Economics*, *41*(1), 32-45.

Ahlfeldt, G. (2011b, 05). If alonso was right: Modeling accessibility and explaining the residential land gradient. *Journal of Regional Science*, *51*(2), 318-338.

Ahlfeldt, G. (2013). Urbanity. *SERC Discussion Paper*, *136*.

Ahlfeldt, G. (in press). If we build it, will they pay? pridicting property price effects of transport innovations. *Environment and Planning A*.

Ahlfeldt, G. & Feddersen, A. (2011, September). *From periphery to core: Economic adjustments to high speed rail* (ERSA conference papers No. ersa11p545). European Regional Science Association.

Ahlfeldt, G. & Kavetsos, G. (2013). Form or function?: the effect of new sports stadia on property prices in London. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Ahlfeldt, G. & Maennig, W. (2007). *Assessing external effects of city airports: Land values in Berlin* (Working Papers No. 011). Chair for Economic Policy, University of Hamburg.

Ahlfeldt, G. & Maennig, W. (2009). Arenas, arena architecture and the impact on location desirability: The case of 'Olympic Arenas' in Prenzlauer Berg, Berlin. *Urban Studies*, *46*(7), 1343-1362.

Ahlfeldt, G. & Maennig, W. (2010a). Impact of sports arenas on land values: evidence from Berlin. *The Annals of Regional Science*, *44*(2), 205-227.

Ahlfeldt, G. & Maennig, W. (2010b). Substitutability and complementarity of urban amenities: External effects of built heritage in Berlin. *Real Estate Economics*, *38*(2), 285-323.

Ahlfeldt, G. & Maennig, W. (2012). Voting on a NIMBY facility: Proximity cost of an "iconic" stadium. *Urban Affairs Review*, *48*(2), 205-237.

Ahlfeldt, G. & Maennig, W. (2013). *Homevoters vs. leasevoters: A spatial analysis of aircraft noise* (Working Paper No. 4301). CESifo.

Ahlfeldt, G., Maennig, W. & Richter, F. (2013). *Urban renewal after the Berlin Wall.*

Ahlfeldt, G. & Mastro, A. (2012). Valuing iconic design: Frank Lloyd Wright architecture in Oak Park, Illinois. *Housing Studies*, *27*(8), 1079-1099.

Ahlfeldt, G., Möller, K., Waights, S. & Wendland, N. (2013, Sep). *Game of zones: The economics of conservation areas* (Tech. Rep. No. SERC Discussion Paper 143). SERC.

Ahlfeldt, G., Nitsch, V. & Wendland, N. (2011). Ease vs. noise: On the conflicting effects of transportation infrastructure. In *51st european congress of the regional science association international.* Barcelona, Spain.

Ahlfeldt, G., Redding, S. J., Sturm, D. M. & Wolf, N. (2012, June). *The economics of density: Evidence from the Berlin Wall* (CEP Discussion Papers No. dp1154). Centre for Economic Performance, LSE.

Ahlfeldt, G. & Wendland, N. (2011). Fifty years of urban accessibility: The impact of urban railway network on the land gradient in Berlin 1890-1936. *Regional Science and Urban Economics*, *41(2)*, 77-88.

Ahlfeldt, G. & Wendland, N. (2013). How polycentric is a monocentric city? centers, spillovers and hysteresis. *Journal of Economic Geography*, *13*(1), 53-83.

Aiken, L., West, S. & Reno, R. (1991). *Multiple regression: Testing and interpreting interactions.* SAGE Publications.

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

Amt für Statistik Berlin-Brandenburg. (2011a). *Arbeitslose in Berlin.* Berlin.

Amt für Statistik Berlin-Brandenburg. (2011b). *Einwohnerstatistik Berlin.* Berlin.

Anderson, S. T. & West, S. E. (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economics*, *36*(6), 773 - 789.

Andreas, A. (1975). *History of Chicago* (No. Bd. 2). Arno Press.

Arellano, M. & Bond, S. (1991, April). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, *58*(2), 277-97.

Arellano, M. & Bover, O. (1995, July). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29-51.

Armstrong, R. J. & Rodríguez, D. A. (2006). An evaluation of the accessibility benefits of commuter rail in Eastern Massachusetts using spatial hedonic price functions. *Transportation*, *33*(1), 21-43.

Arzaghi, M. & Henderson, J. V. (2008, October). Networking off Madison Avenue. *Review of Economic Studies*, *75*(4), 1011-1038.

Asabere, P. K., Hachey, G. & Grubaugh, S. (1989). Architecture, historic zoning, and the value of homes. *The Journal of Real Estate Finance and Economics*, *2*(3), 181-195.

Asabere, P. K., Huffman, F. & Mehdian, S. (1994). The adverse impacts of local historic designation: The case of small apartment buildings in Philadelphia. *The Journal of Real Estate Finance and Economics*, *8*(3), 225-234.

Asabere, P. K. & Huffman, F. E. (1994, July). Historic designation and residential market values. *The Appraisal Journal*, *62*(3), 396-401.

Assenmacher-Wesche, K. & Gerlach, S. (2008, October). *Monetary policy, asset prices and macroeconomic conditions : a panel-VAR study* (Working Paper Research No. 149). National Bank of Belgium.

Aust, B. (1986/1987). Die städtebauliche entwicklung Berlins - Berlin um 1850, 1880, 1910, 1940 [Computer software manual].

Avrami, E., Mason, R. & de la Torre, M. (2000). *Values and heritage conservation Los Angeles* (Tech. Rep.). The Getty Conservation Institute.

Baedeker, K. (1936). *Berlin und Potsdam.* Karl Baedeker, Leipzig.

Bajic, V. (1983). The effects of a new subway line on housing prices in metropolitan Toronto. *Urban Studies*, *20*(2), 147-158.

Bauernschuster, S., Falck, O., Heblich, S. & Suedekum, J. (2012, September). *Why are educated and risk-loving persons more mobile across regions?* (IZA Discussion Papers No. 6860). Institute for the Study of Labor (IZA).

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, *122*(2), 775-805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A. & Zhang, Q. (2012). *Roads, railroads and decentralization of Chinese cities* (Tech. Rep.).

Baum-Snow, N. & Kahn, M. E. (2000). The effects of new public projects to expand urban rail transit. *Journal of Public Economics*, *77*(2), 241 - 263.

Bento, A. M., Kaffine, D. & Roth, K. (2010). *Clearing the air by promoting hybrids? the unintended consequences of allowing solo-hybrid drives in high* (Tech. Rep.). Working Paper.

Berlin Business Location Center. (2012a). *Berlin Startup Map.* Retrieved from `http://www.businesslocationcenter.de/web2zero` (retrieved: 27/06/2012)

Berlin Business Location Center. (2012b). *Informations- und Kommunikationstechnik.* Retrieved from `http://www.businesslocationcenter.de/ikt` (retrieved: 01/11/2012)

Berlin Senate Department for Urban Development. (2007). *Lärmkartierung nach Umgebungsrichtlinie - Ballungsraum Berlin.* Berlin.

Berliner Parlament. (2008). *Berliner Landeswahlgesetzs.* Retrieved from `http://www.parlament-berlin.de/pari/web/wdefault.nsf/vFiles/D14/$FILE/LWahlG_Times%2014.pdf` (retrieved: 06/02/2013)

Bertrand, M., Duflo, E. & Mullainathan, S. (2004, February). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249-275.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics*, *114*(2), pp. 577-599.

Blomquist, G. C., Berger, M. C. & Hoehn, J. P. (1988, March). New estimates of quality of life in urban areas. *American Economic Review*, *78*(1), 89-107.

Blundell, R. & Bond, S. (1998, August). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115-143.

Bogart, W. T. & Cromwell, B. A. (2000). How much is a neighborhood school worth? *Journal of Urban Economics*, *47*(2), 280 - 305.

Borck, R. (2007). Consumption and social life in cities: Evidence from Germany. *Urban Studies*, *44*(11), 2105-2121.

Bortrill, C. (2005). *Homes in historic conservation areas in Great Britain: Calculating the proportion of residential dwellings in conservation areas* (Tech. Rep.). Environmental Change Institute.

Borzo, G. (2007). *The Chicago "L".* Chicago: Arcadia Publishing.

Bowes, D. R. & Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics*, *50*(1), 1 - 25.

Brand, S. & Price, R. (2000). *The economic and social costs of crime* (Vol. 217; Tech. Rep. No. 217). Home Office Research Development and Statistics Directorate.

Brasington, D. & Haurin, D. R. (2006). Educational outcomes and house values: A test of the value added approach. *Journal of Regional Science*, *46*(2), 245-268.

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

Brülhart, M., Jametti, M. & Schmidheiny, K. (2012, 09). Do agglomeration economies reduce the sensitivity of firm location to tax differentials? *Economic Journal*, *122*(563), 1069-1093.

Brunner, E. & Sonstelie, J. (2003, September). Homeowners, property values, and the political economy of the school voucher. *Journal of Urban Economics*, *54*(2), 239-257.

Brunner, E., Sonstelie, J. & Thayer, M. (2001, November). Capitalization and the voucher: An analysis of precinct returns from California's proposition 174. *Journal of Urban Economics*, *50*(3), 517-536.

BVG. (2006). *Berliner Verkehrsbetriebe Liniennetz.* Berlin.

Canova, F. & Ciccarelli, M. (2004, June). Forecasting and turning point predictions in a Bayesian panel VAR model. *Journal of Econometrics*, *120*(2), 327-359.

Canova, F. & Ciccarelli, M. (2013, March). *Panel vector autoregressive models: A survey* (CEPR Discussion Papers No. 9380). C.E.P.R. Discussion Papers.

Carlino, G. & Coulson, N. E. (2004). Compensating differentials and the social benefits oftheNFL. *Journal of Urban Economics*, *56*(1), 25-50.

Carlino, G. & Saiz, A. (2008). *Beautiful city: Leisure amenities and urban growth* (Tech. Rep. No. 08-22). FRB of Philadelphia Working Paper.

Carlsson, F., Lampi, E. & Martinsson, P. (2004). The marginal values of noise disturbance from air traffic: Does the time of the day matter? *Transportation Research Part D: Transport and Environment*, *9*(5), 373 - 385.

Carlton, D. W. (1983). The location and employment choices of new firms: An econometric model with discrete and continious endogenous variables. *Review of Econo*, *65*, 440-449.

Cellini, S. R., Ferreira, F. & Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, *125*(1), 215-261.

Cervero, R. & Hansen, M. (2002). Induced travel demand and induced road investment: A simultaneous equation analysis. *Journal of Transport Economics and Policy*, *36*(3), 469-490.

Cervero, R. & Landis, J. (1995). *Development impacts of urban transport: A US perspective* (D. Banister, Ed.). London: E & F. N. Spon.

Cervero, R. & Landis, J. (1997). Twenty years of the Bay Area Rapid Transit system: Land use and development impacts. *Transportation Research Part A: Policy and Practice*, *31*(4), 309 - 333.

Chatterji, A., Glaeser, E. L. & Karr, W. (2013). Clusters of entrepreneurship and innovation. *NBER Working Paper Series*, *19013*.

Chen, Y. & Rosenthal, S. (2008, November). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics*, *64*(3), 519-537.

Cheshire, P. & Hilber, C. (2008, 06). Office space supply restrictions in Britain: The political economy of market revenge. *Economic Journal*, *118*(529), F185-F221.

Cheshire, P., Hilber, C. & Kaplanis, I. (2011). *Evaluating the effects of planning policies on the retail sector: Or do town centre first policies deliver the goods?* (Discussion Paper No. 66). SERC.

Cheshire, P. & Magrini, S. (2006). Population growth in European cities: Weather matters - but only nationally. *Regional Studies*, *40*(1), 23-37.

Cheshire, P. & Sheppard, S. (1995, May). On the price of land and the value of amenities. *Economica*, *62*(246), 247-67.

Cheshire, P. & Sheppard, S. (2004). Capitalising the value of free schools: The impact of supply characteristics and uncertainty. *The Economic Journal*, *114*(499), F397–F424.

Chicago Transit Authority. (2012). *CTA - 'L' (rail) stations - shapefile.*

Chicago Transit Authority. (2013, July). *Our services.* Retrieved from `http://www.transitchicago.com/riding_cta/service_overview.aspx`

Chicago"L".org. (2013a, July). *"L" system history.* Retrieved from `http://www.chicago-l.org/history/index.html`

Chicago"L".org. (2013b, July). *Stations.* Retrieved from `http://www.chicago-l.org/stations/`

Choi, I. (2001, April). Unit root tests for panel data. *Journal of International Money and Finance*, *20*(2), 249-272.

City of Chicago. (2010). *Boundaries - Central Business District.*

Clapp, J. M., Nanda, A. & Ross, S. L. (2008). Which school attributes matter? the influence of school district performance and demographic composition on property values. *Journal of Urban Economics*, *63*(2), 451 - 466.

Clark, D. E. & Cosgrove, J. C. (1990). Hedonic prices, identification, and the demand for the public safety. *Journal of Regional Science*, *30*(1), 105–121.

Clark, T. N. (2004). The city as an entertainment machine. In T. N. Clark (Ed.), (p. 103-140). Amsterdam: Elsevier.

Claussen, J., Falck, O. & Grohsjean, T. (2010, August). *The strength of direct ties: Evidence from the electronic game industry* (Discussion Papers in Business Administration No. 11745). University of Munich, Munich School of Management.

Clay, P. L. (1979). *Neighborhood renewal. middle-class resettlement and incumbent up-grading in American neighbourhood.* Lexingtion, MA: Lexington Books.

Coates, D. & Humphreys, B. R. (2006). Proximity benefits and voting on stadium and arena subsidies. *Journal of Urban Economics*, *59*(2), 285 - 299.

Coulson, N. E. & Lahr, M. L. (2005, 09). Gracing the land of Elvis and Beale Street: Historic designation and property values in Memphis. *Real Estate Economics*, *33*(3), 487-507.

Coulson, N. E. & Leichenko, R. (2001). The internal and external impact of historical designation on property values. *The Journal of Real Estate Finance and Economics*, *23*(1), 113-124.

Crozet, M., Mayer, T. & Mucchielli, J.-L. (2004, January). How do firms agglomerate? a study of FDI in France. *Regional Science and Urban Economics*, *34*(1), 27-54.

Cudahy, B. J. (1982). *Destination Loop: The story of rapid transit railroading in and around Chicago.* Brattleboro, VT: S. Greene Press.

Cullen, J. B. & Levitt, S. D. (1999). Crime, urban flight, and the consequences for cities. *The Review*, *81*(2), 159-169.

Dalmazzo, A. & de Blasio, G. (2011). Amenities and skill-biased agglomeration effects: Some results on Italian cities. *Papers in Regional Science*, *90*(3), 503–527.

Davidoff, I. & Leigh, A. (2008, 06). How much do public schools really cost? estimating the relationship between house prices and school quality. *The Economic Record*, *84*(265), 193-206.

Davis, M. A. & Ortalo-Magne, F. (2011, April). Household expenditures, wages, rents. *Review of Economic Dynamics*, *14*(2), 248-261.

Dehring, C. A., II, C. A. D. & Ward, M. R. (2008). A direct test of the homevoter hypothesis. *Journal of Urban Economics*, *64*(1), 155 - 170.

Deodhar, V. (2004, March). *Does the housing market value heritage? some empirical evidence* (Research Papers No. 0403). Macquarie University, Department of Economics.

Department for Transport. (2009). *Transport statistics bulletin, road statistics 2008: Traffic, speeds and congestion.* London.

Devereux, M. P., Griffith, R. & Simpson, H. (2007, April). Firm location decisions, regional grants and agglomeration externalities. *Journal of Public Economics*, *91*(3-4), 413-435.

Dewees, D. (1976). The effect of a subway on residential property values in Toronto. *Journal of Urban Economics*, *3*(4), 357 - 369.

Dickens, M. (2013). *Public transport ridership report* (Tech. Rep.). Washington, DC: Amerian Public Transport Association.

DJ Mag. (2009). *Top 100 clubs.* Retrieved from `http://www.djmag.com/top100clubs/` (retrieved: 01/11/2012)

Do, A. & Grudnitski, G. (1995). Golf courses and residential house prices: An empirical examination. *The Journal of Real Estate Finance and Economics*, *10*(3), 261-270.

Drakos, K. & Konstantinou, P. T. (2011). *Terrorism shocks and public spending: Panel VAR evidence from Europe* (Economics of Security Working Paper Series No. 48). DIW Berlin, German Institute for Economic Research.

Dudczak, J. & Dudczak, U. (2012). *Berliner-bahnen.de – Bahnen im Berliner Raum.* Retrieved from `http://www.berliner-bahnen.de` (Retrieved August 14th, 2012)

Duranton, G. & Turner, M. A. (2011, September). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, *101*(6), 2616-52.

Duranton, G. & Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, *79*(4), 1407-1440.

Eickelpasch, A., Brenke, G. K., K. & Gornig, M. (2009). Wachstums- und Beschäftigungschancen in wissensintensiven Dienstleistungsmärkten in Ostdeutschland. *DIW Politikberatung kompakt*, *54*.

Elkins, T. & Hofmeister, B. (1988). *Berlin: The spatial structure of a divided city.* Methuen.

Emmett, D. (2012). *Fabulous Chicago: A great city's history and people.* Garrett County Press.

Fack, G. & Grenet, J. (2010). When do better schools raise housing prices? evidence from Paris public and private schools. *Journal of Public Economics*, *94*(1–2), 59 - 77.

Falck, O., Fritsch, M. & Heblich, S. (2011). The phantom of the opera: Cultural amenities, human capital, and regional economic growth. *Labour Economics*, *18*(6), 755-766.

Feitelson, E. I., Hurd, R. E. & Mudge, R. R. (1996). The impact of airport noise on willingness to pay for residences. *Transportation Research Part D: Transport and Environment*, *1*(1), 1 - 14.

Feng, X. & Humphreys, B. R. (2008, August). *Assessing the economic impact of sports facilities on residential property values: A spatial hedonic approach* (Working Papers No. 0812). International Association of Sports Economists & North American Association of Sports Economists.

Figlio, D. N. & Lucas, M. E. (2004, September). What's in a grade? school report cards and the housing market. *American Economic Review*, *94*(3), 591-604.

Figueiredo, O., Guimarães, P. & Woodward, D. (2002, September). Home-field advantage: location decisions of Portuguese entrepreneurs. *Journal of Urban Economics*, *52*(2), 341-361.

Fischel, F. A. (2001a). *The homevoter hypothesis.* Cambridge, MA.: Harvard University Press.

Fischel, F. A. (2001b). Why are there NIMBYs? *Land Economics*, *77*(1), 144-152.

Florida, R. (2002). *The rise of the creative class - and how it's transforming work, leisure, community and everyday life.* New York: Basic Books.

Florida, R., Mellander, C. & Stolarick, K. (2011). Beautiful places: The role of perceived aesthetic beauty in community satisfaction. *Regional Studies*, *45*(1), 33 - 48.

Ford, D. A. (1989). The effect of historic district designation on single-family home prices. *Real Estate Economics*, *17*(3), 353–362.

Friedrichs, J. (2000). *Gentrification* (Vol. Großstadt. SoziologischeStichwort; H. Häußermann, Ed.). Opladen: Leske + Budrich.

Fujita, M. (1988). A monopolistic competition model of spatial agglomeration: Differentiated product approach. *Regional Science and Urban Economics*, *18*(1), 87 - 124.

Fujita, M. & Ogawa, H. (1982, May). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics*, *12*(2), 161-196.

Gabriel, S. A. & Rosenthal, S. (2004, February). Quality of the business environment versus quality of life: Do firms and households like the same cities? *The Review of Economics and Statistics*, *86*(1), 438-444.

Gat, D. (1998). Urban focal points and design quality influence rents: The Tel Aviv office market. *Journal of Real Estate Research*, *16*(2), 229-247.

Gelbe Seiten Deutschland. (2012). *Gelbe Seiten.*

Gibbons, S. (2004). The costs of urban property crime. *The Economic Journal*, *114*(499), pp. F441-F463.

Gibbons, S. & Machin, S. (2003, March). Valuing english primary schools. *Journal of Urban Economics*, *53*(2), 197-219.

Gibbons, S. & Machin, S. (2005, January). Valuing rail access using transport innovations. *Journal of Urban Economics*, *57*(1), 148-169.

Gibbons, S. & Machin, S. (2006). Paying for primary schools: Admission constraints, school popularity or congestion? *The Economic Journal*, *116*(510), C77–C92.

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.

Gibbons, S., Machin, S. & Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, *75*(0), 15 - 28.

Gibbons, S., Mourato, S. & Resende, G. (2011, March). *The amenity value of English nature: A hedonic price approach* (SERC Discussion Papers No. 0074). Spatial Economics Research Centre, LSE.

Gibbons, S. & Silva, O. (2008). Urban density and pupil attainment. *Journal of Urban Economics*, *63*(2), 631 - 650.

Gilchrist, S. & Himmelberg, C. (1999, July). Investment: Fundamentals and finance. In *Nber macroeconomics annual 1998, volume 13* (p. 223-274). National Bureau of Economic Research, Inc.

Gillanders, R. (2011, August). *The effects of foreign aid in Sub-Saharan africa* (Working Papers No. 201116). School Of Economics, University College Dublin.

Glaeser, E. L. (2011). *Triumph of the city: How our greatest invention makes US richer, smarter, greener, healthier, and happier.* Penguin Group US.

Glaeser, E. L. & Gottlieb, J. D. (2006). Urban resurgence and the consumer city. *Urban Studies*, *43*(8), 1275-1299.

Glaeser, E. L., Kerr, W. R. & Ponzetto, G. A. (2010, January). Clusters of entrepreneurship. *Journal of Urban Economics*, *67*(1), 150-168.

Glaeser, E. L., Kolko, J. & Saiz, A. (2001, January). Consumer city. *Journal of Economic Geography*, *1*(1), 27-50.

Glaeser, E. L. & Sacerdote, B. (1999). Why is there more crime in cities? *Journal of Political Economy*, *107*(S6), pp. S225-S258.

Glazer, A., Gradstein, M. & Ranjan, P. (2003). Consumption variety and urban agglomeration. *Regional Science and Urban Economics*, *33*(6), 653 - 661.

Gottlieb, P. D. (1995). Residential amenities, firm location and economic development. *Urban Studies*, *32*(9), 1413-1436.

Gottwaldt, A. (1994). *Das Berliner U- und S-Bahnnetz: eine Geschichte in Streckenplänen.* Argon.

Gourieroux, C., Monfort, A. & Trognon, A. (1984, May). Pseudo maximum likelihood methods: Applications to poisson models. *Econometrica*, *52*(3), 701-20.

Gravier-Rymaszewska, J. (2012). *How aid supply responds to economic crises: A panel VAR approach* (Working Papers No. UNU-WIDER Research Paper). World Institute for Development Economic Research (UNU-WIDER).

Greenwood, M. J., Hunt, G. L., Rickman, D. S. & Treyz, G. I. (1991, December). Migration, regional equilibrium, and the estimation of compensating differentials. *American Economic Review*, *81*(5), 1382-90.

Grossman, J., Keating, A., Reiff, J., Library, N. & Society, C. H. (2004). *The encyclopedia of Chicago.* University of Chicago Press.

Gründerszene. (2013). *Datenbank.* Retrieved from `http://www.gruenderszene.de/datenbank/unternehmen` (retrieved : 25/04/2013)

Guimarães, P., Figueirdo, O. & Woodward, D. (2003, February). A tractable approach to the firm location decision problem. *The Review of Economics and Statistics*, *85*(1), 201-204.

Guimarães, P., Figueiredo, O. & Woodward, D. (2000, January). Agglomeration and the location of foreign direct investment in Portugal. *Journal of Urban Economics*, *47*(1), 115-135.

Gutachterausschuss für Grundstückswerte in Berlin. (1992). *Bodenrichtwertatlas.*

Gyourko, J. & Tracy, J. (1991, August). The structure of local public finance and the quality of life. *Journal of Political Economy*, *99*(4), 774-806.

Hausbesetzungs Geschichte Berlin. (2010). *Über 200 besetzte Häuser in Berlin seit 1987.* Blog. (retrieved: 29/06/2012)

Hellman, D. A. & Naroff, J. L. (1979). The impact of crime on urban residential property values. *Urban Studies*, *16*, 105-112.

Herdegen, G. (1992). *Einstellungen zur Deutschen Einheit* (Vol. Handwörterbuch zur deutschen Einheit; W. Weidenfeld & K.-R. Korte, Eds.). Frankfurt am Main: Campus.

Hilber, C. & Mayer, C. (2009, January). Why do households without children support local public schools? linking house price capitalization to school spending. *Journal of Urban Economics*, *65*(1), 74-90.

Hilber, C. & Vermeulen, W. (2010). *The impacts of restricting housing supply on house prices and affordability* (Tech. Rep.). Department for Communities and Local Governmnet Report.

HM Government. (2010). *The government's statement on the historic environment for England.* London: DCMS.

Holl, A. (2004, May). Manufacturing location and impacts of road transport infrastructure: Empirical evidence from Spain. *Regional Science and Urban Economics*, *34*(3), 341-363.

Holl, A. & Viladecans-Marsal, E. (2011). *Infrastructure and cities: The impact of new highways on urban growth* (Tech. Rep.). Universitat de Barcelona d'Institut d'Economia de Barcelona Working Paper.

Holman, N. & Ahlfeldt, G. (2012). *No escape? the coordination problem in heritage preservation.*

Holtz-Eakin, D., Newey, W. & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica*, *56*(6), pp. 1371-1395.

Hong, J. (2009). Firm heterogeneity and location choices: Evidence from foreign manufacturing investments in China. *Urban Studies*, *46*(10), 2143-2157.

Hough, D. E. & Kratz, C. G. (1983). Can "good" architecture meet the market test? *Journal of Urban Economics*, *14*(1), 40 - 54.

Hsu, W.-T. & Zhang, H. (2011). *The fundamental law of highway congestion: Evidence from Japanese expressways* (Tech. Rep.). Chinese University of Hong Kong Working Paper.

Im, K. S., Pesaran, M. H. & Shin, Y. (2003, July). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, *115*(1), 53-74.

Imbens, G. W. & Lemieux, T. (2008, February). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615-635.

Immobilien Scout. (2012). *Immobilienpreise in Kartenform.* Retrieved from `http://www.immobilienscout24.de/de/immobilienbewertung/immobilienpreise.jsp` (retrieved: 26/06/2012)

Inoue, A., Kilian, L. & Kiraz, F. B. (2009, October). Do actions speak louder than words? household expectations of inflation based on micro consumption data. *Journal of Money, Credit and Banking*, *41*(7), 1331-1363.

Isaksen, A. (2004). Knowledge-based clusters and urban location: The clustering of software consultancy in Oslo. *Urban Studies*, *41*(5-6), 1157-1174.

Kane, T. J., Staiger, D. O. & Riegg, S. K. (2005, May). *School quality, neighborhoods and housing prices: The impacts of school desegregation* (NBER Working Papers No. 11347). National Bureau of Economic Research, Inc.

Kavetsos, G. (2012). The impact of the London Olympics announcement on property prices. *Urban Studies*, *49*(7), 1453-1470.

Kennedy, P. (2003). *A guide to econometrics.* The MIT Press.

Kiebert, W. (2004). *Die Berliner S-Bahn 1924 bis heute: Geschichte, Technik, Hintergründe.* Transpress-Verlag.

Kiebert, W. (2008). *Die Berliner Nordsüd-S-Bahn.* Transpress Verlag.

Kilian, L. (2011, August). *Structural vector autoregressions* (CEPR Discussion Papers No. 8515). C.E.P.R. Discussion Papers.

Klünner, H. (1985). *S- und U-Bahnarchitektur in Berlin.* Berlin (West): Der Senator für Bau- und Wohnungswesen.

Kolko, J. (1999, November). *Can i get some service here? information technology, service industries, and the future of cities.*

Kolko, J. (2001). Silicon mountain, Silicon molehills. *Discussion Paper*, *2001/2*.

Koster, H. R. A., Van Ommeren, J. N. & Rietveld, P. (2012). *Upscale neighbourhoods: Historic amenities, income and spatial sorting of households* (Tech. Rep.). VU Unversity Amsterdam.

Kotkin, J. (2000). *The New Geography: How the digital revolution is reshaping the American landscape.* Random House.

Krugman, P. (2010). Increasing returns in a comparative advantage world. *IPC Working Paper Series*, *91*.

Kurpjuweit, K. & Meyer-Kronthaler, J. (2009). *Berliner U-bahn: in Fahrt seit über hundert Jahren.* Edition Q.

Lazrak, F., Nijkamp, P., Rietveld, P. & Rouwendal, J. (2013). The market value of cultural heritage in urban areas: an application of spatial hedonic pricing. *Journal of Geographical Systems*, 1-26.

Lee, C.-I. (2007, March). Does provision of public rental housing crowd out private housing investment? a panel VAR approach. *Journal of Housing Economics*, *16*(1), 1-20.

Lee, N. & Nathan, M. (2010). Knowledge workers, cultural diversity and innovation: Evidence from London. *Int. J. Knowledge-Based Development*, *1*, 53-78.

Leichenko, R. M., Coulson, N. E. & Listokin, D. (2001). Historic preservation and residential property values: An analysis of Texas cities. *Urban Studies*, *38*(11), 1973-1987.

Levin, A., Lin, C.-F. & James Chu, C.-S. (2002, May). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, *108*(1), 1-24.

Levinson, D. (2008, January). Density and dispersion: the co-development of land use and rail in London. *Journal of Economic Geography*, *8*(1), 55-77.

Levinson, D. & Karamalaputi, R. (2003, September). Induced supply: A model of highway network expansion at the microscopic level. *Journal of Transport Economics and Policy*, *37*(3), 297-318.

Leyden, F. (1933). *Gross-Berlin. Geographie der Weltstadt.* Breslau: Ferdinand Hirt.

List, J. A. (2001, May). US county-level determinants of inbound FDI: evidence from a two-step modified count data model. *International Journal of Industrial Organization*, *19*(6), 953-973.

Loop, A. (2012). *Untergrundbahn.* Retrieved from `http://www.untergrundbahn.de` (Retrieved August 14th, 2012)

Love, I. & Zicchino, L. (2006, May). Financial development and dynamic investment behavior: Evidence from panel VAR. *The Quarterly Review of Economics and Finance*, *46*(2), 190-210.

Lucas, R. E. J. & Rossi-Hansberg, E. (2002). On the internal structure of cities. *Econometrica*, *70*(4), pp. 1445-1476.

Luisenstädtischer Bildungsverein e.V. (2012). *Berlin von A bis Z.* (Retrieved August 14th, 2012)

Lynch, A. K. & Rasmussen, D. W. (2001). Measuring the impact of crime on house prices. *Applied Economics*, *33*(15), 1981.

Machin, S. (2011). Houses and schools: Valuation of school quality through the housing market. *Labour Economics*, *18*(6), 723 - 729. (European Association of Labour Economists, 3rd World Conference EALE/SOLE, London UK, 17-19 June2010)

Machin, S. & Salvanes, K. (2007). *Valuing school choice and social interactions: Evidence from an admissions reform.*

Mahan, B. L., Polasky, S. & Adams, R. M. (2000). Valuing urban wetlands: A property price approach. *Land Economics*, *76*(1), pp. 100-113.

M'Amanja, D. & Morrissey, O. (2006). *Foreign aid, investment, and economic growth in Kenya: A time series approach.* University of Nottingham. Centre for research in economic development and international trade (CREDIT).

Mauruszat, A. (2010). *Bahnstrecken im Land Brandenburg - interaktive Karte.* Retrieved from `http://www.bahnstrecken.de` (Retrieved February 25th, 2011)

McCullagh, P. & Nelder, J. A. (1989). *Generalized linear models.* London: Chapman and Hall.

McFadden. (1974). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in econometrics.* Academic Press.

McGrane, S. (2000). *GO TO: Berlin.* (retrieved: 01/11/2012)

McMillen, D. P. (1996). One hundred fifty years of land values in Chicago: A nonparametric approach. *Journal of Urban Economics*, *40*(1), 100 - 124.

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.

Michaels, G. (2008, November). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, *90*(4), 683-701.

Miller, N. & Peng, L. (2006, August). Exploring metropolitan housing price volatility. *The Journal of Real Estate Finance and Economics*, *33*(1), 5-18.

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

Mills, E. S. (1969). *The value of urban land* (H. Perloff, Ed.). Baltimore: Resources for the Future, Inc.

Minnesota Population Center. (2011). *National historical geographic information system: Version 2.0.* Minneapolis, MN.

Moorhouse, J. C. & Smith, M. S. (1994). The market for residential architecture: 19th century row houses in Bostons South End. *Journal of Urban Economics*, *35*(3), 267 - 277.

Mourato, S., Atkinson, G., Collins, M., Gibbons, S., MacKerron, G. & Resende, G. (2010). *Economic analysis of cultural services* (Tech. Rep.). U. N. E. A. Report.

Müller, G. (1881-1914). *Karte zur Berechnung des Grund- und Bodenwertes in Berlin, Charlottenburg und Westend.* Map. Berlin.

Muth, R. F. (1969). *Cities and housing: The spatial pattern of urban residential land use.* Chicago: University of Chicago Press.

Narwold, A., Sandy, J. & Tu, C. (2008). Historic designation and residential property values. *International Real Estate Review*, *11*(1), 83-95.

Nelson, J. (2004). Meta-analysis of airport noise and hedonic property values. *Journal of Transport Economics and Policy*, *38*(1), 1-27.

Nelson, J. P. (2008). Hedonic property value studies of transportation noise: Aircraft and road traffic. In A. Baranzini, J. Ramirez, C. Schaerer & P. Thalmann (Eds.), *Hedonic methods in housing markets* (p. 57-82). Springer New York.

Nguyen-Hoang, P. & Yinger, J. (2011). The capitalization of school quality into house values: A review. *Journal of Housing Economics*, *20*(1), 30 - 48.

Nitsch, V. (2009). *Fly or cry: Is airport noise costly?* (CESifo Working Paper Series No. 2732). CESifo Group Munich.

Noonan, D. S. (2007). Finding an impact of preservation policies: Price effects of historic landmarks on attached homes in Chicago, 1990-1999. *Economic Development Quarterly*, *21*(1), 17-33.

Noonan, D. S. & Krupka, D. J. (2011). Making - or picking - winners: Evidence of internal and external price effects in historic preservation policies. *Real Estate Economics*, *39*(2), 379–407.

Oates, W. E. (1969, Nov./Dec.). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis. *Journal of Political Economy*, *77*(6), 957-71.

OpenStreetMap. (2013). *OpenStreetMap-Projekt.* (Data provided as download provided by Geofabrik GmbH; retrieved: 22/01/2013)

Osei, R., Morrissey, O. & Lloyd, T. (2005). The fiscal effects of aid in Ghana. *Journal of International Development*, *17*(8), 1037-1053.

O'Sullivan, A. (2009). *Urban economics.* New York: McGraw-Hill.

Pendlebury, J. (2009). *Pendlebury.* London: Routledge.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265-312.

Phillips, P. C. B. & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrik*, *75*, 335-346.

Pierce, B. L. (1937). *The beginning of a city, 1673 - 1848. a history of Chicago.* New York, London: Knopf.

Pierce, B. L. (1940). *From town to city, 1848 - 1871. a history of Chicago.* New York, London: Knopf.

Rapp, T. (2009). *Lost and sound: Berlin, Techno und der Easyjetset.* Frankfurt (Main): Suhrkamp Verlage.

Rappaport, J. (2007, May). Moving to nice weather. *Regional Science and Urban Economics*, *37*(3), 375-398.

Redding, S. J. & Sturm, D. M. (2008, December). The costs of remoteness: Evidence from German division and reunification. *American Economic Review*, *98*(5), 1766-97.

Redding, S. J., Sturm, D. M. & Wolf, N. (2011, August). History and industry location: Evidence from German airports. *The Review of Economics and Statistics*, *93*(3), 814-831.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, *90*(6), pp. 1257-1278.

Rodríguez-Pose, A. & Ketterer, T. D. (2012, October). Do local amenities affect the appeal of regions in Europe for migrants? *Journal of Regional Science*, *52*(4), 535-561.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, *82*(1), pp. 34-55.

Rosen, S. (1979). *Wage-based indexes of urban quality of life* (S. Miezkowski, Ed.). Baltimore: Johns Hopkins University Press.

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rosenthal, L. (2003). The value of secondary school quality. *Oxford Bulletin of Economics and Statistics*, *65*(3), 329–355.

Rosenthal, S. & Strange, W. C. (2003, May). Geography, industrial organization, and agglomeration. *The Review of Economics and Statistics*, *85*(2), 377-393.

Rosenthal, S. & Strange, W. C. (2005). The geography of entrepreneurship in the New York metropolitan area. *Economic Policy Review*(Dec), 29-53.

Rossi-Hansberg, E., Sarte, P.-D. & Owens, R. (2010, 06). Housing externalities. *Journal of Political Economy*, *118*(3), 485-535.

Santos Silva, J. & Tenreyro, S. (2006, November). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641-658.

Santos Silva, J. & Tenreyro, S. (2011, August). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters*, *112*(2), 220-222.

Sarafidis, V. & Robertson, D. (2009, 03). On the impact of error cross-sectional dependence in short dynamic panel estimation. *Econometrics Journal*, *12*(1), 62-81.

Schaeffer, P. V. & Millerick, C. A. (1991). The impact of historic district designation on property values: An empirical study. *Economic Development Quarterly*, *5*(4), 301-312.

Schiff, N. (2013, October). Cities and product variety [Computer software manual].

Schomacker, M. (2009). *Berliner-Untergrundbahn.de.* (Retrieved February 25th, 2011)

Schwannhäußer, A. (2007). *Kosmonauten des Underground, Ethnografie einer Berliner Szene.* Berlin: Campus Verlag. (Diss. Humboldt-Universität Berlin)

Schwartz, A. E., Susin, S. & Voicu, I. (2003). Has falling crime driven New York City's real estate boom? *Journal of Housing Research*, *14*(1), 101-135.

Senatsverwaltung für Stadtentwicklung und Umwelt. (2012). *Denkmalliste, Denkmalkarte, Denkmaldatenbank.* (Retrieved August 14th, 2012)

Shapiro, J. M. (2006). Smart cities: Quality of life, productivity, and the growth effects of human capital. *The Review of Economics and Statistics*, *88*(2), pp. 324-335.

Siebenhaar, K., Blisse, M., Golz, R. & Lehmann, U. (1998). *Berlin für junge Leute* (H. V. G. im Auftrag des Landes Berlin, Ed.). Druckhaus Berlin-Mitte.

Silverman, B. (1986). *Density estimation for statistics and data analysis.* New York: Chap.

Sims, C. A. (1980, January). Macroeconomics and reality. *Econometrica*, *48*(1), 1-48.

Sinus Sociovision GmbH. (2011). Sinus-Milieus [Computer software manual]. Heidelberg.

Statistisches Amt der Stadt Berlin. (1920). *Statistisches Jahrbuch der Stadt Berlin.* Berlin.

Steven, H. (2005). *Investigations on noise emission of motor vehicles in road traffic: Research project 200 54 135, final report* (Tech. Rep.). RWTUEV Fahrzeug, Inst. for Vehicle Technology, Division noise, measurement technique, modelling.

Stock, J. H., Wright, J. H. & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, *20*(4), 518-529.

Storper, M. & Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, *4*(4), 351-370.

Strange, I. & Whitney, D. (2003). The changing roles and purposes of heritage conservation in the UK. *Planning Practice and Research*, *18*(2-3), 219-229.

Straschewski, M. (2011). *Geschichte und Geschichten rund um die Berliner S-Bahn.* (Retrieved February 25th, 2011)

Suedekum, J., Wolf, K. & Blien, U. (2012). Cultural diversity and local labour markets. *Regional Studies*, *0*(0), 1-19.

Tenreyro, S. (2007, March). On the trade impact of nominal exchange rate volatility. *Journal of Development Economics*, *82*(2), 485-508.

Thaler, R. (1978). A note on the value of crime control: Evidence from the property market. *Journal of Urban Economics*, *5*(1), 137 - 145.

Thill, J.-C. & Kim, M. (2005, 06). Trip making, induced travel demand, and accessibility. *Journal of Geographical Systems*, *7*(2), 229-248.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, *64*, 416.

tip. (2011, 09). Berlin [Computer software manual]. Berlin.

Vandell, K. D. & Lane, J. S. (1989). The economics of architecture and urban design: Some preliminary findings. *Real Estate Economics*, *17*(2), 235–260.

van Duijn, M. & Rouwendal, J. (2013). Cultural heritage and the location choice of Dutch households in a residential sorting model. *Journal of Economic Geography*, *13*(3), 473-500.

Verblichene Locations. (1999). (retrieved: 16/01/2012)

Vetter, R., Spell, C. & Ward, C. (1994, oct.). Mosaic and the world wide web. *Computer*, *27*(10), 49 -57.

von Thünen, J. H. (1826). *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie, oder Untersuchungen über den Einfluss, den die Getreidepreise, der Reichtum des Bodens und die Abgaben auf den Ackerbau ausüben.* Hamburg: Perthes.

Waldfogel, J. (2008). The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics*, *63*(2), 567 - 582.

Wu, J., Adams, R. M. & Plantinga, A. J. (2004). Amenities in an urban equilibrium model: Residential development in Portland, Oregon. *Land Economics*, *80*(1), 19-32.

Xie, F. & Levinson, D. (2010, May). How streetcars shaped suburbanization: a Granger causality analysis of land use and transit in the Twin Cities. *Journal of Economic Geography*, *10*(3), 453-470.

Young, D. (1998). *Chicago Transit: An illustrated history.* Northern Illinois University Press.

Academic career

Kristoffer Möller

Technical University Darmstadt
International Economics
Marktplatz 15 - Residenzschloss
D-64283 Darmstadt
moeller@vwl.tu-darmstadt.de | www.kristoffer-moeller.de

Research interests

Urban Economics, Economic Geography, International Economics

Education

08/2010 - 05/2014	PhD in Economics, Technical University Darmstadt, Germany Thesis: The Role of Amenities in the Location Decision of Households and Firms
10/2007 - 03/2007	Master of Arts in International Economics, Georg-August-University of Göttingen, Germany Thesis: Determinants of Export Diversification in Brazil
10/2004 - 03/2007	Bachelor of Arts in International Economics, Georg-August-University of Göttingen, Germany Thesis: Development and Perspectives of the ASEAN

Profess. experience

03/2013 -	Research assistant at the Centre for Economic Performance, London School of Economics and Political Science, UK
08/2010 - 08/2013	Research assistant at the Chair for International Economics, Technical University Darmstadt, Germany
04/2008 - 04/2010	Student research assistant at the Chair of Economic Policy and SME Research, Georg-August-University of Göttingen, Germany

Stays abroad

01/2013 - 04/2013	Visiting PhD, Spatial Economic Research Centre, London School of Economics and Political Science, UK
09/2007 - 04/2008	Studies abroad at the University of Sevilla, Spain

Presentations

08/2013	53rd European Congress of the Regional Science Association, (ERSA), Urban Economics Association, Palermo, Italy
06/2013	6th Summer Conference in Regional Science, Dortmund, Germany
05/2013	5th Annual SERC Conference, LSE, London, UK
03/2013	SERC work in progress seminar, LSE, London, UK
02/2013	Research Workshop Economic Geography and International Trade (EGIT), Berlin, Germany
08/2012	52nd European Congress of the Regional Science Association, (ERSA), Bratislava, Slovakia
02/2012	Research Workshop Economic Geography and International Trade (EGIT), Duisburg, Germany
09/2011	51st European Congress of the Regional Science Association, (ERSA), Barcelona, Spain
03/2011	Research Workshop Economic Geography and International Trade (EGIT), Munich, Germany

Affidavit

I hereby declare that the dissertation entitled

The role of amenities in the location decision of households and firms

is my own work. I have used only the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given. I have not presented this thesis or parts thereof to a university as part of an examination or degree.