Econometrica Supplementary Material

SUPPLEMENT TO: "THE ECONOMICS OF DENSITY: EVIDENCE FROM THE BERLIN WALL"

(*Econometrica*, Vol. 83, No. 6, NOVEMBER 2015, 2127–2189)

BY GABRIEL M. AHLFELDT, STEPHEN J. REDDING, DANIEL M. STURM, AND NIKOLAUS WOLF

S.1. INTRODUCTION

THIS SUPPLEMENT CONTAINS DETAILED DERIVATIONS and supplementary material for the paper. A separate Technical Data Appendix, located together with the replication data and code, collects together additional empirical results and robustness tests.

Section S.2 of this supplement presents a more detailed analysis of the theoretical model. We report the technical derivations of the expressions reported in the paper. We also establish a number of results about the properties of the general equilibrium with exogenous location characteristics and endogenous agglomeration forces.

Section S.3 calibrates the model for known parameter values and shows that there is a one-to-one mapping from these known parameters and the observed data to unobserved location characteristics. Therefore, these unobserved location characteristics correspond to structural residuals that are functions of the parameters and the observed data.

Section S.4 turns to the structural estimation of the model, where both the parameters and unobserved location characteristics are unknown and to be estimated. We derive the moment conditions used in the estimation and review the Generalized Method of Moments (GMM) estimator as applied to our setting. We discuss the computational algorithms used to estimate the model and report the results of a grid search over the parameter space that we use to characterize the properties of the GMM objective function.

Section S.5 uses the model to undertake counterfactuals for the effects of division and reunification. Section S.6 contains further information about the data sources and definitions.

S.2. THEORETICAL MODEL

In this section, we develop in further detail the theoretical model outlined in the paper. We present the complete technical derivations for all the expressions and results reported in the paper. In the interests of clarity and to ensure that this section of the supplement is self-contained, we reproduce some material from the paper, but also include the intermediate steps for the derivation of expressions.

We consider a city embedded within a wider economy. The city consists of a set of discrete locations or blocks, which are indexed by $i = 1, \dots, S$. The city is

 DOI: 10.3982/ECTA10876

populated by an endogenous measure of H workers, who are perfectly mobile within the city and the larger economy. 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 θ_i and $(1-\theta_i)$, respectively.

Workers decide whether or not to move to the city before observing idiosyncratic utility shocks for each possible pair of residence and employment locations within the city. If a worker decides to move to the city, she observes these realizations for idiosyncratic utility, and picks the pair of residence and employment locations within the city that maximizes her utility. Population mobility between the city and the wider economy implies that the expected utility from moving to the city equals the reservation level of utility in the wider economy $\bar{U}$. Firms produce a single final good, which is costlessly traded within the city and larger economy, and is chosen as the numeraire $(p = 1)$.[1]

Locations differ in terms of their final goods productivity (A_i), residential amenities (B_i), supply of floor space (L_i), and access to the transport network (τ_{ij}). We first develop the model with exogenous values of these location characteristics, before endogenizing them below.

S.2.1. *Preferences*

Workers are risk neutral such that the utility of worker o residing in block i and working in block j is linear in an aggregate consumption index (C_{ijo}):[2]

$$U_{ijo} = C_{ijo}.$$

This aggregate consumption index depends on consumption of the single final good (c_{ijo}), consumption of residential floor space (ℓ_{ijo}), and three other components: first, 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); second, the disutility from commuting from residence block i to workplace block j $(d_{ij} \geq 1)$; third, there is 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

[1]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, though the model can be extended to allow for the consumption of nontraded goods at both workplace and residence.

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

of the city. In particular, the aggregate consumption index is assumed to take the Cobb–Douglas form[3]

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

where the iceberg commuting cost $d_{ij} = e^{\kappa \tau_{ij}} \in [1, \infty)$ increases with the travel time between blocks i and j (τ_{ij}). Travel time is measured in minutes and is computed based on the transport network, as discussed further in the data section of this supplement (Section S.6). The parameter κ controls the size of commuting costs.

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 (S.5) below 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 (S.5) multiplicatively with the wage.

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:

$$F(z_{ijo}) = e^{-T_i E_j z_{ijo}^{-\varepsilon}}, \quad T_i, E_j > 0, \varepsilon > 1, \tag{S.2}$$

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

After observing her realizations for idiosyncratic utility for each pair of residence and employment locations, each worker chooses her blocks of residence and employment to maximize her utility, taking as given residential amenities, goods prices, factor prices, and the location decisions of other workers and firms. Each worker is endowed with one unit of labor that is supplied inelastically with zero disutility. Combining our choice of the final good as numeraire ($p_i = p = 1$ for all i) with the first-order conditions for consumer equilibrium,

[3]For empirical evidence using U.S. 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.

we obtain the following demands for the final good and residential land for worker o residing in block i and working in block j:

$$c_{ijo} = \beta w_j, \tag{S.3}$$

$$\ell_{ijo} = (1-\beta)\frac{w_j}{Q_i}, \tag{S.4}$$

where w_j is the wage received by the worker at her block of employment j (recall that commuting costs are incurred in terms of utility); Q_i is the price of residential land at her block of residence i. We make the standard assumption 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. Substituting equilibrium consumption of the final good (S.3) and residential land use (S.4) into utility (S.1), we obtain the following expression for the indirect utility function:

$$U = \frac{B_i z_{ijo} w_j Q_i^{\beta-1}}{d_{ij}}, \tag{S.5}$$

from which the isomorphic formulation of commuting costs in terms of a reduction in effective units of labor is apparent.

S.2.2. *Distribution of Utility*

Using the monotonic relationship between the aggregate consumption index (S.1) and the idiosyncratic component of utility, the distribution of utility for a worker living in block i and working in block j is also Fréchet distributed:

$$\begin{aligned} G_{ij}(u) &= \Pr[U \le u] = F\left(\frac{u d_{ij} Q_i^{1-\beta}}{B_i w_j}\right), \\ G_{ij}(u) &= e^{-\Phi_{ij} u^{-\varepsilon}}, \quad \Phi_{ij} = T_i E_j \left(d_{ij} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_j)^{\varepsilon}. \end{aligned} \tag{S.6}$$

From all possible pairs of blocks of residence and employment, each worker chooses the bilateral commute that offers the maximum utility. Since the maximum of a sequence of Fréchet distributed random variables is itself Fréchet distributed, the distribution of utility across all possible pairs of blocks of residence and employment is

$$1 - G(u) = 1 - \prod_{r=1}^{S} \prod_{s=1}^{S} e^{-\Phi_{rs} u^{-\varepsilon}},$$

where the left-hand side is the probability that a worker has a utility greater than u, and the right-hand side is one minus the probability that the worker has a utility less than u for all possible pairs of blocks of residence and employment. Therefore, we have

$$G(u)=e^{-\Phi u^{-\varepsilon}}, \quad \Phi=\sum_{r=1}^{S}\sum_{s=1}^{S}\Phi_{rs}. \tag{S.7}$$

Given this Fréchet distribution for utility, the expected utility from moving to the city is

$$\mathbb{E}[u]=\int_0^{\infty}\varepsilon\Phi u^{-\varepsilon}e^{-\Phi u^{-\varepsilon}}\,du. \tag{S.8}$$

Now define the following change of variables:

$$y=\Phi u^{-\varepsilon}, \quad dy=-\varepsilon\Phi u^{-(\varepsilon+1)}\,du. \tag{S.9}$$

Using this change of variables, the expected utility from moving to the city can be written as

$$\mathbb{E}[u]=\int_0^{\infty}\Phi^{1/\varepsilon}y^{-1/\varepsilon}e^{-y}\,dy, \tag{S.10}$$

which can be in turn written as

$$\mathbb{E}[u]=\gamma\Phi^{1/\varepsilon}, \quad \gamma=\Gamma\left(\frac{\varepsilon-1}{\varepsilon}\right), \tag{S.11}$$

where $\Gamma(\cdot)$ is the Gamma function; $\mathbb{E}$ is the expectations operator and the expectation is taken over the distribution for idiosyncratic utility. Population mobility implies that this expected utility must equal the reservation level of utility in the wider economy:

$$\mathbb{E}[u]=\gamma\Phi^{1/\varepsilon}=\gamma\left[\sum_{r=1}^{S}\sum_{s=1}^{S}T_rE_s\left(d_{rs}Q_r^{1-\beta}\right)^{-\varepsilon}(B_rw_s)^{\varepsilon}\right]^{1/\varepsilon}=\bar{U}. \tag{S.12}$$

S.2.3. *Residence and Workplace Choices*

Using the distribution of utility for pairs of blocks of residence and employment (S.6), the probability that a worker chooses the bilateral commute from i to j out of all possible bilateral commutes within the city is

$$\begin{aligned}\pi_{ij}&=\Pr\left[u_{ij}\geq\max\{u_{rs}\};\forall r,s\right]\\&=\int_0^{\infty}\prod_{s\neq j}G_{is}(u)\left[\prod_{r\neq i}\prod_{s}G_{rs}(u)\right]g_{ij}(u)\,du\end{aligned}$$

$$= \int_0^\infty \prod_{r=1}^S \prod_{s=1}^S \varepsilon \Phi_{ij} u^{-(\varepsilon+1)} e^{-\Phi_{rs} u^{-\varepsilon}} \, du$$

$$= \int_0^\infty \varepsilon \Phi_{ij} u^{-(\varepsilon+1)} e^{-\Phi u^{-\varepsilon}} \, du.$$

Note that

$$\text{(S.13)} \qquad \frac{d}{du}\left[-\frac{1}{\Phi} e^{-\Phi u^{-\varepsilon}}\right] = \varepsilon u^{-(\varepsilon+1)} e^{-\Phi u^{-\varepsilon}}.$$

Using this result to evaluate the integral above, the probability that the worker chooses to live in block i and commute to work in block j is

$$\text{(S.14)} \qquad \pi_{ij} = \frac{T_i E_j \left(d_{ij} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_j)^\varepsilon}{\sum_{r=1}^S \sum_{s=1}^S T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_s)^\varepsilon} \equiv \frac{\Phi_{ij}}{\Phi}.$$

Therefore, workers sort across residence and employment locations depending on their idiosyncratic preferences and the characteristics of these locations. As discussed above, although we interpret the idiosyncratic shock as affecting utility, there is an isomorphic interpretation of the model in which the idiosyncratic shock applies to effective units of labor. Therefore, the endogenous sorting of workers across locations implies that both residence and employment locations differ in the composition of workers in terms of idiosyncratic draws for utility or effective units of labor. Residential locations with higher values of T_i have higher average draws of utility (or effective units of labor). Similarly, employment locations with higher values of E_j have higher average draws of utility (or effective units of labor). To ensure that the general equilibrium of the model remains tractable, and because we do not observe worker characteristics in our data, we abstract from other dimensions of worker heterogeneity besides the idiosyncratic shock to preferences or effective units of labor.

Summing across all possible employment locations s, we obtain the probability that a worker chooses to live in block i out of all possible locations within the city:

$$\text{(S.15)} \qquad \pi_{Ri} = \frac{\sum_{s=1}^S T_i E_s \left(d_{is} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_s)^\varepsilon}{\sum_{r=1}^S \sum_{s=1}^S T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_s)^\varepsilon} \equiv \frac{\Phi_i}{\Phi}.$$

Similarly, summing across all possible residence locations r, we obtain the probability that a worker chooses to work in block j out of all possible locations within the city:

$$\pi_{Mj} = \frac{\sum_{r=1}^{S} T_r E_j \left(d_{rj} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_j)^{\varepsilon}}{\sum_{r=1}^{S} \sum_{s=1}^{S} T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_s)^{\varepsilon}} \equiv \frac{\Phi_j}{\Phi}. \tag{S.16}$$

For the measure of workers within block j (H_{Mj}), we can evaluate the conditional probability that they commute from block i (conditional on having chosen to work in block j):

$$\begin{aligned}\pi_{ij|j} &= \Pr\left[u_{ij} \geq \max\{u_{sj}\}; \forall s\right] \\ &= \int_0^\infty \prod_{s \neq i} G_{sj}(u) g_{ij}(u)\, du = \int_0^\infty e^{-\Phi_j u^{-\varepsilon}} \varepsilon \Phi_{ij} u^{-(\varepsilon+1)}\, du.\end{aligned}$$

Using the result (S.13) to evaluate the integral above, the probability that a worker commutes from block i conditional on having chosen to work in block j is

$$\pi_{ij|j} = \frac{T_i E_j \left(d_{ij} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_j)^{\varepsilon}}{\sum_{r=1}^{S} T_r E_j \left(d_{rj} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_j)^{\varepsilon}} = \frac{\Phi_{ij}}{\Phi_j},$$

which simplifies to

$$\pi_{ij|j} = \frac{T_i \left(d_{ij} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i)^{\varepsilon}}{\sum_{r=1}^{S} T_r \left(d_{rj} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r)^{\varepsilon}}. \tag{S.17}$$

For the measure of residents within block i (H_{Ri}), we can evaluate the conditional probability that they commute to block j (conditional on having chosen to live in block i):

$$\begin{aligned}\pi_{ij|i} &= \Pr\left[u_{ij} \geq \max\{u_{is}\}; \forall s\right] \\ &= \int_0^\infty \prod_{s \neq j} G_{is}(u) g_{ij}(u)\, du = \int_0^\infty e^{-\Phi_i u^{-\varepsilon}} \varepsilon \Phi_{ij} u^{-(\varepsilon+1)}\, du.\end{aligned}$$

Using the result (S.13) to evaluate the integral above, the probability that a worker commutes to block j conditional on having chosen to live in block i is

$$\pi_{ij|i} = \frac{T_i E_j \left(d_{ij} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_j)^{\varepsilon}}{\sum_{s=1}^{S} T_i E_s \left(d_{is} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_s)^{\varepsilon}} = \frac{\Phi_{ij}}{\Phi_i},$$

which simplifies to

$$\pi_{ij|i} = \frac{E_j (w_j/d_{ij})^{\varepsilon}}{\sum_{s=1}^{S} E_s (w_s/d_{is})^{\varepsilon}}. \tag{S.18}$$

These conditional commuting probabilities provide microeconomic foundations for the reduced-form gravity equations estimated in the empirical literature on commuting patterns. The probability that a resident of block i commutes to block j depends on the adjusted wage and commuting costs for block j in the numerator ("bilateral resistance"), but also on the adjusted wage and commuting costs for all other possible employment locations s in the denominator ("multilateral resistance").

Commuting market clearing requires that the measure of workers employed in each location j (H_{Mj}) equals the sum across all locations i of their measures of residents (H_{Ri}) times their conditional probabilities of commuting to j ($\pi_{ij|i}$):

$$\begin{aligned} H_{Mj} &= \sum_{i=1}^{S} \pi_{ij|i} H_{Ri} \\ &= \sum_{i=1}^{S} \frac{E_j (w_j/d_{ij})^{\varepsilon}}{\sum_{s=1}^{S} E_s (w_s/d_{is})^{\varepsilon}} H_{Ri}, \end{aligned} \tag{S.19}$$

where, 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.

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:

$$\text{(S.20)} \qquad \mathbb{E}[w_s|i] = \sum_{s=1}^{S} \pi_{is|i} w_s = \sum_{s=1}^{S} \frac{E_s(w_s/d_{is})^{\varepsilon}}{\sum_{r=1}^{S} E_r(w_r/d_{ir})^{\varepsilon}} w_s,$$

where $\mathbb{E}$ denotes the expectations operator and the expectation is taken over the distribution for the idiosyncratic component of utility. Intuitively, expected worker income is high in blocks that have low commuting costs (low d_{is}) to high-wage employment locations.[4]

Finally, another implication of the Fréchet distribution of utility is that the distribution of utility conditional on residing in block i and commuting to block j is the same across all bilateral pairs of blocks with positive residents and employment, and is equal to the distribution of utility for the city as a whole. To establish this result, note that the distribution of utility conditional on residing in block i and commuting to block j is given by

$$\text{(S.21)} \qquad \frac{1}{\pi_{ij}} \int_0^u \prod_{s \neq j} G_{is}(v) \left[\prod_{r \neq i} \prod_{s} G_{rs}(v) \right] g_{ij}(v)\, dv = \frac{1}{\pi_{ij}} \int_0^u \left[\prod_{r=1}^{S} \prod_{s=1}^{S} e^{-\Phi_{rs} v^{-\varepsilon}} \right] \varepsilon \Phi_{ij} v^{-(\varepsilon+1)}\, dv = \frac{\Phi}{\Phi_{ij}} \int_0^u e^{-\Phi v^{-\varepsilon}} \varepsilon \Phi_{ij} v^{-(\varepsilon+1)}\, dv = e^{-\Phi u^{\varepsilon}}.$$

On the one hand, more attractive residential fundamentals in location i or a higher wage in location j raise the utility of a worker with a given realization of idiosyncratic utility z, and hence increase the expected utility of residing in i and working in j. On the other hand, more attractive residential fundamentals or a higher wage induce workers with lower realizations of idiosyncratic utility z to reside in i and work in j, which reduces the expected utility of residing in i and working in j. With a Fréchet distribution of utility, these two effects exactly

[4]For simplicity, we model agents and workers as synonymous, which implies 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, and/or to allow agents to have a constant amount of nonlabor income.

offset one another. Pairs of residence and employment locations with attractive fundamentals attract more commuters on the extensive margin until expected utility is the same across all pairs of residence and employment locations within the city.

S.2.4. *Production*

We follow the canonical urban model in assuming a single final good that is costlessly traded within the city and the larger economy.[5] Final goods production occurs under conditions of perfect competition and constant returns to scale. For simplicity, we assume that the production technology takes the Cobb–Douglas form, so that output of the final good in block j (y_j) is

$$y_j = A_j (H_{Mj})^{\alpha} (L_{Mj})^{1-\alpha},$$

where A_j is final goods productivity, H_{Mj} is workplace employment, and L_{Mj} is 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. From the first-order conditions for profit maximization, we obtain

$$H_{Mj} = \left(\frac{\alpha A_j}{w_j} \right)^{1/(1-\alpha)} L_{Mj}, \tag{S.22}$$

$$L_{Mj} = \left(\frac{(1-\alpha) A_j}{q_j} \right)^{1/\alpha} H_{Mj}. \tag{S.23}$$

Therefore, employment in block j is increasing in productivity (A_j), decreasing in the wage (w_j), and increasing in commercial land use (L_{Mj}). Similarly, commercial land use in block j is increasing in productivity, decreasing in the commercial floor price (q_j), and increasing in employment (H_{Mj}).

To determine the equilibrium commercial floor price, q_j, we use the requirement that profits are zero if the final good is produced:

$$A_j (H_{Mj})^{\alpha} (L_{Mj})^{1-\alpha} - w_j H_{Mj} - q_j L_{Mj} = 0,$$

[5]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)).

which, together with profit maximization (S.22), yields the following expression for the equilibrium commercial floor price:

$$q_j = (1-\alpha)\left(\frac{\alpha}{w_j}\right)^{\alpha/(1-\alpha)} A_j^{1/(1-\alpha)}. \tag{S.24}$$

Intuitively, blocks that have higher productivity (A_j) or lower wages (w_j) are more attractive production locations, and hence must be characterized by higher commercial floor prices in an equilibrium in which firms make zero profits in all locations with positive production.

S.2.5. *Land Market Clearing*

Land market equilibrium requires no-arbitrage between commercial and residential land use after taking into account the tax equivalent of land use regulations:

$$\begin{aligned} &\theta_i = 1 \quad \text{if} \quad q_i > \xi_i Q_i, \\ &\theta_i \in [0,1] \quad \text{if} \quad q_i = \xi_i Q_i, \\ &\theta_i = 0 \quad \text{if} \quad q_i < \xi_i Q_i, \end{aligned} \tag{S.25}$$

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.

Therefore, floor space in each block is either allocated entirely to commercial use ($q_i > \xi_i Q_i$ and $\theta_i = 1$), allocated entirely to residential use ($q_i < \xi_i Q_i$ and $\theta_i = 0$), or allocated to both uses ($q_i = \xi_i Q_i$ and $\theta_i \in (0,1)$). 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

$$\begin{aligned} &\mathbb{Q}_i = q_i, \quad q_i > \xi_i Q_i, \quad \theta_i = 1, \\ &\mathbb{Q}_i = q_i, \quad q_i = \xi_i Q_i, \quad \theta_i \in (0,1), \\ &\mathbb{Q}_i = Q_i, \quad q_i < \xi_i Q_i, \quad \theta_i = 0. \end{aligned} \tag{S.26}$$

We follow the standard approach in the urban literature of assuming that floor space L is supplied by a competitive construction sector that uses geographic 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: $L_i = M_i^{\mu} K_i^{1-\mu}$.[6]

[6]Empirically, we find that this Cobb–Douglas assumption is consistent with the micro data on property transactions for Berlin from 2000 to 2012, as shown in Section S.6.6 of this supplement.

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, and $\mathbb{R}_i$ is the price for geographic land. Since the price for capital is the same across all locations, the relationships between the quantities and prices of floor space and geographical land area can be summarized as

$$L_i = \varphi_i K_i^{1-\mu}, \tag{S.27}$$

$$\mathbb{Q}_i = \chi \mathbb{R}_i^{1-\mu}, \tag{S.28}$$

where $\varphi_i = M_i^{\mu}$ determines the density of development (the ratio of floor space to land area) and χ 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

$$\mathbb{E}[\ell_i]H_{Ri} = (1-\beta)\frac{\mathbb{E}[w_s|i]H_{Ri}}{Q_i} = (1-\theta_i)L_i. \tag{S.29}$$

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

$$\left(\frac{(1-\alpha)A_j}{q_j}\right)^{1/\alpha} H_{Mj} = \theta_j L_j. \tag{S.30}$$

When both residential and commercial land market clearing ((S.29) and (S.30), respectively) are satisfied, total demand for floor space equals the total supply of floor space:

$$(1-\theta_i)L_i + \theta_i L_i = L_i = \varphi_i K_i^{1-\mu}. \tag{S.31}$$

S.2.6. *Properties of General Equilibrium With Exogenous Location Characteristics*

In this subsection, we characterize the properties of general equilibrium with *exogenous* location characteristics. In the next subsection, we relax these assumptions to allow for endogenous agglomeration forces.

We start with a benchmark case in which all locations have strictly positive, finite, and exogenous location characteristics. In this benchmark case, we show that all locations are incompletely specialized with positive values of both workplace and residence employment and positive shares of land allocated to commercial and residential use. We prove the existence of a unique general equilibrium for this benchmark case of incomplete specialization.

We next allow some blocks to have zero workplace and/or residence employment, as observed empirically. We retain the assumption that location characteristics are exogenous. But we extend the analysis to allow for zero final goods productivity and/or zero residential amenities. We show that a necessary and sufficient condition for zero workplace employment in a block is zero final goods productivity in that block. Similarly, a necessary and sufficient condition for zero residence employment in a block is zero residential amenities in that block. We extend our proof of the existence of a unique general equilibrium to allow for these empirically relevant cases. Therefore, with exogenous location characteristics, the model has a unique general equilibrium.

Definition of Equilibrium: We now formally define the general equilibrium of the model. Throughout the following, we use bold math font to denote vectors or matrices.

DEFINITION S.1: Given the model's parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$, the reservation level of utility in the wider economy $\bar{U}$, and exogenous location-specific characteristics $\{\mathbf{T}, \mathbf{E}, \mathbf{A}, \mathbf{B}, \boldsymbol{\varphi}, \mathbf{K}, \boldsymbol{\xi}, \boldsymbol{\tau}\}$, the general equilibrium of the model is referenced by the vector $\{\boldsymbol{\pi}_{\mathbf{M}}, \boldsymbol{\pi}_{\mathbf{R}}, H, \mathbf{Q}, \mathbf{q}, \mathbf{w}, \boldsymbol{\theta}\}$.

The seven elements of the equilibrium vector are determined by the following system of seven equations:

$$\gamma\left[\sum_{r=1}^{S}\sum_{s=1}^{S} T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_s)^{\varepsilon}\right]^{1/\varepsilon} = \bar{U}, \tag{S.32}$$

$$\pi_{Ri} = \frac{\sum_{s=1}^{S} T_i E_s \left(d_{is} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_s)^{\varepsilon}}{\sum_{r=1}^{S}\sum_{s=1}^{S} T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_s)^{\varepsilon}}, \tag{S.33}$$

$$\pi_{Mi} = \frac{\sum_{r=1}^{S} T_r E_i \left(d_{ri} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_i)^{\varepsilon}}{\sum_{r=1}^{S}\sum_{s=1}^{S} T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_s)^{\varepsilon}}, \tag{S.34}$$

$$\theta_i L_i = \left(\frac{(1-\alpha) A_i}{q_i}\right)^{1/\alpha} H_{Mi}, \tag{S.35}$$

$$(1-\theta_i)L_i = (1-\beta)\left[\sum_{s=1}^{S} \frac{E_s(w_s/d_{is})^{\varepsilon}}{\sum_{r=1}^{S} E_r(w_r/d_{ir})^{\varepsilon}} w_s\right] \frac{H_{Ri}}{Q_i}, \tag{S.36}$$

$$q_i = (1-\alpha)\left(\frac{\alpha}{w_i}\right)^{\alpha/(1-\alpha)} A_i^{1/(1-\alpha)}, \tag{S.37}$$

$$\begin{aligned} &\theta_i = 1 \quad \text{if} \quad q_i > \xi_i Q_i, \\ &\theta_i \in [0, 1] \quad \text{if} \quad q_i = \xi_i Q_i, \\ &\theta_i = 0 \quad \text{if} \quad q_i < \xi_i Q_i, \end{aligned} \tag{S.38}$$

where recall $L_i = \varphi_i K_i^{1-\mu}$; (S.32) is population mobility with the wider economy; (S.33) corresponds to the residential choice probabilities; (S.34) corresponds to the workplace choice probabilities; (S.35) is commercial land market clearing; (S.36) is residential land market clearing; (S.37) corresponds to profit maximization and zero profits; (S.38) corresponds to no-arbitrage between alternative uses of land.

Strictly Positive and Finite Exogenous Location Characteristics: We begin by considering a benchmark case, in which all blocks have strictly positive, finite, and exogenous location characteristics $\{\mathbf{T}, \mathbf{E}, \mathbf{A}, \mathbf{B}, \boldsymbol{\varphi}, \mathbf{K}, \boldsymbol{\xi}, \boldsymbol{\tau}\}$. We allow some blocks to be more attractive than others in terms of these characteristics. But workers draw idiosyncratic preferences from a Fréchet distribution for pairs of residence and workplace locations. Therefore, since the support of the Fréchet distribution is unbounded from above, any block with strictly positive characteristics has a positive measure of workers that prefer that location as a residence or workplace at a positive and finite price. Hence, all blocks with finite positive wages attract a positive measure of workers, and all blocks with finite positive floor prices attract a positive measure of residents.

LEMMA S.1: *Assuming strictly positive*, *finite*, *and exogenous location characteristics* ($T_i \in (0,\infty)$, $E_i \in (0,\infty)$, $A_i \in (0,\infty)$, $B_i \in (0,\infty)$, $\varphi_i \in (0,\infty)$, $K_i \in (0,\infty)$, $\xi_i \in (0,\infty)$, $\tau_{ij} \in (0,\infty) \times (0,\infty)$), *each location i with a strictly positive and finite wage* ($w_i \in (0,\infty)$) *attracts a strictly positive measure of workers* ($H_{Mi} \in (0,\infty)$), *and each location with a strictly positive and finite floor price* ($Q_i \in (0,\infty)$) *attracts a strictly positive measure of residents* ($H_{Ri} \in (0,\infty)$).

PROOF: Both properties follow immediately from the support of the Fréchet distribution being unbounded from above. For a strictly positive and finite wage ($w_i \in (0,\infty)$) for location i, there is a positive measure of workers who draw a large enough value of the idiosyncratic shock z_{ri} for each residence location r that their preferred workplace is i. Hence, from (S.18), the conditional probabilities of commuting from each residence location r to workplace i are strictly positive for $w_i \in (0,\infty)$. Additionally, for a strictly positive and finite floor price ($Q_i \in (0, \infty)$) for location i, there is a positive measure of workers who draw a large enough value of the idiosyncratic shock z_{is} that their preferred residence is i for each workplace s. Therefore, from (S.17), the conditional probabilities of commuting from i to each workplace s are strictly positive for $Q_i \in (0, \infty)$. *Q.E.D.*

We next show that blocks with strictly positive, finite, and exogenous location characteristics $\{\mathbf{T}, \mathbf{E}, \mathbf{A}, \mathbf{B}, \boldsymbol{\varphi}, \mathbf{K}, \boldsymbol{\xi}, \boldsymbol{\tau}\}$ must have strictly positive and finite values of both wages and floor prices in equilibrium. The reason is that the utility and production function satisfy the Inada conditions. Therefore, given a positive measure of workers, the return to commercial land use becomes large as the fraction of land allocated to commercial use becomes small. Similarly, given a positive measure of residents, the return to residential land use becomes large as the fraction of land allocated to residential use becomes small. Since locations attract positive measures of workers and residents at any finite positive wage and floor price, it follows that positive fractions of land must be allocated to both commercial and residential use.

LEMMA S.2: *Assuming strictly positive, finite, and exogenous location characteristics* ($T_i \in (0, \infty)$, $E_i \in (0, \infty)$, $A_i \in (0, \infty)$, $B_i \in (0, \infty)$, $\varphi_i \in (0, \infty)$, $K_i \in (0, \infty)$, $\xi_i \in (0, \infty)$, $\tau_{ij} \in (0, \infty) \times (0, \infty)$), *all locations are incompletely specialized and allocate positive fractions of land to commercial and residential use*: $\theta_i \in (0, 1)$.

PROOF: This property follows from the support of the Fréchet distribution being unbounded from above and from the utility and production functions both satisfying the Inada conditions. Lemma S.1 implies that each location with strictly positive and finite wages ($w_i \in (0, \infty)$) attracts a strictly positive measure of workers ($H_{Mi} \in (0, \infty)$). But profit maximization and commercial land market clearing imply

$$q_i = (1 - \alpha) A_i \left(\frac{H_{Mi}}{\theta_i L_i} \right)^{\alpha},$$

which in turn implies (i) $\lim_{\theta_i \to 0} q_i = \infty$ for $A_i \in (0, \infty)$ and $H_{Mi} \in (0, \infty)$; (ii) $q_i \in (0, \infty)$ for all $\theta_i \in (0, 1]$, $A_i \in (0, \infty)$, and $H_{Mi} \in (0, \infty)$. Therefore, a positive fraction of land must be allocated to commercial use: $\theta_i > 0$. Additionally, Lemma S.1 implies that each location with strictly positive and finite values of both amenities ($B_i \in (0, \infty)$) and floor prices ($Q_i \in (0, \infty)$) attracts a strictly positive measure of residents ($H_{Ri} \in (0, \infty)$). But utility maximization and residential land market clearing imply

$$Q_i = (1 - \beta) \left[\sum_{s=1}^{S} \frac{E_s (w_s / d_{is})^{\varepsilon}}{\sum_{r=1}^{S} E_r (w_r / d_{ir})^{\varepsilon}} w_s \right] \frac{H_{Ri}}{(1 - \theta_i) L_i},$$

which in turn implies (i) $\lim_{(1-\theta_i)\to 0} Q_i = \infty$ for $H_{Ri} \in (0, \infty)$; (ii) $Q_i \in (0, \infty)$ for all $(1-\theta_i) \in (0, 1]$ and $H_{Ri} \in (0, \infty)$. Therefore, a positive fraction of land must be allocated to residential use: $(1-\theta_i) > 0$. *Q.E.D.*

Having shown that the assumption of strictly positive, finite, and exogenous location characteristics implies incomplete specialization, we are now in a position to establish the following proposition.

PROPOSITION S.1: *Assuming strictly positive, finite, and exogenous location characteristics* ($T_i \in (0, \infty)$, $E_i \in (0, \infty)$, $A_i \in (0, \infty)$, $B_i \in (0, \infty)$, $\varphi_i \in (0, \infty)$, $K_i \in (0, \infty)$, $\xi_i \in (0, \infty)$, $\tau_{ij} \in (0, \infty) \times (0, \infty)$), *there exists a unique general equilibrium vector* $\{\boldsymbol{\pi}_{\mathbf{M}}, \boldsymbol{\pi}_{\mathbf{R}}, H, \mathbf{Q}, \mathbf{q}, \mathbf{w}, \boldsymbol{\theta}\}$.

PROOF: With strictly positive, finite, and exogenous location characteristics ($T_i \in (0, \infty)$, $E_i \in (0, \infty)$, $A_i \in (0, \infty)$, $B_i \in (0, \infty)$, $\varphi_i \in (0, \infty)$, $K_i \in (0, \infty)$, $\xi_i \in (0, \infty)$, $\tau_{ij} \in (0, \infty) \times (0, \infty)$), locations are incompletely specialized and the no-arbitrage condition between alternative uses of land (S.25) holds, which implies that commercial floor prices can be expressed in terms of residential floor prices: $q_i = \xi_i Q_i$. Using this result together with the probability of residing in a location (S.15), the probability of working in a location (S.16), the zero-profit condition (S.24), and the indifference condition between the city and the larger economy (S.12), the fraction of the city's population residing in location i can be written as

$$
\begin{aligned}
\pi_{Ri} &= \frac{H_{Ri}}{H} \\
&= \left(\frac{\gamma}{\bar{U}}\right)^{\varepsilon} \sum_{s=1}^{S} T_i E_s \left(d_{is} Q_i^{1-\beta}\right)^{-\varepsilon} \left(B_i (1-\alpha)^{(1-\alpha)/\alpha} \alpha A_s^{1/\alpha}\right)^{\varepsilon} \\
&\quad \times (\xi_s Q_s)^{-(\varepsilon(1-\alpha))/\alpha},
\end{aligned}
$$

while the fraction of the city's population working in location i can be written as

$$
\begin{aligned}
\pi_{Mi} &= \frac{H_{Mi}}{H} \\
&= \left(\frac{\gamma}{\bar{U}}\right)^{\varepsilon} \sum_{s=1}^{S} T_s E_i \left(d_{si} Q_s^{1-\beta}\right)^{-\varepsilon} \left(B_s (1-\alpha)^{(1-\alpha)/\alpha} \alpha A_i^{1/\alpha}\right)^{\varepsilon} \\
&\quad \times (\xi_i Q_i)^{-(\varepsilon(1-\alpha))/\alpha},
\end{aligned}
$$

and expected worker income conditional on residing in block i (S.20) can be written as

$$\mathbb{E}[w_s|i] = \sum_{s=1}^{S} \frac{E_s\big(A_s^{1/\alpha}(\xi_s Q_s)^{-(1-\alpha)/\alpha}/d_{is}\big)^{\varepsilon}}{\sum_{r=1}^{S} E_r\big(A_r^{1/\alpha}(\xi_r Q_r)^{-(1-\alpha)/\alpha}/d_{ir}\big)^{\varepsilon}} \times \big[(1-\alpha)^{(1-\alpha)/\alpha}\alpha A_s^{1/\alpha}(\xi_s Q_s)^{-(1-\alpha)/\alpha}\big].$$

Using commercial land market clearing (S.30) and residential land market clearing (S.29), the requirement that the land market clears can be written as

$$\left(\frac{(1-\alpha)A_i}{\xi_i Q_i}\right)^{1/\alpha} \pi_{Mi} + (1-\beta)\frac{\mathbb{E}[w_s|i]}{Q_i}\pi_{Ri} = \frac{L_i}{H}.$$

Combining the above relationships, the land market clearing condition can be written as

$$\begin{aligned} D_i(\mathbf{Q}) = {} & \left(\frac{(1-\alpha)A_i}{\xi_i Q_i}\right)^{1/\alpha} \sum_{s=1}^{S} T_s E_i \left(\frac{B_s(1-\alpha)^{(1-\alpha)/\alpha}\alpha A_i^{1/\alpha}}{d_{si} Q_s^{1-\beta}(\xi_i Q_i)^{(1-\alpha)/\alpha}}\right)^{\varepsilon} \\ & + \frac{(1-\beta)}{Q_i} \sum_{s=1}^{S} \frac{E_s\big(A_s^{1/\alpha}(\xi_s Q_s)^{-(1-\alpha)/\alpha}/d_{is}\big)^{\varepsilon}}{\sum_{r=1}^{S} E_r\big(A_r^{1/\alpha}(\xi_r Q_r)^{-(1-\alpha)/\alpha}/d_{ir}\big)^{\varepsilon}} \\ & \times \left[\left(\frac{1-\alpha}{\xi_s Q_s}\right)^{(1-\alpha)/\alpha} \alpha A_s^{1/\alpha}\right] \\ & \times \sum_{s=1}^{S} T_i E_s \left(\frac{B_i(1-\alpha)^{(1-\alpha)/\alpha}\alpha A_s^{1/\alpha}}{d_{is} Q_i^{1-\beta}(\xi_s Q_s)^{(1-\alpha)/\alpha}}\right)^{\varepsilon} \\ = {} & L_i, \end{aligned}$$

for all i, where we have chosen units in which to measure utility so that $(\bar{U}/\gamma)^{\varepsilon}/H = 1$. The above land market clearing condition provides a system of S equations in the S unknown residential floor prices Q_i for each location i, which has the following properties:

$$\lim_{Q_i \to 0} D_i(\mathbf{Q}) = \infty > L_i, \qquad \lim_{Q_i \to \infty} D_i(\mathbf{Q}) = 0 < L_i,$$

$$\frac{dD_i(\mathbf{Q})}{dQ_i} < 0, \qquad \frac{dD_i(\mathbf{Q})}{dQ_j} < 0, \qquad \left|\frac{dD_i(\mathbf{Q})}{dQ_i}\right| > \left|\frac{dD_i(\mathbf{Q})}{dQ_j}\right|.$$

It follows that there exists a unique vector of residential floor prices $\mathbf{Q}$ that solves this system of land market clearing conditions. Commercial floor prices follow immediately from $\mathbf{q} = \boldsymbol{\xi}\mathbf{Q}$. Having solved for the vectors of floor prices $\{\mathbf{Q}, \mathbf{q}\}$, the vector of wages $\mathbf{w}$ follows immediately from the zero-profit condition for production (S.24). Given floor prices $\{\mathbf{Q}, \mathbf{q}\}$ and wages $(\mathbf{w})$, the probability of residing in a location $(\boldsymbol{\pi}_{\mathbf{R}})$ follows immediately from (S.15), and the probability of working in a location $(\boldsymbol{\pi}_{\mathbf{M}})$ follows immediately from (S.16). Having solved for $\{\boldsymbol{\pi}_{\mathbf{M}}, \boldsymbol{\pi}_{\mathbf{R}}, \mathbf{Q}, \mathbf{q}, \mathbf{w}\}$, the total measure of workers residing in the city can be recovered from our choice of units in which to measure utility $((\bar{U}/\gamma)^{\varepsilon}/H = 1)$, which, together with population mobility (S.12), implies

$$H = \left[\sum_{r=1}^{S}\sum_{s=1}^{S} T_r E_s \left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon} (B_r w_s)^{\varepsilon}\right].$$

We therefore obtain $\mathbf{H}_{\mathbf{M}} = \boldsymbol{\pi}_{\mathbf{M}} H$ and $\mathbf{H}_{\mathbf{R}} = \boldsymbol{\pi}_{\mathbf{R}} H$. Given floor prices $\{\mathbf{Q}, \mathbf{q}\}$ and employments $\{\mathbf{H}_{\mathbf{M}}, \mathbf{H}_{\mathbf{R}}\}$, the fraction of land that is used commercially $(\boldsymbol{\theta})$ follows immediately from commercial and residential land market clearing. This completes the determination of the equilibrium vector $\{\boldsymbol{\pi}_{\mathbf{M}}, \boldsymbol{\pi}_{\mathbf{R}}, H, \mathbf{Q}, \mathbf{q}, \mathbf{w}, \boldsymbol{\theta}\}$. *Q.E.D.*

Allowing for Zero Workplace and/or Residence Employment: A corollary of Lemmas S.1 and S.2 is that a necessary and sufficient condition for zero workplace employment and zero commercial land use in a block is zero final goods productivity. Similarly, a necessary and sufficient condition for zero residence employment and zero residential land use in a block is zero amenities.

LEMMA S.3: *Assuming strictly positive, finite, and exogenous location characteristics* ($T_i \in (0, \infty)$, $E_i \in (0, \infty)$, $\varphi_i \in (0, \infty)$, $K_i \in (0, \infty)$, $\xi_i \in (0, \infty)$, $\tau_{ij} \in (0, \infty) \times (0, \infty)$):

(i) *a necessary and sufficient condition for zero workplace employment* ($H_{Mi} = 0$) *and zero commercial land use* ($\theta_i = 0$) *is zero final goods productivity* ($A_i = 0$) *for location* i,

(ii) *a necessary and sufficient condition for zero residence employment* ($H_{Ri} = 0$) *and zero residential land use* ($(1 - \theta_i) = 0$) *is zero amenities* ($B_i = 0$) *for location* i.

PROOF: From Lemma S.1 and the conditional probability of commuting to location i conditional on living in each residence location r (S.18), a necessary and sufficient condition for $H_{Mi} = 0$ for workplace i is $w_i = 0$. From the first-order conditions for profit maximization (S.22) and (S.23), a necessary and sufficient condition for $w_i = 0$ and $\theta_i = 0$ is $A_i = 0$. From Lemma S.1 and the conditional probability of commuting from location i conditional on working in each workplace s (S.17), a necessary and sufficient condition for $H_{Ri} = 0$ is

$B_i = 0$. From residential land market clearing (S.29), a necessary and sufficient condition for $(1-\theta_i) = 0$ is $H_{Ri} = 0$, which is ensured by $B_i = 0$. *Q.E.D.*

From Lemma S.3, a necessary and sufficient condition for a block to have both no commercial activity and no residential activity is $A_i = 0$ and $B_i = 0$. Such blocks with no economic activity play no direct role in the model, but affect the general equilibrium insofar as they affect travel times (τ_{ij}) between blocks with positive economic activity. We now use the results from Lemma S.3 to generalize Proposition S.1 to prove that there exists a unique equilibrium given exogenous location characteristics once we allow blocks to have no commercial activity and/or no residential activity.

From (S.26), Lemmas S.1–S.3, and no-arbitrage between alternative uses of land, we can summarize the relationships between the observed price of floor space ($\mathbb{Q}_i$), the price of commercial floor space (q_i), the residential price of floor space (Q_i), and land use as

$$\text{(S.39)}\qquad \mathbb{Q}_i = \begin{cases} \zeta_{Mi} q_i, & \zeta_{Mi} = 1, \quad i \in \Im_M = \{A_i > 0, B_i = 0\}, \\ \zeta_{Mi} q_i, & \zeta_{Mi} = 1, \quad i \in \Im_S = \{A_i > 0, B_i > 0\}, \end{cases}$$

$$\mathbb{Q}_i = \begin{cases} \zeta_{Ri} Q_i, & \zeta_{Ri} = 1, \quad i \in \Im_R = \{A_i = 0, B_i > 0\}, \\ \zeta_{Ri} Q_i, & \zeta_{Ri} = \xi_i, \quad i \in \Im_S = \{A_i > 0, B_i > 0\}, \end{cases}$$

where ζ_{Mi} and ζ_{Ri} relate observed floor prices to commercial and residential floor prices, respectively; $\Im_M$ is the set of locations specialized in commercial activity ($\theta_i = 1$); $\Im_S$ is the set of locations with both commercial and residential activity ($\theta_i \in (0, 1)$); and $\Im_R$ is the set of locations specialized in residential activity ($\theta_i = 0$).

From (S.39), these relationships between the observed, commercial, and residential prices of floor space $\{\mathbb{Q}_i, q_i, Q_i\}$, and the allocation of land between commercial and residential use $\{\theta_i, 1-\theta_i\}$, are a function solely of the exogenous locational characteristics $\{\mathbf{A}, \mathbf{B}, \boldsymbol{\xi}\}$. We now use this property to generalize Proposition S.1 to allow blocks to have no commercial activity and/or residential activity.

PROPOSITION S.2: *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 nonnegative final goods productivity* $A_i \in [0, \infty)$ *and residential amenities* $B_i \in [0, \infty)$, *there exists a unique general equilibrium vector* $\{\boldsymbol{\pi}_\mathbf{M}, \boldsymbol{\pi}_\mathbf{R}, H, \mathbf{Q}, \mathbf{q}, \mathbf{w}, \boldsymbol{\theta}\}$.

PROOF: The proof follows a similar structure as for Proposition S.1. For locations that are completely specialized in commercial activity ($i \in \Im_M$), the land market clearing condition can be written

$$
\begin{aligned}
D_i(\mathbb{Q}) &= \left(\frac{(1-\alpha)A_i}{\mathbb{Q}_i}\right)^{1/\alpha} \sum_{s\in\Im_S\cup\Im_R} T_s E_i \left(\frac{B_s(1-\alpha)^{(1-\alpha)/\alpha}\alpha A_i^{1/\alpha}}{d_{si}(\mathbb{Q}_s/\zeta_{Rs})^{1-\beta}(\mathbb{Q}_i)^{(1-\alpha)/\alpha}}\right)^{\varepsilon} \\
&= L_i.
\end{aligned}
$$

For locations that are incompletely specialized in commercial and residential activity ($i \in \Im_S$), the land market clearing condition can be written

$$
\begin{aligned}
D_i(\mathbb{Q}) &= \left(\frac{(1-\alpha)A_i}{\mathbb{Q}_i}\right)^{1/\alpha} \sum_{s\in\Im_S\cup\Im_R} \left(\frac{T_s^{1/\varepsilon}E_i^{1/\varepsilon}B_s(1-\alpha)^{(1-\alpha)/\alpha}\alpha A_i^{1/\alpha}}{d_{si}(\mathbb{Q}_s/\zeta_{Rs})^{1-\beta}(\mathbb{Q}_i)^{(1-\alpha)/\alpha}}\right)^{\varepsilon} \\
&\quad + \frac{(1-\beta)}{\mathbb{Q}_i/\zeta_{Ri}} \\
&\quad \times \sum_{s\in\Im_M\cup\Im_S} \frac{E_s\left(A_s^{1/\alpha}(\mathbb{Q}_s)^{-(1-\alpha)/\alpha}/d_{is}\right)^{\varepsilon}\left[\left(\frac{(1-\alpha)}{\mathbb{Q}_s}\right)^{(1-\alpha)/\alpha}\alpha A_s^{1/\alpha}\right]}{\sum_{r\in\Im_M\cup\Im_S} E_r\left(A_r^{1/\alpha}(\mathbb{Q}_r)^{-(1-\alpha)/\alpha}/d_{ir}\right)^{\varepsilon}} \\
&\quad \times \sum_{s\in\Im_M\cup\Im_S} \left(\frac{T_i^{1/\varepsilon}E_s^{1/\varepsilon}B_i(1-\alpha)^{(1-\alpha)/\alpha}\alpha A_s^{1/\alpha}}{d_{is}(\mathbb{Q}_i/\zeta_{Ri})^{1-\beta}(\mathbb{Q}_s)^{(1-\alpha)/\alpha}}\right)^{\varepsilon} \\
&= L_i.
\end{aligned}
$$

For locations that are completely specialized in residential activity ($i \in \Im_R$), the land market clearing condition can be written

$$
\begin{aligned}
D_i(\mathbb{Q}) &= \frac{(1-\beta)}{\mathbb{Q}_i/\zeta_{Ri}} \\
&\quad \times \sum_{s\in\Im_M\cup\Im_S} \frac{E_s\left(A_s^{1/\alpha}(\mathbb{Q}_s)^{-(1-\alpha)/\alpha}/d_{is}\right)^{\varepsilon}\left[\left(\frac{(1-\alpha)}{\mathbb{Q}_s}\right)^{(1-\alpha)/\alpha}\alpha A_s^{1/\alpha}\right]}{\sum_{r=1}^{S} E_r\left(A_r^{1/\alpha}(\mathbb{Q}_r)^{-(1-\alpha)/\alpha}/d_{ir}\right)^{\varepsilon}} \\
&\quad \times \sum_{s\in\Im_M\cup\Im_S} \left(\frac{T_i^{1/\varepsilon}E_s^{1/\varepsilon}B_i(1-\alpha)^{(1-\alpha)/\alpha}\alpha A_s^{1/\alpha}}{d_{is}(\mathbb{Q}_i/\zeta_{Ri})^{1-\beta}(\mathbb{Q}_s)^{(1-\alpha)/\alpha}}\right)^{\varepsilon} \\
&= L_i.
\end{aligned}
$$

We have again chosen units in which to measure utility so that $(\bar{U}/\gamma)^{\varepsilon}/H = 1$. The above land market clearing conditions provide a system of S equations in the S unknown observed floor prices $\mathbb{Q}_i$ for each location i, which has the following properties:

$$\lim_{\mathbb{Q}_i \to 0} D_i(\mathbb{Q}) = \infty > L_i, \quad \lim_{\mathbb{Q}_i \to \infty} D_i(\mathbb{Q}) = 0 < L_i,$$
$$\frac{dD_i(\mathbb{Q})}{d\mathbb{Q}_i} < 0, \quad \frac{dD_i(\mathbb{Q})}{d\mathbb{Q}_j} < 0, \quad \left|\frac{dD_i(\mathbb{Q})}{d\mathbb{Q}_i}\right| > \left|\frac{dD_i(\mathbb{Q})}{d\mathbb{Q}_j}\right|.$$

It follows that there exists a unique vector of observed floor prices $\mathbb{Q}$ that solves this system of land market clearing conditions. Having determined $\mathbb{Q}$, commercial floor prices ($\mathbf{q}$) and residential floor prices ($\mathbf{Q}$) follow immediately from the relationship between floor prices (S.39) as a function of the exogenous location characteristics $\{\mathbf{A}, \mathbf{B}, \boldsymbol{\xi}\}$. The remainder of the equilibrium vector follows from exactly the same arguments as for Proposition S.1. *Q.E.D.*

We use Proposition S.2 to undertake counterfactuals for division and reunification, in which we treat location characteristics as exogenous and hold them constant at their values before division or reunification. Since the model features a unique equilibrium with exogenous location characteristics, these counterfactuals yield determinate predictions for the impact of division and reunification on the organization of economic activity within the city.

S.2.7. *Properties of General Equilibrium With Agglomeration Forces*

We now relax the assumption that productivity (A_i) and amenities (B_i) are exogenous. We examine how the introduction of endogenous agglomeration forces affects the properties of the general equilibrium of the model. We decompose productivity (A_i) and amenities (B_i) into two components, one of which is exogenous and captures location fundamentals, and the other of which is endogenous to the surrounding concentration of economic activity and captures agglomeration forces.

Agglomeration Forces: We allow final goods productivity to depend on production fundamentals (a_j) and production externalities (Y_j). Production fundamentals capture features of physical geography that make a location more or less productive independently of the surrounding density of economic activity (e.g., 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:[7]

$$A_j = a_j Y_j^{\lambda}, \quad Y_j \equiv \sum_{s=1}^{S} e^{-\delta\tau_{js}} \left(\frac{H_{Ms}}{K_s}\right), \tag{S.40}$$

where H_{Ms}/K_s is workplace employment density per unit of geographical land area; production externalities decline with travel time (τ_{js}) through the iceberg factor $e^{-\delta\tau_{js}} \in (0, 1]$; δ determines their rate of spatial decay; and λ controls their relative importance in determining overall productivity.[8]

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 (Ω_i). Residential fundamentals capture features of physical geography that make a location a more or less attractive place to live independently of the surrounding density of economic activity (e.g., 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:

$$B_i = b_i \Omega_i^{\eta}, \quad \Omega_i \equiv \sum_{r=1}^{S} e^{-\rho\tau_{ir}} \left(\frac{H_{Rr}}{K_r}\right), \tag{S.41}$$

where H_{Rr}/K_r is residence employment density per unit of geographical land area; residential externalities decline with travel time (τ_{ir}) through the iceberg factor $e^{-\rho\tau_{ir}} \in (0, 1]$; ρ determines their rate of spatial decay; and η controls their relative importance in overall residential amenities.

Equilibrium Properties With Agglomeration Forces: We begin by establishing some properties of the general equilibrium of the model with agglomeration forces. Production externalities (Y_j) are modeled as the travel time weighted sum of workplace employment density throughout the city. Therefore, since travel time within Berlin is finite, production externalities are strictly positive for all blocks for a finite spatial decay of production externalities (δ), as long as workplace employment is positive somewhere within Berlin: $Y_j > 0$ for all

[7]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).

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

$j \in \{1, \dots, S\}$ if $H_{Ms} > 0$ for some $s \in \{1, \dots, S\}$ and $0 < \delta < \infty$. Similarly, residential externalities (Ω_i) are modeled as the travel time weighted sum of residence employment density throughout the city. Therefore, since travel time within Berlin is finite, residential externalities are strictly positive for all blocks for a finite spatial decay of residential externalities (ρ), as long as residence employment is positive somewhere within Berlin: $\Omega_i > 0$ for all $i \in \{1, \dots, S\}$ if $H_{Rs} > 0$ for some $s \in \{1, \dots, S\}$ and $0 < \rho < \infty$.

We now combine this result that production and residential externalities are strictly positive and finite with the properties that (i) the support of the Fréchet distribution is unbounded from above and (ii) the utility and production functions satisfy the Inada conditions. From these results, if all location characteristics are strictly positive for all blocks, it follows that all blocks will be incompletely specialized with positive fractions of land allocated to commercial and residential use. We therefore have the following generalization of Lemmas S.1 and S.2 to the case of endogenous agglomeration forces.

LEMMA S.4: *Assume* (i) *strictly positive*, *finite*, *and exogenous location fundamentals* ($T_i \in (0, \infty)$, $E_i \in (0, \infty)$, $a_i \in (0, \infty)$, $b_i \in (0, \infty)$, $\varphi_i \in (0, \infty)$, $K_i \in (0, \infty)$, $\xi_i \in (0, \infty)$, $\tau_{ij} \in (0, \infty) \times (0, \infty)$), (ii) *endogenous agglomeration forces* ($\lambda, \eta > 0$), (iii) *finite spatial decays of agglomeration externalities* ($0 < \delta < \infty$ *and* $0 < \rho < \infty$):

(i) *Each location i with a strictly positive and finite wage* ($w_i \in (0, \infty)$) *attracts a strictly positive measure of workers* ($H_{Mi} \in (0, \infty)$), *and each location with a strictly positive and finite floor price* ($Q_i \in (0, \infty)$) *attracts a strictly positive measure of residents* ($H_{Ri} \in (0, \infty)$).

(ii) *Any equilibrium with positive workplace and residence employment somewhere in the city* ($H_{Mj}, H_{Ri} > 0$ *for some* $j, i \in \{1, \dots, S\}$) *is characterized by incomplete specialization*, *with all locations allocating positive fractions of land to commercial and residential use*: $\theta_i \in (0, 1)$.

PROOF: The proof of the lemma follows exactly the same structure as the proof of Lemmas S.1 and S.2 above. Since the support of the Fréchet distribution is unbounded from above, any location with strictly positive and finite wages ($w_i \in (0, \infty)$) attracts a strictly positive measure of workers ($H_{Mi} \in (0, \infty)$). But profit maximization and commercial land market clearing imply

$$q_i = (1 - \alpha) a_i Y_i^{\lambda} \left(\frac{H_{Mi}}{\theta_i L_i} \right)^{\alpha},$$

which in turn implies (i) $\lim_{\theta_i \to 0} q_i = \infty$ for $a_i \in (0, \infty)$, $Y_i \in (0, \infty)$, and $H_{Mi} \in (0, \infty)$; (ii) $q_i \in (0, \infty)$ for all $\theta_i \in (0, 1]$, $a_i \in (0, \infty)$, $Y_i \in (0, \infty)$, and $H_{Mi} \in (0, \infty)$. Therefore, a positive fraction of land is allocated to commercial use: $\theta_i > 0$. Additionally, each location with strictly positive and finite values of residential fundamentals ($b_i \in (0, \infty)$), residential externalities ($\Omega_i \in (0, \infty)$), and floor prices ($Q_i \in (0, \infty)$) attracts a strictly positive measure of residents ($H_{Ri} \in (0, \infty)$). But utility maximization and residential land market clearing imply

$$Q_i = (1-\beta)\left[\sum_{s=1}^{S} \frac{E_s(w_s/d_{is})^{\varepsilon}}{\sum_{r=1}^{S} E_r(w_r/d_{ir})^{\varepsilon}} w_s\right] \frac{H_{Ri}}{(1-\theta_i)L_i},$$

which in turn implies (i) $\lim_{(1-\theta_i)\to 0} Q_i = \infty$ for $H_{Ri} \in (0, \infty)$; (ii) $Q_i \in (0, \infty)$ for all $(1-\theta_i) \in (0, 1]$ and $H_{Ri} \in (0, \infty)$. Therefore, a positive fraction of land is allocated to residential use: $(1-\theta_i) > 0$. *Q.E.D.*

Since production externalities are strictly positive ($Y_i > 0$), an immediate corollary of Lemma S.4 is that a necessary and sufficient condition for zero workplace employment and zero commercial land use in a block is zero production fundamentals ($a_i = 0$). Similarly, since residential externalities are strictly positive ($\Omega_i > 0$), an immediate corollary of Lemma S.4 is that a necessary and sufficient condition for zero residence employment and zero residential land use in a block is zero residential fundamentals ($b_i = 0$).

LEMMA S.5: *Assume* (i) *strictly positive, finite, and exogenous location characteristics* ($T_i \in (0, \infty)$, $E_i \in (0, \infty)$, $\varphi_i \in (0, \infty)$, $K_i \in (0, \infty)$, $\xi_i \in (0, \infty)$, $\tau_{ij} \in (0, \infty) \times (0, \infty)$), (ii) *endogenous agglomeration forces* ($\lambda, \eta > 0$), (iii) *finite spatial decays of agglomeration externalities* ($0 < \delta < \infty$ *and* $0 < \rho < \infty$). *In any equilibrium with positive workplace and residence employment somewhere in the city* ($H_{Mj}, H_{Ri} > 0$ *for some* $j, i \in \{1, \dots, S\}$):

(i) *a necessary and sufficient condition for zero workplace employment* ($H_{Mi} = 0$) *and zero commercial land use* ($\theta_i = 0$) *is zero production fundamentals* ($a_i = 0$) *for location* i,

(ii) *a necessary and sufficient condition for zero residence employment* ($H_{Ri} = 0$) *and zero residential land use* ($(1-\theta_i) = 0$) *is zero residential fundamentals* ($b_i = 0$) *for location* i.

PROOF: From Lemma S.4 and the conditional probability of commuting to location i conditional on living in each residence location r (S.18), a necessary and sufficient condition for $H_{Mi} = 0$ for workplace i is $w_i = 0$. From the first-order conditions for profit maximization (S.22) and (S.23), a necessary and sufficient condition for $w_i = 0$ and $\theta_i = 0$ is $A_i = 0$. From the productivity specification (S.40), a necessary and sufficient condition for $A_i = 0$ is $a_i = 0$ since $Y_i > 0$. From Lemma S.4 and the conditional probability of commuting from location i conditional on working in each workplace s (S.17), a necessary and sufficient condition for $H_{Ri} = 0$ is $B_i = 0$. From residential land market clearing (S.29), a necessary and sufficient condition for $(1-\theta_i)=0$ is $H_{Ri}=0$, which is ensured by $B_i=0$. From the amenities specification (S.41), a necessary and sufficient condition for $B_i=0$ is $b_i=0$ since $\Omega_i>0$. *Q.E.D.*

Potential for Multiple Equilibria: As is standard in urban models, the presence of endogenous agglomeration forces introduces the potential for multiple equilibria into the model: Each agent's location decision depends on productivity and amenities, but productivity and amenities in turn depend on the location decisions of all agents. Whether or not multiple equilibria exist depends on the strength of these agglomeration forces relative to the size of the exogenous differences in location characteristics ($\mathbf{T}$, $\mathbf{E}$, $\mathbf{a}$, $\mathbf{b}$, $\boldsymbol{\varphi}$, $\mathbf{K}$, $\boldsymbol{\xi}$, $\boldsymbol{\tau}$). The strength of agglomeration forces depends on both their contribution to productivity and amenities (λ, η) and their spatial decay with travel times (δ, ρ). An important feature of our empirical approach is that it explicitly addresses the potential for multiple equilibria, as discussed further in Sections S.3 and S.4 of this supplement (see, in particular, Propositions S.3 and S.4 and Section S.4.5).

S.3. CALIBRATION

We now show that there is a unique mapping from the observed variables to unobserved values of 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:

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

where we use i to index all blocks; the function $\tilde{\varphi}_i(\cdot)$ is defined below; $\zeta_{Ri}=1$ 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 ($\tilde{w}_i$) captures the wage (w_i) and the Fréchet scale parameter for each employment location ($E_i^{1/\varepsilon}$), because these both affect the relative attractiveness of an employment location to workers. On the production side, adjusted productivity ($\tilde{A}_i$) captures productivity (A_i) and the Fréchet scale parameter for each employment location ($E_i^{\alpha/\varepsilon}$), because these both affect the adjusted wage consistent with zero profits. Adjusted production fundamentals are defined analogously. On the consumption side, adjusted amenities ($\tilde{B}_i$) capture amenities (B_i), the Fréchet scale parameter for each residence location ($T_i^{1/\varepsilon}$), and the wedge between commercial and residential floor prices (ξ_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 (φ_i) and other production and residential parameters that affect the equality of the demand and supply for floor space, as shown below.

In the remainder of this section, we show how the model can be calibrated to recover unique adjusted location characteristics given known values of the model's parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$ and the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$. We show that there is a one-to-one mapping from these known parameters and the observed data to the adjusted location characteristics $\{\tilde{a}_i, \tilde{b}_i, \tilde{\varphi}_i\}$. Therefore, these unobserved adjusted location characteristics correspond to *structural residuals* of the model that are one-to-one functions of the parameters and the observed data. We use the resulting closed-form solutions for these structural residuals to construct moment conditions in the structural estimation of the model in Section S.4, where both the parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$ and the unobserved adjusted location characteristics $\{\tilde{a}_i, \tilde{b}_i, \tilde{\varphi}_i\}$ are unknown and to be estimated.

In addition to establishing the one-to-one mapping from the parameters and observables to the unobservables, we show that the model has a recursive structure. Given a subset of the model's parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$, there is a one-to-one mapping from these parameters and the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$ to the unobserved adjusted location characteristics $\{\tilde{A}_i, \tilde{B}_i, \tilde{\varphi}_i\}$. Therefore, overall adjusted productivity ($\tilde{A}_i$), overall adjusted amenities ($\tilde{B}_i$), and the adjusted density of development ($\tilde{\varphi}_i$) can be uniquely determined irrespective of whether they are exogenous or endogenous. Furthermore, overall adjusted productivity ($\tilde{A}_i$) and amenities ($\tilde{B}_i$) can be determined irrespective of the relative importance of their components of externalities $\{Y_i, \Omega_i\}$ and adjusted fundamentals $\{\tilde{a}_i, \tilde{b}_i\}$.

S.3.1. *Determining* $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\boldsymbol{\varphi}}\}$ *From* $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ *and the Observed Data*

We begin by establishing the one-to-one mapping from the subset of the parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ and the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$ to adjusted final goods productivity, residential amenities, and the density of development $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\boldsymbol{\varphi}}\}$. To do so, we use the recursive structure of the model:

1. Given $\{\varepsilon, \kappa\}$ and the observed data $\{\mathbf{H_R}, \mathbf{H_M}, \boldsymbol{\tau}\}$, the equilibrium adjusted wage vector $\{\tilde{\mathbf{w}}\}$ can be uniquely determined from the commuting market clearing condition alone independently of the other equilibrium conditions of the model.

2. Given $\{\varepsilon, \kappa, \beta, \mu\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \boldsymbol{\tau}\}$, and adjusted wages $\{\tilde{\mathbf{w}}\}$, adjusted residential amenities $\{\tilde{\mathbf{B}}\}$ can be uniquely determined from residential choice probabilities.

3. Given $\{\varepsilon, \kappa, \alpha, \mu\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \boldsymbol{\tau}\}$, and adjusted wages $\{\tilde{\mathbf{w}}\}$, adjusted final goods productivity $\{\tilde{\mathbf{A}}\}$ can be uniquely determined from profit maximization and zero profits.

4. Given $\{\varepsilon, \kappa, \alpha, \beta, \mu\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$, and adjusted wages and productivity $\{\tilde{\mathbf{w}}, \tilde{\mathbf{A}}\}$, the adjusted density of development $\{\tilde{\boldsymbol{\varphi}}\}$ can be uniquely determined from land market clearing.

In the remainder of this subsection, we consider each of these steps in turn.

S.3.1.1. *Wages*

Given the parameters $\{\varepsilon, \kappa\}$ and observed data $\{\mathbf{H_M}, \mathbf{H_R}, \boldsymbol{\tau}\}$, commuting market clearing (S.19) provides a system of equations in observed workplace and residence employment that determines a unique adjusted wage vector $(\tilde{\mathbf{w}})$ up to a normalization (a choice of units in which to measure wages):

$$(S.42)\qquad H_{Mj} = \sum_{i=1}^{S} \frac{\left(E_j^{1/\varepsilon} w_j / e^{\kappa\tau_{ij}}\right)^{\varepsilon}}{\sum_{s=1}^{S}\left(E_s^{1/\varepsilon} w_s / e^{\kappa\tau_{is}}\right)^{\varepsilon}} H_{Ri}$$

$$= \sum_{i=1}^{S} \frac{\left(\tilde{w}_j / e^{\kappa\tau_{ij}}\right)^{\varepsilon}}{\sum_{s=1}^{S}\left(\tilde{w}_s / e^{\kappa\tau_{is}}\right)^{\varepsilon}} H_{Ri}.$$

Adjusted wages $(\tilde{w}_j = E_j^{1/\varepsilon} w_j)$ capture (i) wages (w_j) in employment location j and (ii) the Fréchet scale parameter that determines the average utility (or effective units of labor) for commuters to that employment location (E_j). Note that $E_j^{1/\varepsilon}$ enters the commuting market clearing condition isomorphically to w_j. Therefore, only the composite adjusted wage $(\tilde{w}_j)$ can be recovered from the data. From Lemmas S.1–S.3, all locations with zero workplace employment have zero adjusted wages.

We now show that this commuting market clearing condition determines a unique adjusted wage $(\tilde{w}_j)$ for each location $j = 1, \dots, S$. Note that the commuting market clearing condition (S.42) can be rewritten as the following excess demand system:

$$(S.43)\qquad D_j(\tilde{\mathbf{w}}) = H_{Mj} - \sum_{i=1}^{S} \frac{(\tilde{w}_j / d_{ij})^{\varepsilon}}{\sum_{s=1}^{S}(\tilde{w}_s / d_{is})^{\varepsilon}} H_{Ri} = 0, \quad d_{ij} = e^{\kappa\tau_{ij}}.$$

We use $\mathbf{H}_M \in \Re_+^S$ to denote the observed nonnegative vector of workplace employment with elements H_{Mj} given by the data; $\mathbf{H}_R \in \Re_+^S$ denotes the observed nonnegative vector of residence employment with elements H_{Ri}, given by the data; $\tau_{ij} \in \Re_+^S \times \Re_+^S$ denotes the observed bilateral travel time between blocks i and j; and $\tilde{\mathbf{w}} \in \Re_+^S$ is the unknown nonnegative adjusted wage vector with elements $\tilde{w}_j$. The system (S.43) captures the requirement of zero excess demand for commuters at the adjusted wage vector $\tilde{\mathbf{w}}$.

LEMMA S.6: *Given the parameters* $\{\varepsilon, \kappa\}$ *and observables* $\{\mathbf{H}_M, \mathbf{H}_R, \boldsymbol{\tau}\}$, *the wage system* (S.43) *exhibits the following properties in* $\mathbf{w}$:
Property (i): $D(\tilde{\mathbf{w}})$ *is continuous*.
Property (ii): $D(\tilde{\mathbf{w}})$ *is homogeneous of degree zero*.
Property (iii): $\sum_{j\in S} D_j(\tilde{\mathbf{w}}) = 0$ *for all* $\tilde{\mathbf{w}} \in \Re_+^S$.
Property (iv): $D(\tilde{\mathbf{w}})$ *exhibits gross substitution*:

$$\frac{\partial D_j(\tilde{\mathbf{w}})}{\partial \tilde{w}_k} > 0 \quad \textit{for all } j, k, j \neq k \quad \textit{for all } \tilde{\mathbf{w}} \in \Re_+^S,$$

$$\frac{\partial D_j(\tilde{\mathbf{w}})}{\partial \tilde{w}_j} < 0 \quad \textit{for all } j \quad \textit{for all } \tilde{\mathbf{w}} \in \Re_+^S.$$

PROOF: Property (i) follows immediately by inspection of (S.43). Property (ii) follows immediately by inspection of (S.43). Property (iii) can be established by noting

$$\begin{aligned}
\sum_{j=1}^{S} D_j(\tilde{\mathbf{w}}) &= \sum_{j=1}^{S} H_{Mj} - \sum_{i=1}^{S} \frac{\sum_{j=1}^{S} (\tilde{w}_j/d_{ij})^{\varepsilon}}{\sum_{s=1}^{S} (\tilde{w}_s/d_{is})^{\varepsilon}} H_{Ri} \\
&= \sum_{j=1}^{S} H_{Mj} - \sum_{i=1}^{S} H_{Ri} \\
&= 0.
\end{aligned}$$

Property (iv) can be established by noting

$$\frac{\partial D_j(\tilde{\mathbf{w}})}{\partial \tilde{w}_k} = \sum_{i\in S} \frac{(\tilde{w}_j/d_{ij})^{\varepsilon} \varepsilon (\tilde{w}_k/d_{ik})^{\varepsilon} \tilde{w}_k^{-1}}{\left[\sum_{s\in S} (\tilde{w}_s/d_{is})^{\varepsilon}\right]^2} H_{Ri} > 0,$$

and using homogeneity of degree zero, which implies

$$\nabla D(\tilde{\mathbf{w}})\tilde{\mathbf{w}} = 0,$$

and hence:

$$\frac{\partial D_j(\tilde{\mathbf{w}})}{\partial \tilde{w}_j} < 0 \quad \text{for all } j \quad \text{for all } \tilde{\mathbf{w}} \in \Re^S_+.$$

Therefore, we have established gross substitution. *Q.E.D.*

LEMMA S.7: *Given the parameters* $\{\varepsilon, \kappa\}$ *and observed data* $\{\mathbf{H}_M, \mathbf{H}_R, \boldsymbol{\tau}\}$, *there exists a unique adjusted wage vector* $\tilde{\mathbf{w}}^* \in \Re^S_+$ *such that* $D(\tilde{\mathbf{w}}^*) = 0$.

PROOF: We first show that the adjusted wage system $D(\tilde{\mathbf{w}})$ has at most one (normalized) solution using the properties established in Lemma S.6. Gross substitution implies that $D(\tilde{\mathbf{w}}) = D(\tilde{\mathbf{w}}')$ cannot occur whenever $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{w}}'$ are two adjusted wage vectors that are not collinear. By homogeneity of degree zero, we can assume $\tilde{\mathbf{w}}' \geq \tilde{\mathbf{w}}$ and $\tilde{w}_j = \tilde{w}'_j$ for some j. Now consider altering the adjusted wage vector $\tilde{\mathbf{w}}'$ to obtain the adjusted wage vector $\tilde{\mathbf{w}}$ in $S-1$ steps, lowering (or keeping unaltered) the adjusted wage of all the other $S-1$ locations $k \neq j$ one at a time. By gross substitution, the excess demand for labor in location j cannot decrease in any step, and because $\tilde{\mathbf{w}} \neq \tilde{\mathbf{w}}'$, it will actually increase in at least one step. Hence $D_j(\tilde{\mathbf{w}}) > D_j(\tilde{\mathbf{w}}')$ and we have a contradiction.

We next establish that there exists an adjusted wage vector $\tilde{\mathbf{w}}^* \in \Re^S_+$ such that $D(\tilde{\mathbf{w}}^*) = 0$. By homogeneity of degree zero, we can restrict our search for an equilibrium adjusted wage vector to the unit simplex $\Delta = \{\tilde{\mathbf{w}} \in \Re^S_+ : \sum_{j=1}^S \tilde{w}_j = 1\}$. Define on Δ the function $D^+(\cdot)$ by $D^+_j(\tilde{\mathbf{w}}) = \max\{D_j(\tilde{\mathbf{w}}), 0\}$. Note that $D^+(\cdot)$ is continuous. Denote $\alpha(\tilde{\mathbf{w}}) = \sum_{j=1}^S [\tilde{w}_j + D^+_j(\tilde{w})]$. We have $\alpha(\tilde{\mathbf{w}}) \geq 1$ for all $\tilde{\mathbf{w}}$.

Define a continuous function $f(\cdot)$ from the closed convex set Δ into itself by

$$f(\tilde{\mathbf{w}}) = \left[1/\alpha(\tilde{\mathbf{w}})\right]\left[\tilde{\mathbf{w}} + D^+(\tilde{\mathbf{w}})\right].$$

Note that this fixed point function tends to increase the wages of locations with excess demand for commuters. By Brouwer's Fixed Point Theorem, there exists $\tilde{\mathbf{w}}^* \in \Delta$ such that $\tilde{w}^* = f(\tilde{\mathbf{w}}^*)$.

Since $\sum_{j=1}^S D_j(\tilde{\mathbf{w}}) = 0$, it cannot be the case that $D_j(\tilde{\mathbf{w}}) > 0$ for all $j = 1, \ldots, S$ or $D_j(\tilde{\mathbf{w}}) < 0$ for all $j = 1, \ldots, S$. Additionally, if $D_j(\tilde{\mathbf{w}}) > 0$ for some j and $D_k(\tilde{\mathbf{w}}) < 0$ for some $k \neq j, \tilde{\mathbf{w}} \neq f(\tilde{\mathbf{w}})$. It follows that at the fixed point for wages, $\tilde{\mathbf{w}}^* = f(\tilde{\mathbf{w}}^*)$, and $D_j(\tilde{\mathbf{w}}) = 0$ for all j. *Q.E.D.*

Homogeneity of degree zero of the commuting market clearing condition (S.42) implies that the equilibrium adjusted wage vector is unique up to a normalization. We impose the normalization that the geometric mean adjusted wage is equal to 1 ($[\prod_{j=1}^S \tilde{w}_{jt}]^{1/S} = 1$), as discussed further in Section S.3.1.5 below.

In our estimation, it proves convenient to rewrite the commuting market clearing condition (S.42) in terms of a composite parameter $\nu = \varepsilon\kappa$ that captures the semi-elasticity of commuting flows with respect to travel times and a transformation of adjusted wages ($\omega_{jt} = \tilde{w}_{jt}^{\varepsilon} = E_{jt} w_{jt}^{\varepsilon}$):

$$H_{Mjt} = \sum_{i=1}^{S} \frac{\omega_{jt}/e^{\nu\tau_{ijt}}}{\sum_{s=1}^{S} \omega_{st}/e^{\nu\tau_{ist}}} H_{Rit}. \tag{S.44}$$

The commuting market clearing condition (S.44) exhibits the same properties in transformed wages (ω_{jt}) as established for adjusted wages ($\tilde{w}_{jt}$) in Lemmas S.6 and S.7 above.

S.3.1.2. *Adjusted Residential Amenities*

Given the parameters $\{\beta, \mu, \varepsilon, \kappa\}$, observed data $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \boldsymbol{\tau}\}$, and the above solutions for adjusted wages $\{\tilde{\mathbf{w}}\}$, the residential choice probabilities (S.15) and population mobility (S.12) determine a unique vector of adjusted residential amenities ($\tilde{\mathbf{B}}$) up to a normalization. From the residential choice probabilities (S.15) and population mobility (S.12), we have

$$\frac{B_i T_i^{1/\varepsilon}}{\bar{U}/\gamma} = \left(\frac{H_{Ri}}{H}\right)^{1/\varepsilon} \frac{Q_i^{1-\beta}}{W_i^{1/\varepsilon}},$$

where W_i is a measure of commuting market access:

$$W_i = \sum_{s=1}^{S} E_s (w_s/d_{is})^{\varepsilon}, \quad d_{is} = e^{\kappa\tau_{is}},$$

and these expressions can be equivalently rewritten as

$$\frac{\tilde{B}_i}{\bar{U}/\gamma} = \left(\frac{H_{Ri}}{H}\right)^{1/\varepsilon} \frac{\mathbb{Q}_i^{1-\beta}}{W_i^{1/\varepsilon}}, \tag{S.45}$$

$$W_i = \sum_{s=1}^{S} (\tilde{w}_s/d_{is})^{\varepsilon}, \quad d_{is} = e^{\kappa\tau_{is}}. \tag{S.46}$$

For locations $s \in \Im_S \cup \Im_R$ with positive residence employment, residential floor prices are related to observed floor prices through $\mathbb{Q}_s = \zeta_{Rs} Q_s$, where (i) $\zeta_{Rs} = \xi_s$ if $s \in \Im_S$ and (ii) $\zeta_{Rs} = 1$ if $s \in \Im_R$. Adjusted residential amenities ($\tilde{B}_i = B_i T_i^{1/\varepsilon} \zeta_{Ri}^{1-\beta}$) include (i) residential amenities (B_i), (ii) the Fréchet scale parameter (T_i) that determines the average utility (or effective units of labor) for commuters from location i, and (iii) the relationship between observed and residential floor prices (ζ_{Ri}). The parameter ζ_{Ri} captures land use regulations that introduce a wedge between commercial and residential floor prices (ξ_i), where we allow this wedge (ξ_i) to vary across blocks. Note that $T_i^{1/\varepsilon}$ and $\zeta_{Ri}^{1-\beta}$ enter residential choice probabilities isomorphically to B_i. Therefore, only the composite adjusted residential amenities ($\tilde{B}_i$) can be recovered from the data. Additionally, from Lemmas S.1–S.3, all locations with zero residential employment have zero adjusted residential amenities.

We choose units in which to measure adjusted residential amenities such that the geometric mean of adjusted residential amenities is equal to 1: $\overline{\tilde{B}}_i = [\prod_{i=1}^{S} \tilde{B}_i]^{1/S} = 1$, where a bar above a variable denotes a geometric mean. Therefore, dividing through by the geometric means in (S.45), adjusted residential amenities can be determined from

$$\text{(S.47)} \qquad \frac{\tilde{B}_i}{\overline{\tilde{B}}_i} = \left(\frac{H_{Ri}}{\overline{H}_{Ri}}\right)^{1/\varepsilon} \left(\frac{\mathbb{Q}_i}{\overline{\mathbb{Q}}_i}\right)^{1-\beta} \left(\frac{W_i}{\overline{W}_i}\right)^{-1/\varepsilon}.$$

S.3.1.3. *Final Goods Productivity*

Given the parameters $\{\alpha, \mu, \varepsilon, \kappa\}$, observed data $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \boldsymbol{\tau}\}$, and the above solutions for adjusted wages $\{\tilde{\mathbf{w}}\}$, profit maximization and zero profits (S.24) determine a unique vector for adjusted productivity ($\tilde{\mathbf{A}}$) up to the normalization chosen for adjusted wages. For all locations with positive workplace employment, we require

$$\text{(S.48)} \qquad \begin{aligned} q_j &= (1-\alpha)\left(\frac{\alpha}{w_j}\right)^{\alpha/(1-\alpha)} A_j^{1/(1-\alpha)}, \\ \mathbb{Q}_j &= (1-\alpha)\left(\frac{\alpha}{\tilde{w}_j}\right)^{\alpha/(1-\alpha)} \tilde{A}_j^{1/(1-\alpha)}, \end{aligned}$$

where q_j denotes the price of commercial floor space and $\mathbb{Q}_i$ denotes the observed price of floor space.

For locations $s \in \Im_M \cup \Im_S$ with positive workplace employment, commercial floor prices are related to observed floor prices through $q_s = \mathbb{Q}_s$. Adjusted final goods productivity ($\tilde{A}_j^{1/(1-\alpha)} = E_j^{\alpha/(\varepsilon(1-\alpha))} A_j^{1/(1-\alpha)}$) captures (i) final goods productivity (A_j) in an employment location and (ii) the Fréchet scale parameter that determines the average utility (or effective units of labor) for commuters to that location (E_j). Note that $E_j^{\alpha/(\varepsilon(1-\alpha))}$ enters the zero-profit condition isomorphically to $A_j^{1/(1-\alpha)}$. Therefore, only the composite adjusted final goods productivity ($\tilde{A}_j$) can be recovered from the data. Additionally, from Lemmas S.1–S.3, all locations with zero workplace employment have zero adjusted final goods productivity.

LEMMA S.8: *Given the parameters* $\{\alpha, \mu, \varepsilon, \kappa\}$, *observed data* $\{\mathbb{Q}, \mathbf{H}_M, \mathbf{H}_R, \boldsymbol{\tau}\}$, *and the solution for adjusted wages* $\{\tilde{\mathbf{w}}^*\}$, *there exists a unique vector of adjusted final goods productivities* $\tilde{\mathbf{A}}^* \in \mathbb{R}_+^S$ *such that the zero-profit condition* (S.48) *holds for all locations with positive workplace employment*.

PROOF: The lemma follows immediately from the zero-profit condition (S.48). *Q.E.D.*

S.3.1.4. *Land Market Clearing*

Given the parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$, the observed data $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$, and the above solutions for adjusted wages and final goods productivity $\{\tilde{\mathbf{w}}, \tilde{\mathbf{A}}\}$, the land market clearing condition determines a unique vector for the adjusted density of development $(\tilde{\boldsymbol{\varphi}})$.

For all locations with positive workplace employment, we can solve for adjusted commercial floor space use from commercial land market clearing (S.30):

$$L_{Mi} = \left(\frac{w_i}{\alpha A_i}\right)^{1/(1-\alpha)} H_{Mi}, \tag{S.49}$$

$$\tilde{L}_{Mi} = \left(\frac{\tilde{w}_i}{\alpha \tilde{A}_i}\right)^{1/(1-\alpha)} H_{Mi},$$

where adjusted commercial land use satisfies $\tilde{L}_{Mi} = E_i^{1/\varepsilon} L_{Mi}$. From Lemmas S.1–S.3, all locations with zero workplace employment have zero adjusted commercial floor space use ($\tilde{L}_{Mi} = L_{Mi} = 0$).

For all locations with positive residence employment, we can solve for adjusted residential floor space use from residential land market clearing (S.29):

$$L_{Ri} = (1-\beta)\left[\sum_{s=1}^{S} \frac{\left(E_s^{1/\varepsilon} w_s / d_{is}\right)^{\varepsilon}}{\sum_{r=1}^{S} \left(E_r^{1/\varepsilon} w_r / d_{ir}\right)^{\varepsilon}} w_s\right] \frac{H_{Ri}}{Q_i}, \tag{S.50}$$

$$\tilde{L}_{Ri} = (1-\beta)\left[\sum_{s=1}^{S} \frac{(\tilde{w}_s / d_{is})^{\varepsilon}}{\sum_{r=1}^{S} (\tilde{w}_r / d_{ir})^{\varepsilon}} \tilde{w}_s\right] \frac{H_{Ri}}{\mathbb{Q}_i},$$

where Q_i denotes the price of residential floor space and $\mathbb{Q}_i$ denotes the observed price of floor space. Adjusted residential floor space use is defined as

$$\tilde{L}_{Ri} = L_{Ri}\zeta_{Ri} \frac{\displaystyle\sum_{s=1}^{R} \frac{(\tilde{w}_s/d_{is})^{\varepsilon}}{\displaystyle\sum_{r=1}^{S}(\tilde{w}_r/d_{ir})^{\varepsilon}} \tilde{w}_s}{\displaystyle\sum_{s=1}^{R} \frac{(\tilde{w}_s/d_{is})^{\varepsilon}}{\displaystyle\sum_{r=1}^{S}(\tilde{w}_r/d_{ir})^{\varepsilon}} \frac{\tilde{w}_s}{E_s^{1/\varepsilon}}}. \tag{S.51}$$

The parameter ζ_{Ri} allows for land use regulations that introduce a wedge between commercial and residential floor prices (ξ_i), where we allow this wedge (ξ_i) to vary across blocks. The fraction in (S.51) controls for the difference between adjusted wages ($\tilde{w}_i = E_i^{1/\varepsilon} w_i$) and actual wages ($w_i$). From Lemmas S.1–S.3, all locations with zero residence employment have zero adjusted residential floor space use ($\tilde{L}_{Ri} = L_{Ri} = 0$).

Combining these solutions for adjusted commercial and residential floor space use, we can solve for the adjusted density of development ($\tilde{\varphi}_i$) from land market clearing:

$$\tilde{L}_i = \tilde{L}_{Mi} + \tilde{L}_{Ri} = \tilde{\varphi}_i K_i^{1-\mu}, \tag{S.52}$$

where the adjusted density of development ($\tilde{\varphi}_i$) relates adjusted floor space ($\tilde{L}_i$) to observed geographical land area (K_i).

LEMMA S.9: *Given the parameters* $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$, *observed data* $\{\mathbb{Q}, \mathbf{H}_M, \mathbf{H}_R, \mathbf{K}, \boldsymbol{\tau}\}$, *and the solutions for adjusted wages and final goods productivity* $\{\tilde{\mathbf{w}}^*, \tilde{\mathbf{A}}^*\}$, *there exists a unique vector of the adjusted density of development* $\tilde{\boldsymbol{\varphi}}^* \in \mathbb{R}_+^S$ *such that the land market clearing condition* (S.52) *holds for all locations*.

PROOF: The lemma follows immediately from the land market clearing condition (S.52) together with commercial land market clearing (S.49) and residential land market clearing (S.50). *Q.E.D.*

From commercial land market clearing (S.49) and residential land market clearing (S.50), we can also solve for the fractions of adjusted floor space ($\tilde{L}$)

allocated to commercial use ($\tilde{\theta}_i$) and residential use ($1-\tilde{\theta}_i$):

$$\tilde{\theta}_i = \begin{cases} 1 & \text{if } H_{Mi} > 0 \text{ and } H_{Ri} = 0, \\ \dfrac{\tilde{L}_{Mi}}{\tilde{L}_{Mi} + \tilde{L}_{Ri}} & \text{if } H_{Mi} > 0 \text{ and } H_{Ri} > 0, \\ 0 & \text{if } H_{Mi} = 0 \text{ and } H_{Ri} > 0. \end{cases} \tag{S.53}$$

S.3.1.5. *Utility, Amenities, and Productivity*

To abstract from changes in currency units over time, we divide observed floor prices in each year by their geometric mean for that year. Therefore, observed floor prices ($\mathbb{Q}$) are normalized to have a geometric mean of 1 in each year. Furthermore, as discussed above, we also normalize adjusted wages ($\tilde{\mathbf{w}}$) and adjusted residential amenities ($\tilde{\mathbf{B}}$) to each have a geometric mean of 1 in each year. We now discuss the implications of these normalizations for the choice of units in which variables are measured.

First, the normalizations of observed floor prices and adjusted wages involve a choice of units in which to measure adjusted final goods productivity ($\tilde{\mathbf{A}}$), as can be seen from the zero-profit condition (S.48). Second, the normalizations of observed land prices, adjusted wages, and adjusted residential amenities also imply a choice of units in which to measure utility. This can be seen from the population mobility condition, which implies that the reservation level of utility in the wider economy ($\bar{U}$) satisfies

$$\gamma\left[\sum_{r=1}^{S}\sum_{s=1}^{S} T_r E_s\left(d_{rs} Q_r^{1-\beta}\right)^{-\varepsilon}(B_r w_s)^{\varepsilon}\right]^{1/\varepsilon} = \bar{U}, \tag{S.54}$$

$$\gamma\left[\sum_{r=1}^{S}\sum_{s=1}^{S}\left(d_{rs} \mathbb{Q}_r^{1-\beta}\right)^{-\varepsilon}(\tilde{B}_r \tilde{w}_s)^{\varepsilon}\right]^{1/\varepsilon} = \bar{U},$$

where $\gamma = \Gamma(\frac{\varepsilon-1}{\varepsilon})$; $\Gamma(\cdot)$ is the Gamma function; and we have used $\tilde{w}_i = E_i^{1/\varepsilon} w_i$ and $\tilde{B}_i = B_i T_i^{1/\varepsilon} \zeta_{Ri}^{1-\beta}$. Third, the choice of units in which to measure floor prices, adjusted wages, and adjusted final goods productivity in turn implies a choice of units in which to measure the adjusted density of development ($\tilde{\boldsymbol{\varphi}}$), as can be seen from commercial land market clearing (S.49), residential land market clearing (S.50), and overall land market clearing (S.52).

We make these normalizations for each year separately. But we recognize that the absolute levels of adjusted amenities ($\tilde{\mathbf{B}}$), adjusted final goods productivity ($\tilde{\mathbf{A}}$), the adjusted density of development ($\tilde{\boldsymbol{\varphi}}$), and the reservation level of utility in the wider economy ($\bar{U}$) could change over time (and in particular could change before and after division or reunification). Therefore, the moment conditions in our estimation use a "difference-in-difference," where the first difference is before and after division (or reunification) and the second difference is between different parts of West Berlin. These moment conditions only exploit *relative changes* in floor prices and other variables between different parts of West Berlin. This use of relative changes implies that our estimation results are invariant to the choice of units in which to measure floor prices, wages, adjusted amenities, adjusted final goods productivity, the adjusted density of development, and utility. Similarly, this use of relative changes implies that our estimation results are unaffected by changes over time in the absolute levels of adjusted amenities, adjusted final goods productivity, the adjusted density of development, or the reservation level of utility in the wider economy.

S.3.1.6. *One-to-One Mapping From* $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ *and the Observed Data to* $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\boldsymbol{\varphi}}\}$

We now combine the results in the above lemmas to establish a one-to-one mapping from the parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ and the observed data $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$ to the unobserved values of adjusted final goods productivity, residential amenities, and the density of development $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\boldsymbol{\varphi}}\}$ for each location.

PROPOSITION S.3: *Given the model's parameters* $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ *and the observed data* $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$, *there are unique vectors of adjusted final goods productivity* $(\tilde{\mathbf{A}}^*)$, *residential amenities* $(\tilde{\mathbf{B}}^*)$, *and the density of development* $(\tilde{\boldsymbol{\varphi}}^*)$ *that are consistent with the data being an equilibrium of the model.*

PROOF: The proposition follows immediately from Lemmas S.6–S.9 above. *Q.E.D.*

To interpret this identification result, note that in models with multiple equilibria, the mapping from the parameters and fundamentals to the endogenous variables is nonunique. In such models, the inverse mapping from the endogenous variables and parameters to the fundamentals can be either unique or nonunique. In the context of our model, Proposition S.3 conditions on the parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ and a combination of observed endogenous variables $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}\}$ and fundamentals $\{\mathbf{K}, \boldsymbol{\tau}\}$, and uses the equilibrium conditions of the model to determine unique values of the location characteristics $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\boldsymbol{\varphi}}\}$. 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 are 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 (S.19). 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 (S.24). 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 (S.15). 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 (S.31).

These relationships for profit maximization, zero profits, utility maximization, population mobility, and land market clearing hold irrespective of whether productivity, amenities, and the density of development are exogenous or endogenous (since agents are atomistic and take the behavior of others as given). Therefore, they hold regardless of whether the model has a single equilibrium (as with exogenous location characteristics) or multiple equilibria (as is possible with endogenous location characteristics). These relationships also hold irrespective of the relative importance of the two components of productivity and amenities (externalities and fundamentals). Therefore, the model has a recursive structure, in which the adjusted values of overall productivity, amenities, and the density of development $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\boldsymbol{\varphi}}\}$ can be determined using the subset of the parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ and the observed data. In the next subsection, we examine how overall adjusted productivity and amenities $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}\}$ can be broken down into their two components of externalities $\{Y_i, \Omega_i\}$ and fundamentals $\{\tilde{a}_i, \tilde{b}_i\}$ using the remaining parameters $\{\lambda, \delta, \eta, \rho\}$.

S.3.2. *Determining* $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}\}$ *From* $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$ *and the Observed Data*

We now establish the one-to-one mapping from the full set of parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$ and the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$ to adjusted production and residential fundamentals $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}\}$. This involves decomposing adjusted productivity and amenities $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}\}$ into their two components of externalities $\{\boldsymbol{Y}, \boldsymbol{\Omega}\}$ and adjusted fundamentals $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}\}$.

We have already established a one-to-one mapping from the subset of parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ and the observed data to adjusted productivity $\{\tilde{\mathbf{A}}\}$. From our specification of productivity (S.40), there is in turn a one-to-one mapping from adjusted productivity $\{\tilde{\mathbf{A}}\}$ and the parameters $\{\lambda, \delta\}$ to production externalities and adjusted production fundamentals $\{\boldsymbol{Y}, \tilde{\mathbf{a}}\}$:

$$a_i = A_i Y_i^{-\lambda}, \quad Y_i = \left[\sum_{s=1}^{S} e^{-\delta\tau_{is}} \frac{H_{Ms}}{K_s}\right], \tag{S.55}$$

$$\tilde{a}_i = \tilde{A}_i Y_i^{-\lambda},$$

where adjusted production fundamentals ($\tilde{a}_i = E_i^{\alpha/\varepsilon} a_i$) captures (i) production fundamentals (a_i) and (ii) the Fréchet scale parameter that determines the average utility from commuting to an employment location (E_i). Note that $E_i^{\alpha/\varepsilon}$ enters adjusted production fundamentals isomorphically to a_i. Therefore, only the composite adjusted production fundamentals ($\tilde{a}_i$) can be recovered from the data. Additionally, from Lemmas S.1–S.5, all locations with zero workplace employment have zero adjusted productivity and hence zero adjusted production fundamentals.

LEMMA S.10: *Given the parameters* $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$, *the observed data* $\{\mathbb{Q}, \mathbf{H}_M, \mathbf{H}_R, \mathbf{K}, \boldsymbol{\tau}\}$, *and the solution for adjusted final goods productivity* $\{\tilde{\mathbf{A}}^*\}$, *there exists a unique vector of adjusted production fundamentals* $\tilde{\mathbf{a}}^* \in \mathbb{R}_+^S$.

PROOF: The lemma follows immediately from productivity (S.55). *Q.E.D.*

We have also already established a one-to-one mapping from the subset of parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ and the observed data to adjusted amenities $\{\tilde{\mathbf{B}}\}$. From our specification of amenities (S.41), there is in turn a one-to-one mapping from adjusted amenities $\{\tilde{\mathbf{B}}\}$ and the parameters $\{\eta, \rho\}$ to residential externalities and adjusted residential fundamentals $\{\boldsymbol{\Omega}, \tilde{\mathbf{b}}\}$:

$$b_i = B_i \Omega_i^{-\eta}, \quad \Omega_i = \left[\sum_{s=1}^{S} e^{-\rho \tau_{is}} \frac{H_{Rs}}{K_s} \right], \tag{S.56}$$
$$\tilde{b}_i = \tilde{B}_i \Omega_i^{-\eta},$$

where adjusted residential fundamentals ($\tilde{b}_i = b_i T_i^{1/\varepsilon} \zeta_{Ri}^{1-\beta}$) include (i) residential fundamentals (b_i), (ii) the Fréchet scale parameter that determines the average utility (or effective units of labor) for commuters from location i (T_i), and (iii) the relationship between observed and residential floor prices (ζ_{Ri}). As discussed above, the parameter ζ_{Ri} includes the effects of land use regulations that introduce a wedge between commercial and residential floor prices (ξ_i), where we allow this wedge (ξ_i) to vary across blocks. Note that $T_i^{1/\varepsilon}$ and $\zeta_{Ri}^{1-\beta}$ enter adjusted residential fundamentals isomorphically to b_i. Therefore, only the composite value of adjusted residential fundamentals ($\tilde{b}_i$) can be recovered from the data. Additionally, from Lemmas S.1–S.5, all locations with zero residential employment have zero adjusted amenities and hence zero adjusted residential fundamentals.

LEMMA S.11: *Given the parameters* $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$, *the observed data* $\{\mathbb{Q}, \mathbf{H}_M, \mathbf{H}_R, \mathbf{K}, \boldsymbol{\tau}\}$, *and the solution for adjusted residential amenities* $\{\tilde{\mathbf{B}}^*\}$, *there exists a unique vector of adjusted residential fundamentals* $\tilde{\mathbf{b}}^* \in \mathbb{R}_+^S$.

PROOF: The lemma follows immediately from residential amenities (S.56). *Q.E.D.*

Combining the above results, we are in a position to establish the following proposition.

PROPOSITION S.4: *Given the model's parameters* $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$ *and the observed data* $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$, *there are unique vectors of adjusted production fundamentals* $(\tilde{\mathbf{a}}^*)$ *and adjusted residential fundamentals* $(\tilde{\mathbf{b}}^*)$ *that are consistent with the data being an equilibrium of the model*.

PROOF: The proposition follows immediately from Proposition S.3 and Lemmas S.10–S.11. *Q.E.D.*

Therefore, the earlier Proposition S.3 established a one-to-one mapping from the known parameters and the observed data to adjusted productivity $(\tilde{\mathbf{A}})$, amenities $(\tilde{\mathbf{B}})$, and the density of development $(\tilde{\boldsymbol{\varphi}})$. Proposition S.4 goes further in establishing a one-to-one mapping from the known parameters and the observed data to the two components of adjusted productivity and amenities: production and residential externalities $\{\boldsymbol{\Upsilon}, \boldsymbol{\Omega}\}$ and adjusted production and residential fundamentals $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}\}$. As for the earlier Proposition S.3, the results in Proposition S.4 hold regardless of whether the model has a single equilibrium or multiple equilibria. Conditional on the parameters and the observed combination of endogenous variables and fundamentals, the structural relationships of the model contain enough information to uniquely determine adjusted production and residential fundamentals $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}\}$, which correspond to structural residuals that ensure the model exactly replaces the observed data as an equilibrium.

In our structural estimation of the model in Section S.4, we use Propositions S.3 and S.4 as an input into our Generalized Method of Moments (GMM) estimation, in which we determine both the parameters and the unobserved adjusted fundamentals.

This completes our characterization of the one-to-one mapping from the known parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$ and the observed data $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$ to the unobserved location characteristics $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}, \tilde{\boldsymbol{\varphi}}\}$.

S.4. STRUCTURAL ESTIMATION

We now turn to the structural estimation of the model, where both the parameters and the unobserved location characteristics are unknown and to be estimated. First, in Section S.4.1, we use the results from Section S.3 to express the unobserved production and residential fundamentals $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}\}$ as one-to-one functions of the observed data $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$ and the parameters

$\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$. Therefore, these unobserved location characteristics correspond to *structural residuals* of the model that are functions of the observed data and parameters.

Second, in Section S.4.2, we use the resulting closed-form solutions for these structural residuals to construct moment conditions using the exogenous variation provided by Berlin's division and reunification. Third, in Section S.4.3, we show how these moment conditions can be used to estimate the model's parameters using the Generalized Method of Moments (GMM). Fourth, in Section S.4.4 we discuss the computationally demanding optimization problem over the parameter vector and the algorithms that we use to solve this problem.

Fifth, in Section S.4.5, we show how the moment conditions identify the parameters, and characterize the properties of the GMM objective function in the parameter space. We show that the GMM objective has a unique global minimum in the parameter space. We therefore find that there is only a single parameter vector that is consistent with the data under our identifying assumptions.

S.4.1. *Structural Residuals*

We first use the results from Section S.3 to write unobserved production and residential fundamentals as structural residuals that are functions of the parameters and observed data. Of the model's eight 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 value of these parameters, we set them equal to central estimates from the existing empirical literature. We set the share of residential floor space in consumer expenditure $(1-\beta)$ equal to 0.25, which is consistent with the estimates in Davis and Ortalo-Magné (2011). We assume that the share of commercial floor space in firm costs $(1-\alpha)$ 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 estimates in Combes, Duranton, and Gobillon (2014) and Epple, Gordon, and Sieg (2010) and with micro data on property transactions for Berlin from 2000 to 2012.

Given these values for $\{\alpha, \beta, \mu\}$, we use the observed data $\mathbf{X} = [\mathbb{Q}\ \mathbf{H_M}\ \mathbf{H_R}\ \mathbf{K}\ \boldsymbol{\tau}]$ and the structure of the model to estimate the six parameters determining the strength of agglomeration forces and commuting costs $\boldsymbol{\Lambda} = [\nu\ \ \varepsilon\ \ \lambda\ \ \delta\ \ \eta\ \ \rho]'$ (where $\nu = \varepsilon\kappa$ is the semi-elasticity of commuting flows with respect to travel times) and the unobserved characteristics for each location $\boldsymbol{\Phi} = [\tilde{\boldsymbol{\varphi}}\ \ \tilde{\mathbf{a}}\ \ \tilde{\mathbf{b}}]$. From profit maximization and zero profits (S.48) and productivity (S.55), the structural residual for adjusted production fundamentals can be written as the following function of the parameters and observed data:

$$\text{(S.57)} \qquad \frac{\tilde{a}_{it}}{\bar{\tilde{a}}_t} = \left(\frac{\mathbb{Q}_{it}}{\overline{\mathbb{Q}}_t}\right)^{1-\alpha} \left(\frac{\tilde{w}_{it}}{\bar{\tilde{w}}_t}\right)^{\alpha} \left(\frac{Y_{it}}{\overline{Y}_t}\right)^{-\lambda},$$

where we now make time explicit with the subscript t and a bar above a variable denotes a geometric mean so that $\overline{\mathbb{Q}}_t = \exp\{\frac{1}{S}\sum_{s=1}^{S}\ln \mathbb{Q}_{st}\}$; adjusted wages ($\tilde{w}_{it}$) are a function of observed workplace employment, residence employment, and travel times $\{H_{Mit}, H_{Rit}, \tau_{ijt}\}$ from commuting market clearing (S.42); production externalities (Y_{it}) are a function of observed workplace employment, geographical land area, and travel times $\{H_{Mit}, K_i, \tau_{ijt}\}$:

$$Y_{it} = \sum_{s=1}^{S} e^{-\delta\tau_{ist}} \frac{H_{Mst}}{K_s},$$

where in the estimation we exclude the own region i from the summation to rule out a mechanical correlation through own-region workplace employment.

From the residential choice probabilities (S.15), the expected utility from moving to the city (S.54), and amenities (S.56), the structural residual for adjusted residential fundamentals can be written as the following function of the parameters and observed data:

$$\text{(S.58)} \qquad \frac{\tilde{b}_{it}}{\bar{\tilde{b}}_t} = \left(\frac{H_{Rit}}{\overline{H}_{Rt}}\right)^{1/\varepsilon} \left(\frac{\mathbb{Q}_{it}}{\overline{\mathbb{Q}}_t}\right)^{1-\beta} \left(\frac{W_{it}}{\overline{W}_t}\right)^{-1/\varepsilon} \left(\frac{\Omega_{it}}{\overline{\Omega}_t}\right)^{-\eta},$$

where a bar above a variable again denotes a geometric mean. Commuting market access (W_{it}) is a function of adjusted wages ($\tilde{w}_{it}$) and observed travel times (τ_{ijt}):

$$W_{it} = \left[\sum_{s=1}^{S} \left(\tilde{w}_{st}/e^{\kappa\tau_{ist}}\right)^{\varepsilon}\right],$$

where adjusted wages ($\tilde{w}_{it}$) are again a function of observed workplace employment, residence employment, and travel times $\{H_{Mit}, H_{Rit}, \tau_{ijt}\}$ from commuting market clearing (S.42). Residential externalities (Ω_{it}) are a function of observed residence employment, geographical land area, and travel times $\{H_{Rit}, K_i, \tau_{ijt}\}$:

$$\Omega_{it} = \sum_{s=1}^{S} e^{-\rho\tau_{ist}} \frac{H_{Rst}}{K_s},$$

where in the estimation we again exclude the own region i from the summation to rule out a mechanical correlation through own-region residence employment.

We solve for these structural residuals for all of Berlin, both before the war and after reunification, and for West Berlin during division. To structurally estimate the model's parameters, we focus on the impact of division and reunification on West Berlin, since it remained a market economy and hence we expect the mechanisms in the model to apply.[9]

We assume that each structural residual consists of a time-invariant fixed effect $\{\tilde{a}_i^F, \tilde{b}_i^F\}$ and a time-varying stochastic shock $\{\tilde{a}_{it}^V, \tilde{b}_{it}^V\}$. Taking differences before and after division (or before and after reunification), the fixed effects are differenced out, and we obtain the following expression for the relative changes in the structural residuals across blocks within West Berlin:

$$\Delta \ln\left(\frac{\tilde{a}_{it}^V}{\bar{\tilde{a}}_t^V}\right) = (1-\alpha)\Delta \ln\left(\frac{\mathbb{Q}_{it}}{\overline{\mathbb{Q}}_t}\right) + \alpha \Delta \ln\left(\frac{\tilde{w}_{it}}{\bar{\tilde{w}}_t}\right) - \lambda \Delta \ln\left(\frac{Y_{it}}{\bar{Y}_t}\right), \tag{S.59}$$

$$\Delta \ln\left(\frac{\tilde{b}_{it}^V}{\bar{\tilde{b}}_t^V}\right) = \frac{1}{\varepsilon}\Delta \ln\left(\frac{H_{Rit}}{\bar{H}_{Rt}}\right) + (1-\beta)\Delta \ln\left(\frac{\mathbb{Q}_{it}}{\overline{\mathbb{Q}}_t}\right) - \frac{1}{\varepsilon}\Delta \ln\left(\frac{W_{it}}{\bar{W}_t}\right) - \eta \ln \Delta\left(\frac{\Omega_{it}}{\bar{\Omega}_t}\right), \tag{S.60}$$

where a bar above a variable again denotes a geometric mean.

The structural residuals in (S.59) and (S.60) correspond to double-differenced adjusted production and residential fundamentals. The first difference is before and after division (or before and after reunification) and is denoted by the time-difference operator (Δ) in (S.59) and (S.60). This first difference eliminates any time-invariant factors with time-invariant effects, where we allow these fixed effects to be correlated with the endogenous variables of the model. The second difference is across blocks within West Berlin and is reflected in the normalization relative to the geometric mean in (S.59) and (S.60). This second difference eliminates variables that are common across blocks within each time period (e.g., the reservation level of utility, $\bar{U}_t$). It also ensures that our results are invariant to the choice of units in which to measure production and residential fundamentals, since this choice of units is common across blocks and hence is differenced out.

[9]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.

S.4.2. *Moment Conditions*

Our moment conditions exploit the exogenous change in the surrounding concentration of economic activity induced by Berlin's division and reunification. Our identifying assumption is that double-differenced log adjusted production and residential fundamentals are uncorrelated with a set of indicator variables ($\mathbb{I}_k$ for $k \in \{1, \dots, K_{\mathbb{I}}\}$) capturing proximity to economic activity in East Berlin prior to division. Based on the results of our reduced-form regressions, we measure proximity to economic activity in East Berlin using distance from the pre-war CBD, and use 50 indicator variables for percentiles of this distance distribution:

$$\mathbb{E}\big[\mathbb{I}_k \times \Delta \ln(\tilde{a}_{it}/\bar{\tilde{a}}_t)\big] = 0, \quad k \in \{1, \dots, K_{\mathbb{I}}\}, \tag{S.61}$$

$$\mathbb{E}\big[\mathbb{I}_k \times \Delta \ln(\tilde{b}_{it}/\bar{\tilde{b}}_t)\big] = 0, \quad k \in \{1, \dots, K_{\mathbb{I}}\}. \tag{S.62}$$

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 (production and residential fundamentals). 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 production and residential fundamentals in West Berlin blocks.

Since the moment conditions (S.61)–(S.62) are based on double differences in adjusted production and residential fundamentals, they only exploit relative variation across different areas within West Berlin. Any changes in the attractiveness of West Berlin relative to the larger economy that are common across locations within West Berlin are differenced out. 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).

We augment these moment conditions for adjusted production and residential fundamentals with two other moment conditions that use data on commuting travel times and wages for West Berlin during division.[10] The first of these moment conditions requires that the total number of workers commuting for

[10] In the paper, we report over-identification checks in which we show that the model using an estimated value of $\nu = \varepsilon\kappa$ 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.

less than 30 minutes in the model is equal to the corresponding number in the data. From the commuting market clearing condition (S.44), this moment condition can be expressed as

$$\text{(S.63)} \qquad \mathbb{E}\left[\psi H_{Mjt} - \sum_{i \in \aleph_j}^{S} \frac{\omega_{jt}/e^{\nu\tau_{ijt}}}{\sum_{s=1}^{S} \omega_{st}/e^{\nu\tau_{ijt}}} H_{Rit}\right] = 0,$$

where $\nu = \varepsilon\kappa$; ψ is the fraction of workers that commute for less than 30 minutes in the data; $\omega_{it} = \tilde{w}_{it}^{\varepsilon}$ is a measure of transformed wages from solving the commuting market clearing condition (S.44); and $\aleph_j$ is the set of residence locations i within 30 minutes travel time of workplace location j.

The second of these moment conditions requires that the variance of log adjusted wages ($\tilde{w}_{it} = \omega_{it}^{1/\varepsilon}$) in the model is equal to the variance of log wages in the data ($\sigma^2_{\ln w_i}$) for West Berlin during division:[11]

$$\text{(S.64)} \qquad \mathbb{E}\left[(1/\varepsilon)^2 \ln(\omega_{jt})^2 - \sigma^2_{\ln w_{it}}\right] = 0,$$

where transformed wages (ω_i) depend solely on ν, workplace employment, residence employment, and travel times from the labor market clearing condition (S.44); ε scales the variance of log adjusted wages ($\tilde{w}_i$) relative to the variance of log transformed wages (ω_i).

S.4.3. *GMM Estimation*

In this subsection, we briefly review the Generalized Method of Moments (GMM) estimator (Hansen (1982), Cameron and Triveldi (2005)) as applied to our setting.

One-Step GMM Estimator: Observations are indexed by $i \in \{1, \dots, N\}$. The observed data are given by the $N \times 5$ vector $\mathbf{X} = [\mathbb{Q} \;\; \mathbf{H_M} \;\; \mathbf{H_R} \;\; \mathbf{K} \;\; \boldsymbol{\tau}]$. There are M moment conditions and P parameters in the $P \times 1$ vector $\boldsymbol{\Lambda} = [\nu \;\; \varepsilon \;\; \kappa \;\; \lambda \;\; \delta \;\; \eta \;\; \rho]'$. Our moment conditions can be written as

$$\text{(S.65)} \qquad \mathbb{M}(\boldsymbol{\Lambda}) = \frac{1}{N} \sum_{i=1}^{N} m(\mathbf{X}_i, \boldsymbol{\Lambda}) = 0,$$

where $m(\mathbf{X}_i, \boldsymbol{\Lambda})$ is the moment function. The one-step GMM estimator solves

$$\text{(S.66)} \qquad \hat{\boldsymbol{\Lambda}}_{\text{GMM}} = \arg\min \left(\frac{1}{N} \sum_{i=1}^{N} m(\mathbf{X_i}, \boldsymbol{\Lambda})'\right) \mathbb{W} \left(\frac{1}{N} \sum_{i=1}^{N} m(\mathbf{X_i}, \boldsymbol{\Lambda})\right),$$

[11] As reliable wage data for pre-war Berlin are 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.

where the weighting matrix $\mathbb{W}$ is the identity matrix. The estimated variance–covariance matrix for the one-step GMM estimates $\hat{V}(\hat{\boldsymbol{\Lambda}}_{\text{GMM}})$ is

$$\hat{\mathbb{V}}(\hat{\boldsymbol{\Lambda}}_{\text{GMM}}) = (\hat{\mathbb{G}}'\mathbb{W}\hat{\mathbb{G}})^{-1}(\hat{\mathbb{G}}'\mathbb{W}\hat{\mathbb{S}}\mathbb{W}'\hat{\mathbb{G}})(\hat{\mathbb{G}}'\mathbb{W}\hat{\mathbb{G}})^{-1},$$

where $\hat{\mathbb{G}}$ is the estimated $M \times P$ Jacobian of the M moment conditions with respect to the P parameters:

$$\hat{\mathbb{G}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial m(\mathbf{X}_i, \boldsymbol{\Lambda})}{\partial \boldsymbol{\Lambda}'}\bigg|_{\hat{\boldsymbol{\Lambda}}_{\text{GMM}}}. \tag{S.67}$$

The White (1980) heteroscedasticity robust estimator of the matrix $\mathbb{S}$ is

$$\hat{\mathbb{S}}_0 = \frac{1}{N}\sum_{i=1}^{N} m(\mathbf{X}_i, \hat{\boldsymbol{\Lambda}})m(\mathbf{X}_i, \hat{\boldsymbol{\Lambda}})'.$$

To allow for spatial correlation of the structural errors, we report standard errors based on the Conley (1999) heteroscedasticity and autocorrelation consistent (HAC) estimator of the matrix $\mathbb{S}$:

$$\begin{aligned}\hat{\mathbb{S}} = \hat{\mathbb{S}}_0 + \frac{1}{N}\sum_{j=1}^{J}\omega(j) \\ \times \sum_{i=j+1}^{N}\big(m(\mathbf{X}_i, \hat{\boldsymbol{\Lambda}})m(\mathbf{X}_{i-j}, \hat{\boldsymbol{\Lambda}})' + m(\mathbf{X}_{i-j}, \hat{\boldsymbol{\Lambda}})m(\mathbf{X}_i, \hat{\boldsymbol{\Lambda}})'\big),\end{aligned} \tag{S.68}$$

which can be written as

$$\hat{\mathbb{S}} = \hat{\mathbb{S}}_0 + \sum_{j=1}^{J}\boldsymbol{\varpi}(j)\big(\hat{\mathbb{S}}_j + \hat{\mathbb{S}}_j'\big), \tag{S.69}$$

where J is the maximum spatial lag between observations and $\boldsymbol{\varpi}(j)$ is a spatial weight that is equal to 1 if the spatial distance is less than the specified maximum spatial lag and zero otherwise. We set the maximum spatial lag equal to 0.5 kilometers.

Two-Step (Efficient) GMM Estimator: The two-step (efficient) GMM estimator uses the efficient (optimal) weighting matrix ($\hat{\mathbb{S}}^{-1}$) and solves

$$\hat{\boldsymbol{\Lambda}}^{E}_{\text{GMM}} = \arg\min\left(\frac{1}{N}\sum_{i=1}^{N} m(\mathbf{X}_{\mathbf{i}}, \boldsymbol{\Lambda})'\right)\hat{\mathbb{S}}^{-1}\left(\frac{1}{N}\sum_{i=1}^{N} m(\mathbf{X}_{\mathbf{i}}, \boldsymbol{\Lambda})\right),$$

where $\hat{\mathbb{S}}$ is computed using the heteroscedasticity robust and autocorrelation consistent (HAC) estimator (S.68) evaluated at the one-step GMM parameter estimates ($\hat{\boldsymbol{\Lambda}}_{\text{GMM}}$). The estimated variance–covariance matrix for the two-step GMM estimates $\hat{V}(\hat{\boldsymbol{\Lambda}}^{E}_{\text{GMM}})$ is

$$\hat{\mathbb{V}}(\hat{\boldsymbol{\Lambda}}^{E}_{\text{GMM}}) = \frac{1}{N}(\hat{\mathbb{G}}'\tilde{\mathbb{S}}^{-1}\hat{\mathbb{G}})^{-1},$$

where $\hat{\mathbb{G}}$ is computed using (S.67) evaluated at the efficient GMM parameter estimates ($\hat{\boldsymbol{\Lambda}}^{E}_{\text{GMM}}$); $\tilde{\mathbb{S}}$ is computed using (S.68) evaluated at the efficient GMM parameter estimates ($\hat{\boldsymbol{\Lambda}}^{E}_{\text{GMM}}$).

S.4.4. *Estimation Algorithms*

The GMM estimator chooses the values of the model's parameters $\boldsymbol{\Lambda} = [\nu\ \ \varepsilon\ \ \lambda\ \ \delta\ \ \eta\ \ \rho]'$ to minimize the GMM objective function (S.66). This optimization routine searches over alternative parameter vectors and evaluates the moment function $m(\mathbf{X}_i, \boldsymbol{\Lambda})$ for each parameter vector. Evaluating the moment function for each parameter vector in turn involves solving a fixed point problem for the transformed wage vector ($\boldsymbol{\omega}$) that solves the commuting market clearing condition (S.44). Solving this fixed point problem is computationally demanding, because it involves solving for transformed wages in 15,937 blocks, where the matrix of commuting probabilities includes $15{,}937 \times 15{,}937 = 254$ million bilateral commuting flows. We now discuss the algorithms that we use to solve these problems and estimate the model's parameters.

We first discuss the algorithm that we use to solve the fixed point problem for transformed wages and hence evaluate the moment function $m(\mathbf{X}_i, \boldsymbol{\Lambda})$ for each parameter vector. We next discuss the algorithms that we use to search over alternative parameter vectors to minimize the GMM objective function.

Algorithms for evaluating the moment function for each parameter vector: To evaluate the moment function for each parameter vector, we use the recursive structure of the model, as characterized in Section S.3 above:

1. Given ν and the observed data $\{\mathbf{H_R}, \mathbf{H_M}, \boldsymbol{\tau}\}$, the equilibrium transformed wage vector $\{\boldsymbol{\omega}\}$ can be uniquely determined (up to a normalization) from the commuting market clearing condition (S.44) alone independently of the other equilibrium conditions of the model.

2. Given $\{\nu, \varepsilon, \beta, \mu\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \boldsymbol{\tau}\}$, and solutions for transformed wages $\{\boldsymbol{\omega}\}$, adjusted amenities $\{\tilde{\mathbf{B}}\}$ can be uniquely determined (up to a normalization) from residential choices (S.47).

3. Given $\{\nu, \varepsilon, \alpha, \mu\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \boldsymbol{\tau}\}$, and solutions for transformed wages $\{\boldsymbol{\omega}\}$, adjusted productivity $\{\tilde{\mathbf{A}}\}$ can be uniquely determined from the zero-profit condition (S.48).

4. Given $\{\nu, \varepsilon, \alpha, \beta, \mu\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$, and solutions for transformed wages and adjusted productivity $\{\boldsymbol{\omega}, \tilde{\mathbf{A}}\}$, the adjusted density of development $\{\tilde{\boldsymbol{\varphi}}\}$ can be uniquely determined from land market clearing (S.52).

5. Given $\{\nu, \varepsilon, \alpha, \beta, \mu, \lambda, \delta\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$, and solutions for adjusted productivity $\{\tilde{\mathbf{A}}\}$, adjusted production fundamentals $\{\tilde{\mathbf{a}}\}$ can be determined from the specification of productivity (S.55).

6. Given $\{\nu, \varepsilon, \alpha, \beta, \mu, \eta, \rho\}$, the observed data $\{\mathbb{Q}, \mathbf{H_R}, \mathbf{H_M}, \mathbf{K}, \boldsymbol{\tau}\}$, and solutions for adjusted residential amenities $\{\tilde{\mathbf{B}}\}$, adjusted residential fundamentals $\{\tilde{\mathbf{b}}\}$ can be determined from the specification of residential amenities (S.56).

We now consider these steps in turn.

1. The commuting market clearing condition (S.44) can be written as

$$\mathbf{H_M} = \mathbb{C}(\boldsymbol{\omega})\mathbf{H_R}, \tag{S.70}$$

where $\mathbb{C}(\boldsymbol{\omega})$ is the matrix of commuting probabilities. Lemma S.3 establishes that locations without positive workplace employment must have zero productivity and a zero wage. Furthermore, locations with zero residence employment supply zero commuters to all workplace locations. Therefore, we set transformed wages equal to zero for all locations with zero workplace employment, and we set commuting probabilities equal to zero for all source locations with zero residence employment and all destination locations with zero workplace employment. Hence, we can reduce the dimensionality of the system of equations (S.70), such that $\mathbf{H_M}$ is a $N_M \times 1$ vector, where N_M is the number of locations with positive workplace employment; $\mathbf{H_R}$ is a $N_R \times 1$ vector, where N_R in the number of locations with positive residence employment; and $\mathbb{C}(\boldsymbol{\omega})$ is a $N_M \times N_R$ matrix. Note that Lemmas S.6–S.7 establish that the wage system (S.44) satisfies gross substitution and has a unique equilibrium. Therefore, we solve for this unique equilibrium using the following iterative procedure. We guess an initial transformed wage vector $\widehat{\boldsymbol{\omega}}^0$ and evaluate the matrix of commuting probabilities $\mathbb{C}(\widehat{\boldsymbol{\omega}}^0)$ to generate predicted workplace employment $(\widehat{\mathbf{H}}_\mathbf{M})$ given observed residence employment $(\mathbf{H_R})$. If predicted workplace employment is less than observed workplace employment for a location, we increase our guess of the transformed wage for that location. If predicted workplace employment is greater than observed workplace employment for a location, we decrease our guess of the transformed wage for that location. To update the transformed wage for a location, we compute an intelligent wage adjustment factor using the numerator of the commuting probabilities in (S.44): $\widehat{\boldsymbol{\omega}}^1 = (\mathbf{H_M}/\widehat{\mathbf{H}}_\mathbf{M})\widehat{\boldsymbol{\omega}}_0$. We use this intelligent adjustment factor to update our guess of the transformed wage vector to $\widehat{\boldsymbol{\omega}}^2 = (0.5 \times \widehat{\boldsymbol{\omega}}^1) + (0.5 \times \widehat{\boldsymbol{\omega}}^0)$. We then repeat the above process using this new guess for the transformed wage vector $\widehat{\boldsymbol{\omega}}^2$ until the wage system converges. Since the wage system (S.44) satisfies gross substitution, this iterative procedure converges rapidly to the unique equilibrium transformed wage vector.

2. Solving for transformed wages is the only computationally intensive component of the evaluation of the moment function $m(\mathbf{X}_i, \boldsymbol{\Lambda})$. Having solved for transformed wages $\{\boldsymbol{\omega}\}$, we can solve for adjusted amenities $\{\tilde{\mathbf{B}}\}$, adjusted productivity $\{\tilde{\mathbf{A}}\}$, the adjusted density of development $\{\tilde{\boldsymbol{\varphi}}\}$, adjusted production fundamentals $\{\tilde{\mathbf{a}}\}$, and adjusted residential fundamentals $\{\tilde{\mathbf{b}}\}$ from simple manipulations of single equations in Steps 2–6 above, as shown in Propositions S.3 and S.4 in Sections S.3.1 and S.3.2 of this supplement, respectively.

Algorithms for minimizing the GMM objective function with respect to the parameter vector: Having discussed our algorithms to evaluate the moment function for each parameter vector $\boldsymbol{\Lambda} = [\nu \ \ \varepsilon \ \ \lambda \ \ \delta \ \ \eta \ \ \rho]'$, the algorithms for minimizing the GMM objective function with respect to the parameter vector are straightforward. The estimation is undertaken in `Matlab`. To minimize the GMM objective function with respect to the parameter vector, we experimented with using both derivative-based constrained optimization routines (e.g., `fmincon` and the `Knitro` plug-in `ktrlink` for `Matlab`) and nonderivative-based constrained optimization methods (e.g., `patternsearch` and `simulannealbnd` from the `Global Optimization Toolbox`). To characterize the properties of the GMM objective in the parameter space, we also undertook a grid search over the parameter space. As discussed in Section S.4.5 below, the GMM objective is well-behaved in the parameter space. Therefore, we obtain similar parameter estimates from these alternative optimization routines and from alternative initial conditions. The results reported in the paper are estimated using `patternsearch`.

Computational Time: Evaluating the moment function for a given parameter vector $\boldsymbol{\Lambda} = [\nu \ \ \varepsilon \ \ \lambda \ \ \delta \ \ \eta \ \ \rho]'$, and hence solving for transformed wages for 15,937 blocks for this parameter vector, takes around 30 seconds of computing time on the latest generation of desktop computers. This process also uses a large amount of RAM to store the $15{,}937 \times 15{,}937 = 254$ million elements of the matrix of commuting probabilities. Minimizing the GMM objective function with respect to the parameter vector takes a few days on the latest generation of desktop computers. Therefore, we trial code on a random sample of 25 percent of blocks on the latest generation of desktop machines. This reduces the number of blocks for which we solve for transformed wages to 3,984 and hence reduces the size of the matrix of commuting probabilities to 1.59 million elements. Once the code is up and running, we estimate the model for the full sample of blocks using the computer cluster of the Humboldt University of Berlin. Minimizing the GMM objective with respect to the parameter vector takes less than one day for the full sample using this computer cluster.

S.4.5. *Identification*

In Propositions S.3 and S.4, 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 $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$ and known values of the model's parameters $\{\nu, \varepsilon, \lambda, \delta, \eta, \rho\}$. Therefore, adjusted production and residential fundamentals are structural residuals that are one-to-one functions of the observed data and parameters, as demonstrated in Section S.4.1. 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 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 Propositions S.3 and S.4 to solve for the implied values of 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 moment conditions in terms of adjusted production and residential fundamentals, commuting flows, and wages, as discussed in Section S.4.2. 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 $\{\nu, \varepsilon, \lambda, \delta, \eta, \rho\}$ and unobserved adjusted fundamentals $\{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}\}$ that are consistent with the same observed data $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \boldsymbol{\tau}\}$. 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. Below in this subsection, 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 A.3 of the separate Technical Data 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 $\{\nu, \varepsilon, \lambda, \delta, \eta, \rho\}$. We begin with the semi-elasticity of commuting flows with respect to travel times (ν). A higher value of ν 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 (S.63). The recursive structure of the model implies that none of the other parameters $\{\varepsilon, \lambda, \delta, \eta, \rho\}$ affect the commuting moment condition (ε only enters through $\nu = \varepsilon\kappa$ and ω_j). To characterize the properties of the commuting moment condition in the parameter space, we undertake a grid search over 20 possible values of ν (from 0.01 to 0.20). In Figure S.1, we

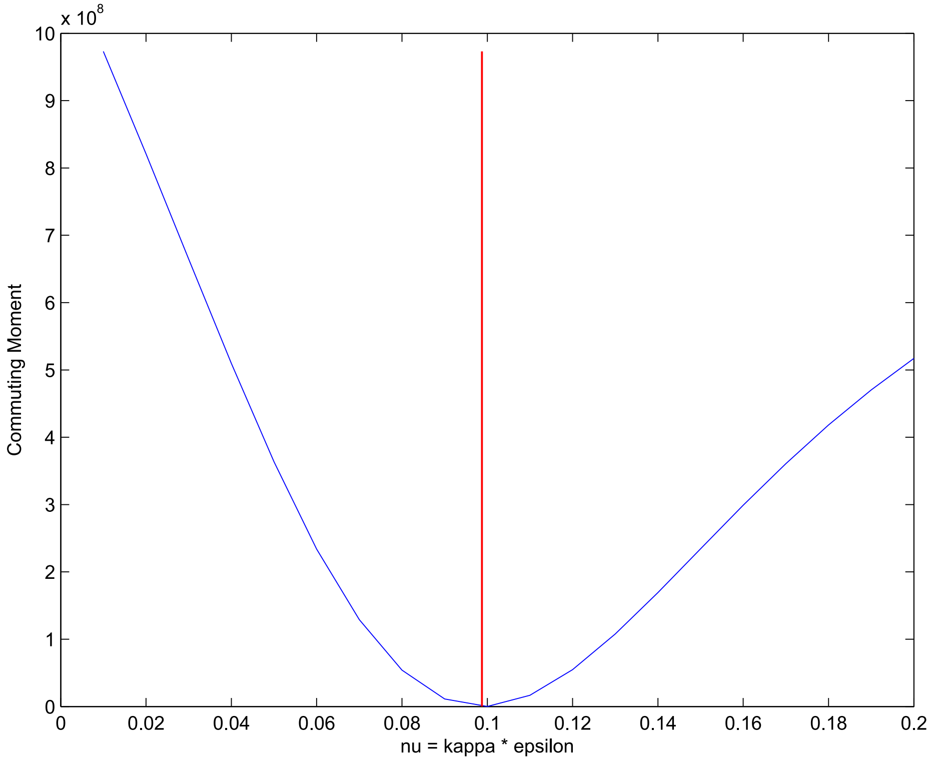

FIGURE S.1.—Commuting moment condition (sum of squared deviations).

display the value of the commuting moment condition for West Berlin for division for each value of ν. As apparent from the figure, the moment condition has a unique global minimum that identifies ν.

We next consider the Fréchet shape parameter determining the heterogeneity of commuting decisions (ε). A higher value of ε implies a smaller dispersion in adjusted wages ($\tilde{w}_{it}$) in the wage moment condition (S.64) for a given dispersion of transformed wages (ω_{it}), since $\sigma^2_{\ln \tilde{w}_{it}} = (1/\varepsilon)^2 \sigma^2_{\ln \omega_{it}}$. From the commuting market clearing condition (S.44), transformed wages (ω_{it}) depend solely on the parameter ν and observed workplace employment, residence employment, and travel times. The recursive structure of the model implies that none of the other parameters $\{\lambda, \delta, \eta, \rho\}$ affect the wage moment condition. To characterize the properties of the wage dispersion moment condition in the parameter space, we undertake a grid search over 20 possible values of ε (from 5 to 10) for our estimated value of ν. In Figure S.2, we display the value of the wage dispersion moment condition for West Berlin during division for each value of ε. As shown in the figure, the moment condition has a unique global minimum that identifies ε.

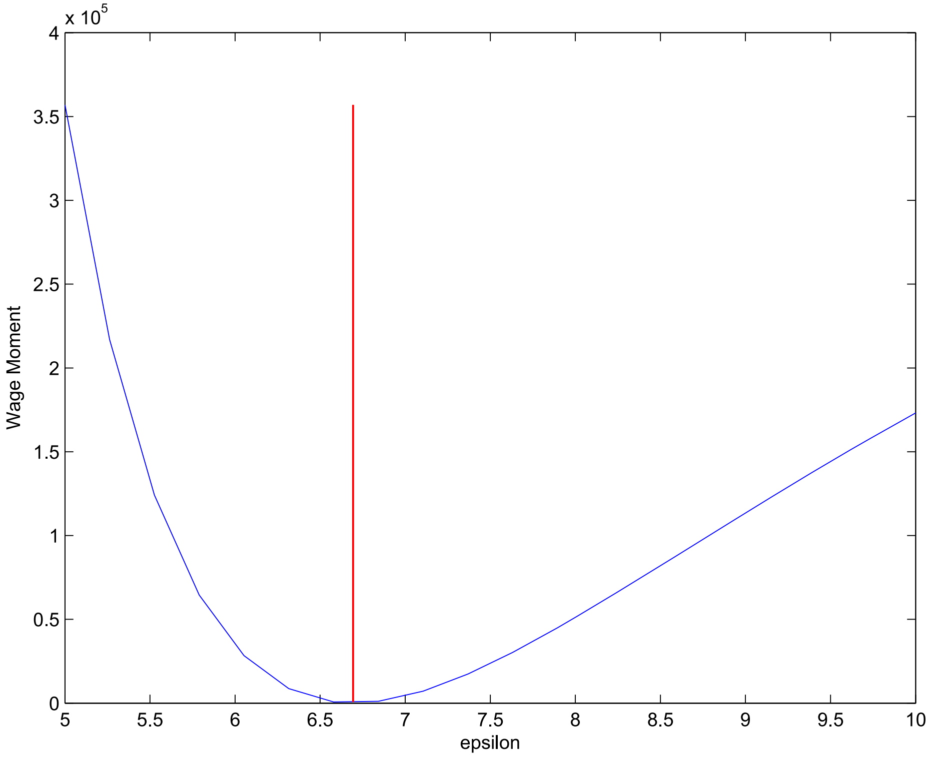

FIGURE S.2.—Wage moment condition (sum of squared deviations).

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 (S.48) and (S.47)). Given these separate measures of adjusted 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 (Υ_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 (Ω_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 Section S.2.7. From the moment conditions for changes in production and residential fundamentals (S.61)–(S.62), our GMM estimator is similar to these instrumental variable regressions, but jointly estimates the parameters $\{\nu, \varepsilon, \lambda, \delta, \eta, \rho\}$ as part of a system that includes our moment conditions for commuting and wages.

The division of Berlin implies a fall in production externalities (Υ_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 from Eastern concentrations of workplace employment. In Figure S.3, we show the mean changes in log adjusted production fundamentals following division across the distance grid cells from the pre-war CBD used in the estimation (i) for the estimated parameters, (ii) for stronger agglomeration forces (larger λ and δ than estimated), and (iii) for weaker agglomeration forces (smaller λ and δ

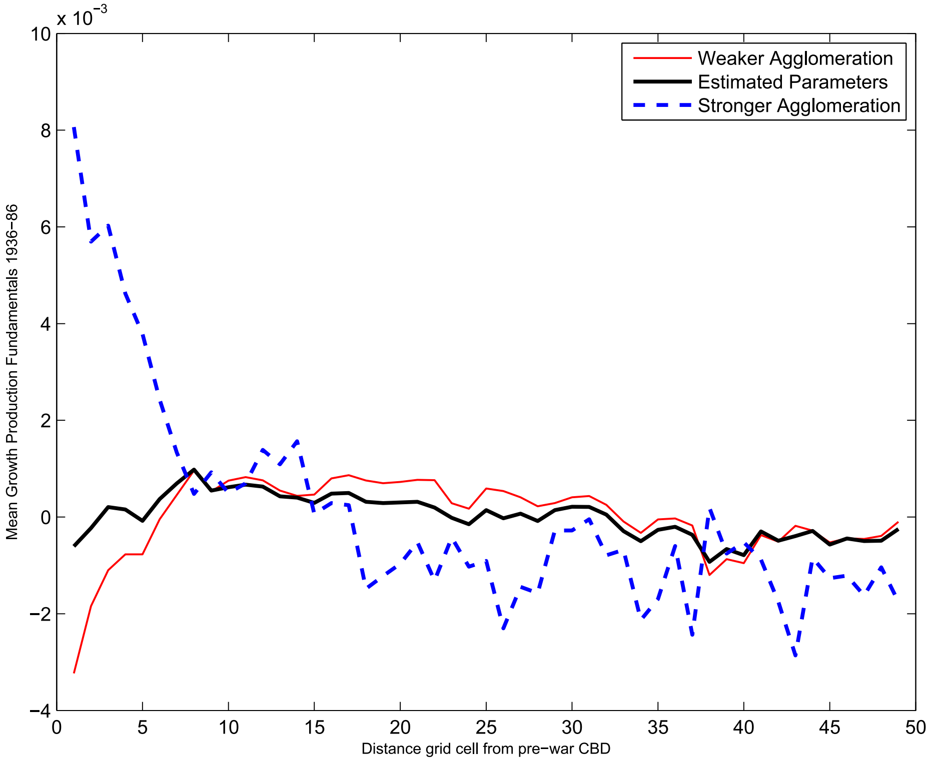

FIGURE S.3.—Production fundamentals distance grid cell moments (division).

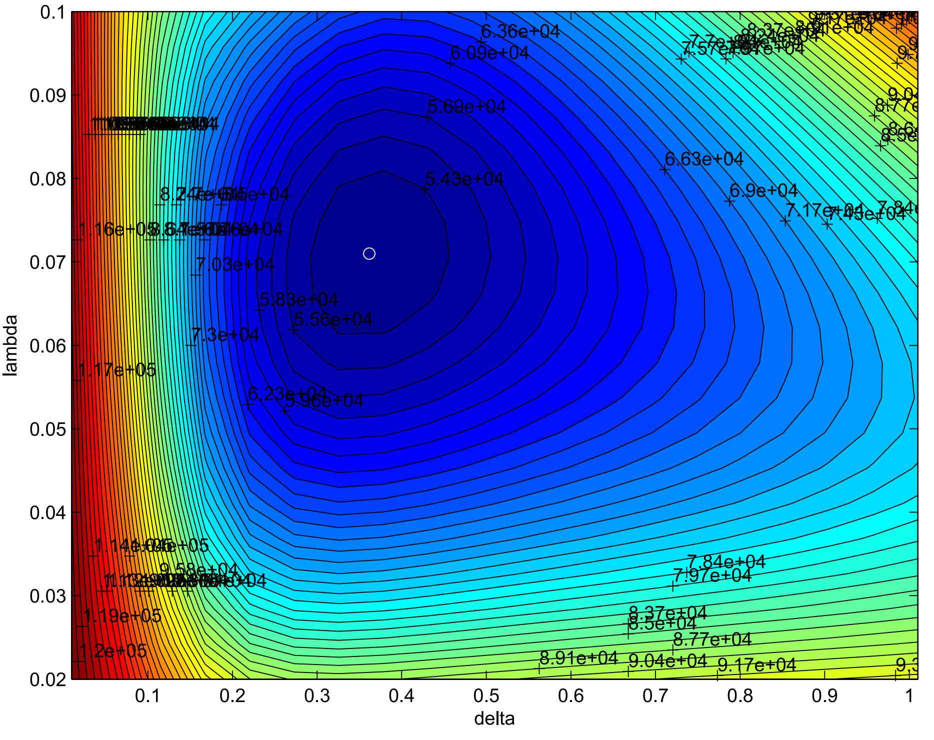

FIGURE S.4.—Production fundamentals moment condition (sum of squared deviations).

than estimated). From the moment condition (S.61), the production spillover parameters $\{\lambda, \delta\}$ are chosen to make the mean changes in log adjusted production fundamentals (S.57) as flat as possible across the distance grid cells from the pre-war CBD.

To characterize the properties of the adjusted production fundamentals moment conditions in the parameter space, we undertake a grid search over 20 possible values of λ (from 0.02 to 0.10) and 20 possible values of δ (from 0.01 to 0.101) for our estimated values of ν and ε (400 parameter configurations). In Figure S.4, we display the sum of the squared mean changes in log adjusted production fundamentals across the distance grid cells for our baseline specification pooling division and reunification. We construct contours through this sum of squared mean changes in log adjusted production fundamentals for the 400 values of $\{\lambda, \delta\}$ in the parameter space (shown on the horizontal and vertical axes). As shown in the figure, the sum of the squared mean changes in log adjusted production fundamentals has a unique global minimum in the parameter space that identifies $\{\lambda, \delta\}$.

Similarly, the division of Berlin implies a fall in residential externalities (Ω_i) for the parts of West Berlin close to the Berlin Wall. If this fall in residential

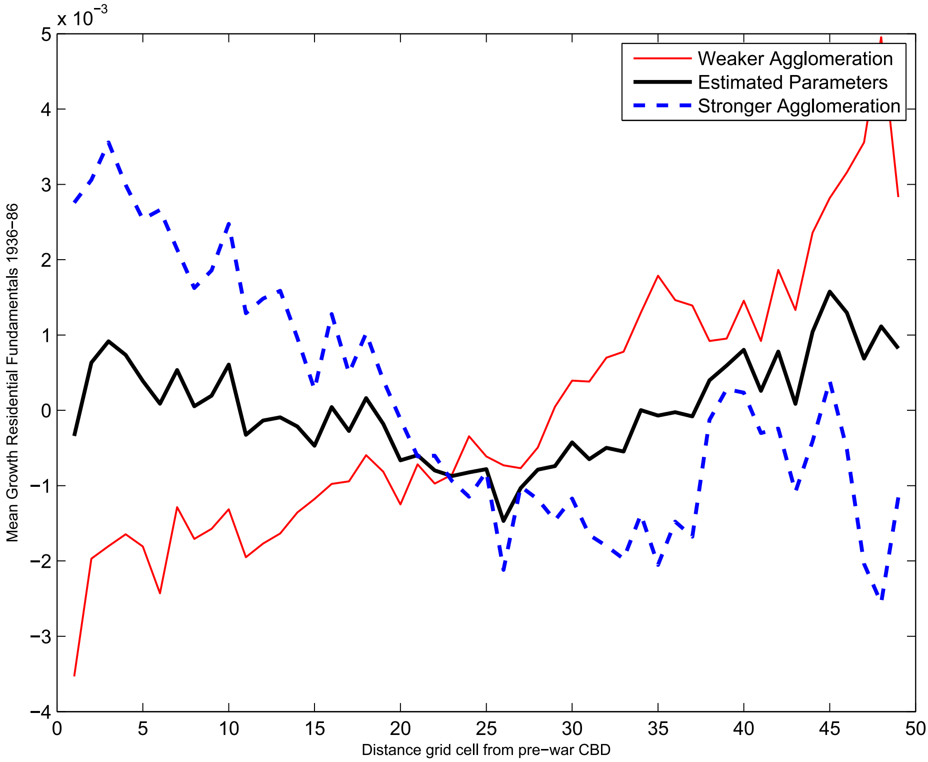

FIGURE S.5.—Residential fundamentals distance grid cell moments (division).

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 from Eastern concentrations of residence employment. In Figure S.5, we show the mean changes in log adjusted residential fundamentals following division across the distance grid cells from the pre-war CBD used in the estimation (i) for the estimated parameters, (ii) for stronger agglomeration forces (larger η and ρ than estimated), and (iii) for weaker agglomeration forces (smaller η and ρ than estimated). From the moment condition (S.62), the residential spillover parameters $\{\eta, \rho\}$ are chosen to make the mean changes in log adjusted residential fundamentals (S.58) as flat as possible across the distance grid cells from the pre-war CBD.

To characterize the properties of the adjusted residential fundamentals moment conditions in the parameter space, we undertake a grid search over 20 possible values of η (from 0.11 to 0.18) and 20 possible values of ρ (from 0.31 to 1.01) for our estimated values of ν and ε (400 parameter configurations).

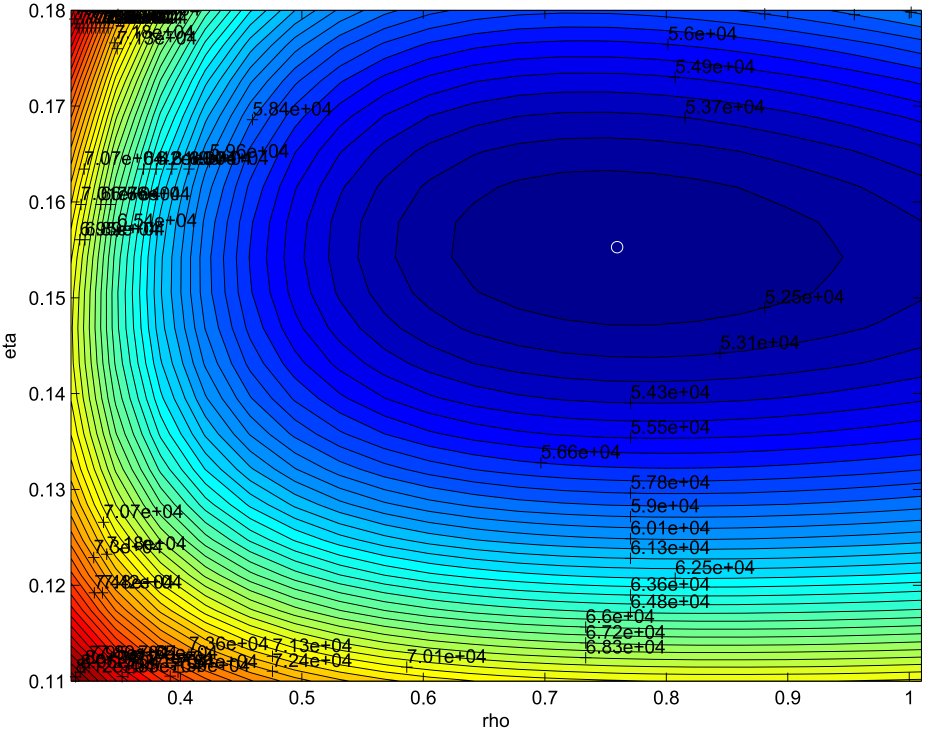

FIGURE S.6.—Residential fundamentals moment condition (sum of squared deviations).

In Figure S.6, we display the sum of the squared mean changes in log adjusted residential fundamentals for our baseline specification pooling division and reunification. We construct contours through this sum of squared mean changes in log adjusted residential fundamentals for the 400 values of $\{\eta, \rho\}$ in the parameter space (shown on the horizontal and vertical axes). As shown in the figure, the sum of the squared mean changes in log adjusted residential fundamentals has a unique global minimum in the parameter space that identifies $\{\eta, \rho\}$.

S.5. COUNTERFACTUALS

Counterfactual Exercises: We consider three sets of counterfactual exercises. First, we simulate the impact of division using the calibrated adjusted fundamentals from 1936. We capture division in the model by assuming infinite costs of trading the final good, infinite commuting costs ($\kappa \to \infty$), infinite rates of decay of production externalities ($\delta \to \infty$), and infinite rates of decay of residential externalities ($\rho \to \infty$) across the Berlin Wall. 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.

Second, we simulate the impact of reunification using calibrated adjusted fundamentals for 1986 for West Berlin and values of these fundamentals for either 2006 or 1936 for East Berlin. We capture reunification in the model by assuming that the costs of trading the final good across the Berlin Wall fall from infinity to zero and by assuming that commuting costs, the decay of production externalities, and the decay of residential externalities across the Berlin Wall fall from infinity to their estimated values. We choose the reservation level of utility in the wider economy following reunification to ensure that the total population of Greater Berlin (H) is equal to its value in the data in 2006.

Third, we simulate the impact of changes in the matrix of bilateral travel times between locations ($\boldsymbol{\tau}$) as a result of a change in transport technology. In particular, we examine the impact of the automobile on the location of economic activity within Berlin, by using the model to solve for the equilibrium distribution of economic activity in 2006 using travel time measures based solely on the public transport network. To focus on the impact of the change in transport technology in Berlin, we hold the reservation utility in the wider economy constant.

We undertake these counterfactuals for the special case of the model with exogenous productivity and amenities ($\lambda = \eta = 0$ and hence $\tilde{\mathbf{A}} = \tilde{\mathbf{a}}$ and $\tilde{\mathbf{B}} = \tilde{\mathbf{b}}$) and for the estimated values of the agglomeration parameters $\{\lambda, \delta, \eta, \rho\}$. We also examine counterfactual changes in the model's agglomeration parameters $\{\lambda, \delta, \eta, \rho\}$, including no production externalities ($\lambda = 0$), no residential externalities ($\eta = 0$), and half the rates of spatial decay of production and residential externalities ($\{\delta, \rho\}$ half their estimated values).

In the special case of the model with exogenous productivity and amenities, there is a unique equilibrium (as shown in Proposition S.2 above). Therefore, these counterfactuals yield determinate predictions for the spatial distribution of economic activity. In contrast, 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. 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. 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.

Our structural estimation exactly explains the observed data, because we solve for unique values of the location fundamentals $\{\tilde{a}_i, \tilde{b}_i, \tilde{\varphi}_i\}$ that replicate the observed data as an equilibrium of the model. In contrast, our counterfactuals assume alternative values for the location fundamentals and agglomeration parameters, and hence the model's predictions are no longer necessarily equal to the observed data. We undertake counterfactuals holding location fundamentals constant at their values before division or reunification to show that the model can explain the observed treatment effects in the data through its agglomeration and dispersion forces rather than through changes in location fundamentals.

Exogenous Location Characteristics: To solve the counterfactuals with exogenous location characteristics, we use the following solution algorithm. We observe land area and travel times $\{K_i, \tau_{ij}\}$. We assume values for the model's parameters $\{\alpha, \beta, \mu, \nu, \varepsilon\}$ and the adjusted location fundamentals $\{\tilde{a}_i, \tilde{b}_i, \tilde{\varphi}_i\}$, where in this special case of the model $\tilde{A}_i = \tilde{a}_i$ and $\tilde{B}_i = \tilde{b}_i$ since $\lambda = \eta = 0$. We assume starting values for floor prices, adjusted wages, and the fraction of adjusted floor space used commercially equal to their values in the observed equilibrium prior to the counterfactual $\{\mathbb{Q}_i^0, \tilde{w}_i^0, \tilde{\theta}_i^0\}$. Given these starting values, we use the equilibrium conditions of the model to solve for new predicted values for these endogenous variables $\{\mathbb{Q}_i^1, \tilde{w}_i^1, \tilde{\theta}_i^1\}$:

$$\pi_{ij}^1 = \frac{\left(d_{ij}\left(\mathbb{Q}_i^0\right)^{1-\beta}\right)^{-\varepsilon}\left(\tilde{B}_i \tilde{w}_j^0\right)^{\varepsilon}}{\sum_{r=1}^{S}\sum_{s=1}^{S}\left(d_{rs}\left(\mathbb{Q}_r^0\right)^{1-\beta}\right)^{-\varepsilon}\left(\tilde{B}_r \tilde{w}_s^0\right)^{\varepsilon}}, \tag{S.71}$$

$$\pi_{ij|i}^1 = \frac{(\tilde{w}_j/d_{ij})^{\varepsilon}}{\sum_{s=1}^{S}(\tilde{w}_s/d_{is})^{\varepsilon}}, \tag{S.72}$$

$$H_{Ri}^1 = \sum_{s=1}^{S}\pi_{is}^1 H, \tag{S.73}$$

$$H_{Mi}^1 = \sum_{r=1}^{S}\pi_{ri}^1 H, \tag{S.74}$$

$$Y_i^1 = \tilde{A}_i\left(H_{Mi}^1\right)^{\alpha}\left(\tilde{\theta}_i^0\tilde{\varphi}_i K_i^{1-\mu}\right)^{1-\alpha}, \tag{S.75}$$

$$\tilde{w}_i^1 = \frac{\alpha Y_i^1}{H_{Mi}^1}, \tag{S.76}$$

$$\bar{\bar{v}}_i^1 = \mathbb{E}\big[\tilde{w}_s^0 | i\big] = \sum_{s=1}^{S} \pi_{is|i}^1 \tilde{w}_s^0, \tag{S.77}$$

$$\mathbb{Q}_i^1 = \frac{(1-\alpha)Y_i^1}{\tilde{\theta}_i^0 \tilde{\varphi}_i K_i^{1-\mu}}, \qquad i \in \Im_M \cup \Im_S = \{\tilde{A}_i > 0, \tilde{B}_i = 0\} \cup \{\tilde{A}_i > 0, \tilde{B}_i > 0\}, \tag{S.78}$$

$$\mathbb{Q}_i^1 = \frac{(1-\beta)\bar{\bar{v}}_i^1 H_{Ri}^1}{\big(1-\tilde{\theta}_i^0\big)\tilde{\varphi}_i K_i^{1-\mu}}, \qquad i \in \Im_R \cup \Im_S = \{\tilde{A}_i = 0, \tilde{B}_i > 0\} \cup \{\tilde{A}_i > 0, \tilde{B}_i > 0\}, \tag{S.79}$$

$$\tilde{\theta}_i^1 = 1, \quad i \in \Im_M = \{\tilde{A}_i > 0, \tilde{B}_i = 0\}, \tag{S.80}$$

$$\tilde{\theta}_i^1 = 0, \quad i \in \Im_R = \{\tilde{A}_i = 0, \tilde{B}_i > 0\}, \tag{S.81}$$

$$\tilde{\theta}_i^1 = \frac{(1-\alpha)Y_i^1}{\mathbb{Q}_i^1 \tilde{\varphi}_i K_i^{1-\mu}}, \quad i \in \Im_S = \{\tilde{A}_i > 0, \tilde{B}_i > 0\}. \tag{S.82}$$

If the new predicted values for the endogenous variables of the model are equal to the starting values:

$$\big\{\mathbb{Q}_i^1, \tilde{w}_i^1, \tilde{\theta}_i^1\big\} = \big\{\mathbb{Q}_i^0, \tilde{w}_i^0, \tilde{\theta}_i^0\big\},$$

we have found the counterfactual equilibrium. If the new predicted values for the endogenous variables of the model are not equal to the starting values, we update the endogenous variables of the model using a weighted average of the starting values and the new predicted values:

$$\begin{aligned} \mathbb{Q}_i^2 &= \varsigma \mathbb{Q}_i^0 + (1-\varsigma)\mathbb{Q}_i^1, \\ \tilde{w}_i^2 &= \varsigma \tilde{w}_i^0 + (1-\varsigma)\tilde{w}_i^1, \\ \tilde{\theta}_i^2 &= \varsigma \tilde{\theta}_i^0 + (1-\varsigma)\tilde{\theta}_i^1, \end{aligned} \tag{S.83}$$

where $0 < \varsigma < 1$. We continue to solve the above system of equations for the equilibrium conditions of the model until the endogenous variables converge to the counterfactual equilibrium.

As shown in Lemmas S.1–S.3, any location $i \in \Im_S = \{\tilde{A}_i > 0, \tilde{B}_i > 0\}$ with strictly positive values of both productivity and amenities remains incompletely specialized; any location $i \in \Im_M = \{\tilde{A}_i > 0, \tilde{B}_i = 0\}$ with strictly positive productivity and zero amenities remains completely specialized in commercial activity; and any location $i \in \Im_R = \{\tilde{A}_i = 0, \tilde{B}_i > 0\}$ with zero productivity and strictly positive amenities remains completely specialized in residential activity. Furthermore, as shown in Proposition S.2, the general equilibrium of the model with exogenous location characteristics is unique. Therefore, the above algorithm converges rapidly to this unique equilibrium, and these counterfactuals yield determinate predictions for the spatial distribution of economic activity.

Endogenous Agglomeration Forces: To solve the counterfactuals with endogenous agglomeration forces, we use a directly analogous solution algorithm. We observe land area and travel times $\{K_i, \tau_{ij}\}$. We assume values for the model's parameters $\{\alpha, \beta, \mu, \nu, \varepsilon, \lambda, \delta, \eta, \rho\}$ and the adjusted location fundamentals $\{\tilde{a}_i, \tilde{b}_i, \tilde{\varphi}_i\}$. We assume starting values for floor prices, adjusted wages, and the fraction of adjusted floor space used commercially equal to their values in the observed equilibrium prior to the counterfactual $\{\mathbb{Q}_i^0, \tilde{w}_i^0, \tilde{\theta}_i^0\}$. Given these starting values, we use the equilibrium conditions of the model to solve for new predicted values for these endogenous variables $\{\mathbb{Q}_i^1, \tilde{w}_i^1, \tilde{\theta}_i^1\}$. We expand the equilibrium conditions of the model (S.71)–(S.82) to include the endogenous determination of adjusted productivity and amenities as a function of production and residential externalities:

$$\text{(S.84)} \qquad \Upsilon_i^1 = \sum_{s=1}^{S} e^{-\delta\tau_{is}} \left(\frac{H_{Ms}^1}{K_s} \right),$$

$$\text{(S.85)} \qquad \tilde{A}_i^1 = \tilde{a}_i \left(\Upsilon_i^1 \right)^{\lambda},$$

$$\text{(S.86)} \qquad \Omega_i^1 = \sum_{s=1}^{S} e^{-\rho\tau_{is}} \left(\frac{H_{Rs}^1}{K_s} \right),$$

$$\text{(S.87)} \qquad \tilde{B}_i^1 = \tilde{b}_i \left(\Omega_i^1 \right)^{\eta}.$$

We continue to solve the system of equations for the equilibrium conditions of the model until the endogenous variables converge to the counterfactual equilibrium. As discussed above, in the presence of endogenous agglomeration forces, there is the potential for multiple equilibria in the model. Our use and updating of the endogenous variables from the observed equilibrium as our initial guess for the counterfactual equilibrium implies that our equilibrium selection rule is to solve for the closest counterfactual equilibrium to the observed equilibrium. Our goal in these counterfactuals is not to determine the unique impact of division or reunification, but rather to examine whether the model with the estimated agglomeration and dispersion forces is capable of generating counterfactual treatment effects close to the observed data.

Transport Technology Counterfactual: 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 the impact of the automobile on the location of economic activity within Berlin, by using 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.

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. Nevertheless, the public transport network is far more extensive in Berlin than in American cities (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 S.I of this supplement 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 from 51 minutes to 70 minutes,

TABLE S.I

TRANSPORT TECHNOLOGY COUNTERFACTUAL[a]

Row	Variable	Transport Technology Counterfactual
	Unweighted Actual Travel Time (mins)	
1.	Mean	50.88
2.	Standard Deviation	12.52
	Unweighted Counterfactual Travel Time (mins)	
3.	Mean	69.65
4.	Standard Deviation	26.19
	Weighted Actual Travel Time (mins)	
5.	Mean (Actual Commuting Weights)	32.03
	Weighted Counterfactual Travel Time (mins)	
6.	Mean (Actual Commuting Weights)	37.81
	Total City Employment	
7.	Counterfactual/Actual	86.46%
	Total City Output	
8.	Counterfactual/Actual	87.55%
	Log Change in Block Floor Prices	
9.	Mean	−20.04%
10.	Mean Above Median Treatment Block	−29.21%
11.	Mean Below Median Treatment Block	−10.87%
	Weighted Counterfactual Travel Time (mins)	
12.	Mean (Counterfactual Commuting Weights)	29.52

[a]Counterfactual using travel time measures based solely on the public transport network in 2006. Unweighted mean and standard deviation travel times across all possible bilateral connections with positive values of either workplace or residence employment. Weighted mean travel times weight these bilateral travel times by commuting probabilities in either the actual or the counterfactual equilibrium. Above median treatment blocks are those with an above median increase in average unweighted travel times across bilateral connections with positive values of either workplace or residence employment.

and its standard deviation rises from 13 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 Section 6.1 of the paper, 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 employment and output fall by around 14 and 12 percent, respectively (rows 7–8 of the table). Output falls by less than employment, 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 20 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 S.7 of this supplement. 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, this change in worker sorting implies that average travel times weighted by commuting flows in the counterfactual equilibrium (row 12 of Table S.I) 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).

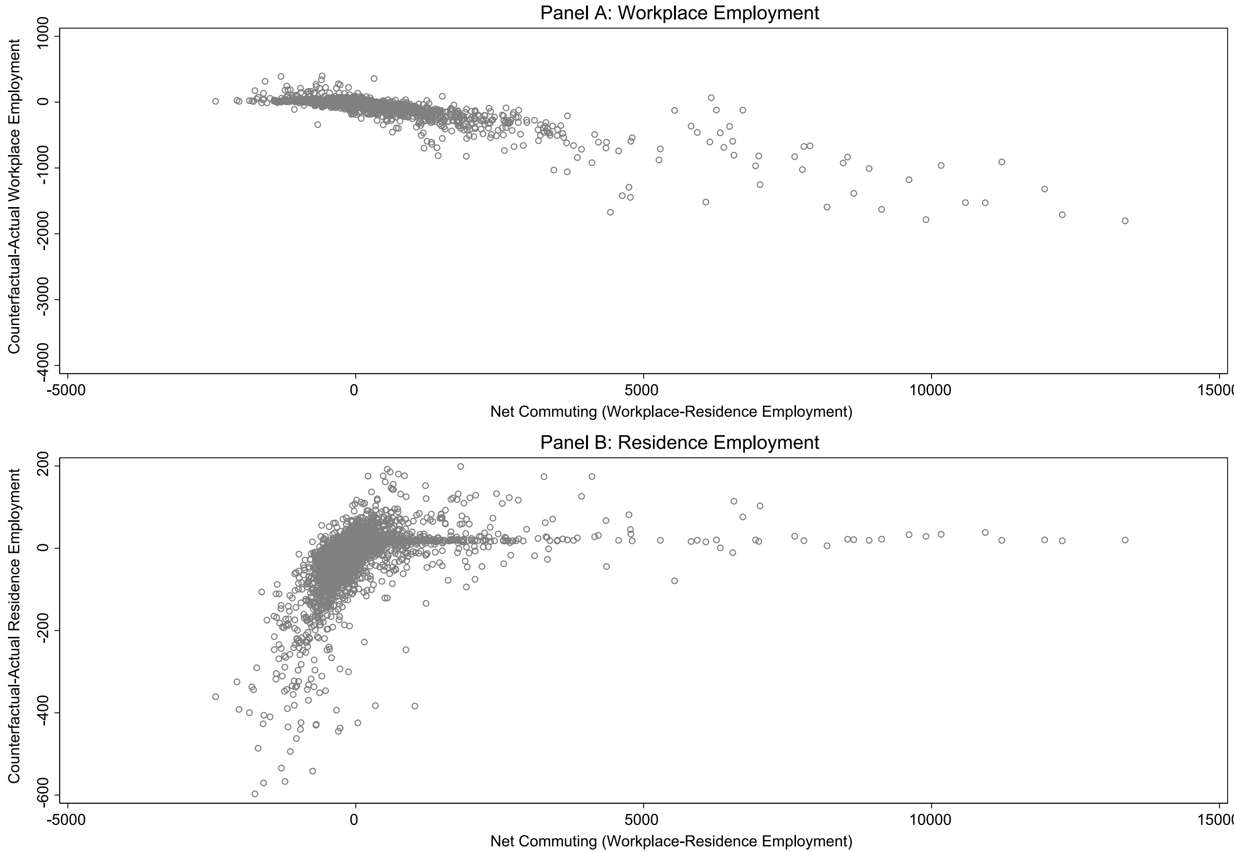

FIGURE S.7.—Simulated change in workplace and residence employment. Note: variables on horizontal and vertical axes normalized to have a mean of zero.

S.6. DATA SOURCES AND DEFINITIONS

The data section of the main paper provides an overview of our various data sources. This section of the supplement provides more detail. Section S.6.1 discusses the construction of the data on employment by residence for the pre-war period. Section S.6.2 discusses the construction of the data on employment by workplace for the pre-war period. Section S.6.3 discusses the construction of the travel times in minutes between blocks. Section S.6.4 discusses the construction of the data on other block characteristics. Section S.6.5 discusses the commuting survey data. Section S.6.6 discusses the comparison of our standard land values data with micro data on property transactions.

S.6.1. *Employment at Place of Residence 1930s*

The 1933 census published data on population for each of the 20 pre-war districts of Berlin and also the population of each street or segment of street in Berlin. We digitized the information on population by street and merged it to the modern block structure. In a first step, we used information on street name changes provided on http://www.luise-berlin.de to convert the historical street names into the modern street names. In a second step, we obtained a data set from the Statistical Office of Berlin ("Senatsverwaltung für Berlin") that contains information on the modern statistical blocks to which each street in Berlin is contiguous. We use this information to distribute the population of each street equally across all blocks which are contiguous with the street. In doing so, we take into account whether blocks are water areas or parks to avoid population being allocated to these blocks. In the case of unmatched streets, we correct misspellings of street names in both data sets by using an algorithm that matches streets within the same district and subdistrict ("Ortsteil") whose names only differ by one letter. In a small number of cases, we are unable to match a street to a block, in which case we spread the street's population equally across all blocks within the same subdistrict that have positive population. Finally, we convert our 1933 estimates of population in each of the modern blocks into employment by residence by using the labor force participation rates for each district ("Bezirk") from the 1933 census.

S.6.2. *Employment at Workplace 1930s*

To estimate the 1933 workplace employment in each modern block, we take a two-step approach. In the first step, we create estimates of 1933 private sector employment in each modern block, and in a second step, we estimate 1933 public sector employment in each block.

For the first step, we use two key data sources. First, the 1933 census published data on employment at workplace in private enterprises in each district of Berlin ("Mitteilungen des Statistischen Amts der Stadt Berlin" 1935). Data at a finer spatial scale were not published in pre-war censuses. Second, we obtained a copy of the 1931 company register of Berlin ("Handelsregister"). The company register contains the name and registered address of each firm in Berlin and in 1931 lists just over 47,000 firms. We have entered the name and address of each of these firms. We use information on street name changes provided on http://www.luise-berlin.de to convert the historical addresses to their modern equivalent. The Statistical Office of Berlin supplied us with a file that lists, for each modern postal address in Berlin, the block in which this address is located. We use this concordance to create a count variable which counts how many firms were registered in each modern block in 1931. Due to incomplete or defective addresses, we managed to allocate 42,818 of the 47,098 firms listed in the company register to a modern block.

To spread total private sector workplace employment across blocks within each district, we first estimate the relationship at the district level between log private sector workplace employment from the 1933 census and the log number of firms from the company register. In Figure S.8, we display the values for these variables for each district as well as the regression relationship between them. As apparent from the figure, we find a close relationship between private sector workplace employment and the number of firms at the district level, with a regression R^2 of over 0.75. We use the estimated coefficients from this regression and the number of firms in each block to construct a predicted share of that block in total district private sector workplace employment. We then use these predicted employment shares to allocate the district totals across blocks within districts.

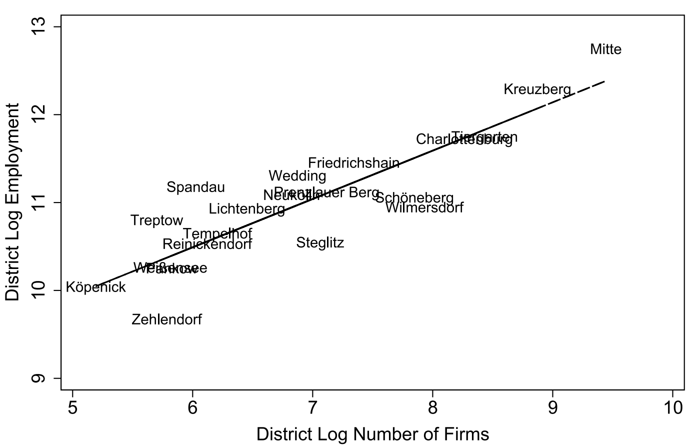

FIGURE S.8.—District employment and number of firms. Note: the graph shows the correlation between the log number of firms in the 1931 company register in each district of Berlin and the log of total private-sector workplace employment in the 1933 census. The R^2 of the regression is 0.77.

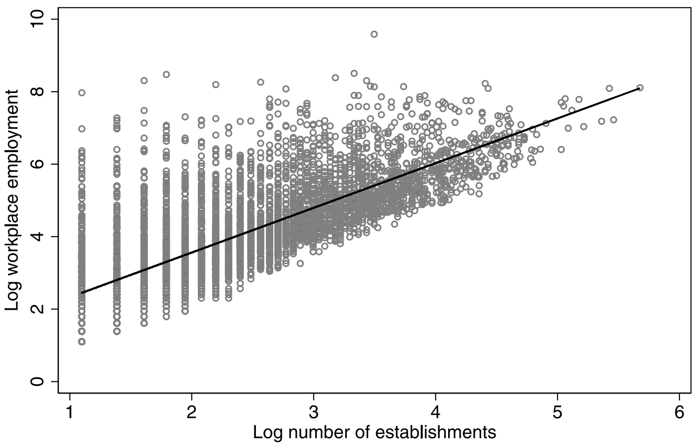

FIGURE S.9.—Block employment and establishments 1987. Note: the graph shows the relationship between log workplace employment and the log number of establishments across blocks in West Berlin (as reported in the 1987 census) and the regression relationship between them. The R^2 of the regression is 0.57.

Since this first step uses predicted employment shares to allocate district totals across blocks within districts, the district totals for private sector workplace employment in our data are the same as in the 1933 census. To further assess the reliability of predicting workplace employment at the block level using information on the number of firms, we use the 1987 census data for West Berlin, which reports both workplace employment and the number of establishments by block. In Figure S.9, we display log workplace employment and the log number of establishments for all blocks with more than three workers and establishments (for confidentiality reasons, the disaggregated totals for blocks are only reported for observations with more than three cases). As apparent from the figure, we also find a close relationship between these two variables at the block level, with a regression R^2 of 0.57.

In the second step, we construct public sector employment in 1933 for each modern block by combining data from the 1933 census with detailed information on the location of public buildings prior to the Second World War. The occupational census of 1933 reports city-level totals of public sector employees and their breakdown into subcategories such as civil servants in the federal and city administration, primary and secondary school teachers, police officers, or clergymen. To allocate these occupation-specific totals for Berlin across blocks, we used a detailed street map of Berlin showing the location of each public building prior to the Second World War and its purpose (e.g., federal government ministries, public utilities, schools). This map was compiled by the Allied occupation authorities in 1945 (War Office (1945)). We allocate the employment of each occupational group (e.g., primary school teachers) across the buildings in which workers from this group are typically employed (e.g., primary schools).

S.6.3. *Travel Times Between Blocks in Berlin*

To determine commuting costs in the model, we need to know the minimum travel time between each of the 15,937 blocks of Berlin in our data, that is, nearly 254 million (15,937$\times$15,937) bilateral connections. We have computed these travel times for 1936, 1986, and 2006. In 1936, commuting to work by car was rare, and hence we construct minimum travel times using the public transport network.[12] In 1986 and 2006, we construct minimum travel times by combining information on the public transport network and driving times by car.

To construct minimum travel times between each pair of blocks i and j by public transport for the three years, we collected information on the underground rail ("U-Bahn"), suburban rail ("S-Bahn"), tram ("Strassenbahn"), and bus ("Bus") network of Berlin in each year. These networks were digitized using ArcGIS and we used the ArcGIS Network Analyst to compute the fastest connection between locations i and j. In this computation, we allow passengers to combine all modes of public transport and walking to minimize the travel time between i and j. We use the following assumed travel speeds for each mode of transport: 5 km/h for walking, 25 km/h for underground and suburban rail travel, 14.5 km/h for trams, and 14.3 km/h for buses. Whenever passengers change between modes of transport (e.g., changing from the suburban rail to a bus), we assume that 3 minutes are lost in waiting time at each connection point. These speeds of travel are taken from Vetter (1928). We assume that these speeds are the same in all three years of our data set, which is supported by comparing these travel speeds to current public transport timetables. Note that these assumptions imply that the travel times from i to j and j to i are the same.

To construct minimum driving times by car between each pair of blocks i and j for 1986 and 2006, we obtained an ArcGIS shapefile of the modern street network of Berlin from a commercial geographical data provider "Geofabrik" (www.geofabrik.de). This shapefile contains information on the maximum and average speed on all streets in and around Berlin, and also restrictions on driving such as one-way streets or prohibited turns. Therefore, the driving times from i to j and from j to i do not have to be the same, because of one-way streets and other similar restrictions on road traffic. Using the ArcGIS Network Analyst, we computed the minimum driving times between all pairs of locations i and j using the average travel speed on each street. As a check on

[12] Leyden (1933) reported 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.

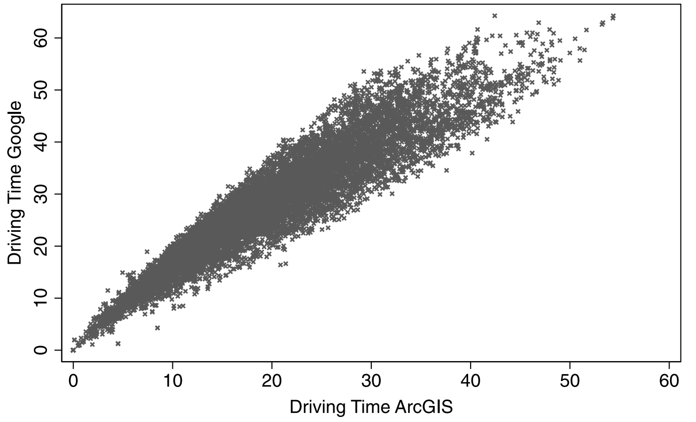

FIGURE S.10.—ArcGIS versus Google. Note: the graph shows bilateral travel times in minutes for 100 randomly selected blocks in Berlin. The travel time for these 10,000 bilateral connections was computed using both Google's public use license and the ArcGIS Network Analyst. The correlation between these two measures of driving time is 0.94.

our ArcGIS calculations, we compared the resulting bilateral minimum driving times for 100 randomly selected blocks to those computed using Google Maps.[13]

Figure S.10 shows a scatter plot of our ArcGIS travel times and the Google travel times for the 10,000 bilateral connections between the 100 randomly chosen blocks. The correlation between the two estimates of bilateral driving times within Berlin is nearly 0.94. Our estimates of the car travel times are slightly lower than Google's, with the median difference being 7.9 minutes. To compute the 1986 car travel times, we restricted the road network to streets in West Berlin. We also adjusted the shapefile to account for the small number of changes in the main road network of West Berlin between 1986 and 2006.[14]

To combine the minimum travel times by public transport and car in 1986 and 2006 into a single travel time measure, we use data on the proportion of journeys undertaken by these two modes of transport in Berlin. In particular, we use information on the modal split of commuting journeys in each of the 12 present-day districts of Berlin from Senatsverwaltung für Stadtentwicklung (2011). In these data, the average share of journeys by car is about one third. We use these data to estimate a simple logit regression that explains the share of journeys undertaken by car in each of the 12 modern districts as a function

[13]Under its public use license, Google restricts users to a small number of requests for travel times per day. Using Google's public use license to compute all 254 million bilateral driving times would have taken several years.

[14]The main changes to the urban motorway system between 1986 and 2006 include the extension of the A113 between Adlershof and Kreuz Schönefeld, two small extensions of the A100, and the construction of the A111.

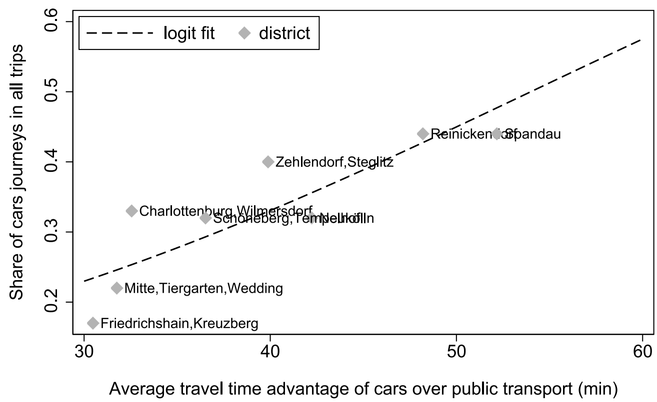

FIGURE S.11.—Car journeys in overall trips. Note: the graph shows the results of a logit regression that relates the share of car journeys in overall journeys in a district in 2010 to the average time saving from undertaking a trip originating in this district by car rather than public transport. See the main text for data sources and details of the estimation.

of the average difference in driving times by car and public transport between blocks in this district and any other block in Berlin. In particular, we estimate the following regression:

$$\ln\left(\frac{\text{car}_d}{1-\text{car}_d}\right) = \beta_1 + \beta_2 \Delta_d + \varepsilon_d, \tag{S.88}$$

where d indexes districts, car_d is the share of journeys undertaken by car, and Δ_d is the average difference in travel times between public transport and driving over all bilateral block connections involving district d. Figure S.11 displays the fitted values from this regression against the actual values of the data.

Using the parameter estimates from this regression, we predict the share of journeys undertaken by car for each bilateral commute between two blocks in Berlin. Our final estimate of the average travel time between two blocks i and j is the weighted average of the car and public transport travel times using the predicted car and public transport shares as weights. We use the same weights to combine the public transport and car travel times for 1986.[15]

S.6.4. *Block Characteristics*

We have collected data on observable block characteristics from a number of sources, as discussed below.

[15]We were unable to find data on the modal split of commuting journeys in Berlin in 1986 by district. However, data in Kloas, Kuhfeld, and Kunert (1988) show that the overall share of journeys by car was very similar to 2006.

Block Land Area and Centroids: We used ArcGIS and a shapefile provided by the Statistical Office of Berlin ("Senatsverwaltung") to compute geographic land area in square meters and the centroid of each block for 2006. As discussed in the data section in the paper, we hold the 2006 block structure constant for all years in our data.

Distance to the Nearest U-Bahn and S-Bahn Station: From each block centroid, we compute the straight-line distance in meters to the nearest underground (U-Bahn) and suburban (S-Bahn) railway station in 1936, 1986, and 2006. Shapefiles showing the exact routing (line shapes) of the rail lines as well as the exact locations of the stations (point shapes) in 2006 were provided by the Statistical Office of Berlin ("Senatsverwaltung"). We used historic network plans to identify those stations that did not exist in 1936 or 1986. Scans of historic network plans are available from the website of Berlin Verkehr (www.berliner-verkehr.de). To create shapefiles of the 1936 and 1986 networks, we start from the 2006 shapefiles and delete those parts of the networks that did not exist in 1936 or 1986. In this backward adjustment process, a small number of stations were added that existed in 1936, but were no longer served in 2006 (and 1986).

Green Areas: We used ArcGIS and a shapefile provided by the Statistical Office of Berlin ("Senatsverwaltung") to compute the straight-line distance in meters from each block centroid to the edge of the nearest green area (public parks, forests, and other green public areas in 2005). We also compute the square meters of green area for each block.

Water Areas: We used ArcGIS and a shapefile provided by the Statistical Office of Berlin ("Senatsverwaltung") to compute the straight-line distance in meters from each block centroid to the edge of the nearest canal, river, or lake. We also create a dummy variable for blocks that are adjacent to a canal, river, or lake.

Schools: We used ArcGIS and a shapefile provided by the Statistical Office of Berlin ("Senatsverwaltung") to compute the straight-line distance in meters from each block centroid to the nearest school in 2006.

Noise: To capture the average noise level within a block, we used ArcGIS and data provided by the Statistical Office of Berlin ("Senatsverwaltung") on the (estimated) average noise levels expressed in decibels (db) for 10×10 meter grid cells. For each block, we compute the average noise level across the grid cells that fall within the block.

Land Use: The 2006 land value map published by the Committee of Valuation Experts ("Gutachterausschuss für Grundstückswerte") defines zones ("Bodenrichtwertzonen") that are homogeneous in terms of land value, building density, and land use. We use this map to create four dummy variables for whether the typical land use in a block is commercial, residential, industrial, or mixed.

Second World War Destruction: We constructed the share of the built-up area in a block that was destroyed during the Second World War as reported on a map from the Agency for Cartography of Berlin in 1945 ("Gebäudeschäden im Gebiet der Stadt Berlin, Stand 1945, Topographische Karte 1:25000," Herausgeber: Hauptamt für Vermessung der Stadt Berlin).

Listed Buildings: The number of listed buildings in a block in 2008 was created based on a shapefile provided by the Statistical Office of Berlin ("Senatsverwaltung").

Urban Regeneration Policies Post-Reunification: We constructed three dummy variables indicating whether a block belongs to a renewal area ("Sanierungsgebiet") designated in 2002, a renewal area designated in 2005, or an area of urban restructuring ("Stadtumbau West") designated in 2005. Two shapefiles showing the exact boundaries of renewal areas and areas of urban restructuring were provided by the Statistical Office of Berlin ("Senatsverwaltung").

Government Buildings Post-Reunification: We construct a dummy variable for whether a major government building was located in a block in 2006 using a map showing all government buildings in 2006 provided by the Statistical Office of Berlin ("Senatsverwaltung").

Wages: The Statistical Yearbook of Berlin reports mean wages for each district of West Berlin in 1986. The data refer to mean annual wages of blue collar workers working in the manufacturing industry.

S.6.5. *Commuting Survey Data*

Micro Commuting Survey Data 2008: Ahrens, Liesske, Wittwer, and Hubrich (2009) reported the results of a representative commuting survey in Berlin and several other large German cities for 2008. The survey records, for each trip that a respondent makes on the day of the survey, the start and end district, time travelled in minutes, and purpose of the trip. In these data, we observe for Berlin 7,984 journeys between a worker's place of residence and her place of work. We use these data to construct a matrix of bilateral commuting probabilities between the 12 districts of Berlin in 2008.

We also use these data to construct the fractions of commuters with travel times in the following eight bins: 0–10, 10–20, 20–30, 30–40, 40–50, 50–60, 60–75, and 75–90 minutes. In constructing the fractions of commuters for these travel time bins, we exclude the negligible fraction of workers that commute for longer than 90 minutes (in one direction from residence to workplace), who are likely to be influenced by factors outside the model. We therefore construct the fractions of commuters for these travel time bins conditional on commuting for 90 minutes or less, so that the fractions add up to 1.

Commuting Survey Data 1982: Brög (1982) reported the results of a representative commuting survey of 27,560 households in West Berlin and West Germany. The data include 291 households in West Berlin and report travel times in minutes between place of residence and place of work. We use these data to construct the fraction of commuters with travel times in the following eight bins: 0–10, 10–20, 20–30, 30–40, 40–50, 50–60, 60–75, and 75–90 minutes.

We again exclude the negligible fraction of workers that commute for longer than 90 minutes (in one direction from residence to workplace) and construct the commuting fractions conditional on commuting for 90 minutes or less so that they add up to 1.

Commuting Survey Data 1930s: The pre-war commuting data were taken from Feder (1939). In the second half of the 1930s, Gottfried Feder carried out a survey of commuting in Berlin. He surveyed a total of 24,336 workers across eight work locations in Berlin that he intended to be representative for the city and which included industry, service, and public sector employers. He asked respondents for the travel time in minutes from their place of residence to their place of work. We use these data to construct the fraction of commuters with travel times in the following six bins: 0–20, 20–30, 30–45, 45–60, 60–75, and 75–90 minutes. We again exclude the negligible fraction of workers that commute for longer than 90 minutes (in one direction from residence to workplace) and construct the commuting fractions conditional on commuting for 90 minutes or less so that they add up to 1.

S.6.6. *Micro Data on Property Transactions*

We follow the standard approach in the urban literature of assuming that floor space L is supplied by a competitive construction sector that uses geographic 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: $L_i = M_i^{\mu} K_i^{1-\mu}$. We now show that these assumptions are consistent with the micro data on property transactions for Berlin from 2000 to 2012.

From the first-order condition for profit maximization in the construction sector, the ratio of capital to land area depends on the price of land relative to the price of capital:

$$\frac{M_i}{K_i} = \frac{\mu}{1-\mu} \frac{\mathbb{R}_i}{\mathbb{P}}, \tag{S.89}$$

where $\mathbb{R}_i$ is the price of land and $\mathbb{P}$ is the common price of capital across all locations. From the zero-profit condition, total revenue from floor space equals total payments to capital and land:

$$\frac{\mathbb{Q}_i L_i}{K_i} = \frac{\mathbb{P} M_i + \mathbb{R}_i K_i}{K_i}. \tag{S.90}$$

Combining these two conditions, total floor space multiplied by the price of floor space and divided by land area is a linear transformation of the price of land:

$$\frac{\mathbb{Q}_i L_i}{K_i} = \frac{1}{1-\mu} \mathbb{R}_i. \tag{S.91}$$

We have obtained access to confidential data from the Committee of Valuation Experts, which contain property transactions in Berlin from 2000 to 2012. In this data set, we observe transaction prices, which correspond to $\mathbb{Q}_i L_i$, as well as the corresponding lot sizes, which correspond to K_i. We compare property prices divided by the lot size in these transactions data to the standard land values reported by the Committee that are used in our empirical analysis. This comparison serves two purposes. First, a strong correlation will indicate that the standard land values provided by the Committee are truly reflective of the market valuation. Second, an approximately linear relationship will suggest that the Cobb–Douglas functional form is a reasonable approximation for the construction sector in Berlin.

We adjust the observed prices in the property transactions data to 2006 prices using a Case–Shiller type repeated sales approach at the block level. Unlike in conventional hedonic analysis of $\mathbb{Q}_i$, there is no need to correct for housing attributes like the number of bathrooms or bedrooms because $(\mathbb{Q}_i L_i)/K_i$ is directly observed in the data. However, since housing is durable and depreciates over time, it is important to control for the age of the building stock. We use the following specification to predict the average (log) price for a newly developed property per unit of geographic area (in meters squared) in a block in 2006 prices:

$$\ln(V_{kt}) = \sum_{m} b_m^{AD} AD_m + \sum_{n \neq 2006} b_n^{YD} YD_n + \Phi_i + \varepsilon_{kt}, \tag{S.92}$$

where k indexes properties; i indexes blocks and t indexes time; V_{kt} is the transaction price of a property k sold in year t divided by its lot size (land area); AD_m is a full set of dummies for ten age cohorts m (0–5 years is the base category); YD_n is a full set of dummies for year n (2006 is the base category); ε_{kt} is a stochastic error; and Φ_i is a time-invariant block specific fixed effect, which we recover for further analysis.

The parameter estimates are reported in Table S.II. The transaction data set has sufficient observations to recover block fixed effects for 8,907 blocks. Figure S.12 provides a comparison of this measure of $(\mathbb{Q}_i L_i)/K_i$ to the 2006 land values assessed by the Committee of Valuation Experts, which correspond to $\mathbb{R}_i$. As predicted by our framework, the two measures are log-linearly related with a slope of approximately 1. The transactions data we obtained also allow for a validation of the ratio of floor space to land area (GFZ) reported by the Committee. Figure S.13 compares the mean floor area divided by the lot size for each block in the property transactions data to the values reported by the Committee. Again, we find that the two variables are closely correlated with a log slope of approximately 1.

TABLE S.II

MICRO DATA ON PROPERTY TRANSACTIONS[a]

	ln(Transaction Price/Lot Size)	
	Coefficient	Standard Error
Year: 2000	−0.073***	(0.020)
Year: 2001	−0.105***	(0.021)
Year: 2002	−0.142***	(0.020)
Year: 2003	−0.176***	(0.019)
Year: 2004	−0.193***	(0.020)
Year: 2005	−0.124***	(0.017)
Year: 2007	0.082***	(0.017)
Year: 2008	−0.004	(0.017)
Year: 2009	−0.014	(0.019)
Year: 2010	0.065***	(0.018)
Year: 2011	0.203***	(0.018)
Year: 2012	0.279***	(0.023)
Age cohort: 5 to 15 years	−0.297***	(0.031)
Age cohort: 15 to 25 years	−0.569***	(0.035)
Age cohort: 25 to 35 years	−0.783***	(0.035)
Age cohort: 35 to 45 years	−1.014***	(0.036)
Age cohort: 45 to 55 years	−1.155***	(0.038)
Age cohort: 55 to 65 years	−1.021***	(0.039)
Age cohort: 65 to 75 years	−0.988***	(0.033)
Age cohort: 75 to 85 years	−1.058***	(0.035)
Age cohort: 85 to 95 years	−0.967***	(0.039)
Age cohort: 95 to 105 years	−0.822***	(0.039)
Age cohort: 105 to 115 years	−0.823***	(0.042)
Age cohort: 115 to 125 years	−0.859***	(0.048)
Age cohort: 125 to 135 years	−0.892***	(0.070)
Age cohort: 135 to 145 years	−0.883***	(0.071)
Age cohort: 145 to 155 years	−0.854***	(0.135)
Age cohort: 155 to 165 years	−1.419***	(0.206)
Age cohort: 165 to 175 years	−1.481***	(0.509)
Age cohort: 175 to 185 years	−0.430	(0.314)
Age cohort: 185 to 195 years	−0.607***	(0.115)
Age cohort: 195 to 205 years	−0.811*	(0.462)
Age cohort: 205 to 215 years	−0.619**	(0.306)
Block fixed effects	Yes	
R^2	0.769	
Observations	51,275	

[a]This table reports the results of estimating a Case–Shiller type repeated sales specification at the block level using the micro data on property transactions from 2000–2012. Standard errors in parentheses are heteroscedasticity robust and clustered by block. * significant at 10%; ** significant at 5%; *** significant at 1%.

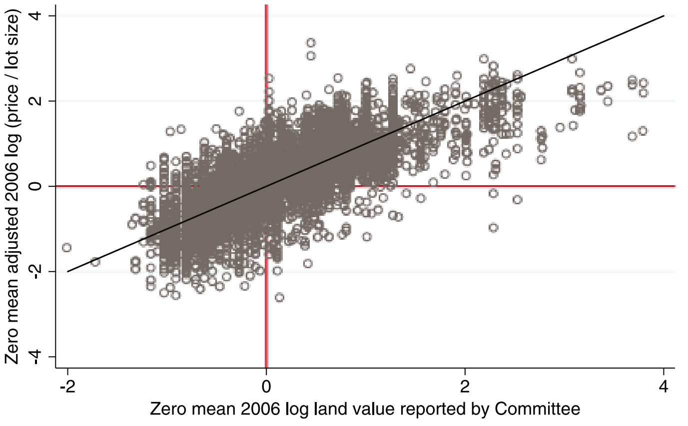

FIGURE S.12.—Transactions prices versus assessed value.

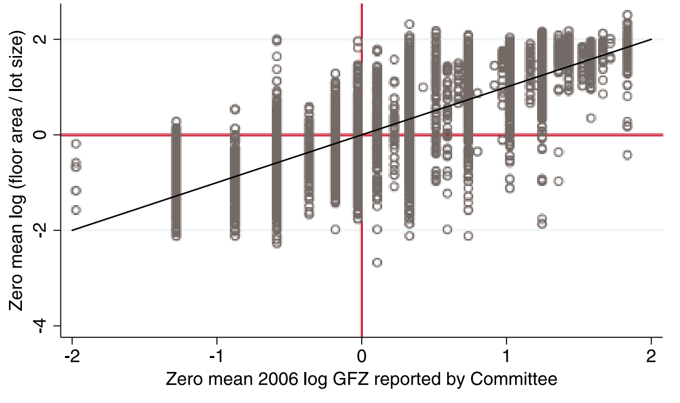

FIGURE S.13.—Density of development versus assessed GFZ.

REFERENCES

AHRENS, G.-A., F. LIESSKE, R. WITTWER, AND S. HUBRICH (2009): "Endbericht zur Verkehrserhebung Mobilität in Städten—SrV 2008 und Auswertungen zum SrV-Städtepegel," Technische Universität Dresden, Dresden. [69]

ALBOUY, D. (2008): "Are Big Cities Bad Places to Live? Estimating Quality of Life Across Metropolitan Areas," Working Paper 14472, NBER. [3]

ALONSO, W. (1964): *Location and Land Use*. Cambridge, MA: Harvard. [22]

AMERICAN EMBASSY (1965): *Handbook of Economic Statistics: Federal Republic of Germany and Western Sectors of Berlin*. Bonn: American Embassy. [10]

BRÖG, W. (1982): "Kontinuierliche Erhebung zum Verkehrsverhalten 1982 (KONTIV 82)," Institut für Verkehrs- und Infrastrukturforschung, München. [69]

CAMERON, A. C., AND P. K. TRIVELDI (2005): *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press. [43]

COMBES, P.-P., G. DURANTON, AND L. GOBILLON (2014): "The Costs of Agglomeration: Land Prices in French Cities," Report, Wharton. [11,39,70]

CONLEY, T. G. (1999): "GMM Estimation With Cross Sectional Dependence," *Journal of Econometrics*, 92 (1), 1–45. [44]

DAVIS, M. A., AND F. ORTALO-MAGNÉ (2011): "Housing Expenditures, Wages, Rents," *Review of Economic Dynamics*, 14 (2), 248–261. [3,39]

DURANTON, G., AND D. PUGA (2004): "Micro-Foundations of Urban Agglomeration Economies," in *Handbook of Regional and Urban Economics*, Vol. 4, ed. by J. V. Henderson and J.-F. Thisse. Amsterdam: Elsevier, 2063–2117. [22]

EATON, J., AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70 (5), 1741–1779. [3]

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. [11,39,70]

FEDER, G. (1939): *Arbeitstätte—Wohnstätte*. Schriftenreihe der Reichsarbeitsgemeinschaft für Raumforschung an der Technischen Hochschule Berlin. Berlin: Julius Springer. [70]

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. [22]

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50 (4), 1029–1054. [43]

KLOAS, J., H. KUHFELD, AND U. KUNERT (1988): "Welche Entwicklung nimmt der Personenverkehr in Berlin-West? Zwei verkehrspolitische Szenarien," *Verkehr und Technik*, 41 (10), 383–395. [67]

LEYDEN, F. (1933): *Gross-Berlin: Geographie einer Weltstadt*. Berlin: Gebr. Mann Verlag. [65]

LUCAS, R. E. (2000): "Externalities and Cities," *Review of Economic Dynamics*, 4 (2), 245–274. [22]

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

MILLS, E. S. (1967): "An Aggregative Model of Resource Allocation in a Metropolitan Centre," *American Economic Review*, 57 (2), 197–210. [22]

MUTH, R. (1969): *Cities and Housing*. Chicago: University of Chicago Press. [22]

ROBACK, J. (1982): "Wages, Rents and the Quality of Life," *Journal of Political Economy*, 90 (6), 1257–1278. [3]

SENATSVERWALTUNG FÜR STADTENTWICKLUNG (2011): "Stadtentwicklungsplan Verkehr," Senatsverwaltung, Berlin. [66]

SVEIKAUSKAS, L. A. (1975): "The Productivity of Cities," *Quarterly Journal of Economics*, 89 (3), 393–413. [22]

VALENTINYI, Á., AND B. HERRENDORF (2008): "Measuring Factor Income Shares at the Sectoral Level," *Review of Economic Dynamics*, 11 (4), 820–835. [39]

VETTER, K. (1928): *Offizieller Führer für Berlin und Umgebung und Potsdam und seine Schlösser*. Berlin: Rotophot. [65]

WAR OFFICE (1945): "Germany Town Plan of Berlin 1945," No. 4480 in Central Intelligence Unit and War Office, Geographical Section. General Staff. [64]

WHITE, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48 (4), 817–838. [44]

Dept. of Geography and Environment, London School of Economics, Houghton Street, London, WC2A 2AE, U.K. and CEPR; g.ahlfeldt@lse.ac.uk,

WWS and Dept. of Economics, Princeton University, Fisher Hall, Princeton, NJ 08540, U.S.A., NBER, and CEPR; reddings@princeton.edu, *Dept. of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, U.K. and CEPR; d.m.sturm@lse.ac.uk,*

and

Dept. of Economics, Humboldt University, Spandauer Strasse 1, 10178 Berlin, Germany and CEPR; nikolaus.wolf@wiwi.hu-berlin.de.

Manuscript received June, 2012; final revision received February, 2015.