**Gabriel M. Ahlfeldt, Stephen J. Redding, Daniel M. Sturm, Nikolaus Wolf**

# The economics of density: evidence from the Berlin Wall

## Article (Accepted version) (Refereed)

**Original citation:**
Ahlfeldt, Gabriel M., Redding, Stephen, Sturm, Daniel M. and Wolf, Nikolaus (2015) *The economics of density: evidence from the Berlin Wall*. Econometrica, 83 (6). pp. 2127-2189. ISSN 0012-9682
DOI: 10.3982/ECTA10876

# The Economics of Density: Evidence from the Berlin Wall*

Gabriel M. Ahlfeldt
*London School of Economics and CEPR*

Stephen J. Redding
*Princeton University, NBER and CEPR*

Daniel M. Sturm
*London School of Economics and CEPR*

Nikolaus Wolf
*Humboldt University and CEPR*

February 26, 2015

**Abstract**

This paper develops a quantitative model of internal city structure that features agglomeration and dispersion forces and an arbitrary number of heterogeneous city blocks. The model remains tractable and amenable to empirical analysis because of stochastic shocks to commuting decisions, which yield a gravity equation for commuting flows. To structurally estimate agglomeration and dispersion forces, we use data on thousands of city blocks in Berlin for 1936, 1986 and 2006 and exogenous variation from the city's division and reunification. We estimate substantial and highly localized production and residential externalities. We show that the model with the estimated agglomeration parameters can account both qualitatively and quantitatively for the observed changes in city structure. We show how our quantitative framework can be used to undertake counterfactuals for changes in the organization of economic activity within cities in response for example to changes in the transport network.

Keywords: agglomeration, cities, commuting, density, gravity

JEL: N34, O18, R12

*We are grateful to the European Research Council (Grant Number 263733), German Science Foundation (Grant Numbers STU 477/1-1 and WO 1429/1-1), the Centre for Economic Performance and Princeton University for financial support. Thilo Albers, Horst Bräunlich, Daniela Glocker, Kristoffer Möller, David Nagy, Utz Pape, Ferdinand Rauch, Claudia Steinwender, Sevrin Waights and Nicolai Wendland provided excellent research assistance. We would like to thank the editor, four anonymous referees, conference and seminar participants at AEA, Arizona, Barcelona, Berkeley, Brussels, CEPR, Chicago, Clemson, Columbia, EEA, Harvard, Mannheim, Marseille, MIT, Munich, NARSC, Luzern, Nottingham, NYU, Oxford, Paris, Princeton, Rutgers, Shanghai, Stanford, Tübingen, UCLA, Virginia and the World Bank for helpful comments. We are also grateful to Dave Donaldson, Cecilia Fieler, Gene Grossman, Bo Honore, Ulrich Müller, Sam Kortum, Eduardo Morales and Esteban Rossi-Hansberg for their comments and suggestions. The usual disclaimer applies.

# 1 Introduction

Economic activity is highly unevenly distributed across space, as reflected in the existence of cities and the concentration of economic functions in specific locations within cities, such as Manhattan in New York and the Square Mile in London. Understanding the strength of the agglomeration and dispersion forces that underlie these concentrations of economic activity is central to a range of economic and policy questions. These forces shape the size and internal structure of cities, with implications for the incomes of immobile factors, congestion costs and city productivity. They also determine the impact of public policy interventions, such as transport infrastructure investments and urban development and taxation policies.

Although there is a long literature on economic geography and urban economics dating back to at least Marshall (1920), a central challenge remains distinguishing agglomeration and dispersion forces from variation in locational fundamentals. While high land prices and levels of economic activity in a group of neighboring locations are consistent with strong agglomeration forces, they are also consistent with shared amenities that make these locations attractive places to live (e.g. leafy streets and scenic views) or common natural advantages that make these locations attractive for production (e.g. access to natural water). This challenge has both theoretical and empirical dimensions. From a theoretical perspective, to develop tractable models of cities, the existing literature typically makes simplifying assumptions such as monocentricity or symmetry, which abstracts from variation in locational fundamentals and limits the usefulness of these models for empirical work. From an empirical perspective, the challenge is to find exogenous sources of variation in the surrounding concentration of economic activity to help disentangle agglomeration and dispersion forces from variation in locational fundamentals.

In this paper, we develop a quantitative theoretical model of internal city structure. This model incorporates agglomeration and dispersion forces and an arbitrary number of heterogeneous locations within the city, while remaining tractable and amenable to empirical analysis. Locations differ in terms of productivity, amenities, the density of development (which determines the ratio of floor space to ground area), and access to transport infrastructure. Productivity depends on production externalities, which are determined by the surrounding density of workers, and production fundamentals, such as topography and proximity to natural supplies of water. Amenities depend on residential externalities, which are determined by the surrounding density of residents, and residential fundamentals, such as access to forests and lakes. Congestion forces take the form of an inelastic supply of land and commuting costs that are increasing in travel time, where travel time in turn depends on the transport network.[1]

We combine this quantitative theoretical model with the natural experiment of Berlin's division in the aftermath of the Second World War and its reunification following the fall of the Iron Curtain. The division of Berlin severed all local economic interactions between East and West Berlin, which corresponds in the model to prohibitive trade and commuting costs and no production and residential externalities between these two parts of the city. We make use of a remarkable and newly-collected dataset for Berlin, which

[1]We use "production fundamental" to refer to a characteristic of a location that directly affects productivity (e.g. natural water) independently of the surrounding economic activity. We use "residential fundamental" to refer to a characteristic of a location that directly affects the utility of residents (e.g. forests) independently of the surrounding economic activity.

includes data on land prices, employment by place of work (which we term "workplace employment") and employment by place of residence (which we term "residence employment") covering the pre-war, division and reunification periods. We first present reduced-form evidence in support of the model's qualitative predictions without imposing the full structure of the model. We show that division leads to a reorientation of the gradient in land prices and employment in West Berlin away from the main pre-war concentration of economic activity in East Berlin, while reunification leads to a reemergence of this gradient. In contrast, there is little effect of division or reunification on land prices or employment along other more economically remote sections of the Berlin Wall. We show that these results are not driven by pre-trends prior to division or reunification. We also show that these results are robust to controlling for a host of observable block characteristics, including controls for access to the transport network, schools, parks and other green areas, lakes and other water areas, Second World War destruction, land use, urban regeneration policies and government buildings post reunification.

We next examine whether the model can account quantitatively for the observed impact of division and reunification. We show that the model implies a gravity equation for commuting flows, which can be used to estimate its commuting parameters. Using these estimates, we determine overall measures of productivity, amenities and the density of development for each block, without making any assumptions about the relative importance or functional form of externalities and fundamentals as components of overall productivity and amenities. In the special case of the model in which overall productivity and amenities are exogenous, the model has a unique equilibrium, and hence can be used to undertake counterfactuals that have determinate predictions for the impact of division and reunification. We use these counterfactuals to show that the model with exogenous productivity and amenities is unable to account quantitatively for the observed impact of division and reunification on the pattern of economic activity within West Berlin.

We next use the exogenous variation from Berlin's division and reunification to structurally estimate both the agglomeration and commuting parameters. This allows us to decompose overall productivity and amenities for each block into production and residential externalities (which capture agglomeration forces) and production and residential fundamentals (which are structural residuals). Our identifying assumption is that changes in these structural residuals are uncorrelated with the exogenous change in the surrounding concentration of economic activity induced by Berlin's division and reunification. This identifying assumption requires that the systematic change in the pattern of economic activity in West Berlin following division and reunification is explained by the mechanisms in the model (the changes in commuting access and production and residential externalities) rather than by systematic changes in the pattern of structural residuals (production and residential fundamentals).

Our structural estimates of the model's parameters imply substantial and highly localized production externalities. Our central estimate of the elasticity of productivity with respect to the density of workplace employment is 0.07, which is towards the high end of the range of existing estimates using variation between cities. In contrast to these existing estimates using data across cities, our analysis makes use of variation within cities. Our structural estimate of the elasticity of productivity with respect to density holds constant the distribution of travel times within the city. In reality, a doubling in total city population is typically achieved by a combination of an increase in the density of economic activity and an expansion in geographical land area, which increases travel times within the city. Since we find that production externalities decay rapidly with travel times, this attenuates production externalities, which has to be taken into account when comparing estimates within and across cities. We also find substantial and highly localized residential externalities. Our central estimate of the elasticity of amenities with respect to the density of residents is 0.14, which is consistent with the view that consumption externalities are an important agglomeration force in addition to production externalities.

In the presence of agglomeration forces, there is the potential for multiple equilibria in the model. An advantage of our estimation approach is that it addresses this potential existence of multiple equilibria. We distinguish between calibrating the model to the observed data given known parameter values and estimating the model for unknown parameters. First, given known values for the model's parameters, we show that there is a unique mapping from these parameters and the observed data to the structural residuals (production and residential fundamentals). This mapping is unique regardless of whether the model has a single equilibrium or multiple equilibria, because the parameters, observed data and equilibrium conditions of the model (including profit maximization, zero profits, utility maximization and population mobility) contain enough information to solve for unique values of these structural residuals. Second, we estimate the model's parameters using the generalized method of moments (GMM) and moment conditions in terms of these structural residuals. Since these structural residuals are closed-form functions of the observed data and parameters, this estimation holds constant the observed endogenous variables of the model at their values in the data. In principle, these moment conditions need not uniquely identify the model parameters, because the objective function defined by them may not be globally concave. For example, the objective function could be flat in the parameter space or there could be multiple local minima corresponding to different combinations of parameters and unobserved fundamentals that explain the data. In practice, we show that the objective function is well behaved in the parameter space, and that these moment conditions determine a unique value for the parameter vector.

We also undertake counterfactuals in the estimated model with agglomeration forces. To address the potential for multiple equilibria in this case, we assume the equilibrium selection rule of searching for the counterfactual equilibrium closest to the observed equilibrium prior to the counterfactual. We show that the model with the estimated agglomeration parameters can generate counterfactual predictions for the treatment effects of division and reunification that are close to the observed treatment effects. We show how our quantitative framework can be used to undertake counterfactuals for changes in the organization of economic activity within cities in response, for example, to changes in the transport network.

Finally, we undertake a variety of over identification checks and robustness tests. First, using our estimates of the model's commuting parameters based on bilateral commuting flows for 2008, we show that the model is successful in capturing the cumulative distribution of commuters across travel times in the pre-war, division and reunification periods. Second, we find that the ratio of floor space to land area in the model is strongly related to separate data on this variable not used in the estimation of the model. Finally, we also find that production and residential fundamentals in the model are correlated in the expected way with observable proxies for these fundamentals.

Our paper builds on the large theoretical literature on urban economics. Much of this literature has analyzed the monocentric city model, in which firms are assumed to locate in a Central Business District (CBD) and workers decide how close to live to this CBD.[2] Lucas and Rossi-Hansberg (2002) were the first to develop a model of a two-dimensional city, in which equilibrium patterns of economic activity can be non-monocentric. In their model, space is continuous and the city is assumed to be symmetric, so that distance from the center is a summary statistic for the organization of economic activity within the city.[3] Empirically cities are, however, not perfectly symmetric because of variation in locational fundamentals, and most data on cities are reported for discrete spatial units such as blocks or census tracts.

Our contribution is to develop a quantitative theoretical model of internal city structure that allows for a large number of discrete locations within the city that can differ arbitrarily in terms of their natural advantages for production, residential amenities, land supply and transport infrastructure. The analysis remains tractable despite the large number of asymmetric locations because we incorporate a stochastic formulation of workers' commuting decisions that follows Eaton and Kortum (2002) and McFadden (1974). This stochastic formulation yields a system of equations that can be solved for unique equilibrium wages given observed workplace and residence employment in each location. It also provides microeconomic foundations for a gravity equation for commuting flows that has been found to be empirically successful.[4]

Our paper is also related to the broader literature on the nature and sources of agglomeration economies, as reviewed in Duranton and Puga (2004) and Rosenthal and Strange (2004). A large empirical literature has regressed wages, land prices, productivity or employment growth on population density.[5] We contribute to the small strand of research within this literature that has sought sources of exogenous variation in the surrounding concentration of economic activity. For example, Rosenthal and Strange (2008) and Combes, Duranton, Gobillon, and Roux (2010) use geology as an instrument for population density, exploiting the idea that tall buildings are easier to construct where solid bedrock is accessible. Greenstone, Hornbeck, and Moretti (2010) provide evidence on agglomeration spillovers by comparing changes in total factor productivity (TFP) among incumbent plants in "winning" counties that attracted a large manufacturing plant and "losing" counties that were the new plant's runner-up choice.[6]

Other related research has examined the effect of historical natural experiments on the location of economic activity, including Hanson (1996, 1997) using Mexican trade liberalization; Davis and Weinstein (2002, 2008) using the wartime bombing of Japan; Bleakley and Lin (2012) using historical portage sites; and Kline and Moretti (2014) using the Tennessee Valley Authority (TVA). Using the division and reunification of

[2]The classic urban models of Alonso (1964), Mills (1967) and Muth (1969) assume monocentricity. While Fujita and Ogawa (1982) and Fujita and Krugman (1995) allow for non-monocentricity, they model one-dimensional cities on the real line.

[3]For an empirical analysis of the symmetric-city model of Lucas and Rossi-Hansberg (2002), see Brinkman (2013).

[4]See Grogger and Hanson (2011) and Kennan and Walker (2011) for analyses of worker migration decisions using stochastic formulations of utility following McFadden (1974).

[5]See, for example, Ciccone and Hall (1996), Dekle and Eaton (1999), Glaeser and Mare (2001), Henderson, Ari, and Turner (1995), Moretti (2004), Rauch (1993), Roback (1982) and Sveikauskas (1975), as surveyed in Moretti (2011).

[6]Another related empirical literature has examined the relationship between economic activity and transport infrastructure, including Donaldson (2014), Baum-Snow (2007), Duranton and Turner (2012), Faber (2014) and Michaels (2008).

Germany, Redding and Sturm (2008) examine the effect of changes in market access on the growth of West German cities, and Redding, Sturm, and Wolf (2011) examine the relocation of Germany's air hub from Berlin to Frankfurt as a shift between multiple steady-states. In contrast to all of the above studies, which exploit variation across regions or cities, our focus is on the determinants of economic activity within cities.[7] Our main contribution is to develop a tractable quantitative model of internal city structure that incorporates agglomeration forces and a rich geography of heterogeneous location characteristics and structurally estimate the model using the exogenous variation of Berlin's division and reunification.

The remainder of the paper is structured as follows. Section 2 discusses the historical background. Section 3 outlines the model. Section 4 introduces our data. Section 5 presents reduced-form empirical results on the impact of Berlin's division and reunification. Section 6 uses the model's gravity equation predictions to determine the commuting parameters and solve for overall values of productivity, amenities and the density of development. Section 7 structurally estimates both the model's agglomeration and commuting parameters; uses these estimated parameters to decompose overall productivity and amenities into the contributions of externalities and fundamentals; and undertakes counterfactuals. Section 8 concludes.

## 2 Historical Background

The city of Berlin in its current boundaries was created in 1920 when the historical city and its surrounding agglomeration were incorporated under the Greater Berlin law ("Gross Berlin Gesetz"). The city comprises 892 square kilometers of land compared for example to 606 square kilometers for Chicago. The city was originally divided into 20 districts ("Bezirke"), which had minimal administrative autonomy.[8] The political process that ultimately led to the construction of the Berlin Wall had its origins in war-time planning during the Second World War. A protocol signed in London in September 1944 delineated zones of occupation in Germany for the American, British and Soviet armies after the eventual defeat of Germany. This protocol also stipulated that Berlin, although around 200 kilometers within the Soviet occupation zone, should be jointly occupied. For this purpose, Berlin was itself divided into separate occupation sectors.

The key principles underlying the drawing of the boundaries of the occupation sectors in Berlin were that the sectors should be geographically orientated to correspond with the occupation zones (with the Soviets in the East and the Western Allies in the West); the boundaries between them should respect the boundaries of the existing administrative districts of Berlin; and the American, British and Soviet sectors should be approximately equal in population (prior to the creation of the French sector from part of the British sector). The final agreement in July 1945 allocated six districts to the American sector (31 percent of the 1939 population and 24 percent of the area), four districts to the British sector (21 percent of the 1939 population and 19 percent of the area), two districts to the French sector (12 percent of the 1939 population

[7] Other research using within-city data includes Arzaghi and Henderson (2008) on the location of advertising agencies in Manhattan and Rossi-Hansberg, Sarte, and Owens (2010) on urban revitalization policies in Richmond, Virginia.

[8] The boundaries of these 20 districts were slightly revised in April 1938. During division, the East Berlin authorities created three new districts (Hellersdorf, Marzahn and Hohenschönhausen), which were created from parts of Weissensee and Lichtenberg. Except for a few other minor changes, as discussed in Elkins and Hofmeister (1988), the district boundaries remained unchanged during the post-war period until an administrative reform in 2001, which reduced the number of districts to twelve.

and 12 percent of the area), and eight districts to the Soviet sector (37 percent of the 1939 population and 46 percent of the area).[9]

The London protocol specifying the occupation sectors also created institutions for a joint administration of Berlin (and Germany more generally). The intention was for Berlin to be governed as a single economic and administrative unit by a joint council ("Kommandatura") with Soviet, American, British and French representatives. However, with the onset of the Cold War, the relationship between the Western allies and the Soviet Union began to deteriorate. In June 1948 the Western allies unilaterally introduced a new currency in their occupation zones and the Western sectors of Berlin. In retaliation the Soviet Union decided to block all road and rail access to the Western sectors of Berlin for nearly eleven months and West Berlin was supplied through the Berlin airlift during this time. The foundation of East and West Germany as separate states in 1949 and the creation of separate city governments in East and West Berlin further cemented the division of Germany and Berlin into Eastern and Western parts.

Following the adoption of Soviet-style policies of command and control in East Germany, the main border between East and West Germany was closed in 1952. While the implementation of these policies in East Berlin limited economic interactions with the Western sectors, the boundary between East and West Berlin remained formally open.[10] This open border resulted in some commuting of workers between East and West Berlin.[11] It also became a conduit for refugees fleeing to West Germany. To stem this flow of refugees, the East German authorities constructed the Berlin Wall in 1961, which ended all local economic interactions between East and West Berlin.

Figure 1 shows the pre-war land price gradient in Berlin and the path of the Berlin Wall. As apparent from the figure, the Berlin Wall consisted of an inner boundary between West and East Berlin and an outer boundary between West Berlin and East Germany. The inner boundary ran along the Western edge of the district Mitte, which contained Berlin's main administrative, cultural and educational institutions and by far the largest pre-war concentration of employment. The Berlin Wall cut through the pre-war transport network, intersecting underground railway ("U-Bahn") and suburban railway ("S-Bahn") lines, which were closed off at the boundaries with East Berlin or East Germany.[12] During the period of division, West Germany introduced a number of policies to support economic activity in West Berlin, such as subsidies to transportation between West Berlin and West Germany, reduced tax rates and an exemption from military service for residents of West Berlin. Whereas our empirical analysis exploits relative variation across

[9]The occupation sectors were based on the April 1938 revision of the boundaries of the 20 pre-war districts. For further discussion of the diplomatic history of the division of Berlin, see Franklin (1963) and Sharp (1975).

[10]While East Berlin remained the main concentration of economic activity in East Germany after division, only around 2 percent of West Berlin's exports from 1957-1967 were to East Germany (including East Berlin) and other Eastern block countries (see Lambrecht and Tischner 1969).

[11]Approximately 122,000 people commuted from West to East Berlin in the fall of 1949, but this number quickly declined after waves of mass redundancies of Western workers in East Berlin and stood at about 13,000 workers in 1961 just before the construction of the Berlin Wall. Commuting flows in the opposite direction are estimated to be 76,000 in 1949 and decline to 31,000 in 1953 before slowly climbing to 63,000 in 1961 (Roggenbuch 2008).

[12]In a few cases, trains briefly passed through East Berlin territory en route from one part of West Berlin to another. These cases gave rise to ghost stations ("Geisterbahnhöfe") in East Berlin, where trains passed through stations patrolled by East German guards without stopping.

locations within West Berlin, these policies applied equally to all of West Berlin.

While the division of Germany and Berlin appeared to be permanent, the Soviet policies of "Glasnost" and "Perestroika" introduced by Mikhail Gorbachev in 1985 started a process of opening up of Eastern Europe.[13] As part of this wider transformation, large-scale demonstrations in East Germany in 1989 led to the fall of the Berlin Wall on 9 November 1989. In the aftermath of these events, the East German system rapidly began to disintegrate. Only eleven months later East and West Germany were formally reunified on 3 October 1990. In June 1991 the German parliament voted to relocate the seat of the parliament and many of the federal ministries back to Berlin. As East and West Berlin again became part of the same city, suburban and underground rail lines and utility networks were rapidly reconnected. The reunification of the city was also accompanied by some urban regeneration initiatives and we include controls for these policies in our empirical analysis below.

# 3 Theoretical Model

To guide our empirical analysis, we develop a model in which the internal structure of the city is driven by a tension between agglomeration forces (in the form of production and residential externalities) and dispersion forces (in the form of commuting costs and an inelastic supply of land).[14]

We consider a city embedded within a wider economy. The city consists of a set of discrete locations or blocks, which are indexed by $i = 1, ..., S$. Each block has an effective supply of floor space $L_i$. Floor space can be used commercially or residentially, and we denote the endogenous fractions of floor space allocated to commercial and residential use by $\theta_i$ and $1 - \theta_i$, respectively.

The city is populated by an endogenous measure of $H$ workers, who are perfectly mobile within the city and the larger economy, which provides a reservation level of utility $\bar{U}$. Workers decide whether or not to move to the city before observing idiosyncratic utility shocks for each possible pair of residence and employment blocks within the city. If a worker decides to move to the city, they observe these realizations for idiosyncratic utility, and pick the pair of residence and employment blocks within the city that maximizes their utility. Firms produce a single final good, which is costlessly traded within the city and the larger economy, and is chosen as the numeraire ($p = 1$).[15]

Blocks differ in terms of their final goods productivity, residential amenities, supply of floor space and access to the transport network, which determines travel times between any two blocks in the city. We first develop the model with exogenous values of these location characteristics, before introducing endogenous agglomeration forces below.

[13] After the signing of the Basic Treaty ("Grundlagenvertrag ") in December 1972, which recognized "two German states in one German Nation", East and West Germany were accepted as full members of the United Nations. West German opinion polls in the 1980s show that less than 10 percent of respondents expected a re-unification to occur during their lifetime (Herdegen 1992).

[14] A more detailed discussion of the model and the technical derivations of all expressions and results reported in this section are contained in a separate web appendix.

[15] We follow the canonical urban model in assuming a single tradable final good and examine the ability of this canonical model to account quantitatively for the observed impact of division and reunification. In the web appendix, we discuss an extension of the model to introduce a non-traded good.

### 3.1 Workers

Workers are risk neutral and have preferences that are linear in a consumption index: $U_{ijo} = C_{ijo}$, where $C_{ijo}$ denotes the consumption index for worker $o$ residing in block $i$ and working in block $j$.[16] This consumption index depends on consumption of the single final good ($c_{ijo}$); consumption of residential floor space ($\ell_{ijo}$); residential amenities ($B_i$) that capture common characteristics that make a block a more or less attractive place to live (e.g. leafy streets and scenic views); the disutility from commuting from residence block $i$ to workplace block $j$ ($d_{ij} \geq 1$); and an idiosyncratic shock that is specific to individual workers and varies with the worker's blocks of employment and residence ($z_{ijo}$). This idiosyncratic shock captures the idea that individual workers can have idiosyncratic reasons for living and working in different parts of the city. In particular, the aggregate consumption index is assumed to take the Cobb-Douglas form:[17]

$$C_{ijo} = \frac{B_i z_{ijo}}{d_{ij}} \left(\frac{c_{ijo}}{\beta}\right)^{\beta} \left(\frac{\ell_{ijo}}{1-\beta}\right)^{1-\beta}, \qquad 0 < \beta < 1, \tag{1}$$

where the iceberg commuting cost $d_{ij} = e^{\kappa \tau_{ij}} \in [1, \infty)$ increases with the travel time ($\tau_{ij}$) between blocks $i$ and $j$. Travel time is measured in minutes and is computed based on the transport network, as discussed further in Section 4 below. The parameter $\kappa$ controls the size of commuting costs.

We model the heterogeneity in the utility that workers derive from living and working in different parts of the city following McFadden (1974) and Eaton and Kortum (2002). For each worker $o$ living in block $i$ and commuting to block $j$, the idiosyncratic component of utility ($z_{ijo}$) is drawn from an independent Fréchet distribution:

$$F(z_{ijo}) = e^{-T_i E_j z_{ijo}^{-\epsilon}}, \qquad T_i, E_j > 0,\ \epsilon > 1, \tag{2}$$

where the scale parameter $T_i > 0$ determines the average utility derived from living in block $i$; the scale parameter $E_j$ determines the average utility derived from working in block $j$; and the shape parameter $\epsilon > 1$ controls the dispersion of idiosyncratic utility.

After observing her realizations for idiosyncratic utility for each pair of residence and employment blocks, each worker chooses where to live and work to maximize her utility, taking as given residential amenities, goods prices, factor prices, and the location decisions of other workers and firms. Therefore workers sort across pairs of residence and employment blocks depending on their idiosyncratic preferences and the characteristics of these locations. The indirect utility from residing in block $i$ and working in block $j$ can be expressed in terms of the wage paid at this workplace ($w_j$), commuting costs ($d_{ij}$), the residential floor price ($Q_i$), the common component of amenities ($B_i$) and the idiosyncratic shock ($z_{ijo}$):[18]

$$u_{ijo} = \frac{z_{ijo} B_i w_j Q_i^{\beta-1}}{d_{ij}}, \tag{3}$$

---

[16] To simplify the exposition, throughout the paper, we index a worker's block of residence by $i$ or $r$ and her block of employment by $j$ or $s$ unless otherwise indicated.

[17] For empirical evidence using US data in support of the constant housing expenditure share implied by the Cobb-Douglas functional form, see Davis and Ortalo-Magné (2011). The role played by residential amenities in influencing utility is emphasized in the literature following Roback (1982). See Albouy (2008) for a recent prominent contribution.

[18] We make the standard assumption in the urban literature that income from land is accrued by absentee landlords and not spent within the city, although it is also possible to consider the case where it is redistributed lump sum to workers.

where we have used utility maximization and the choice of the final good as numeraire.

Although we model commuting costs in terms of utility, there is an isomorphic formulation in terms of a reduction in effective units of labor, because the iceberg commuting cost $d_{ij} = e^{\kappa \tau_{ij}}$ enters the indirect utility function (3) multiplicatively. As a result, commuting costs are proportional to wages, and this specification captures changes over time in the opportunity cost of travel time. Similarly, although we model the heterogeneity in commuting decisions in terms of an idiosyncratic shock to preferences, there is an isomorphic interpretation in terms of a shock to effective units of labor, because this shock $z_{ijo}$ enters indirect utility (3) multiplicatively with the wage.

Since indirect utility is a monotonic function of the idiosyncratic shock ($z_{ijo}$), which has a Fréchet distribution, it follows that indirect utility for workers living in block $i$ and working in block $j$ also has a Fréchet distribution. Each worker chooses the bilateral commute that offers her the maximum utility, where the maximum of Fréchet distributed random variables is itself Fréchet distributed. Using these distributions of utility, the probability that a worker chooses to live in block $i$ and work in block $j$ is:

$$\pi_{ij} = \frac{T_i E_j \left( d_{ij} Q_i^{1-\beta} \right)^{-\epsilon} (B_i w_j)^{\epsilon}}{\sum_{r=1}^{S} \sum_{s=1}^{S} T_r E_s \left( d_{rs} Q_r^{1-\beta} \right)^{-\epsilon} (B_r w_s)^{\epsilon}} \equiv \frac{\Phi_{ij}}{\Phi}. \tag{4}$$

Summing these probabilities across workplaces for a given residence, we obtain the overall probability that a worker resides in block $i$ ($\pi_{Ri}$), while summing these probabilities across residences for a given workplace, we obtain the overall probability that a worker works in block $j$ ($\pi_{Mi}$):

$$\pi_{Ri} = \sum_{j=1}^{S} \pi_{ij} = \frac{\sum_{j=1}^{S} \Phi_{ij}}{\Phi}, \qquad \pi_{Mj} = \sum_{i=1}^{S} \pi_{ij} = \frac{\sum_{i=1}^{S} \Phi_{ij}}{\Phi}. \tag{5}$$

These residential and workplace choice probabilities have an intuitive interpretation. The idiosyncratic shock to preferences $z_{ijo}$ implies that individual workers choose different bilateral commutes when faced with the same prices $\{Q_i, w_j\}$, commuting costs $\{d_{ij}\}$ and location characteristics $\{B_i, T_i, E_j\}$. Other things equal, workers are more likely to live in block $i$, the more attractive its amenities $B_i$, the higher its average idiosyncratic utility as determined by $T_i$, the lower its residential floor prices $Q_i$, and the lower its commuting costs $d_{ij}$ to employment locations. Other things equal, workers are more likely to work in block $j$, the higher its wage $w_j$, the higher its average idiosyncratic utility as determined by $E_j$, and the lower its commuting costs $d_{ij}$ from residential locations.

Conditional on living in block $i$, the probability that a worker commutes to block $j$ is:

$$\pi_{ij|i} = \frac{E_j \left( w_j / d_{ij} \right)^{\epsilon}}{\sum_{s=1}^{S} E_s \left( w_s / d_{is} \right)^{\epsilon}}, \tag{6}$$

where the terms in $\{Q_i, T_i, B_i\}$ have cancelled from the numerator and denominator. Therefore the probability of commuting to block $j$ conditional on living in block $i$ depends on the wage ($w_j$), average utility draw ($E_j$) and commuting costs ($d_{ij}$) of employment location $j$ in the numerator ("bilateral resistance") as well as the wage ($w_s$), average utility draw ($E_s$) and commuting costs ($d_{is}$) for all other possible employment locations $s$ in the denominator ("multilateral resistance").

Using these conditional commuting probabilities, we obtain the following commuting market clearing condition that equates the measure of workers employed in block $j$ ($H_{Mj}$) with the measure of workers choosing to commute to block $j$:

$$H_{Mj} = \sum_{i=1}^{S} \frac{E_j \left(w_j/d_{ij}\right)^{\epsilon}}{\sum_{s=1}^{S} E_s \left(w_s/d_{is}\right)^{\epsilon}} H_{Ri}, \tag{7}$$

where $H_{Ri}$ is the measure of residents in block $i$. Since there is a continuous measure of workers residing in each location, there is no uncertainty in the supply of workers to each employment location. Our formulation of workers' commuting decisions implies that the supply of commuters to each employment location $j$ in (7) is a continuously increasing function of its wage relative to other locations.[19]

Expected worker income conditional on living in block $i$ is equal to the wages in all possible employment locations weighted by the probabilities of commuting to those locations conditional on living in $i$:

$$\mathbb{E}\left[w_j | i\right] = \sum_{j=1}^{S} \frac{E_j \left(w_j/d_{ij}\right)^{\epsilon}}{\sum_{s=1}^{S} E_s \left(w_s/d_{is}\right)^{\epsilon}} w_j, \tag{8}$$

Therefore expected worker income is high in blocks that have low commuting costs (low $d_{is}$) to high-wage employment locations.[20]

Finally, population mobility implies that the expected utility from moving to the city is equal to the reservation level of utility in the wider economy ($\bar{U}$):

$$\mathbb{E}\left[u\right] = \gamma \left[\sum_{r=1}^{S} \sum_{s=1}^{S} T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\epsilon} \left(B_r w_s\right)^{\epsilon}\right]^{1/\epsilon} = \bar{U}, \tag{9}$$

where $\mathbb{E}$ is the expectations operator and the expectation is taken over the distribution for the idiosyncratic component of utility; $\gamma = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right)$ and $\Gamma(\cdot)$ is the Gamma function.

## 3.2 Production

Production of the tradeable final good occurs under conditions of perfect competition and constant returns to scale.[21] For simplicity, we assume that the production technology takes the Cobb-Douglas form, so that

[19]This feature of the model is not only consistent with the gravity equation literature on commuting flows discussed above but also greatly simplifies the quantitative analysis of the model. In the absence of heterogeneity in worker productivity, small changes in wages can induce all workers residing in one location to start or stop commuting to another location, which is both empirically implausible and complicates the determination of general equilibrium with asymmetric locations.

[20]For simplicity, we model agents and workers as synonymous and assume that labor is the only source of income. More generally, it is straightforward to extend the analysis to introduce families, where each worker has a fixed number of dependents that consume but do not work. Similarly, we can allow agents to have a constant amount of non-labor income.

[21]Even during division, there was substantial trade between West Berlin and West Germany. In 1963, the ratio of exports to GDP in West Berlin was around 70 percent, with West Germany the largest trade partner. Overall, industrial production accounted for around 50 percent of West Berlin's GDP in this year (American Embassy 1965).

output of the final good in block $j$ ($y_j$) is:

$$y_j = A_j H_{Mj}^{\alpha} L_{Mj}^{1-\alpha}, \tag{10}$$

where $A_j$ is final goods productivity and $L_{Mj}$ is the measure of floor space used commercially.

Firms choose their block of production and their inputs of workers and commercial floor space to maximize profits, taking as given final goods productivity $A_j$, the distribution of idiosyncratic utility, goods and factor prices, and the location decisions of other firms and workers. Profit maximization implies that equilibrium employment in block $j$ is increasing in productivity ($A_j$), decreasing in the wage ($w_j$), and increasing in commercial floor space ($L_{Mj}$):

$$H_{Mj} = \left(\frac{\alpha A_j}{w_j}\right)^{\frac{1}{1-\alpha}} L_{Mj}, \tag{11}$$

where the equilibrium wage is determined by the requirement that the demand for workers in each employment location (11) equals the supply of workers choosing to commute to that location (7).

From the first-order conditions for profit maximization and zero profits, equilibrium commercial floor prices ($q_j$) in each block with positive employment must satisfy:

$$q_j = (1-\alpha)\left(\frac{\alpha}{w_j}\right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}}. \tag{12}$$

Intuitively, firms in blocks with higher productivity ($A_j$) and/or lower wages ($w_j$) are able to pay higher commercial floor prices and still make zero profits.

### 3.3 Land Market Clearing

Land market equilibrium requires no-arbitrage between the commercial and residential use of floor space after the tax equivalent of land use regulations. The share of floor space used commercially ($\theta_i$) is:

$$\begin{array}{llll} \theta_i & = 1 & \text{if} & q_i > \xi_i Q_i, \\ \theta_i & \in [0,1] & \text{if} & q_i = \xi_i Q_i, \\ \theta_i & = 0 & \text{if} & q_i < \xi_i Q_i, \end{array} \tag{13}$$

where $\xi_i \geq 1$ captures one plus the tax equivalent of land use regulations that restrict commercial land use relative to residential land use. We allow this wedge between commercial and residential floor prices to vary across blocks. We assume that the observed price of floor space in the data is the maximum of the commercial and residential price of floor space: $\mathbb{Q}_i = \max\{q_i, Q_i\}$. Hence the relationship between observed, commercial and residential floor prices can be summarized as:

$$\begin{array}{lll} \mathbb{Q}_i = q_i, & q_i > \xi_i Q_i, & \theta_i = 1, \\ \mathbb{Q}_i = q_i, & q_i = \xi_i Q_i, & \theta_i \in [0,1], \\ \mathbb{Q}_i = Q_i, & q_i < \xi_i Q_i, & \theta_i = 0. \end{array} \tag{14}$$

We follow the standard approach in the urban literature of assuming that floor space $L$ is supplied by a competitive construction sector that uses land $K$ and capital $M$ as inputs. Following Combes, Duranton, and Gobillon (2014) and Epple, Gordon, and Sieg (2010), we assume that the production function takes the Cobb-Douglas form: $L_i = M_i^{\mu} K_i^{1-\mu}$.[22] Therefore the corresponding dual cost function for floor space is $\mathbb{Q}_i = \mu^{-\mu}(1-\mu)^{-(1-\mu)}\mathbb{P}^{\mu}\mathbb{R}_i^{1-\mu}$, where $\mathbb{Q}_i = \max\{q_i, Q_i\}$ is the price for floor space, $\mathbb{P}$ is the common price for capital across all blocks, and $\mathbb{R}_i$ is the price for land. Since the price for capital is the same across all locations, the relationships between the quantities and prices of floor space and land can be summarized as:

$$L_i = \varphi_i K_i^{1-\mu} \tag{15}$$

$$\mathbb{Q}_i = \chi \mathbb{R}_i^{1-\mu}, \tag{16}$$

where we refer to $\varphi_i = M_i^{\mu}$ as the density of development (since it determines the relationship between floor space and land area) and $\chi$ is a constant.

Residential land market clearing implies that the demand for residential floor space equals the supply of floor space allocated to residential use in each location: $(1-\theta_i)\, L_i$. Using utility maximization for each worker and taking expectations over the distribution for idiosyncratic utility, this residential land market clearing condition can be expressed as:

$$\mathbb{E}\left[\ell_i\right] H_{Ri} = (1-\beta)\frac{\mathbb{E}\left[w_s|i\right] H_{Ri}}{Q_i} = (1-\theta_i)\, L_i. \tag{17}$$

Commercial land market clearing requires that the demand for commercial floor space equals the supply of floor space allocated to commercial use in each location: $\theta_j L_j$. Using the first-order conditions for profit maximization, this commercial land market clearing condition can be written as:

$$\left(\frac{(1-\alpha)\, A_j}{q_j}\right)^{\frac{1}{\alpha}} H_{Mj} = \theta_j L_j. \tag{18}$$

When both residential and commercial land market clearing ((17) and (18) respectively) are satisfied, total demand for floor space equals the total supply of floor space:

$$(1-\theta_i)\, L_i + \theta_i L_i = L_i = \varphi_i K_i^{1-\mu}. \tag{19}$$

## 3.4 General Equilibrium with Exogenous Location Characteristics

We begin by characterizing the properties of a benchmark version of the model in which location characteristics are exogenous, before relaxing this assumption to introduce endogenous agglomeration forces. Given the model's parameters $\{\alpha, \beta, \mu, \epsilon, \kappa\}$, the reservation level of utility in the wider economy $\bar{U}$ and vectors of exogenous location characteristics $\{\boldsymbol{T}, \boldsymbol{E}, \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\varphi}, \boldsymbol{K}, \xi, \boldsymbol{\tau}\}$, the general equilibrium of the model is referenced by the six vectors $\{\boldsymbol{\pi_M}, \boldsymbol{\pi_R}, \boldsymbol{Q}, \boldsymbol{q}, \boldsymbol{w}, \boldsymbol{\theta}\}$ and total city population $H$.[23] These seven components of the equilibrium vector are determined by the following system of seven equations: population mobility (9), the residential choice probability ($\pi_{Ri}$ in (5)), the workplace choice probability ($\pi_{Mj}$ in (5)), commercial land market clearing (18), residential land market clearing (17), profit maximization and zero profits (12), and no-arbitrage between alternative uses of land (13).

[22] Empirically, we find that this Cobb-Douglas assumption is consistent with confidential micro data on property transactions for Berlin from 2000-2012, as discussed in the web appendix.

[23] Throughout the following we use bold math font to denote vectors or matrices.

**Proposition 1** *Assuming exogenous, finite and strictly positive location characteristics ($T_i \in (0,\infty)$, $E_i \in (0,\infty)$, $\varphi_i \in (0,\infty)$, $K_i \in (0,\infty)$, $\xi_i \in (0,\infty)$, $\tau_{ij} \in (0,\infty) \times (0,\infty)$), and exogenous, finite and non-negative final goods productivity $A_i \in [0,\infty)$ and residential amenities $B_i \in [0,\infty)$, there exists a unique general equilibrium vector {$\boldsymbol{\pi_M}$, $\boldsymbol{\pi_R}$, $H$, $\boldsymbol{Q}$, $\boldsymbol{q}$, $\boldsymbol{w}$, $\boldsymbol{\theta}$}.*

**Proof.** See Lemmas A1-A3 and the proofs of Propositions A1-A2 in Section A2 of the web appendix. ■

In this case of exogenous location characteristics, there are no agglomeration forces, and hence the model's congestion forces of commuting costs and an inelastic supply of land ensure the existence of a unique equilibrium. We establish a number of other properties of the general equilibrium with exogenous location characteristics in the web appendix. Assuming that all other location characteristics {$\boldsymbol{T}$, $\boldsymbol{E}$, $\boldsymbol{\varphi}$, $\boldsymbol{K}$, $\boldsymbol{\xi}$, $\boldsymbol{\tau}$} are exogenous, finite and strictly positive, a necessary and sufficient condition for zero residents is $B_i = 0$. Similarly, a necessary and sufficient condition for zero employment is $w_j = 0$, which in turn requires zero final goods productivity $A_j = 0$. Therefore the model rationalizes zero workplace employment with zero productivity ($A_i$) and zero residence employment with zero amenities ($B_i$).

## 3.5 Introducing Agglomeration Forces

Having established the properties of the model with exogenous location characteristics, we now introduce endogenous agglomeration forces. We allow final goods productivity to depend on production fundamentals ($a_j$) and production externalities ($\Upsilon_j$). Production fundamentals capture features of physical geography that make a location more or less productive independently of the surrounding density of economic activity (for example access to natural water). Production externalities impose structure on how the productivity of a given block is affected by the characteristics of other blocks. Specifically, we follow the standard approach in urban economics of modeling these externalities as depending on the travel-time weighted sum of workplace employment density in surrounding blocks:[24]

$$A_j = a_j \Upsilon_j^{\lambda}, \qquad \Upsilon_j \equiv \sum_{s=1}^{S} e^{-\delta \tau_{js}} \left( \frac{H_{Ms}}{K_s} \right), \tag{20}$$

where $H_{Ms}/K_s$ is workplace employment density per unit of land area; production externalities decline with travel time ($\tau_{js}$) through the iceberg factor $e^{-\delta \tau_{js}} \in (0,1]$; $\delta$ determines their rate of spatial decay; and $\lambda$ controls their relative importance in determining overall productivity.[25]

We model the externalities in workers' residential choices analogously to the externalities in firms' production choices. We allow residential amenities to depend on residential fundamentals ($b_i$) and residential externalities ($\Omega_i$). Residential fundamentals capture features of physical geography that make a location a

[24]While the canonical interpretation of these production externalities in the urban economics literature is knowledge spillovers, as in Alonso (1964), Fujita and Ogawa (1982), Lucas (2000), Mills (1967), Muth (1969), and Sveikauskas (1975), other interpretations are possible, as considered in Duranton and Puga (2004).

[25]We make the standard assumption that production externalities depend on employment density per unit of land area $K_i$ (rather than per unit of floor space $L_i$) to capture the role of higher ratios of floor space to land area in increasing the surrounding concentration of economic activity.

more or less attractive place to live independently of the surrounding density of economic activity (for example green areas). Residential externalities again impose structure on how the amenities in a given block are affected by the characteristics of other blocks. Specifically, we adopt a symmetric specification as for production externalities, and model residential externalities as depending on the travel-time weighted sum of residential employment density in surrounding blocks:

$$B_i = b_i \Omega_i^{\eta}, \qquad \Omega_i \equiv \sum_{r=1}^{S} e^{-\rho \tau_{ir}} \left( \frac{H_{Rr}}{K_r} \right), \tag{21}$$

where $H_{Rr}/K_r$ is residence employment density per unit of land area; residential externalities decline with travel time ($\tau_{ir}$) through the iceberg factor $e^{-\rho\tau_{ir}} \in (0,1]$; $\rho$ determines their rate of spatial decay; and $\eta$ controls their relative importance in overall residential amenities. The parameter $\eta$ captures the net effect of residence employment density on amenities, including negative spillovers such as air pollution and crime, and positive externalities through the availability of urban amenities. Although $\eta$ captures the direct effect of higher population density on utility through amenities, there are clearly other general equilibrium effects through floor prices, commuting times and wages.

The introduction of these agglomeration forces generates the potential for multiple equilibria in the model if these agglomeration forces are sufficiently strong relative to the exogenous differences in characteristics across locations. An important feature of our empirical approach is that it explicitly addresses the potential for multiple equilibria, as discussed further in the next subsection.

### 3.6 Recovering Location Characteristics

We now show that there is a unique mapping from the observed variables to unobserved location characteristics. These unobserved location characteristics include production and residential fundamentals and several other unobserved variables. Since a number of these unobserved variables enter the model isomorphically, we define the following composites denoted by a tilde:

$$\begin{aligned} \tilde{A}_i &= A_i E_i^{\alpha/\epsilon}, & \qquad \tilde{a}_i &= a_i E_i^{\alpha/\epsilon}, \\ \tilde{B}_i &= B_i T_i^{1/\epsilon} \zeta_{Ri}^{1-\beta}, & \qquad \tilde{b}_i &= b_i T_i^{1/\epsilon} \zeta_{Ri}^{1-\beta}, \\ \tilde{w}_i &= w_i E_i^{1/\epsilon}, & & \\ \tilde{\varphi}_i &= \tilde{\varphi}_i \left( \varphi_i, E_i^{1/\epsilon}, \xi_i \right), & & \end{aligned}$$

where we use $i$ to index all blocks; the function $\tilde{\varphi}_i(\cdot)$ is defined in the web appendix; $\zeta_{Ri} = 1$ for completely specialized residential blocks; and $\zeta_{Ri} = \xi_i$ for residential blocks with some commercial land use.

In the labor market, the adjusted wage for each employment location ($\tilde{w}_i$) captures the wage ($w_i$) and the Fréchet scale parameter for that location ($E_i^{1/\epsilon}$), because these both affect the relative attractiveness of an employment location to workers. On the production side, adjusted productivity for each employment location ($\tilde{A}_i$) captures productivity ($A_i$) and the Fréchet scale parameter for that location ($E_i^{\alpha/\epsilon}$), because these both affect the adjusted wage consistent with zero profits. Adjusted production fundamentals are defined analogously. On the consumption side, adjusted amenities for each residence location ($\tilde{B}_i$) capture amenities ($B_i$), the Fréchet scale parameter for that location ($T_i^{1/\epsilon}$), and the relationship between observed and residential floor prices ($\zeta_{Ri} \in \{1, \xi_i\}$), because these all affect the relative attractiveness of a location consistent with population mobility. Adjusted residential fundamentals are defined analogously. Finally, in the land market, the adjusted density of development ($\tilde{\varphi}_i$) includes the density of development ($\varphi_i$) and other production and residential parameters that affect land market clearing.

**Proposition 2** ***(i)*** *Given known values for the parameters* $\{\alpha, \beta, \mu, \epsilon, \kappa\}$ *and the observed data* $\{\mathbb{Q}, \boldsymbol{H_M}, \boldsymbol{H_R}, \boldsymbol{K}, \boldsymbol{\tau}\}$*, there exist unique vectors of the unobserved location characteristics* $\{\tilde{\boldsymbol{A}}^*, \tilde{\boldsymbol{B}}^*, \tilde{\boldsymbol{\varphi}}^*\}$ *that are consistent with the data being an equilibrium of the model.*
***(ii)*** *Given known values for the parameters* $\{\alpha, \beta, \mu, \epsilon, \kappa, \lambda, \delta, \eta, \rho\}$ *and the observed data* $\{\mathbb{Q}, \boldsymbol{H_M}, \boldsymbol{H_R}, \boldsymbol{K}, \boldsymbol{\tau}\}$*, there exist unique vectors of the unobserved location characteristics* $\{\tilde{\boldsymbol{a}}^*, \tilde{\boldsymbol{b}}^*, \tilde{\boldsymbol{\varphi}}^*\}$ *that are consistent with the data being an equilibrium of the model.*

**Proof.** See the proofs of Propositions A3-A4 in Section A3 of the web appendix. ■

To interpret this identification result, note that in models with multiple equilibria, the mapping from the parameters and fundamentals to the endogenous variables is non-unique. In such models, the inverse mapping from the endogenous variables and parameters to the fundamentals in principle can be either unique or non-unique. In the context of our model, Proposition 2 conditions on the parameters $\{\alpha, \beta, \mu, \epsilon, \kappa, \lambda, \delta, \eta, \rho\}$ and a combination of observed endogenous variables $\{\mathbb{Q}, \boldsymbol{H_M}, \boldsymbol{H_R}\}$ and fundamentals $\{\boldsymbol{K}, \boldsymbol{\tau}\}$, and uses the equilibrium conditions of the model to determine unique values of the unobserved adjusted fundamentals $\{\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{\varphi}}\}$. This identification result hinges on the data available. In the absence of any one of the five observed variables (floor prices, workplace employment, residence employment, land area and travel times), these unobserved adjusted fundamentals would be under-identified, and could not be determined without making further structural assumptions.

The economics underlying this identification result is as follows. Given observed workplace and residence employment, and our measures of travel times, worker commuting probabilities can be used to solve for unique adjusted wages consistent with commuting market clearing (7). Given adjusted wages and observed floor prices, the firm cost function can be used to solve for the unique adjusted productivity consistent with zero profits (12). Given adjusted wages, observed floor prices and residence employment shares, worker utility maximization and population mobility can be used to solve for the unique adjusted amenities consistent with residential choice probabilities (5). Hence the model has a recursive structure, in which overall adjusted productivity and amenities $\{\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}\}$ can be determined without making assumptions about the functional form or relative importance of externalities $\{\boldsymbol{\Upsilon}, \boldsymbol{\Omega}\}$ and adjusted fundamentals $\{\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}\}$. Having recovered overall adjusted productivity and amenities, we can use our spillovers specification to decompose these variables into their two components of externalities and adjusted fundamentals ((20) and (21)). Finally, given observed land area, the implied demands for commercial and residential floor space can be used to solve for the unique adjusted density of development consistent with market clearing for floor space (19). Therefore the observed data, parameters and equilibrium conditions of the model can be used to determine unique values of the unobserved adjusted fundamentals regardless of whether the model has a single equilibrium or multiple equilibria.

In our structural estimation of the model in Section 7, we use Proposition 2 as an input into our generalized method of moments (GMM) estimation, in which we determine both the parameters and the unobserved adjusted fundamentals.

## 3.7 Berlin's Division and Reunification

We focus in our empirical analysis on West Berlin, since it remained a market-based economy after division and we therefore expect the mechanisms in the model to apply.[26] We capture the division of Berlin in the model by assuming infinite costs of trading the final good, infinite commuting costs ($\kappa \rightarrow \infty$), infinite rates of decay of production externalities ($\delta \rightarrow \infty$), and infinite rates of decay of residential externalities ($\rho \rightarrow \infty$) across the Berlin Wall.

The model points to four key channels through which division affects the distribution of economic activity within West Berlin: a loss of employment opportunities in East Berlin, a loss of commuters from East Berlin, a loss of production externalities from East Berlin, and a loss of residential externalities from East Berlin. Each of these four effects reduces the expected utility from living in West Berlin, and hence reduces its overall population, as workers out migrate to West Germany. As both commuting and externalities decay with travel time, each of these effects is stronger for parts of West Berlin close to employment and residential concentrations in East Berlin, reducing floor prices, workplace employment and residence employment in these parts of West Berlin relative to those elsewhere in West Berlin. The mechanisms that restore equilibrium in the model are changes in wages and floor prices. Workplace and residence employment reallocate across locations within West Berlin and to West Germany, until wages and floor prices have adjusted such that firms make zero profits in all locations with positive production, workers are indifferent across all populated locations, and there are no-arbitrage opportunities in reallocating floor space between commercial and residential use.

Since reunification involves a re-integration of West Berlin with employment and residential concentrations in East Berlin, we would expect to observe the reverse pattern of results in response to reunification. But reunification need not necessarily have exactly the opposite effects from division. As discussed above, if agglomeration forces are sufficiently strong relative to the differences in fundamentals across locations, there can be multiple equilibria in the model. In this case, division could shift the distribution of economic activity in West Berlin between multiple equilibria, and reunification need not necessarily reverse the impact of division. More generally, the level and distribution of economic activity within East Berlin could have changed between the pre-war and division periods, so that reunification is a different shock from division. Notwithstanding these points, reintegration with employment and residential concentrations in East Berlin is predicted to raise relative floor prices, workplace employment and residence employment in the areas of West Berlin close to those concentrations.

[26] In contrast, the distribution of economic activity in East Berlin during division was heavily influenced by central planning, which is unlikely to mimic market forces.

# 4 Data Description

The quantitative analysis of our model requires four key sets of data: workplace employment, residence employment, the price of floor space and commuting times between locations. We have compiled these variables for Berlin for the pre-war and reunification periods and for West Berlin for the division period. For simplicity we generally refer to the three years for which we have data as 1936, 1986, and 2006 even though some of the data are from the closest available neighboring year. In addition to these main variables we have compiled data on a wide range of other block characteristics: commuting behavior, the dispersion of wages across districts, and also the price of floor space in 1928 and 1966. Below we briefly describe the data definitions and sources. A more detailed discussion is included in the web appendix.

Data for Berlin is available at a number of different levels of spatial disaggregation. The finest available disaggregation is statistical blocks ("Blöcke"). In 2006 the surface of Berlin was partitioned into 15,937 blocks, of which just under 9,000 are in the former West Berlin. We hold this block structure constant for all years in our data. These blocks have a mean area of about 50,000 square meters and an average 2005 population of 274 for the 12,192 blocks with positive 2005 population.[27] Blocks can be aggregated up to larger spatial units including statistical areas ("Gebiete") and districts ("Bezirke").[28]

Our measure of employment at the place of work for the reunification period is a count of the 2003 social security employment ("Sozialversicherungspflichtig Beschäftigte") in each block, which was provided by the Statistical Office of Berlin ("Senatsverwaltung für Berlin") in electronic form. We scale up social security employment in each block by the ratio of social security employment to total employment for Berlin as a whole. Data for the division period come from the 1987 West German census, which reports total workplace employment by block.[29] We construct comparable data for the pre-war period by combining data on district total private-sector workplace employment published in the 1933 census with the registered addresses of all firms on the Berlin company register ("Handelsregister") in 1931. As described in detail in the web appendix, we use the number of firms in each block to allocate the 1933 district totals for private-sector workplace employment across blocks within districts. Finally, we allocate 1933 public-sector workplace employment across blocks using detailed information on the location of public administration buildings (including ministries, utilities and schools) immediately prior to the Second World War.

To construct employment at the place of residence for the reunification period, we use data on the population of each block in 2005 from the Statistical Office of Berlin and scale the population data using district-level information on labor force participation.[30] Employment at residence for the division period is

[27]There are a number of typically larger blocks that only contain water areas, forests, parks and other uninhabited areas. Approximately 29 percent of the area of Berlin in 2006 is covered by forests and parks, while another 7 percent is accounted for by lakes, rivers and canals (Statistical Yearbook of Berlin 2007).

[28]As discussed in Section 2, we use the 1938 district boundaries upon which the occupation sectors were based unless otherwise indicated.

[29]For 2003, only social security employment and not total employment is available at the block level. The main difference between these two measures of employment is self-employment. Empirically, we find that the ratio of social security to total employment in 1987 is relatively constant across districts (the correlation coefficient between the two variables is over 0.98), which supports our approach for 2003 of scaling up social security employment to total employment.

[30]Empirically, labor force participation is relatively constant across districts within Berlin in all years of our dataset.

reported by block in the 1987 West German census. To construct pre-war data on employment at residence, we use a tabulation in the 1933 census that lists the population of each street or segment of street in Berlin. As described in more detail in the web appendix, we use a concordance between streets and blocks to allocate the population of streets to individual blocks. We then again use labor force participation rates at the district level to scale the population data to obtain employment at residence by block.

Berlin has a long history of providing detailed assessments of land values, which have been carried out by the independent Committee of Valuation Experts ("Gutachterausschuss für Grundstückswerte") in the post-war period. The committee currently has 50 members who are building surveyors, real estate practitioners and architects. Our land price data for 1986 and 2006 are the land values ("Bodenrichtwerte") per square meter of land published by the Committee on detailed maps of Berlin which we have digitized and merged with the block structure. The Committee's land values capture the fair market value of a square meter of land if it was undeveloped. While the Committee does not publish the details of its valuation procedure, the land values are based on recent market transactions. As a check on the Committee's land values, we compare them to confidential micro data on property transactions from 2000-2012. As shown in the web appendix, we find a high correlation between the land values reported by the Committee for 2006 and the land values that we compute from the property transactions data. Finally, the land value data also includes information on the typical density of development, measured as the ratio of floor space to ground area ("GFZ").[31]

Our source of land price data for the pre-war period is Kalweit (1937). Kalweit was a chartered building surveyor ("Gerichtlich Beeideter Bausachverständiger"), who received a government commission for the assessment of land values in Berlin ("Baustellenwerte") for 1936. These land values were intended to provide official and representative guides for private and public investors in Berlin's real estate market. As with the modern land value data, they capture the fair market price of a square meter of undeveloped land and are reported for each street or segment of street in Berlin. Using ArcGIS, we matched the streets or segments of streets in Kalweit (1937) to blocks, and aggregated the street-level land price data to the block-level.[32] To convert land prices ($\mathbb{R}_i$) to floor prices ($\mathbb{Q}_i$), we use the assumption of a competitive construction sector with a Cobb-Douglas technology, as discussed in subsection 3.3 above.

Travel times are measured in minutes based on the transport network available in each year and assumed average travel speeds for each mode of transport. To determine travel times between each of the 15,937 blocks in our data, i.e. nearly 254 million (15,937$\times$15,937) bilateral connections, we distinguish between travel times by public transport and car. As described in more detail in the web appendix, we construct minimum travel times by public transport for the three years using information on the underground railway ("U-Bahn"), suburban railway ("S-Bahn"), tram ("Strassenbahn") and bus ("Bus") network of Berlin in each

[31]Note that the Committee's land values are completely different from the unit values ("Einheitswerte") used to calculate property taxes. The current unit values are still based on an assessment ("Hauptfeststellung") that took place as early as 1964 for the former West Germany and 1935 for the former East Germany. In contrast, the Committee's land values are based on contemporaneous market transactions and are regularly updated.

[32]In robustness checks, we also use land value data for 1928 from Kalweit (1929) (which has the same structure as Kalweit 1937), and for 1966 from the Committee of Valuation Experts (which has the same structure as the 1986 and 2006 data).

year. We use ArcGIS to compute the fastest connection between each pair of blocks allowing passengers to combine all modes of public transport and walking to minimize travel time. We also construct minimum driving times by car in 1986 and 2006 using an ArcGIS shape file of the street network of Berlin. For 1986 and 2006, we measure overall travel times by weighting public transport and car minimum travel times using district-level data on the proportion of journeys undertaken with these two modes of transport. For 1936, commuting to work by car was rare, and hence we use public transport minimum travel times.[33]

In addition to our main variables, we have compiled a number of other data, which are described in detail in the web appendix. First, we have data on observable block characteristics including, the location of parks and other green spaces, proximity to lakes, rivers and canals, proximity to schools, land use, average noise level, the number of listed buildings, the extent of destruction during the Second World War, and urban regeneration programs and government buildings post reunification. Second, we have obtained survey data on commuting flows in Berlin in 1936, 1982 and 2008. Third, we have obtained data on average wages by workplace for each district of West Berlin in 1986.

# 5 Reduced-Form Results

In this section, we provide reduced-form evidence in support of the model's qualitative predictions that complements our later structural estimation of the model. First, we use this reduced-form analysis to establish reorientations of land prices, workplace employment and residence employment within West Berlin following division and reunification without imposing the full structure of the model. Second, this reduced-form analysis enables us to demonstrate the robustness of these reorientations to the inclusion of a wide range of controls and provide evidence against alternative possible explanations.

## 5.1 Evolution of the Land Price Gradient over Time

In Figure 2, we display the spatial distribution of land prices across blocks for each year as a three-dimensional map. The main public parks and forests are shown in green and the main bodies of water are shown in blue. White areas correspond to other undeveloped areas including railways. Since we use the same vertical scale for each figure, and land prices are normalized to have a mean of one in each year, the levels of the land price surfaces in each figure are comparable.

As apparent from Panel A of Figure 2, Berlin's land price gradient in 1936 was in fact approximately monocentric, with the highest values concentrated in the district Mitte. We measure the center of the pre-war Central Business District (CBD) as the intersection of Friedrich Strasse and Leipziger Strasse, close to the U-Bahn station "Stadtmitte." Around this central point, there are concentric rings of progressively lower land prices surrounding the pre-war CBD. Towards the Western edge of these concentric rings is the Kudamm ("Kurfürstendamm") in Charlottenburg and Wilmersdorf, which had developed into a fashionable shopping area in the decades leading up to the Second World War. This area lies to the West of the Tiergarten

[33] Leyden (1933) reports data on travel by mode of transport in pre-war Berlin, in which travel by car accounts for less than 10 percent of all journeys.

Park, which explains the gap in land prices between the Kudamm and Mitte. Panel A also shows the future line of the Berlin Wall (shown in gray font), including the inner boundary between East and West Berlin and the outer boundary that separated West Berlin from its East German hinterland.

To show relative land values in locations that subsequently became part of West Berlin, Panel B displays the 1936 distribution of land prices for only these locations. The two areas of West Berlin with the highest pre-war land prices were parts of a concentric ring around the pre-war CBD: the area around the Kudamm discussed above and a second area just West of Potsdamer Platz and the future line of the Berlin Wall. This second area was a concentration of commercial and retail activity surrounding the "Anhalter Bahnhof" mainline and suburban rail station. Neither of these areas contained substantial government administration, which was instead concentrated in Mitte in the future East Berlin, particularly around Wilhelmstrasse.

In Panel C, we examine the impact of division by displaying the 1986 distribution of land prices for West Berlin. Comparing Panels B and C, three main features stand out. First, land prices exhibit less dispersion and smaller peak values in West Berlin during division than in Berlin during the pre-war period. Second, one of the pre-war land price peaks in West Berlin – the area just West of Potsdamer Platz – is entirely eliminated following division, as this area ceased to be an important center of commercial and retail activity. Third, West Berlin's CBD during division coincided with the other area of high pre-war land values in West Berlin around the Kudamm, which was relatively centrally located within West Berlin.

To examine the impact of reunification, Panel D displays the 2006 distribution of land prices across blocks within Berlin as a whole, while Panel E shows the same distribution but only for blocks in the former West Berlin. Comparing these two figures with the previous two figures, three main features are again apparent. First, land prices are more dispersed and have higher peak values following reunification than during division. Second, the area just West of Potsdamer Platz is re-emerging as a concentration of office and retail development with high land values. Third, Mitte is also re-emerging as a center of high land values. As in the pre-war period, the main government ministries are either concentrated in Mitte in the former East Berlin or around the Federal parliament ("Reichstag").

Figures A1 and A2 in the web appendix display the log difference in land prices from 1936-1986 and 1986-2006 for each block. As evident from these figures, the largest declines in land prices following division and the largest increases in land prices following reunification are along those segments of the Berlin Wall around the pre-war CBD. In contrast, there is little evidence of comparable declines in land prices along other sections of the Berlin Wall. Therefore these results provide some first evidence that it is not proximity to the Berlin Wall *per se* that matters but the loss of access to the pre-war CBD.[34]

[34]Regressing the growth in West Berlin floor prices from 1986-2006 on their growth from 1936-1986, we find an estimated coefficient (Conley 1999 standard error) of -0.262 (0.017) and an R-squared of 0.29, suggesting that the areas that experienced the largest decline in floor prices after division also experienced the largest growth in floor prices after reunification.

## 5.2 Difference-in-Difference Estimates

To establish the statistical significance of these findings and their robustness to the inclusion of controls, we estimate the following "difference-in-difference" specification for division and reunification separately:

$$\triangle \ln O_i = \alpha + \sum_{k=1}^{K} \mathbb{I}_{ik}\beta_k + \ln M_i \gamma + u_i, \tag{22}$$

where $i$ denotes blocks; $\triangle O_i$ is the change in an economic outcome of interest (floor prices, workplace employment, residence employment); $\alpha$ is a constant; $\mathbb{I}_{ik}$ is an indicator variable for whether block $i$ lies within a distance grid cell $k$ from the pre-war CBD; $\beta_k$ are coefficients to be estimated; $M_i$ are time-invariant observable block characteristics (such as proximity to parks and lakes) and $\gamma$ captures changes over time in the premium to these time-invariant observable block characteristics; and $u_i$ is a stochastic error. This specification allows for time-invariant factors that have constant effects over time, which are differenced out before and after division (or reunification). It also allows for a common time effect of division or reunification across all blocks, which is captured in the constant $\alpha$.

We begin by considering distance grid cells of 500 meter intervals. Since the minimum distance to the pre-war CBD in West Berlin is around 0.75 kilometers, our first distance grid cell is for blocks with distances less than 1.25 kilometers. We include grid cells for blocks with distances up to 3.25-3.75 kilometers, so that the excluded category is blocks more than 3.75 kilometers from the pre-war CBD.[35] This grid cells specification allows for a flexible functional form for the relationship between changes in block economic outcomes and distance from the pre-war CBD. In these reduced-form regressions, we take the location of the pre-war CBD as given, whereas in the structural model its location is endogenously determined. In subsection 5.3, we show that we find similar results using other non-parametric approaches that do not require us to specify grid cells, such as locally-weighted linear least squares.

We show that our results are robust to two alternative approaches to controlling for spatial correlation in the error term $u_i$. As our baseline specification throughout the paper, we report Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors following Conley (1999), which allow for spatial correlation in the errors across neighboring blocks with distances less than a specified threshold.[36] As a robustness check, the web appendix reports standard errors clustered on statistical areas ("Gebiete"), which allows for a general correlation structure in the errors across blocks within areas, but assumes that the errors are independent across areas (see for example Bertrand, Duflo, and Mullainathan 2004).

Table 1 reports the results of estimating our baseline specification (22) for division.[37] The dependent variable in Columns (1)-(5) is the log difference in the price of floor space from 1936-86. In Column (1) we include only the distance grid cells, and find a negative and statistically significant effect of proximity to the pre-war CBD, which declines monotonically with distance from the pre-war CBD. On average, West

[35] The number of West Berlin blocks with floor price data in all three years in each grid cell (from nearest to furthest from the pre-war CBD) are: 32, 48, 60, 111, 171 and 195. The maximum distance to the pre-war CBD in West Berlin is around 23 kilometers.

[36] We use a threshold of 0.5 kilometers, where the median block in Berlin has 19 other blocks within 0.5 kilometers.

[37] Table A1 in the web appendix reports the robustness test using standard errors clustered on statistical areas ("Gebiete") instead of HAC standard errors following Conley (1999).

Berlin blocks within the first grid cell experience around a 55 percent reduction in the price of floor space between 1936 and 1986 (since $1 - e^{-0.800} = 0.55$) relative to those more than 3.75 kilometers away from the pre-war CBD. Together the six grid cells alone explain around one quarter of the variation in the change in the price of floor space following division ($R^2 = 0.26$), suggesting a powerful effect of proximity to the pre-war concentration of economic activity in East Berlin.

In Column (2), we show that these results are robust to including district fixed effects, which focuses solely on within-district variation in proximity to the pre-war CBD. Column (3) examines whether it is really proximity to the pre-war CBD that matters by including analogous 500 meter grid cells for distance to the closest point on (a) the inner boundary between East and West Berlin and (b) the outer boundary between West Berlin and its East German hinterland (see Table A2 of the web appendix for the coefficients on these distance grid cells).[38] Again we find a negative and statistically significant effect of proximity to the pre-war CBD that remains of around the same magnitude. In contrast, the coefficients for the inner boundary grid cells are close to zero and typically statistically insignificant, while the coefficients for the outer boundary grid cells are positive and statistically significant (although substantially smaller in magnitude than those for the pre-war CBD).[39]

Our finding that there is little evidence of a negative treatment effect of division along segments of the Berlin Wall far from the pre-war CBD suggests that our results are indeed capturing a loss of access to the pre-war CBD rather than other considerations associated with being close to the Berlin Wall such as its disamenity value. But by themselves these reduced-form regressions do not distinguish between different explanations for why access to the pre-war CBD matters, such as loss of access to employment opportunities, production externalities and/or residential externalities. In our structural estimation of the model below, we use the structure of the model to separate out these different explanations.

Column (4) shows that we find a similar pattern of results if we also include analogous 500 meter grid cells for distance to the Kudamm, providing further evidence that our results are indeed capturing a loss of access to the pre-war CBD (see Table A2 of the web appendix for the coefficients on these distance grid cells).[40] In Column (5), we further augment the specification from Column (4) with a wide range of

[38]The number of West Berlin blocks with floor price data in all three years in each grid cell for the inner boundary (from nearest to furthest from the inner boundary) are: 355, 406 431, 379, 313 and 326. The corresponding numbers for the outer boundary are 574, 646, 605, 594, 488 and 335. Comparing these numbers with those for the pre-war CBD grid cells in footnote 35, it is clear that the intersection of observations in each pair of grid cells is small relative to the union of observations in that pair of grid cells, which enables us to separately identify the coefficients for each grid cell.

[39]The small positive effects for the outer boundary could reflect a number of considerations. First, the areas beyond the outer boundary of Berlin are relatively undeveloped, implying little loss of access to surrounding economic activity following division. In 1933, total workplace and residence employment in Berlin were 1,628,622 and 1,591,723, respectively, implying small net inward commuting of 36,899. Second, there is a general equilibrium shift in economic activity within West Berlin following division. As a result, locations along the outer boundary of West Berlin become closer to the center of economic activity. Third, peak floor prices are lower relative to mean floor prices following division (compare Panels B and C of Figure 2). Since mean floor prices are constant by construction, this raises floor prices in peripheral locations relative to central locations. Fourth, there is the usual pattern of new residential developments appearing around the fringes of an existing city. Consistent with this, we below find positive effects along the outer boundary for residence employment but not workplace employment.

[40]Division has several opposing effects on floor prices for locations close to the Kudamm. First, they lose access to the pre-war CBD to which they were relatively close. Second, the Kudamm becomes the new center of economic activity in West Berlin. Third, peak floor prices are lower relative to mean floor prices following division (compare Panels B and C of Figure 2). Since

controls for block characteristics. Although some of these controls are potentially endogenous to division, we demonstrate that our results are not driven by their omission by reporting results both with and without the controls. Our block characteristics include log distance to the nearest school in 2006, log distance to the nearest lake, river or canal in 2006, log distance to the nearest park in 2006, log block area, Second World War destruction, and indicator variables for land use in 2006, whether a block qualified for urban regeneration policies post-reunification and government buildings post-reunification.[41]

In the next two Columns, we report results for employment residence. While Column (6) includes only our distance grid cells for proximity to the pre-war CBD and district fixed effects, Column (7) includes all the controls from Column (5). In both cases, we find that West Berlin blocks close to the pre-war CBD experienced a decline in employment residence relative to other parts of West Berlin following division. Columns (8) and (9) demonstrate a similar pattern of results for employment workplace.[42]

In Table 2, we report analogous specifications for reunification.[43] Consistent with the predictions of the model, we observe the reverse pattern of results for reunification. In Column (1), we include only the distance grid cells. We find that West Berlin blocks within the first distance grid cell experience around a 49 percent increase in the price of floor space between 1986 and 2006 ($e^{0.398} - 1 = 0.49$) relative to those more than 3.75 kilometers away from the pre-war CBD. Together the six distance grid cells now explain around 8 percent of the observed variation of the change in the price of floor space ($R^2 = 0.08$). Columns (2)-(5) show that these results are robust to including the same set of controls as for division (see Table A4 of the web appendix for the coefficients on the other distance grid cells). In Column (6)-(9), we report results for employment residence and employment workplace. Again we find statistically significant treatment effects, although these effects are less precisely estimated than for division.

## 5.3 Further Evidence

In Figure 3, we provide additional evidence on the timing of the estimated treatment effects, and demonstrate the absence of pre-trends. Panels A and B confirm our distance grid cell results by displaying the log difference in floor prices for each West Berlin block against distance from the pre-war CBD for 1936-86 and 1986-2006 respectively. We also show the locally-weighted linear least squares regression relationships between the two variables. From comparing Panels A and B, the effects of division are substantially larger and extend much further into West Berlin than the effects of reunification.

Panels C and D examine the timing of the division treatment by reporting results for 1936-66 and 1966-86 respectively. Consistent with the rapid disintegration of economic activity between West and East Berlin

mean floor prices are constant by construction, this raises floor prices in peripheral locations relative to central locations. The net effect is small negative coefficients for the Kudamm distance grid cells.

[41] Our three indicator variables for urban regeneration programs are for whether a block qualified for "Sanierungsgebiet" in 2002, "Sanierungsgebiet" in 2006, and the "Stadtumbau West" program that was initiated in 2005.

[42] Although our identification strategy exploits relative changes across different parts of West Berlin, we also find that West Berlin's overall population declines from 2,683,099 to 2,012,709 between the 1933 and 1987 censuses respectively, consistent with the predictions of the model discussed in subsection 3.7 above.

[43] Table A3 in the web appendix reports the robustness test using standard errors clustered on statistical areas ("Gebiete") instead of HAC standard errors following Conley (1999).

in the immediate aftermath of the Second World War, Panel C shows that most of the treatment effect of division on the price of floor space had already occurred by 1966. Therefore, as shown in Panel D, we find only a small negative treatment effect of division for 1966-86.[44] These results in Panel D also demonstrate that the positive treatment effect of reunification in Panel B is not driven by pre-trends in floor prices in the parts of West Berlin close to the pre-war CBD prior to reunification. Finally, Panel E displays results for 1928-36 and shows that the negative treatment effect of division is not driven by pre-trends in the parts of West Berlin close to the pre-war CBD prior to the Second World War.

In the web appendix, we provide further evidence that the estimated treatment effects of division and reunification are capturing a loss of access to the surrounding concentration of economic activity using a different source of variation in the data based on proximity to U/S-Bahn stations. Taken together, the results of this subsection provide further evidence in support of the model's qualitative predictions of a reallocation of economic activity within West Berlin in response to both division and reunification.

# 6 Gravity, Productivity and Amenities

In this section, we take a first step towards examining the extent to which the model can account quantitatively for the observed variation in the data. In particular, we use the recursive structure of the model discussed in subsection 3.6 to recover overall productivity, amenities and the density of development just using the model's gravity equation predictions for commuting flows. This approach has three advantages. First, we can determine the commuting parameters $\{\epsilon, \kappa\}$ using only information on commuting probabilities and wages and without taking a stand on the values of the agglomeration parameters $\{\lambda, \delta, \eta, \rho\}$. Hence we can solve for overall adjusted productivity, amenities and the density of development $\{\tilde{A}_i, \tilde{B}_i, \tilde{\varphi}_i\}$ regardless of the relative importance or functional form of externalities $\{\Upsilon_{it}, \Omega_{it}\}$ and fundamentals $\{\tilde{a}_{it}, \tilde{b}_{it}\}$. Second, we can determine the commuting parameters $\{\epsilon, \kappa\}$ without imposing the full set of identifying assumptions used in the structural estimation in section 7 below. Third, we use the solutions for overall adjusted productivity, amenities and the density of development $\{\tilde{A}_i, \tilde{B}_i, \tilde{\varphi}_i\}$ to show that the model with exogenous location characteristics is unable to explain the observed impact of division and reunification.

## 6.1 Gravity

From the commuting probabilities (4), one of the model's key predictions is a semi-log gravity equation for commuting flows from residence $i$ to workplace $j$:

$$\ln \pi_{ij} = -\nu \tau_{ij} + \vartheta_i + \varsigma_j, \tag{23}$$

where the residence fixed effects ($\vartheta_i$) capture residence characteristics $\{B_i, T_i, Q_i\}$; the workplace fixed effects ($\varsigma_j$) capture workplace characteristics $\{w_j, E_j\}$; the denominator in (4) is a constant that is absorbed

[44] Tables A.5-A.8 in the web appendix corroborate these findings by re-estimating the "difference-in-difference" regression specification (22) for these sub-periods. While the division treatment is large and statistically significant for 1936-66 in Tables A.5-A.6, it is small and typically statistically insignificant for 1966-86 in Tables A.7-A.8.

into the fixed effects; commuting costs are $d_{ij} = e^{\kappa\tau_{ij}}$; and travel times $\tau_{ij}$ are measured in minutes. The parameter $\nu = \epsilon\kappa$ is the semi-elasticity of commuting flows with respect to travel times and is a combination of the commuting cost parameter $\kappa$ and the commuting heterogeneity parameter $\epsilon$.

To provide empirical evidence on these gravity equation predictions, we use micro data on a representative survey of individual commuters in Berlin for 2008, which report district of residence, district of workplace and individual bilateral travel times in minutes for 7,948 commuters. We use these micro survey data to compute the probability that a worker commutes between any of the 12 districts of Berlin in 2008, which yields $12 \times 12 = 144$ pairs of bilateral commuting probabilities.[45] We observe positive commuting probabilities for all bilateral district pairs, although some district pairs have a small number of commuters in these micro survey data. While the model uses measures of bilateral travel times that we construct based on the transport network, the micro survey data includes self-reported travel times for each commuter. Therefore we augment the gravity equation derived from the model (23) with a stochastic error that captures measurement error in travel times:

$$\ln \pi_{ij} = -\nu\tau_{ij} + \vartheta_i + \varsigma_j + e_{ij}, \tag{24}$$

where we assume that this measurement error is uncorrelated with self-reported travel times.

The gravity equation (24) yields predictions for commuting probabilities between pairs of blocks, whereas the commuting survey data reports commuting probabilities between pairs of districts. Taking means across pairs of blocks within pairs of districts in (24), we estimate the following district-level gravity equation:

$$\ln \pi_{IJ} = -\nu\tau_{IJ} + \vartheta_I + \varsigma_J + e_{IJ}, \tag{25}$$

where $I$ denotes district of residence; $J$ denotes district of employment; $\tau_{IJ}$ is the average of the travel times $\tau_{ij}$; and we approximate the unobserved mean of the log block commuting probabilities (the mean of $\ln \pi_{ij}$) with the observed log district commuting probabilities ($\ln \pi_{IJ}$).[46]

In Column (1) of Table 3, we estimate (25) using a linear fixed effects estimator, and find a semi-elasticity of commuting with respect to travel time of -0.0697 that is statistically significant at the one percent level. This estimate implies that each additional minute of travel time reduces the flow of commuters by around 7 percent. From the regression R-squared, this gravity equation specification explains around 83 percent of the variation in bilateral commuting patterns. To address concerns about sampling error for bilateral pairs with small numbers of commuters in these micro survey data, Column (2) re-estimates the same specification restricting attention to bilateral pairs with 10 or more commuters. We find a semi-elasticity of a similar magnitude of -0.0702, which is now more precisely estimated, and the regression R-squared rises to 91 percent.

[45] The districts reported in the micro survey data are post-2001 districts, as discussed in footnote 8.

[46] As shown in the web appendix, this approximation involves approximating a mean of logs with the log of a mean. If the bilateral commuting probabilities were the same for all pairs of blocks within pairs of districts, these two variables would take the same value. More generally, they differ from one another because of Jensen's inequality. In the web appendix, we examine the quantitative relevance of this difference, by using the calibrated model to compare the results of gravity equations estimated at the block and district level using data generated from the model. In practice, we find that this discrepancy is small. Estimating the gravity equation using district-level log commuting probabilities and data generated from the model under the assumption of $\nu = 0.07$, we find a semi-elasticity of $\nu = 0.0726$.

The remaining two columns of Table 3 report additional robustness checks suggested by the international trade literature on gravity equations (see in particular Head and Mayer 2014). In Column (3), we estimate the fixed effects specification from Column (2) using a Poisson Pseudo Maximum Likelihood estimator, and find a semi-elasticity of -0.0771. In Column (4), we re-estimate the same specification using a Gamma Pseudo Maximum Likelihood estimator, and find a semi-elasticity of -0.0723. Therefore, across a range of different specifications, we find a precisely estimated value of $\nu = \epsilon\kappa$ of around $0.07$. Taken together, these results suggest that the gravity equation predicted by the model provides a good approximation to observed commuting behavior.

In Panel A of Figure 4, we provide further evidence on the fit of the semi-elasticity functional form implied by the model using the specification from Column (2) of Table 3. We regress both the log bilateral commuting probabilities and travel times on workplace and residence fixed effects and graph the residuals from these two regressions against one another. As apparent from the figure, the semi-elasticity functional form provides a good fit to the data, with an approximately linear relationship between the two residuals.[47]

Using our estimate for $\nu = \epsilon\kappa = 0.07$, the model's labor market clearing condition (7) can be solved for a transformation of wages ($\omega_{jt} = \tilde{w}_{jt}^{\epsilon} = E_{jt}w_{jt}^{\epsilon}$) in each location in each year using observed workplace employment ($H_{Mjt}$), residence employment ($H_{Rit}$) and bilateral travel times ($\tau_{ijt}$):

$$H_{Mjt} = \sum_{i=1}^{S} \frac{\omega_{jt}/e^{\nu\tau_{ijt}}}{\sum_{s=1}^{S} \omega_{st}/e^{\nu\tau_{ist}}} H_{Rit}. \tag{26}$$

Using these solutions for transformed wages, we obtain bilateral commuting flows in each year (from (6)).

In Panel B of Figure 4, we compare log commuting probabilities in the model and micro survey data at the district level. Again we focus on the sample of bilateral pairs with 10 or more commuters in the micro survey data. The model's predictions and micro survey data can differ for two sets of reasons. First, the reduced-form gravity equation does not perfectly fit the micro survey data (the R-squared in Table 3 is around 0.90). Second, the bilateral travel times in the model are estimated based on minimum travel time calculations using the transport network, which need not equal the self-reported travel times in the micro survey data. Nonetheless, we find a strong relationship between the two sets of commuting probabilities.

We now undertake a number of additional overidentification checks. Given our estimate for $\nu = \epsilon\kappa = 0.07$ from the gravity equation estimation for 2008, we use the model to predict commuting flows and construct a cumulative distribution function of commuters across travel time bins (for example, 20-30 minutes) for all three years of our sample. We compare these predictions of the model to the corresponding cumulative distribution functions in the data. For reunification, we use the micro survey data for individual commuters for 2008 used in the gravity equation estimation above. For division, we use separate data on the fractions of workers in discrete travel time bins from a representative sample of commuters in West Berlin in 1982. For the pre-war period, we use the data reported in Feder (1939).

[47] The use of reduced-form gravity equations for commuting flows has a long tradition in urban and regional economics, as reviewed in McDonald and McMillen (2010). Fortheringham and O'Kelly (1989) argues that the consensus in the literature is that a semi-log specification provides the best fit to commuting data within cities. A recent contribution to this literature using a semi-log specification and travel times is McArthur, Kleppe, Thorsen, and Uboe (2011), which finds a similar semi-elasticity of commuting flows with respect to travel times as we find for Berlin.

In Panel C of Figure 4, we show the cumulative distribution functions of commuters across the travel time bins in the model and micro survey data for Berlin in 2008. Although these moments were not used in the gravity equation estimation of $\nu = \epsilon\kappa$ above, we find that the model approximates the cumulative distribution function in the data well. In Panel D of Figure 4, we undertake the same exercise for West Berlin in 1986. Although the smaller geographic area of West Berlin ensures that it has a quite different distribution of workplace employment, residence employment and travel times from Berlin as a whole, we again find that the model approximates the relationship in the data. Finally, in Panel E of Figure 4, we repeat the exercise for Berlin in 1936. Although the distribution of travel times for Berlin differs between 1936 (based on public transport) and 2006 (based on public transport and private automobiles), we again find that the model has explanatory power for the data.[48]

Therefore, despite our model necessarily being an abstraction, we find that it is successful in capturing the key features of commuting patterns in Berlin during our sample period, and successfully predicts moments not used in the estimation of the commuting parameters.

## 6.2 Productivity and Amenities

We now use the gravity equation estimation to recover overall adjusted productivity, amenities and the density of development $\{\tilde{A}_{it}, \tilde{B}_{it}, \tilde{\varphi}_i\}$. We use the model to recover these objects without taking a stand of the relative importance of externalities and fundamentals. From profit maximization and zero profits (12), log adjusted final goods productivity relative to its geometric mean is:

$$\ln\left(\frac{\tilde{A}_{it}}{\overline{\tilde{A}}_t}\right) = (1-\alpha)\ln\left(\frac{\mathbb{Q}_{it}}{\overline{\mathbb{Q}}_t}\right) + \frac{\alpha}{\epsilon}\ln\left(\frac{\omega_{it}}{\overline{\omega}_t}\right), \tag{27}$$

where a bar above a variable denotes a geometric mean so that $\overline{\tilde{A}}_t = \exp\{\frac{1}{S}\sum_{s=1}^{S}\ln\tilde{A}_{st}\}$. Intuitively high floor prices and wages in (27) require high final goods productivity in order for zero profits to be satisfied.

From the residential choice probabilities (5) and population mobility with the larger economy (9), log residential amenities relative to their geometric mean are:

$$\ln\left(\frac{\tilde{B}_{it}}{\overline{\tilde{B}}_t}\right) = \frac{1}{\epsilon}\ln\left(\frac{H_{Rit}}{\overline{H}_{Rt}}\right) + (1-\beta)\ln\left(\frac{\mathbb{Q}_{it}}{\overline{\mathbb{Q}}_t}\right) - \frac{1}{\epsilon}\ln\left(\frac{W_{it}}{\overline{W}_t}\right), \tag{28}$$

where $W_{it}$ is a measure of commuting market access that can be written in terms of the transformed wages ($\omega_{it}$) from the commuting market clearing condition (26):

$$W_{it} = \sum_{s=1}^{S}\omega_{st}/e^{\nu\tau_{ist}}, \qquad \omega_{st} = \tilde{w}_{st}^{\epsilon} = E_{st}w_{st}^{\epsilon}. \tag{29}$$

Intuitively, high residence employment and high floor prices in (28) must be explained either by high commuting market access or attractive residential amenities. Using residential land market clearing (17) and

[48] Consistent with the results in Duranton and Turner (2011) for U.S. metropolitan areas, we find that the majority of commuters in Berlin have travel times of less than forty-five minutes in all three years of our sample.

commercial land market clearing (18), we can also recover the adjusted density of development relative to its geometric mean.

To solve for adjusted productivity, amenities and the density of development, we require values for a subset of the model's parameters: $\{\alpha, \beta, \mu, \epsilon, \kappa\}$. Of these parameters, the share of residential floor space in consumer expenditure $(1-\beta)$, the share of commercial floor space in firm costs $(1-\alpha)$, and the share of land in construction costs $(1-\mu)$ are hard to determine from our data, because information on consumer expenditures and factor payments at the block level is not available over our long historical sample period. As there is a degree of consensus about the values of these parameters, we set them equal to central estimates from the existing empirical literature. We set the share of consumer expenditure on residential floor space $(1-\beta)$ equal to 0.25, which is consistent with the estimates in Davis and Ortalo-Magné (2011). We assume that the share of firm expenditure on commercial floor space $(1-\alpha)$ is 0.20, which is in line with the findings of Valentinyi and Herrendorf (2008). We set the share of land in construction costs $(1-\mu)$ equal to 0.25, which is consistent with the values in Combes, Duranton, and Gobillon (2014) and Epple, Gordon, and Sieg (2010) and with micro data on property transactions that is available for Berlin from 2000-2012, as discussed in the web appendix.

We use our estimate of $\nu = \epsilon\kappa = 0.07$ from the gravity equation estimation above. To calibrate the value of the Fréchet shape parameter ($\epsilon$), we use our data on the dispersion of log wages by workplace across the districts of West Berlin for 1986. From the labor market clearing condition (26), transformed wages ($\omega_{it}$) are determined independently of $\epsilon$ from workplace employment, residence employment and travel times. Therefore $\epsilon$ merely determines the monotonic transformation that maps transformed wages ($\omega_{it}$) into adjusted wages ($\tilde{w}_{it} = \omega_{it}^{1/\epsilon}$). Hence $\epsilon$ merely scales the dispersion of log adjusted wages relative to the dispersion of log transformed wages: $\sigma^2_{\ln \tilde{w}_{it}} = (1/\epsilon)^2 \sigma^2_{\ln \omega_{it}}$. We choose $\epsilon$ to minimize the squared difference between the variances across districts of log adjusted wages in the model and log wages in the data, which yields a value of $\epsilon = 6.83$.[49] This value of $\epsilon = 6.83$ for commuting decisions is broadly in line with the range of estimates for the Fréchet shape parameter for international trade flows (the range of estimates in Eaton and Kortum 2002 is from 3.60 to 12.86 with a preferred value of 8.28).

From $\nu = \epsilon\kappa = 0.07$ and $\epsilon = 6.83$, we obtain $\kappa = 0.01$. Using these assumed parameter values, we solve for adjusted productivity ($\tilde{A}_{it}$) from (27) and adjusted amenities ($\tilde{B}_{it}$) from (28). We treat these solutions of the model as data and examine the changes in productivity and amenities underlying the impact of division and reunification in our reduced-form "difference-in-difference" specification (22). In Column (1) of Table 4, we estimate our baseline specification for the impact of division for adjusted productivity ($\tilde{A}_i$) including our six grid cells for distance to the pre-war CBD. We find substantial and statistically significant negative treatment effects of division. For example, for the first distance grid cell, we estimate a reduction in productivity of -0.207 log points. In Column (2), we estimate the same specification for adjusted amenities ($\tilde{B}_i$). Again we find substantial and statistically significant negative treatment effects of division. For ex-

[49] Wages across West Berlin districts in 1986 differ by a maximum value of 26 percent, which is in line with the maximum difference in mean residual wages (after controlling for worker observables) across areas of Boston and Minneapolis of 15 and 18 percent reported in Timothy and Wheaton (2001).

ample, for the first distance grid cell, we estimate a reduction in amenities of -0.347 log points. Columns (3) and (4) demonstrate a similar pattern of results for reunification, although the estimated effects are smaller and more localized. These results provide a first piece of evidence that a model in which productivity and amenities are exogenous and unaffected by division and reunification is inconsistent with the data.

### 6.3 Counterfactuals with Exogenous Location Characteristics

To provide further evidence on the ability of a model with exogenous productivity, amenities and the density of development to explain the data, we now undertake counterfactuals for the effects of division and reunification for this special case of the model. Even in the absence of production and residential externalities, the model predicts treatment effects from division, because residents in West Berlin lose access to employment opportunities in East Berlin, and firms in West Berlin lose access to commuters from East Berlin. In response to this shock, workers and residents reallocate across locations, and land is reallocated between commercial and residential use, until wages and floor prices adjust to satisfy zero profits and population mobility. As shown in Proposition 1, the model has a unique equilibrium with exogenous location characteristics, and hence these counterfactuals yield determinate predictions.

In our first counterfactual, we simulate the impact of division on West Berlin, holding productivity, amenities and the density of development constant at their 1936 values. In Column (5) of Table 4, we re-estimate our baseline "difference-in-difference" specification using the counterfactual changes in floor prices predicted by the model with exogenous location characteristics instead of the actual changes in floor prices. We find that the counterfactual treatment effect of division is negative and statistically significant, but substantially smaller than the actual treatment effect of division (-0.408 log points as compared to -0.800 log points for the first distance grid cell in Column (1) of Table 1).

In our second counterfactual, we simulate the impact of reunification on West Berlin, holding productivity, amenities and the density of development constant at their 1986 values in West Berlin, and using the 2006 values of these location characteristics for East Berlin. In Column (6) of Table 4, we re-estimate our baseline "difference-in-difference" specification for the impact of reunification using the counterfactual changes in floor prices. We again find that the counterfactual treatment effect is smaller than the actual treatment effect and is now sometimes statistically insignificant (close to zero as compared to 0.398 log points for the first distance grid cell in Column (1) of Table 2). Therefore the results of these counterfactuals provide further evidence that a model in which productivity and amenities are exogenous and unaffected by division and reunification is unable to explain the data.

## 7 Structural Estimation

In the previous section, we used the model's gravity equation predictions to determine the commuting parameters $\{\nu, \epsilon\}$ without taking a stand on the agglomeration parameters $\{\lambda, \delta, \eta, \rho\}$. In this section, we use the exogenous variation from Berlin's division and reunification to structurally estimate the model's parameters for both agglomeration and dispersion forces $\{\nu, \epsilon, \lambda, \delta, \eta, \rho\}$, where $\nu = \epsilon\kappa$. This enables us to decompose overall adjusted productivity and amenities $\{\tilde{A}_i, \tilde{B}_i\}$ into their two components of externalities $\{\Upsilon_i, \Omega_i\}$ and adjusted fundamentals $\{\tilde{a}_i, \tilde{b}_i\}$. We continue to assume the same central values for the share of floor space in consumer expenditure $(1-\beta)$, the share of floor space in firm costs $(1-\alpha)$, and the share of land in construction costs $(1-\mu)$ as in section 6.2 above.

First, we use the results from Proposition 2 to show that adjusted production and residential fundamentals $\{\tilde{a}_{it}, \tilde{b}_{it}\}$ are structural residuals of the model that are one-to-one functions of the observed data and parameters. Second, we develop moment conditions in terms of these structural residuals that use the exogenous variation induced by Berlin's division and reunification. Third, we discuss the Generalized Method of Moments (GMM) estimation. Fourth, we show that the moment conditions uniquely identify the estimated parameters $\{\nu, \epsilon, \lambda, \delta, \eta, \rho\}$. Fifth, we report the GMM estimation results. Sixth, we compare our results to findings from the existing literature. Seventh, we report additional over identification checks on the model's predictions. Finally, we use the model to undertake counterfactuals in the presence of the estimated production and residential externalities.

## 7.1 Structural Residuals

In this section, we use Proposition 2 to obtain closed-form solutions for adjusted production and residential fundamentals in terms of the observed data and parameters. From profit maximization and zero profits (12) and productivity (20), the proportional change in adjusted production fundamentals in each block relative to the geometric mean can be written as the following function of observed data and parameters:

$$\triangle \ln\left(\frac{\tilde{a}_{it}}{\overline{\overline{\tilde{a}}}_t}\right) = (1-\alpha)\triangle \ln\left(\frac{\mathbb{Q}_{it}}{\overline{\overline{\mathbb{Q}}}_t}\right) + \frac{\alpha}{\epsilon}\triangle \ln\left(\frac{\omega_{it}}{\overline{\overline{\omega}}_t}\right) - \lambda\triangle \ln\left(\frac{\Upsilon_{it}}{\overline{\overline{\Upsilon}}_t}\right), \tag{30}$$

where production externalities $\{\Upsilon_{it}\}$ depend on the travel-time weighted sum of observed workplace employment densities (from (20)); $\omega_{it}$ can be solved from observed workplace employment and residence employment from labor market clearing (26); and a bar above a variable denotes a geometric mean such that $\overline{\overline{\tilde{a}}}_t = \exp\left\{\frac{1}{S}\sum_{s=1}^{S} \ln \tilde{a}_{it}\right\}$.

From population mobility and utility maximization (5) and amenities (21), the proportional change in adjusted residential fundamentals in each block relative to the geometric mean can be written as the following function of observed data and parameters:

$$\triangle \ln\left(\frac{\tilde{b}_{it}}{\overline{\overline{\tilde{b}}}_t}\right) = \frac{1}{\epsilon}\triangle \ln\left(\frac{H_{Rit}}{\overline{\overline{H}}_{Rt}}\right) + (1-\beta)\triangle \ln\left(\frac{\mathbb{Q}_{it}}{\overline{\overline{\mathbb{Q}}}_t}\right) - \frac{1}{\epsilon}\triangle \ln\left(\frac{W_{it}}{\overline{\overline{W}}_t}\right) - \eta\triangle \ln\left(\frac{\Omega_{it}}{\overline{\overline{\Omega}}_t}\right), \tag{31}$$

where residential externalities $\{\Omega_{it}\}$ depend on the travel-time weighted sum of observed residence employment densities (from (21)); commuting market access $\{W_{it}\}$ can be solved from observed workplace employment and residence employment (see (26) and (29)); and a bar above a variable again denotes a geometric mean.

The structural residuals in (30) and (31) difference out any time-invariant factors with time-invariant effects, because of the differencing before and after division (as denoted by the time-difference operator $\triangle$).

These structural residuals also difference out any common fixed effect across all blocks in each year (e.g. changes in the reservation level of utility $\bar{U}_t$ or the choice of units in which to measure production and residential fundamentals), because we divide by the geometric mean of each variable in each year before taking logs. Therefore the mean changes in log adjusted production and residential fundamentals in (30) and (31) are necessarily equal to zero.

## 7.2 Moment Conditions

Our first set of moment conditions impose that the changes in adjusted production and residential fundamentals in (30) and (31) are uncorrelated with the exogenous change in the surrounding concentration of economic activity induced by Berlin's division and reunification. Based on the results of our reduced-form regressions, we capture this exogenous change in the surrounding concentration of economic activity using distance grid cells from the pre-war CBD. Therefore our first set of moment conditions are:

$$\mathbb{E}\left[\mathbb{I}_k \times \triangle \ln\left(\tilde{a}_{it}/\bar{\tilde{a}}_t\right)\right] = 0, \qquad k \in \{1, \dots, K_{\mathbb{I}}\}, \tag{32}$$

$$\mathbb{E}\left[\mathbb{I}_k \times \triangle \ln\left(\tilde{b}_{it}/\bar{\tilde{b}}_t\right)\right] = 0, \qquad k \in \{1, \dots, K_{\mathbb{I}}\}. \tag{33}$$

where $\mathbb{I}_k$ for $k \in \{1, \dots, K_{\mathbb{I}}\}$ are indicator variables for distance grid cell $k$ from the pre-war CBD. We use 50 indicator variables based on percentiles of distance to the pre-war CBD. Therefore the moment conditions (32) and (33) impose that the mean change in log adjusted production and residential fundamentals is zero for each of the distance grid cells.

This identifying assumption requires that the systematic change in the gradient of economic activity in West Berlin relative to the pre-war CBD following division is explained by the mechanisms in the model (the changes in commuting access and production and residential externalities) rather than by systematic changes in the pattern of structural residuals (adjusted production and residential fundamentals).[50] Since Berlin's division stemmed from military considerations during the Second World War and its reunification originated in the wider collapse of Communism, the resulting changes in the surrounding concentration of economic activity are plausibly exogenous to changes in adjusted production and residential fundamentals in West Berlin blocks.

In addition to the above moment conditions for adjusted production and residential fundamentals, we use two other moment conditions for division and reunification based on commuting travel times and wage dispersion for West Berlin during division.[51] The first of these moment conditions requires that the total number of workers commuting for less than 30 minutes in the model is equal to the corresponding number in the data. From the commuting market clearing condition (26), this moment condition can be expressed

[50] We do not use moment conditions in the adjusted density of development ($\tilde{\varphi}_i$) in our estimation, because the density of development could in principle respond to changes in the relative demand for floor space across locations within West Berlin as a result of the mechanisms in the model (the changes in commuting access and production and residential externalities).

[51] In section 6.1, we reported over identification checks in which we showed that the model using an estimated value of $\nu = \epsilon\kappa$ for one year is successful in capturing the pattern of commuting flows in other years of the data, suggesting that the commuting parameters are stable over our sample period.

as the following expectation:

$$\mathbb{E}\left[\psi H_{Mj}-\sum_{i\in\aleph_j}^{S}\frac{\omega_j/e^{\nu\tau_{ij}}}{\sum_{s=1}^{S}\omega_s/e^{\nu\tau_{is}}}H_{Ri}\right]=0, \tag{34}$$

where $\psi$ in the first term inside the square parentheses is the fraction of workers that commute for less than 30 minutes in the data; $\omega_j = \tilde{w}_j^{\epsilon}$; $\aleph_j$ is the set of residence locations $i$ within 30 minutes travel time of workplace location $j$; hence the second term inside the square parentheses captures the model's predictions for commuting flows with travel times less than 30 minutes.

The second of these moment conditions requires that the variance of log adjusted wages in the model (var $(\tilde{w}_i)$) is equal to the variance of log wages in the data ($\sigma^2_{\ln w_i}$) for West Berlin during division:[52]

$$\mathbb{E}\left[(1/\epsilon)^2\ln(\omega_j)^2-\sigma^2_{\ln w_i}\right]=0, \tag{35}$$

where transformed wages ($\omega_i = \tilde{w}_i^{\epsilon}$) depend solely on $\nu$, workplace employment, residence employment and travel times from the labor market clearing condition (26). The parameter $\epsilon$ scales the variance of log adjusted wages ($\tilde{w}_i$) relative to the variance of log transformed wages ($\omega_i$).

### 7.3 GMM Estimation

We use the above moment conditions and the Generalized Method of Moments (GMM) to estimate the model's full set of parameters for agglomeration and dispersion forces $\boldsymbol{\Lambda}$ ={$\nu$, $\epsilon$, $\lambda$, $\delta$, $\eta$, $\rho$}. Stacking our moment conditions together, we obtain:

$$\mathbb{M}(\boldsymbol{\Lambda})=\frac{1}{N}\sum_{i=1}^{N}m(\boldsymbol{X}_i,\boldsymbol{\Lambda})=0. \tag{36}$$

where $m(\boldsymbol{X}_i,\boldsymbol{\Lambda})$ is the moment function for observation $i$.

We estimate the model separately for the difference between the pre-war and division periods and for the difference between the division and reunification periods. The efficient GMM estimator solves:

$$\hat{\boldsymbol{\Lambda}}_{\text{GMM}}=\arg\min\left(\frac{1}{N}\sum_{i=1}^{N}m(\boldsymbol{X_i},\boldsymbol{\Lambda})'\right)\mathbb{W}\left(\frac{1}{N}\sum_{i=1}^{N}m(\boldsymbol{X_i},\boldsymbol{\Lambda})\right) \tag{37}$$

where $\mathbb{W}$ is the efficient weighting matrix. As in our reduced-form estimation, we report Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors that allow for spatial correlation in the errors following Conley (1999).

This minimization problem involves evaluating the moment conditions (36) for each parameter vector and searching over alternative parameter vectors ($\boldsymbol{\Lambda}$). We briefly discuss here the algorithms that we use to solve these problems and include a more detailed discussion in the web appendix. First, evaluating

[52] As reliable wage data for pre-war Berlin is unavailable, we use wages by workplace for West Berlin during division in our moment conditions, which is consistent with our use of the commuting data above.

the moment conditions for each parameter vector involves solving a fixed point problem for the vector of transformed wages that solves the labor market clearing condition (26). In the web appendix, we show analytically that transformed wages are gross substitutes in the labor market clearing condition and that this system of equations has a unique solution (see Lemmas A6 and A7). Therefore, we solve for transformed wages using an iterative fixed point procedure that converges rapidly to this unique solution. Second, this iterative fixed point problem is nested within an optimization routine over the parameter vector ($\mathbf{\Lambda}$). We use standard optimization algorithms to search over alternative possible values for the parameter vector.

## 7.4 Identification

In Proposition 2 we show that we can use the equilibrium conditions of the model to exactly identify adjusted production and residential fundamentals $\{\tilde{a}_i, \tilde{b}_i\}$ from the observed data $\{\mathbb{Q}, \boldsymbol{H_M}, \boldsymbol{H_R}, \boldsymbol{K}, \boldsymbol{\tau}\}$ and known values of the model's parameters $\{\nu, \epsilon, \lambda, \delta, \eta, \rho\}$. Therefore adjusted production and residential fundamentals are structural residuals that are one-to-one functions of the observed data and parameters, as demonstrated in subsection 7.1. We now show how our moment conditions in terms of these structural residuals can be used to identify the model's parameters (and hence recover both the unknown parameters and unobserved adjusted fundamentals).

An important feature of our GMM estimation is that we have closed-form solutions for the structural residuals of adjusted production and residential fundamentals in terms of the observed data and parameters. Therefore, when we consider alternative parameter vectors, we always condition on the *same observed* endogenous variables, and use Proposition 2 to solve for the implied values of adjusted production and residential fundamentals. In contrast, in simulation methods such as simulated method of moments (SMM) or indirect inference, these closed-form solutions are typically not available. Hence these simulation methods are required to solve for alternative values of the endogenous variables for each parameter vector.

We identify the model's parameters using the moment conditions from subsection 7.2. In principle, these moment conditions need not uniquely identify the model's parameters, because the objective function defined by them may not be globally concave. For example, the objective function could be flat in the parameter space or there could be multiple local minima corresponding to different combinations of the parameters $\{\nu, \epsilon, \lambda, \delta, \eta, \rho\}$ and unobserved fundamentals $\{\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}\}$ that are consistent with the same observed data $\{\mathbb{Q}, \boldsymbol{H_M}, \boldsymbol{H_R}, \boldsymbol{K}, \boldsymbol{\tau}\}$. However, in practice, we find that the objective function is well behaved in the parameter space, and that our moment conditions determine a unique parameter vector. In subsection A4.5 of the web appendix, we report the results of a grid search over the parameter space, in which we show that the GMM objective has a unique global minimum that identifies the parameters. In section A6 of the web appendix, we report the results of a Monte Carlo simulation, in which we generate data for a hypothetical city using known parameters, and show that our estimation approach recovers the correct values of these known parameters.

We now consider each of the moment conditions in turn and show how they identify the parameters $\{\nu, \epsilon, \lambda, \delta, \eta, \rho\}$. We begin with the semi-elasticity of commuting flows with respect to travel times ($\nu$). A higher value of $\nu$ implies that commuting flows decline more rapidly with travel times, which implies that a larger fraction of workers commute for less than thirty minutes in the commuting moment condition (34). The recursive structure of the model implies that none of the other parameters $\{\epsilon, \lambda, \delta, \eta, \rho\}$ affect the commuting moment condition ($\epsilon$ only enters through $\nu = \epsilon\kappa$ and $\omega_j$).

We next consider the Fréchet shape parameter determining the heterogeneity of workers' commuting decisions ($\epsilon$). A higher value of $\epsilon$ implies a smaller dispersion in adjusted wages ($\tilde{w}_{it}$) in the wage moment condition (35) given the dispersion in transformed wages ($\omega_{it}$) determined by the commuting parameter $\nu$ and the commuting market clearing condition (26). The recursive structure of the model implies that none of the other parameters $\{\lambda, \delta, \eta, \rho\}$ affect the wage moment condition.

We now turn to the parameters for production spillovers $\{\lambda, \delta\}$ and residential spillovers $\{\eta, \rho\}$. Although the division of Berlin provides a single shock, we can separately identify these two sets of spillover parameters. The reason is that adjusted productivity and amenities $\{\tilde{A}_i, \tilde{B}_i\}$ can be separately recovered from the observed data using the equilibrium conditions of the model (see (27) and (28)). Given these separate measures of productivity and amenities, the productivity spillover parameters $\{\lambda, \delta\}$ could be estimated from a regression of changes in productivity ($\tilde{A}_i$) on changes in production externalities ($\Upsilon_i$), instrumenting changes in production externalities with indicator variables for distance grid cells from the pre-war CBD. Similarly, the residential spillover parameters $\{\eta, \rho\}$ could be estimated from a regression of changes in amenities ($\tilde{B}_i$) on changes in residential externalities ($\Omega_i$), instrumenting changes in residential externalities with indicator variables for distance grid cells from the pre-war CBD. The exclusion restrictions are that: (i) workplace employment affects adjusted productivity but not adjusted amenities, (ii) residence employment affects adjusted amenities but not adjusted productivities. Assumption (i) is the standard specification of production externalities in urban economics and assumption (ii) models residential externalities symmetrically to production externalities, as discussed in subsection 3.5. From the moment conditions for changes in production and residential fundamentals (32)-(33), our GMM estimator is similar to these instrumental variable regressions, but jointly estimates the parameters $\{\nu, \epsilon, \lambda, \delta, \eta, \rho\}$ as part of a system that includes our moment conditions for commuting and wages.

In subsection A4.5 of the web appendix, we show how changes in the spillover parameters $\{\lambda, \delta, \eta, \rho\}$ affect our moment conditions for adjusted production and residential fundamentals. The division of Berlin implies a fall in production externalities ($\Upsilon_i$) for the parts of West Berlin close to the Berlin Wall. If this fall in production externalities does not fully explain the changes in adjusted productivity ($\tilde{A}_{it}$) close to the Berlin Wall, the remainder will be explained by a change in adjusted production fundamentals ($\tilde{a}_{it}$). The parameters $\{\lambda, \delta\}$ control the magnitude of the fall in production externalities and its rate of decay with travel time to Eastern concentrations of workplace employment. From the moment condition (32), the production spillover parameters $\{\lambda, \delta\}$ are chosen to make the mean changes in log adjusted production fundamentals (30) as flat as possible across the distance grid cells from the pre-war CBD.

Similarly, the division of Berlin implies a fall in residential externalities ($\Omega_i$) for the parts of West Berlin close to the Berlin Wall. If this fall in residential externalities does not fully explain the changes in adjusted amenities ($\tilde{B}_i$) close to the Berlin Wall, the remainder will be explained by a change in adjusted residential fundamentals ($\tilde{b}_i$). The parameters $\{\eta, \rho\}$ control the magnitude of the fall in residential externalities and its rate of decay with travel time to Eastern concentrations of residence employment. From the moment condition (33), the residential spillover parameters $\{\eta, \rho\}$ are chosen to make the mean changes in log adjusted residential fundamentals (31) as flat as possible across the distance grid cells from the pre-war CBD.

## 7.5 GMM Estimation Results

In Table 5, we report efficient GMM estimation results for the division and reunification experiments, both separately and pooling the two experiments. In Column (1), we report the results for division. We find substantial and statistically significant agglomeration forces, with an estimated elasticity of productivity with respect to the surrounding concentration of workplace employment of $\lambda = 0.07$, and an estimated elasticity of amenities with respect to the surrounding concentration of residence employment of $\eta = 0.14$. Both production and residential externalities are highly localized, with exponential rates of decay of $\delta = 0.36$ and $\rho = 0.89$ respectively. We find similar commuting parameters as in our earlier estimation based on the gravity equation, with a semi-elasticity of commuting flows with respect to travel time of $\nu = \epsilon\kappa = 0.10$ compared to $\nu = 0.07$. Together our estimates of $\nu$ and $\epsilon$ imply a spatial decay parameter for commuting costs of $\kappa = \nu/\epsilon = 0.01$.

In Column (2) of Table 5, we report the efficient GMM estimation results for reunification. We find a broadly similar pattern of results, although the estimates are smaller and less precisely estimated than for division. We find an elasticity of productivity with respect to production externalities of $\lambda = 0.04$ and an elasticity of amenities with respect to residential externalities of $\eta = 0.07$, which are both significant at conventional levels. Production and residential externalities are again highly localized with rates of spatial decay of 0.89 and 0.55 respectively (although the spatial decay of residential externalities is not significant at conventional levels). Our estimates of both commuting parameters are again similar to our earlier estimates based on the gravity equation.

In Column (3) of Table 5, we report the efficient GMM results pooling the division and reunification experiments, which exploits both sources of variation in the data. To illustrate the magnitude of the spatial decays implied by our parameter estimates, Columns (1) and (2) of Table 6 report the proportional reductions in production and residential externalities with travel time, using the pooled efficient GMM parameter estimates. After around 10 minutes of travel time, both production and residential externalities fall to close to zero. Given our estimated travel speeds for each mode of transport, 10 minutes of travel time corresponds to around 0.83 kilometers by foot (at an average speed of five kilometers per hour) and about 4 kilometers by U-Bahn or S-Bahn (at an average speed of 25 kilometers per hour).

In Column (3) of Table 6, we report the proportional increase in commuting costs with travel time, again using the pooled efficient GMM parameter estimates. Commuting costs are much less responsive to travel times than production or residential externalities. Nonetheless, consistent with the rapid observed decline in commuting with travel time, the implied commuting costs are still substantial. Other things equal, after around 10 minutes of travel time, utility falls by 12 percentage points ($(1 - 0.88) \times 100$). In interpreting this result, one has to take into account that workers self-select across bilateral commutes. Intuitively, workers will only choose to take an extremely long bilateral commute if they have a high draw for the idiosyncratic utility derived from that pair of workplace and residence locations. More formally, an implication of the Fréchet distribution for idiosyncratic utility is that average utility conditional on choosing a bilateral commute is the same for all bilateral commutes, as shown in section A.2.3 in the web appendix.

To the extent that the spillover parameters $\{\lambda, \eta, \delta, \rho\}$ are deep structural parameters, we would expect the estimates to be the same for division and reunification. On the one hand, production technologies, industry composition and the nature of urban amenities could have changed between the division and reunification periods, in such a way as to affect both the magnitude $\{\lambda, \eta\}$ and localization $\{\delta, \rho\}$ of production and residential externalities. On the other hand, as shown in the reduced-form regressions and in Figure 3, reunification is a smaller shock than division, which provides less variation to identify the parameters, as reflected in the larger standard errors on the spillover parameters for reunification than for division. Therefore, to exploit all of the variation in the data, we focus in what follows on the parameter estimates pooling both division and reunification. Although our model (like any model) is necessarily an abstraction, we show below in subsection 7.8 that our pooled parameter estimates generate counterfactual treatment effects for both division and reunification that provide a good approximation to the observed data.

## 7.6 Comparison with Existing Estimates of Agglomeration Economies

Our estimate of the elasticity of productivity with respect to production externalities ($\lambda = 0.07$) is towards the high end of the 3-8 percent range stated in the survey by Rosenthal and Strange (2004), but less than the elasticities from some quasi-experimental studies (see for example Kline and Moretti 2014 and Greenstone, Hornbeck, and Moretti 2010).[53] In comparing our results to the existing literature, a number of points must be taken into account. First, the zero-profit condition in production (12) links overall productivity to wages and floor prices. Therefore lower estimates of agglomeration economies using data on wages, for example, could reflect that higher productivity leads to both higher wages and higher floor prices.[54] Second, our estimate of $\lambda = 0.07$ captures the effect of doubling workplace employment density holding constant travel times. In reality, a doubling in total city population is typically achieved by a combination of an increase in the density of workplace employment and an expansion in geographical land area, with the accompanying increase in average travel times within the city. Therefore the elasticity of productivity with respect to such a doubling of total city population is less than $\lambda = 0.07$, because an increase in average travel times reduces production externalities at a rate determined by the spatial decay parameter $\delta$.[55]

Our findings of substantial and highly localized production externalities are also consistent with recent

[53] In a recent meta-analysis of estimates of urban agglomeration economies, Melo, Graham, and Noland (2009) report a mean estimate of 0.058 across 729 estimates from 34 studies, consistent with Rosenthal and Strange (2004).

[54] To further explore this point, we estimate elasticities of floor prices, wages and productivity with respect to distance from the pre-war CBD in 2006, which are -0.20, -0.11 and -0.13 respectively. Therefore some of the higher productivity close to the CBD is reflected in higher floor prices as well as in higher wages. The wage and productivity elasticities weighted by their coefficients in zero-profits (12) add up to the floor price elasticity (up to rounding): $-((0.8/0.2) \times -0.11) + ((1/0.2) \times -0.13) = -0.21$.

[55] To illustrate the quantitative relevance of this point, we have used the model to estimate the median impact on the productivity of a West Berlin block within 3.75 kilometers of the pre-war CBD from production externalities from East Berlin in 1936, which corresponds to $0.0115$ log points. In contrast, if one could distribute the workplace employment of East Berlin proportionately to each West Berlin block, this would increase workplace employment in each West Berlin block by a factor of 1.7556, and hence from (20) would increase the productivity of each West Berlin block by $0.0738 \times \ln(1.7556) = 0.0415$ log points.

research using within-city data. Using data on the location of advertising agencies in Manhattan, Arzaghi and Henderson (2008) find little evidence of knowledge spillovers beyond 500 meters straight-line distance. To compare straight-line distances to travel times, we computed the mean travel time in 2006 across all bilateral connections in our data that cover a straight-line distance between 450 to 550 meters, which is approximately 9 minutes.[56] After 9 minutes of travel time, our estimates suggest production externalities have declined to around 5 percent. Our estimates of positive residential externalities are in line with the idea that urban amenities are endogenous to the surrounding concentration of economic activity, as in Glaeser, Kolko, and Saiz (2001) and Diamond (2013). Our findings of large and highly localized residential externalities are consistent with other evidence on spillovers in residential choices. Rossi-Hansberg, Sarte, and Owens (2010) uses data on urban revitalization programs in Richmond, Virginia, and finds find that each dollar of home improvement spending generated between \$2 and \$6 in land value by way of externalities in the targeted neighborhoods, but housing externalities fall by approximately one half every 1,000 feet.

### 7.7 Overidentification Checks

In addition to the over identification checks using commuting data discussed above, we now examine the model's predictions for other variables not used in the estimation. We begin with the ratio of floor space to land area. In our structural estimation, we use the equilibrium relationships of the model to solve for the adjusted density of development ($\tilde{\varphi}_i$) that equates the demand for floor space to the supply of floor space. The resulting measure of the ratio of adjusted floor space to land area ($\tilde{L}_i/K_i = \tilde{\varphi}_i K_i^{-\mu}$) is implicitly quality adjusted and captures other variables besides $\varphi_i$ that enter the model isomorphically through $\tilde{\varphi}_i$. In contrast, our measure of the ratio of floor space to land area in the data is coarse, because it is based on a number of discrete categories (e.g. greater than 4.5), and it does not control for the quality of floor space. Nonetheless, as reported in Table A.15 of the web appendix, we find a strong, statistically significant and approximately log linear relationship between the two variables. For example, the estimated coefficient in 2006 is close to one (0.960 with Conley standard error 0.018), with an R-squared of 0.37. Given the caveats noted above, and the fact that this is a univariate regression using cross-sectional micro data, the strength of this empirical relationship provides further support for the model's predictions.

As an additional external validity check, we examine whether our estimates of adjusted production and residential fundamentals $\{\tilde{a}_i, \tilde{b}_i\}$ in 2006 are correlated with observable block characteristics that plausibly affect their suitability for production or residence. As reported in Table A.16 of the web appendix, we find that adjusted residential fundamentals are positively correlated with green areas, proximity to water and listed buildings, and negatively correlated with noise and the level of destruction in the Second World War. In contrast, adjusted production fundamentals are uncorrelated with the level of noise, are less negatively correlated with the level of war-time destruction, and are positively correlated with the other observable block characteristics. Therefore our estimates of adjusted production and residential fundamentals are related in the expected way to separate data on observable correlates for these variables.

[56] The standard deviation of travel times in the 450 to 550 meter straight-line distance bin is 3.4 minutes, which illustrates that straight-line distances are an imperfect proxy for actual travel times in cities, particularly over shorter distances.

## 7.8 Counterfactuals with Endogenous Location Characteristics

In subsection 6.3, we reported counterfactuals for the special case of the model with exogenous location characteristics. In this subsection, we undertake counterfactuals for the model with agglomeration forces, in which productivity and amenities depend on endogenous production and residential externalities. We assume alternative values of location characteristics $\{\tilde{a}_i, \tilde{b}_i, \tau_{ij}\}$ or spillover parameters $\{\lambda, \delta, \eta, \rho\}$ and solve for the model's counterfactual equilibrium. We first use these counterfactuals to provide further evidence on the model's fit by examining the extent to which the observed treatment effects of division and reunification can be explained by the model's agglomeration and dispersion forces rather than by changes in location fundamentals. We next examine the relative importance of production and residential externalities. Finally, we report an out of sample counterfactual, in which we show that the model provides a framework that can be used to examine the impact of a change in transport technology.

As discussed above, in the presence of agglomeration forces, there is the potential for multiple equilibria in the model. We assume the equilibrium selection rule of solving for the closest counterfactual equilibrium to the observed equilibrium prior to the counterfactual. In particular, we use the values of the endogenous variables from the observed equilibrium as our initial guess for the counterfactual equilibrium.[57] Our goal in these counterfactuals is not to determine the unique impact on economic activity, but rather to examine whether the model with the estimated agglomeration parameters is capable of generating counterfactual treatment effects for division and reunification close to the observed treatment effects. In our structural estimation, the model exactly replicates the observed data, because we solve for the values of the structural residuals for which the observed floor price and employment data are an equilibrium of the model. In contrast, in these counterfactuals, the model's predictions need not necessarily replicate the observed data, because we assume alternative (counterfactual) values of location characteristics or parameters.

In our first counterfactual, we simulate division using our pooled parameter estimates and holding location fundamentals $\{\tilde{a}_i, \tilde{b}_i, \tilde{\varphi}_i\}$ constant at their 1936 values. We choose the reservation level of utility in the wider economy following division to ensure that the total population of West Berlin ($H$) is equal to its value in the data in 1986. We estimate our baseline "difference-in-difference" specification (22) using the counterfactual change in floor prices following division. As reported in Column (1) of Table 7, we find counterfactual treatment effects of division close to the observed treatment effects (e.g. -0.781 for the first distance grid cell compared to -0.800 in Column (1) of Table 1). Therefore the estimated model with agglomeration forces can explain quantitatively the observed impact of division, which suggests that our equilibrium selection rule is selecting an equilibrium following division that is close to the observed equilibrium.

In Column (2), we set residential externalities to zero ($\eta = 0$). In Column (3), we set production externalities to zero ($\lambda = 0$). As apparent from the table, both production and residential externalities make substantive contributions to the overall impact of division. In Column (4), we half the rates of spatial decay

[57] Using these initial values, we solve the model's system of equations for a new value of the endogenous variables. We then update our guess for the counterfactual equilibrium based on a weighted average of these new values and the initial values. Finally, we repeat this process until the new values and initial values converge.

of both production externalities and residential externalities $\{\delta, \rho\}$. In this case, we find somewhat larger counterfactual treatment effects for distance grids cells further from the pre-war CBD, consistent with the effect of division extending further into West Berlin.

In our next three counterfactuals, we simulate reunification, choosing the reservation level of utility in the wider economy following reunification to ensure that the total population of Berlin ($H$) is equal to its value in the data in 2006. We again estimate our baseline "difference-in-difference" specification (22) using the counterfactual change in floor prices. In Column (5), we use our pooled parameter estimates, 1986 values of location fundamentals for West Berlin, and 2006 values of location fundamentals for East Berlin. As shown in the table, we find counterfactual treatment effects close to the observed treatment effects (e.g. 0.345 for the first distance grid cell compared to 0.398 in Column (1) of Table 2). Therefore, consistent with our results for division, the model with the estimated agglomeration forces can explain the observed impact of reunification.

In Column (6), we consider the same specification as in Column (5), but use 1936 (instead of 2006) values of location fundamentals for East Berlin. We now find counterfactual treatment effects substantially larger than the observed treatment effects (e.g. 1.097 versus 0.398 for the first distance grid cell). This pattern of results is consistent with the idea that a recovery of East Berlin to the relative levels of economic development prior to the Second World War would lead to a further reorientation of economic activity within West Berlin. In Column (7), we consider the same specification as in Column (5), but use our division (instead of pooled) parameter estimates. We find counterfactual treatment effects for reunification close to those in Column (5), which is consistent with the similarity of the division and pooled parameter estimates.[58]

Although the focus of our analysis is on the division and reunification of Berlin, our quantitative model provides a tractable platform for undertaking a range of counterfactuals. As an illustration of the model's potential, our final counterfactual examines a change in transport technology. In particular, we examine the impact of the automobile on the location of economic activity within Berlin. We use the model to solve for the counterfactual equilibrium distribution of economic activity in 2006 using travel time measures based solely on the public transport network in 2006. To focus on the impact of the change in transport technology in Berlin, we hold the reservation utility in the wider economy constant.

Our 2006 travel time measures using only public transport are typically higher than our baseline 2006 measures that weight public transport and the automobile by their modal shares. In comparison with American cities, the public transport network is far more extensive in Berlin (on average public transport, including walking and cycling, accounts for around two thirds of journeys in our 2006 data) and is relatively more important for commuting into the central city. Table A.18 of the web appendix compares the actual and counterfactual travel times. As shown in rows 1-4 of the table, the unweighted average travel time across all possible bilateral connections with positive values of either workplace or residence employment rises

[58] While for brevity we concentrate on the model's counterfactual predictions for the gradient of economic activity with respect to pre-war CBD, we find that it is also successful in accounting for other features of the observed data, such as the gradient of economic activity with respect to the Kudamm. For example, regressing floor prices in West Berlin in 1986 and 2006 on grid cells of 500 meter intervals for distance to the Kudamm (and including grid cells for distance to the pre-war CBD as controls), we find similar gradients of economic activity with respect to the Kudamm for counterfactual floor prices as for actual floor prices.

from 51 minutes to 70 minutes, and its standard deviation rises from 12 minutes to 26 minutes (since remote locations with high actual travel times and poor public transport connections are most affected). As implied by our gravity equation estimation in subsection 6.1, commuting flows are higher on average for shorter travel times. Therefore, as shown in rows 5-6 of the table, if we weight travel times by the actual bilateral commuting flows in the 2006 equilibrium, average travel times rise from 32 minutes to 38 minutes.

The commuting technology facilitates a separation of workplace and residence, enabling people to work in relatively high productivity locations (typically in more central locations) and live in high amenity locations (typically in suburban locations). The deterioration of the commuting technology triggers an outflow of workers from Berlin, until floor prices fall such that expected utility in Berlin is again equal to the unchanged reservation level of utility in the wider economy. Total city population and output fall by around 11 and 10 percent respectively (rows 7-8 of the table). Output falls by less than population, because labor is only one of the two factors of production and the total supply of floor space is held constant. On average floor prices decline by 16 percent (row 9 of the table). This decline in floor prices is substantially larger for blocks experiencing above median increases in average unweighted travel times (typically in remote locations) than for blocks experiencing below median increases in these travel times (typically in more central locations), as shown in rows 10-11 of the table.

The general equilibrium response of the economy to the deterioration in the commuting technology is that locations become less specialized in workplace and residence activity, as shown in Figure A.16 of the web appendix. Panel A shows that blocks that are larger importers of commuters before the change in transport technology (larger net commuting on the horizontal axis) experience larger declines in workplace employment (on the vertical axis). Panel B shows that blocks that are larger exporters of commuters before the change in transport technology (smaller net commuting on the horizontal axis) experience larger declines in residence employment (on the vertical axis). A corollary of this decline in block specialization is a change in the pattern of worker sorting across bilateral pairs of workplace and residence locations. Even though travel times for a typical bilateral pair have increased, we find that this change in worker sorting results in average travel times weighted by commuting flows in the counterfactual equilibrium (row 12 of Table A.18) that marginally decline relative to average travel times weighted by commuting flows in the actual equilibrium (row 5 of the table). Taken together, these results highlight that the model provides a framework that can be used to analyze the endogenous change in the organization of economic activity within cities in response to changes in the transport network and other interventions (such as planning regulations).

# 8 Conclusions

In this paper, we develop a quantitative theoretical model of city structure that incorporates agglomeration and dispersion forces, allows for asymmetries in locational fundamentals, and remains tractable and amenable to empirical analysis. To separate out agglomeration and dispersion forces from heterogeneity in locational fundamentals, we combine the model with the exogenous source of variation in the surrounding concentration of economic activity provided by Berlin's division and reunification.

The model implies a gravity equation for bilateral commuting flows, which is successful in accounting for observed commuting patterns. We find substantial differences in productivity and amenities across locations within cities that are endogenous to the surrounding concentration of economic activity. While our estimates of the elasticity of productivity with respect to density using within-city data are somewhat higher than those using across-city data, we highlight the importance of taking into account the rapid spatial decay of production externalities when comparing estimates at different levels of spatial aggregation. We also find residential externalities of a comparable magnitude to production externalities.

We show that the special case of the model in which productivity and amenities are exogenous is unable to account quantitatively for the observed treatment effects of division and reunification. In contrast, for the estimated values of agglomeration forces, the model is successful in accounting for these observed treatment effects. Although we use the natural experiment of Berlin to estimate the model's parameters, our quantitative framework can be used to undertake counterfactuals for changes in the organization of economic activity within cities in response to other interventions, such as changes in the transport network.

# References


Albouy, D. (2008): "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas," *NBER Working Paper*, 14472.

Alonso, W. (1964): *Location and Land Use.* Harvard, Cambridge MA.

American Embassy (1965): *Handbook of Economic Statistics: Federal Republic of Germany and Western Sectors of Berlin.* American Embassy, Bonn.

Arzaghi, M., and J. V. Henderson (2008): "Networking Off Madison Avenue," *Review of Economic Studies*, 75(4), 1011–1038.

Baum-Snow, N. (2007): "Did Highways Cause Suburbanization?," *Quarterly Journal of Economics*, 122(2), 775–780.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How Much Should We Trust Differences-in-Differences Estimators?," *Quarterly Journal of Economics*, 119(1), 249–275.

Bleakley, H., and J. Lin (2012): "Portage: Path Dependence and Increasing Returns in U.S. History," *Quarterly Journal of Economics*, 127(2), 587–644.

Brinkman, J. (2013): "Congestion, Agglomeration, and the Structure of Cities," *Philadelphia Fed Working Paper*, 13-25.

Ciccone, A., and R. E. Hall (1996): "Productivity and the Density of Economic Activity," *American Economic Review*, 86(1), 54–70.

Combes, P.-P., G. Duranton, and L. Gobillon (2014): "The Costs of Agglomeration: Land Prices in French Cities," Wharton, mimeograph.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010): "Estimating Agglomeration Economies with History, Geology, and Worker Effects," in *Agglomeration Economies*, ed. by E. L. Glaeser. Chicago University Press, Chicago.

Conley, T. G. (1999): "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 92(1), 1–45.

Davis, D. R., and D. E. Weinstein (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92(5), 1269–1289.

——— (2008): "A Search for Multiple Equilibria in Urban Industrial Structure," *Journal of Regional Science*, 48(1), 29–65.

Davis, M. A., and F. Ortalo-Magné (2011): "Housing Expenditures, Wages, Rents," *Review of Economic Dynamics*, 14(2), 248–261.

Dekle, R., and J. Eaton (1999): "Agglomeration and Land Rents: Evidence from the Prefectures," *Journal of Urban Economics*, 46(2), 200–214.

Diamond, R. (2013): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," Stanford Graduate School of Business, mimeo.

Donaldson, D. (2014): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, forthcoming.

Duranton, G., and D. Puga (2004): "Micro-Foundations of Urban Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J.-F. Thisse, vol. 4, pp. 2063–2117. Elsevier, Amsterdam.

Duranton, G., and M. A. Turner (2011): "The Fundamental Law of Road Congestion: Evidence from US Cities," *American Economic Review*, 101(6), 2616–52.

——— (2012): "Urban Growth and Transportation," *Review of Economic Studies*, 79(4), 1407–1440.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Elkins, T. H., and B. Hofmeister (1988): *Berlin: Spatial Structure of a Divided City*. Methuen & Co., London and New York.

Epple, D., B. Gordon, and H. Sieg (2010): "A New Approach to Estimating the Production Function for Housing," *American Economic Review*, 100(3), 905–924.

Faber, B. (2014): "Integration and the Periphery: The Unintended Effects of New Highways in a Developing Country," *Review of Economic Studies*, forthcoming.

Feder, G. (1939): *Arbeitstätte - Wohnstätte*, Schriftenreihe der Reichsarbeitsgemeinschaft für Raumforschung an der Technischen Hochschule Berlin. Julius Springer, Berlin.

Fortheringham, S., and M. O'Kelly (1989): *Spatial Interaction Models: Formulations and Applications.* Kluwer, Dordrecht.

Franklin, W. M. (1963): "Zonal Boundaries and Access to Berlin," *World Politics*, 16(1), 1–31.

Fujita, M., and P. Krugman (1995): "When is the Economy Monocentric? Von Thünen and Chamberlain Unified," *Regional Science and Urban Economics*, 25(4), 505–528.

Fujita, M., and H. Ogawa (1982): "Multiple Equilibria and Structural Transformation of Non-monocentric Urban Configurations," *Regional Science and Urban Economics*, 12(2), 161–196.

Glaeser, E. L., J. Kolko, and A. Saiz (2001): "Consumer City," *Journal of Economic Geography*, 1(1), 27–50.

Glaeser, E. L., and D. C. Mare (2001): "Cities and Skills," *Journal of Labor Economics*, 19(2), 316–342.

Greenstone, M., R. Hornbeck, and E. Moretti (2010): "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 118(3), 536–598.

Grogger, J., and G. Hanson (2011): "Income Maximization and the Selection and Sorting of International Migrants," *Journal of Development Economics*, 95(1), 42–57.

Hanson, G. H. (1996): "Localization Economies, Vertical Organization, and Trade," *American Economic Review*, 86(5), 1266–1278.

——— (1997): "Increasing Returns, Trade, and the Regional Structure of Wages," *Economic Journal*, 107(1), 113–133.

Head, K., and T. Mayer (2014): "Gravity Equations: Workhorse, Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by E. Helpman, G. Gopinath, and K. Rogoff, vol. 4, pp. 131–196. Elsevier, North Holland, Amsterdam.

Henderson, J. V., K. Ari, and M. Turner (1995): "Industrial Development in Cities," *Journal of Political Economy*, 103(5), 1067–1085.

Herdegen, G. (1992): "Einstellungen zur Deutschen Einheit," in *Handwörterbuch zur Deutschen Einheit*, ed. by W. Weidenfeld, and K.-R. Korte, vol. 4. Campus Verlag, Frankfurt am Main.

Kalweit, F. (1929): *Die Baustellenwerte in Berlin 1928*. Emro, Berlin.

——— (1937): *Die Baustellenwerte in Berlin: Strassen-Verzeichnis, Neubearbeitung Abgeschlossen am 30. Juni 1936*. Carl Heymanns, Berlin.

Kennan, J., and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211–251.

Kline, P., and E. Moretti (2014): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 129(1), 275–331.

Lambrecht, H., and H. Tischner (1969): *Der Handel West-Berlins mit der DDR und Osteuropa: Probleme und Möglichkeiten*. Deutsches Institut für Wirtschaftsforschung, Deutschland.

Leyden, F. (1933): *Gross-Berlin: Geographie einer Weltstadt*. Gebr. Mann Verlag, Berlin.

Lucas, R. E. (2000): "Externalities and Cities," *Review of Economic Dynamics*, 4(2), 245–274.

Lucas, R. E., and E. Rossi-Hansberg (2002): "On the Internal Structure of Cities," *Econometrica*, 70(4), 1445–1476.

Marshall, A. (1920): *Principles of Economics*. Macmillan, London.

McArthur, D. P., G. Kleppe, I. Thorsen, and J. Uboe (2011): "The Spatial Transferability of Parameters in a Gravity Model of Commuting Flows," *Journal of Transport Geography*, 19(4), 596–605.

McDonald, J., and D. McMillen (2010): *Urban Economics and Real Estate: Theory and Policy*. John Wiley & Sons, Hoboken, NJ.

McFadden, D. (1974): "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3(4), 303–328.

Melo, P. C., D. J. Graham, and R. B. Noland (2009): "A Meta-analysis of Estimates of Urban Agglomeration Economies," *Regional Science and Urban Economics*, 39(3), 332–342.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *Review of Economics and Statistics*, 90(4), 683–701.

Mills, E. S. (1967): "An Aggregative Model of Resource Allocation in a Metropolitan Centre," *American Economic Review*, 57(2), 197–210.

Moretti, E. (2004): "Workers' Education, Spillovers and Productivity: Evidence from Plant-level Production Functions," *American Economic Review*, 94, 656–690.

——— (2011): "Local Labor Markets," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 4b, pp. 1238–1303. Elsevier North Holland, Amsterdam.

Muth, R. (1969): *Cities and Housing*. University of Chicago Press, Chicago.

Rauch, J. E. (1993): "Productivity Gains from Geographic Concentration of Human Capital," *Journal of Urban Economics*, 34(3), 380–400.

Redding, S. J., and D. M. Sturm (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5), 1766–1797.

Redding, S. J., D. M. Sturm, and N. Wolf (2011): "History and Industry Location: Evidence from German Airports," *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), 1257–1278.

Roggenbuch, F. (2008): *Das Berliner Grenzgängerproblem. Verflechtung und Systemkonkurrenz vor dem Mauerbau*. Walter de Gruyter, Berlin and New York.

Rosenthal, S. S., and W. C. Strange (2004): "Evidence on the Nature and Sources of Agglomeration Economics," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. Thisse, vol. 4. Elsevier North Holland, Amsterdam.

——— (2008): "The Attenuation of Human Capital Spillovers," *Journal of Urban Economics*, 64(2), 373–389.

Rossi-Hansberg, E., P.-D. Sarte, and R. I. Owens (2010): "Housing Externalities," *Journal of Political Economy*, 118(3), 485–535.

Sharp, T. (1975): *The Wartime Alliance and the Zonal Division of Germany*. Clarendon Press, Oxford.

Sveikauskas, L. A. (1975): "The Productivity of Cities," *Quarterly Journal of Economics*, 89(3), 393–413.

Timothy, D., and W. C. Wheaton (2001): "Intra-Urban Wage Variation, Employment Location, and Commuting Times," *Journal of Urban Economics*, 50(2), 338–366.

Valentinyi, A., and B. Herrendorf (2008): "Measuring Factor Income Shares at the Sectoral Level," *Review of Economic Dynamics*, 11(4), 820–835.

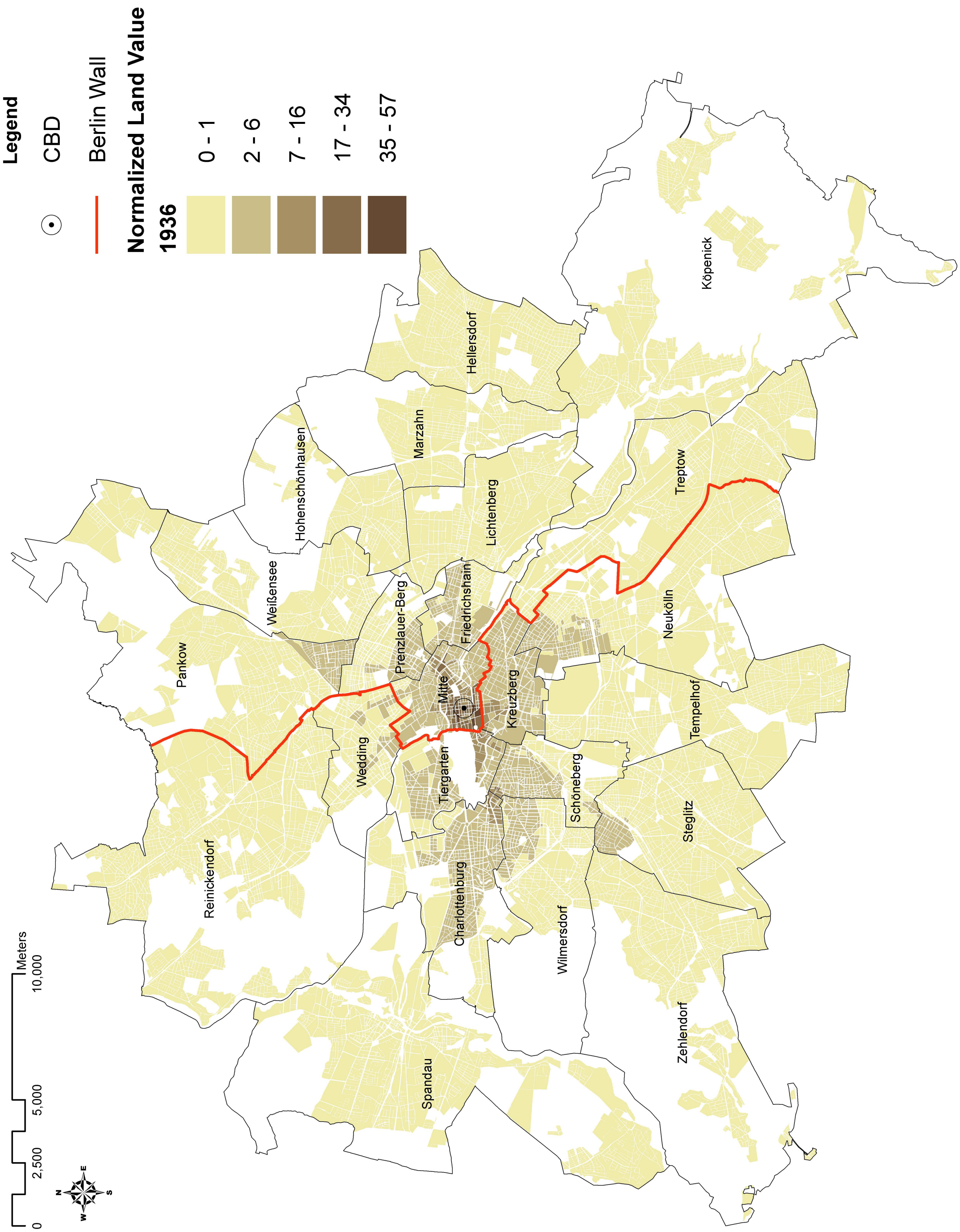



Figure 1: Land Values in Berlin in 1936

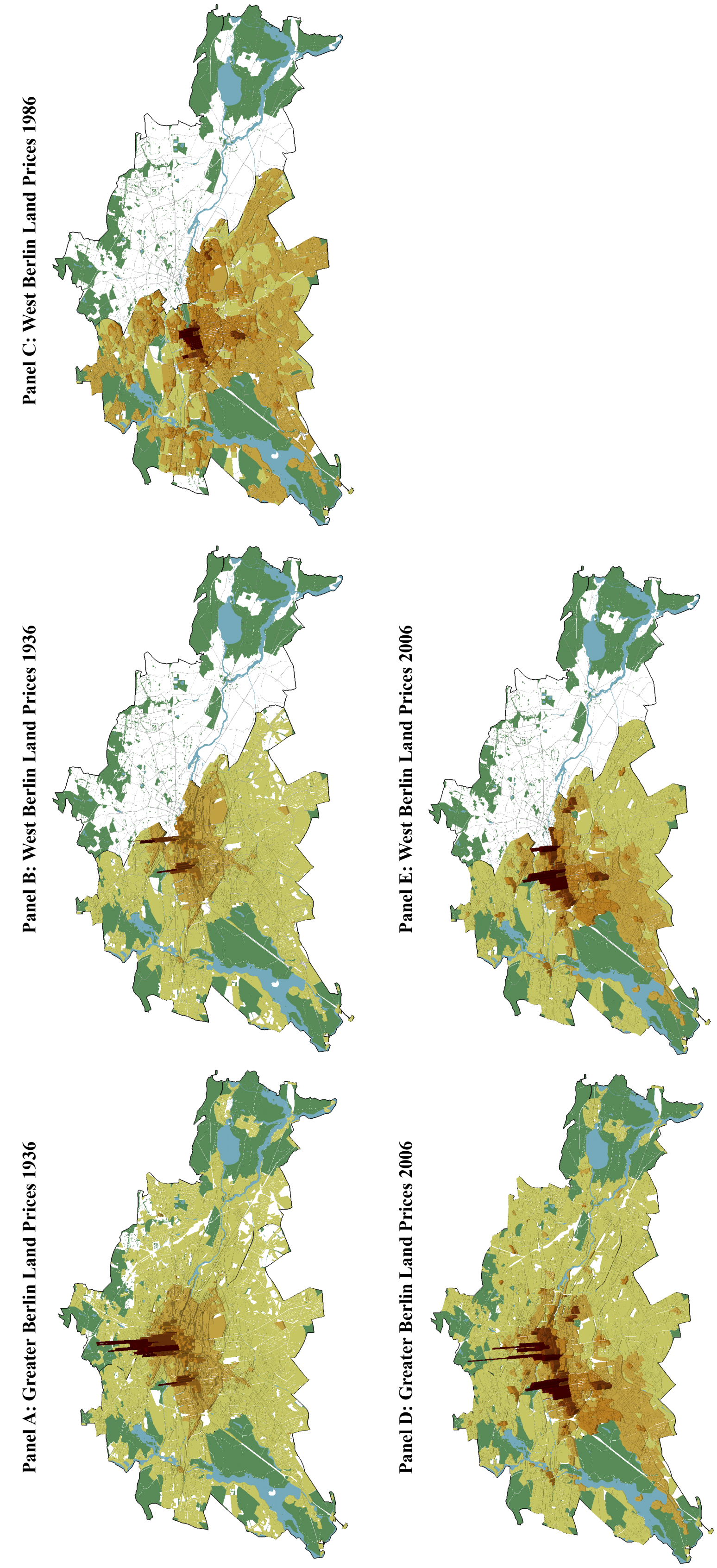
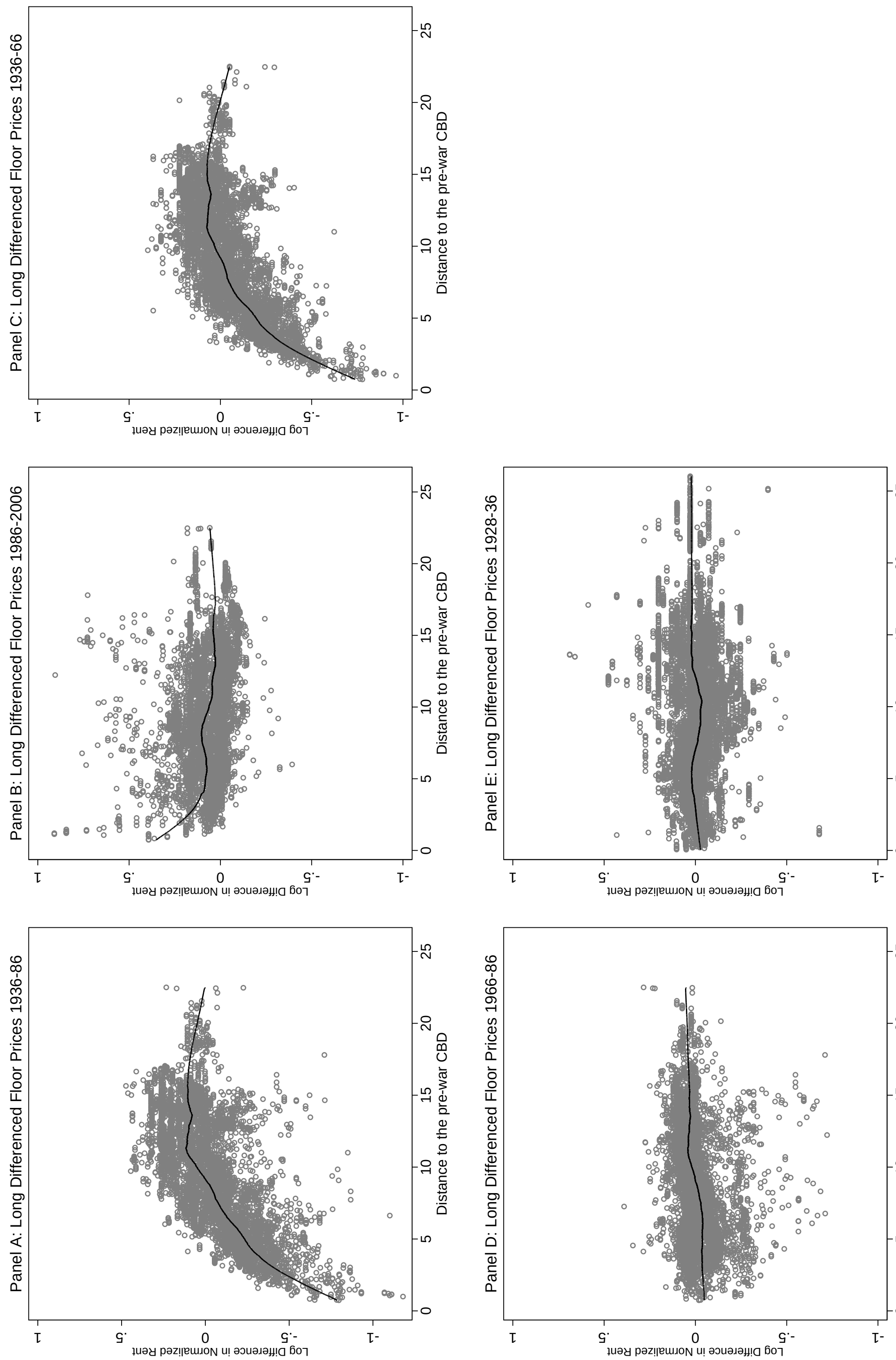



Figure 2: The Evolution of Land Prices in Berlin Over Time




Note: Log floor prices are normalized to have a mean of zero in each year before taking the long difference. Solid lines are fitted values from locally-weighted linear least squares regressions.

Figure 3: Division and Reunification Treatments and Placebos

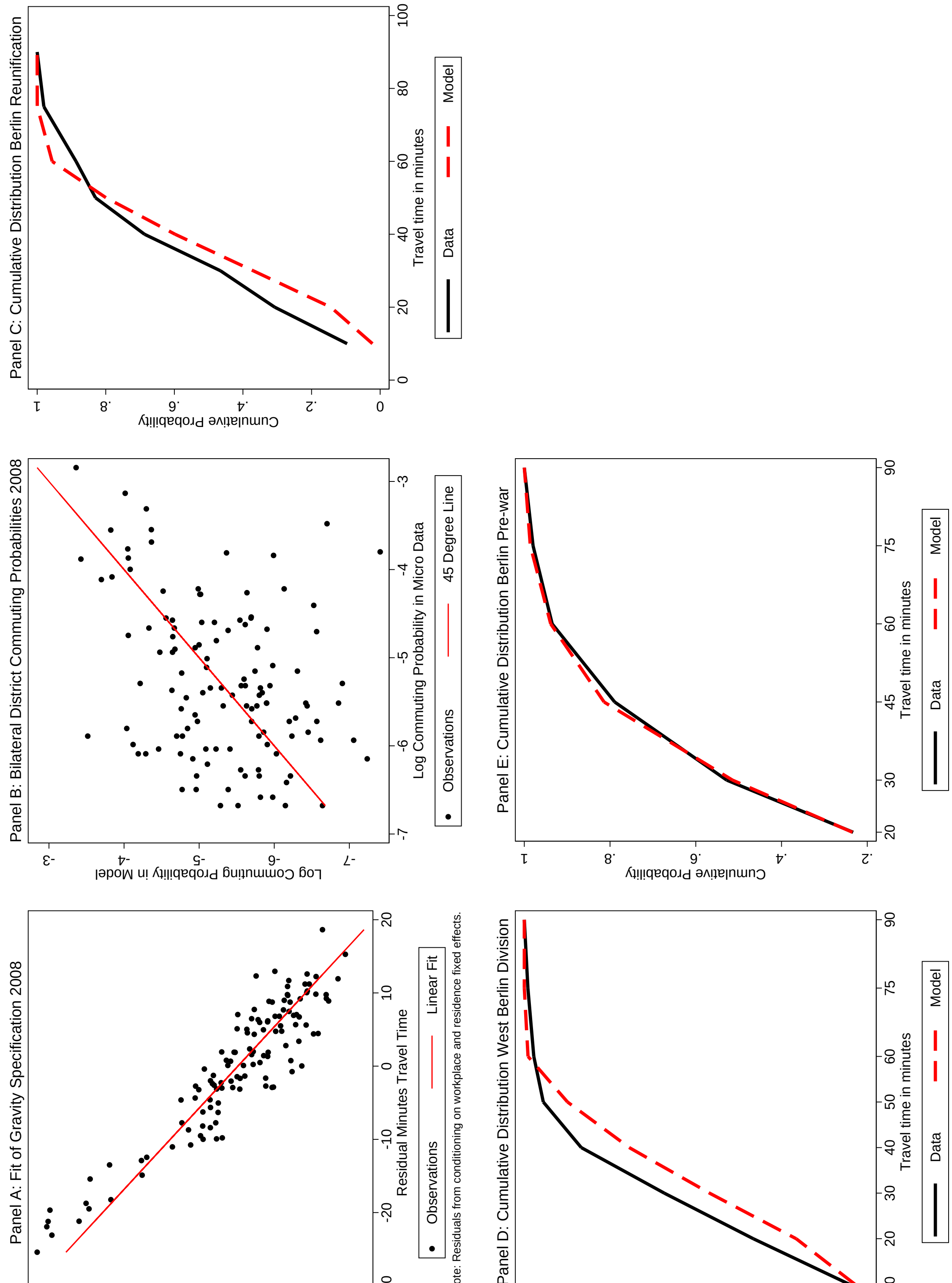



Figure 4: Commuting Parameter Estimation and Overidentification

| | (1) $\Delta$ ln Q | (2) $\Delta$ ln Q | (3) $\Delta$ ln Q | (4) $\Delta$ ln Q | (5) $\Delta$ ln Q | (6) $\Delta$ ln EmpR | (7) $\Delta$ ln EmpR | (8) $\Delta$ ln EmpW | (9) $\Delta$ ln EmpW |
|---|---|---|---|---|---|---|---|---|---|
| CBD 1 | -0.800*** | -0.567*** | -0.524*** | -0.503*** | -0.565*** | -1.332*** | -0.975*** | -0.691* | -0.639* |
| | (0.071) | (0.071) | (0.071) | (0.071) | (0.077) | (0.383) | (0.311) | (0.408) | (0.338) |
| CBD 2 | -0.655*** | -0.422*** | -0.392*** | -0.360*** | -0.400*** | -0.715** | -0.361 | -1.253*** | -1.367*** |
| | (0.042) | (0.047) | (0.046) | (0.043) | (0.050) | (0.299) | (0.280) | (0.293) | (0.243) |
| CBD 3 | -0.543*** | -0.306*** | -0.294*** | -0.258*** | -0.247*** | -0.911*** | -0.460** | -0.341 | -0.471** |
| | (0.034) | (0.039) | (0.037) | (0.032) | (0.034) | (0.239) | (0.206) | (0.241) | (0.190) |
| CBD 4 | -0.436*** | -0.207*** | -0.193*** | -0.166*** | -0.176*** | -0.356** | -0.259 | -0.512*** | -0.521*** |
| | (0.022) | (0.033) | (0.033) | (0.030) | (0.026) | (0.145) | (0.159) | (0.199) | (0.169) |
| CBD 5 | -0.353*** | -0.139*** | -0.123*** | -0.098*** | -0.100*** | -0.301*** | -0.143 | -0.436*** | -0.340*** |
| | (0.016) | (0.024) | (0.024) | (0.023) | (0.020) | (0.110) | (0.113) | (0.151) | (0.124) |
| CBD 6 | -0.291*** | -0.125*** | -0.094*** | -0.077*** | -0.090*** | -0.360*** | -0.135 | -0.280** | -0.142 |
| | (0.018) | (0.019) | (0.017) | (0.016) | (0.016) | (0.100) | (0.089) | (0.130) | (0.116) |
| Inner Boundary 1-6 | | | Yes | Yes | Yes | | Yes | | Yes |
| Outer Boundary 1-6 | | | Yes | Yes | Yes | | Yes | | Yes |
| Kudamm 1-6 | | | | Yes | Yes | | Yes | | Yes |
| Block Characteristics | | | | | Yes | | Yes | | Yes |
| District Fixed Effects | | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6260 | 6260 | 6260 | 6260 | 6260 | 5978 | 5978 | 2844 | 2844 |
| R-squared | 0.26 | 0.51 | 0.63 | 0.65 | 0.71 | 0.19 | 0.43 | 0.12 | 0.33 |

Note: Q denotes the price of floor space. EmpR denotes employment by residence. EmpW denotes employment by workplace. CBD1-CBD6 are six 500m distance grid cells for distance from the pre-war CBD. Inner Boundary 1-6 are six 500m grid cells for distance to the Inner Boundary between East and West Berlin. Outer Boundary 1-6 are six 500m grid cells for distance to the outer boundary between West Berlin and East Germany. Kudamm 1-6 are six 500m grid cells for distance to Breitscheid Platz on the Kurfürstendamm. The coefficients on the other distance grid cells are reported in Table A2 of the web appendix. Block characteristics include the logarithm of distance to schools, parks and water, the land area of the block, the share of the block's built-up area destroyed during the Second World War, indicators for residential, commercial and industrial land use, and indicators for whether a block includes a government building and urban regeneration policies post-reunification. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley 1999). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1: Baseline Division Difference-in-Difference Results (1936-86)

| | (1) $\Delta \ln Q$ | (2) $\Delta \ln Q$ | (3) $\Delta \ln Q$ | (4) $\Delta \ln Q$ | (5) $\Delta \ln Q$ | (6) $\Delta \ln EmpR$ | (7) $\Delta \ln EmpR$ | (8) $\Delta \ln EmpW$ | (9) $\Delta \ln EmpW$ |
|---|---|---|---|---|---|---|---|---|---|
| CBD 1 | 0.398*** | 0.408*** | 0.368*** | 0.369*** | 0.281*** | 1.079*** | 1.025*** | 1.574*** | 1.249** |
| | (0.105) | (0.090) | (0.083) | (0.081) | (0.088) | (0.307) | (0.297) | (0.479) | (0.517) |
| CBD 2 | 0.290*** | 0.289*** | 0.257*** | 0.258*** | 0.191** | 0.589* | 0.538* | 0.684** | 0.457 |
| | (0.111) | (0.096) | (0.090) | (0.088) | (0.087) | (0.315) | (0.299) | (0.326) | (0.334) |
| CBD 3 | 0.122*** | 0.120*** | 0.110*** | 0.115*** | 0.063** | 0.340* | 0.305* | 0.326 | 0.158 |
| | (0.037) | (0.033) | (0.032) | (0.032) | (0.028) | (0.180) | (0.158) | (0.216) | (0.239) |
| CBD 4 | 0.033*** | 0.031 | 0.030 | 0.034 | 0.017 | 0.110 | 0.034 | 0.336** | 0.261 |
| | (0.013) | (0.023) | (0.022) | (0.021) | (0.020) | (0.068) | (0.066) | (0.161) | (0.185) |
| CBD 5 | 0.025*** | 0.018 | 0.020 | 0.020 | 0.015 | -0.012 | -0.056 | 0.114 | 0.066 |
| | (0.010) | (0.015) | (0.014) | (0.014) | (0.013) | (0.056) | (0.057) | (0.118) | (0.131) |
| CBD 6 | 0.019** | -0.000 | -0.000 | -0.003 | 0.005 | 0.060 | 0.053 | 0.049 | 0.110 |
| | (0.009) | (0.012) | (0.012) | (0.012) | (0.011) | (0.039) | (0.041) | (0.095) | (0.098) |
| Inner Boundary 1-6 | | | Yes | Yes | Yes | | Yes | | Yes |
| Outer Boundary 1-6 | | | Yes | Yes | Yes | | Yes | | Yes |
| Kudamm 1-6 | | | | Yes | Yes | | Yes | | Yes |
| Block Characteristics | | | | | Yes | | Yes | | Yes |
| District Fixed Effects | | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7050 | 7050 | 7050 | 7050 | 7050 | 6718 | 6718 | 5602 | 5602 |
| R-squared | 0.08 | 0.32 | 0.34 | 0.35 | 0.43 | 0.04 | 0.07 | 0.03 | 0.06 |

Note: Q denotes the price of floor space. EmpR denotes employment by residence. EmpW denotes employment by workplace. CBD1-CBD6 are six 500m distance grid cells for distance from the pre-war CBD. Inner Boundary 1-6 are six 500m grid cells for distance to the Inner Boundary between East and West Berlin. Outer Boundary 1-6 are six 500m grid cells for distance to the outer boundary between West Berlin and East Germany. Kudamm 1-6 are six 500m grid cells for distance to Breitscheid Platz on the Kurfürstendamm. The coefficients on the other distance grid cells are reported in Table A4 of the web appendix. Block characteristics include the log distance to schools, parks and water, the land area of the block, the share of the block's built-up area destroyed during the Second World War, indicators for residential, commercial and industrial land use, and indicators for whether a block includes a government building and urban regeneration policies post-reunification. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley 1999).* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: Baseline Reunification Difference-in-Difference Results (1986-2006)

| | (1) ln Bilateral Commuting Probability 2008 | (2) ln Bilateral Commuting Probability 2008 | (3) ln Bilateral Commuting Probability 2008 | (4) ln Bilateral Commuting Probability 2008 |
|---|---|---|---|---|
| Travel Time ($-\kappa\varepsilon$) | -0.0697*** | -0.0702*** | -0.0771*** | -0.0706*** |
| | (0.0056) | (0.0034) | (0.0025) | (0.0026) |
| Estimation | OLS | OLS | Poisson PML | Gamma PML |
| More than 10 Commuters | | Yes | Yes | Yes |
| Fixed Effects | Yes | Yes | Yes | Yes |
| Observations | 144 | 122 | 122 | 122 |
| R-squared | 0.8261 | 0.9059 | - | - |

Note: Gravity equation estimates based on representative micro survey data on commuting for Greater Berlin for 2008. Observations are bilateral pairs of 12 workplace and residence districts (post 2001 Bezirke boundaries). Travel time is measured in minutes. Fixed effects are workplace district fixed effects and residence district fixed effects. The specifications labelled more than 10 commuters restrict attention to bilateral pairs with 10 or more commuters. Poisson PML is Poisson Pseudo Maximum Likelihood estimator. Gamma PML is Gamma Pseudo Maximum Likelihood Estimator. Standard errors in parentheses are heteroscedasticity robust. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Commuting Gravity Equation

**Table 4: Floor Prices, Productivity and Amenities**

| | (1) Δ ln A 1936-1986 | (2) Δ ln B 1936-1986 | (3) Δ ln A 1986-2006 | (4) Δ ln B 1986-2006 | (5) Δ ln QC 1936-1986 | (6) Δ ln QC 1986-2006 |
|---|---|---|---|---|---|---|
| CBD 1 | -0.207*** | -0.347*** | 0.261*** | 0.203*** | -0.408*** | -0.010 |
| | (0.049) | (0.070) | (0.073) | (0.054) | (0.038) | (0.020) |
| CBD 2 | -0.260*** | -0.242*** | 0.144** | 0.109* | -0.348*** | 0.079** |
| | (0.032) | (0.053) | (0.056) | (0.058) | (0.017) | (0.036) |
| CBD 3 | -0.138*** | -0.262*** | 0.077*** | 0.059** | -0.353*** | 0.036 |
| | (0.021) | (0.037) | (0.024) | (0.026) | (0.022) | (0.031) |
| CBD 4 | -0.131*** | -0.154*** | 0.057*** | 0.010 | -0.378*** | 0.093*** |
| | (0.016) | (0.023) | (0.015) | (0.008) | (0.021) | (0.026) |
| CBD 5 | -0.095*** | -0.126*** | 0.028** | -0.014* | -0.380*** | 0.115*** |
| | (0.014) | (0.013) | (0.013) | (0.007) | (0.022) | (0.033) |
| CBD 6 | -0.061*** | -0.117*** | 0.023** | 0.001 | -0.354*** | 0.066*** |
| | (0.015) | (0.015) | (0.010) | (0.005) | (0.018) | (0.023) |
| Counterfactuals | | | | | Yes | Yes |
| Agglomeration Effects | | | | | No | No |
| Observations | 2844 | 5978 | 5602 | 6718 | 6260 | 7050 |
| R-squared | 0.09 | 0.06 | 0.02 | 0.03 | 0.07 | 0.03 |

Note: Columns (1)-(4) based on calibrating the model for $\nu=\varepsilon\kappa=0.07$ and $\varepsilon=6.83$ from the gravity equation estimation. Columns (5)-(6) report counterfactuals for these parameter values. A denotes adjusted overall productivity. B denotes adjusted overall amenities. QC denotes counterfactual floor prices (simulating the effect of division on West Berlin). Column (5) simulates division holding A and B constant at their 1936 values. Column (6) simulates reunification holding A and B for West Berlin constant at their 1986 values and using 2006 values of A and B for East Berlin. CBD1-CBD6 are six 500m distance grid cells for distance from the pre-war CBD. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley 1999). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Productivity, Amenities and Counterfactual Floor Prices

| | (1) Division Efficient GMM | (2) Reunification Efficient GMM | (3) Division and Reunification Efficient GMM |
|---|---|---|---|
| Commuting Travel Time Elasticity (κε) | 0.0951*** | 0.1011*** | 0.0987*** |
| | (0.0016) | (0.0016) | (0.0016) |
| Commuting Heterogeneity (ε) | 7.6278*** | 7.7926*** | 7.7143*** |
| | (0.1085) | (0.1152) | (0.1049) |
| Productivity Elasticity (λ) | 0.0738*** | 0.0449*** | 0.0657*** |
| | (0.0056) | (0.0071) | (0.0048) |
| Productivity Decay (δ) | 0.3576*** | 0.8896*** | 0.3594*** |
| | (0.0945) | (0.3339) | (0.0724) |
| Residential Elasticity (η) | 0.1441*** | 0.0740*** | 0.1444*** |
| | (0.0080) | (0.0287) | (0.0073) |
| Residential Decay (ρ) | 0.8872*** | 0.5532 | 0.7376*** |
| | (0.2774) | (0.3699) | (0.1622) |

Note: Generalized Method of Moments (GMM) estimates. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley 1999). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Generalized Method of Moments (GMM) Estimation Results

| | (1) Production Externalities ($1 \times e^{-\delta\tau}$) | (2) Residential Externalities ($1 \times e^{-\rho\tau}$) | (3) Utility after Commuting ($1 \times e^{-\kappa\tau}$) |
|---|---|---|---|
| 0 minutes | 1.000 | 1.000 | 1.000 |
| 1 minute | 0.698 | 0.478 | 0.987 |
| 2 minutes | 0.487 | 0.229 | 0.975 |
| 3 minutes | 0.340 | 0.109 | 0.962 |
| 5 minutes | 0.166 | 0.025 | 0.938 |
| 7 minutes | 0.081 | 0.006 | 0.914 |
| 10 minutes | 0.027 | 0.001 | 0.880 |
| 15 minutes | 0.005 | 0.000 | 0.825 |
| 20 minutes | 0.001 | 0.000 | 0.774 |
| 30 minutes | 0.000 | 0.000 | 0.681 |

Note: Proportional reduction in production and residential externalities with travel time and proportional reduction in utility from commuting with travel time. Travel time is measured in minutes. Results are based on the pooled efficient GMM parameter estimates: δ=0.3594, ρ=0.7376, κ=0.0128.

Table 6: Externalities and Commuting Costs

**Table 7: Counterfactuals**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Δ ln QC | Δ ln QC | Δ ln QC | Δ ln QC | Δ ln QC | Δ ln QC | Δ ln QC |
| | 1936-1986 | 1936-1986 | 1936-1986 | 1936-1986 | 1986-2006 | 1986-2006 | 1986-2006 |
| CBD 1 | -0.781*** | -0.612*** | -0.433*** | -0.766*** | 0.345*** | 1.097*** | 0.375*** |
| | (0.050) | (0.030) | (0.058) | (0.048) | (0.041) | (0.047) | (0.042) |
| CBD 2 | -0.516*** | -0.396*** | -0.335*** | -0.580*** | 0.222*** | 0.745*** | 0.226*** |
| | (0.032) | (0.024) | (0.018) | (0.027) | (0.027) | (0.042) | (0.026) |
| CBD 3 | -0.414*** | -0.306*** | -0.308*** | -0.489*** | 0.153*** | 0.568*** | 0.174*** |
| | (0.036) | (0.029) | (0.030) | (0.034) | (0.029) | (0.041) | (0.029) |
| CBD 4 | -0.386*** | -0.273*** | -0.312*** | -0.476*** | 0.127*** | 0.422*** | 0.131*** |
| | (0.025) | (0.018) | (0.022) | (0.029) | (0.019) | (0.042) | (0.019) |
| CBD 5 | -0.379*** | -0.251*** | -0.320*** | -0.472*** | 0.161*** | 0.375*** | 0.166*** |
| | (0.030) | (0.022) | (0.026) | (0.037) | (0.029) | (0.038) | (0.029) |
| CBD 6 | -0.314*** | -0.207*** | -0.275*** | -0.394*** | 0.090*** | 0.312*** | 0.094*** |
| | (0.023) | (0.015) | (0.021) | (0.028) | (0.022) | (0.034) | (0.021) |
| Counterfactuals | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Agglomeration Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6260 | 6260 | 6260 | 6260 | 7050 | 6260 | 7050 |
| R-squared | 0.11 | 0.15 | 0.07 | 0.13 | 0.12 | 0.24 | 0.13 |

Note: Columns (1)-(6) are based on the parameter estimates pooling division and reunification from Table 5. Column (7) is based on the parameter estimates for division from Table 5. QC denotes counterfactual floor prices. Column (1) simulates division using our estimates of production and residential externalities and 1936 fundamentals. Column (2) simulates division using our estimates of production externalities and 1936 fundamentals but setting residential externalities to zero. Column (3) simulates division using our estimates of residential externalities and 1936 fundamentals but setting production externalities to zero. Column (4) simulates division using our estimates of production and residential externalities and 1936 fundamentals but halving their rates of spatial decay with travel time. Column (5) simulates reunification using our estimates of production and residential externalities, 1986 fundamentals for West Berlin, and 2006 fundamentals for East Berlin. Column (6) simulates reunification using our estimates of production and residential externalities, 1986 fundamentals for West Berlin and 1936 fundamentals for East Berlin. Column (7) simulates reunification using division rather than pooled parameter estimates, 1986 fundamentals for West Berlin, and 2006 fundamentals for East Berlin. CBD1-CBD6 are six 500m distance grid cells for distance from the pre-war CBD. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley 1999). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Counterfactuals