Online Appendix for “Evaluating Transport Improvements in Spatial Equilibrium” (Not for Publication)

Stephen J. Redding*
Princeton University, NBER and CEPR

February 2025

Table of Contents

A Introduction

This Online Appendix reports the derivations of results in the main paper and provides additional quantitative results for our numerical example of a city. Section B contains the derivation of results for our baseline quantitative urban model in Section 2 of the paper. Section C gives the derivations for the market access sufficient statistics in Section 4 of the paper. Section D presents additional quantitative results for our numerical example of a city from Section 5 of the paper.

B Baseline Quantitative Urban Model

In this Section of the Online Appendix, we report additional derivations for the baseline quantitative urban model following Ahlfeldt et al. (2015) from Section 2 of the paper.

B.1 Derivation of Residence and Workplace Choices

B.1.1 Distribution of Utility

From the indirect utility function in equation (1) in the paper, we have the following monotonic relationship between idiosyncratic preferences ($b_{ni}(\omega)$) and utility ($u_{ni}(\omega)$):

$$b_{ni}(\omega) = \frac{u_{ni}(\omega)\, \kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}}{B_n w_i}. \tag{B.1}$$

*Dept. Economics and SPIA, JRR Building, Princeton, NJ 08544. Email: reddings@princeton.edu.

We assume that idiosyncratic preferences ($b_{ni}(\omega)$) are drawn from an independent extreme value (Fréchet) distribution for each residence-workplace pair and each worker:

$$G_{ni}(b) = e^{-b^{-\epsilon}}, \qquad \epsilon > 1, \tag{B.2}$$

where we normalize the Fréchet scale parameter in equation (B.2) to one, because it enters worker choice probabilities isomorphically to residential amenities B_n.

Together equations (B.1) and (B.2) imply that the distribution of utility for residence n and workplace i is:

$$G_{ni}(u) = e^{-\Psi_{ni}u^{-\epsilon}}, \qquad \Psi_{ni} \equiv (B_n w_i)^{\epsilon} \left(\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon}. \tag{B.3}$$

From all possible pairs of residence and workplace, 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 residence and workplace is:

$$1 - G(u) = 1 - \prod_{k \in \mathbb{N}} \prod_{\ell \in \mathbb{N}} e^{-\Psi_{k\ell} u^{-\epsilon}},$$

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 residence and employment locations. Therefore we have:

$$G(u) = e^{-\Psi_{\mathbb{N}} u^{-\epsilon}}, \qquad \Psi_{\mathbb{N}} = \sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \Psi_{k\ell}. \tag{B.4}$$

Given this Fréchet distribution for utility, expected utility is:

$$\mathbb{E}\left[u\right] = \int_0^{\infty} \epsilon \Psi_{\mathbb{N}} u^{-\epsilon} e^{-\Psi_{\mathbb{N}} u^{-\epsilon}} du. \tag{B.5}$$

Now define the following change of variables:

$$y = \Psi_{\mathbb{N}} u^{-\epsilon}, \qquad dy = -\epsilon \Psi_{\mathbb{N}} u^{-(\epsilon+1)} du. \tag{B.6}$$

Using this change of variables, expected utility can be written as:

$$\mathbb{E}\left[u\right] = \int_0^{\infty} \Psi_{\mathbb{N}}^{1/\epsilon} y^{-1/\epsilon} e^{-y} dy, \tag{B.7}$$

which can be in turn written as:

$$\mathbb{E}\left[u\right] = \vartheta \Psi_{\mathbb{N}}^{1/\epsilon}, \qquad \vartheta = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right), \tag{B.8}$$

where $\Gamma(\cdot)$ is the Gamma function. Therefore we obtain the expression in equation (7) in the main text:

$$U = \mathbb{E}\left[u\right] = \vartheta \Psi_{\mathbb{N}}^{1/\epsilon} = \vartheta \left[\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} (B_k w_\ell)^{\epsilon} \left(\kappa_{k\ell} P_k^{\alpha} Q_k^{1-\alpha}\right)^{-\epsilon}\right]^{1/\epsilon}. \tag{B.9}$$

B.1.2 Commuting Probabilities

Using the distribution of utility for pairs of residence and employment locations, the probability that a worker chooses the bilateral commute from n to i out of all possible bilateral commutes is:

$$\begin{aligned}\lambda_{ni} &= \Pr\left[u_{ni} \geq \max\{u_{k\ell}\}; \forall k, \ell\right], \\ &= \int_0^\infty \prod_{\ell \neq i} G_{n\ell}(u) \left[\prod_{k \neq n} \prod_{\ell \in \mathbb{N}} G_{k\ell}(u)\right] g_{ni}(u) du, \\ &= \int_0^\infty \prod_{k \in \mathbb{N}} \prod_{\ell \in \mathbb{N}} \epsilon \Psi_{ni} u^{-(\epsilon+1)} e^{-\Psi_{k\ell} u^{-\epsilon}} du, \\ &= \int_0^\infty \epsilon \Psi_{ni} u^{-(\epsilon+1)} e^{-\Psi_{\mathbb{N}} u^{-\epsilon}} du.\end{aligned} \tag{B.10}$$

Note that:

$$\frac{d}{du}\left[\frac{1}{\Psi_{\mathbb{N}}} e^{-\Psi_{\mathbb{N}} u^{-\epsilon}}\right] = \epsilon u^{-(\epsilon+1)} e^{-\Psi_{\mathbb{N}} u^{-\epsilon}}. \tag{B.11}$$

Using this result to evaluate the integral above, the probability that the worker chooses to live in location n and work in location i is:

$$\lambda_{ni} = \frac{L_{ni}}{L_{\mathbb{N}}} = \frac{\Psi_{ni}}{\Psi_{\mathbb{N}}} = \frac{\left(B_n w_i\right)^\epsilon \left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B_k w_\ell\right)^\epsilon \left(\kappa_{k\ell} P_k^\alpha Q_k^{1-\alpha}\right)^{-\epsilon}}, \tag{B.12}$$

where L_{ni} is the measure of commuters from residence n to workplace i; $L_{\mathbb{N}}$ is the total measure of workers in the city; and this expression corresponds to equation (4) in the main text.

Summing across workplaces i in equation (B.12), we obtain the probability that a worker chooses to live in residence n (λ_n^R):

$$\lambda_n^R = \frac{R_n}{L_{\mathbb{N}}} = \frac{\sum_{i \in \mathbb{N}} \left(B_n w_i\right)^\epsilon \left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B_k w_\ell\right)^\epsilon \left(\kappa_{k\ell} P_k^\alpha Q_k^{1-\alpha}\right)^{-\epsilon}}, \tag{B.13}$$

where R_n denotes employment by residence in location n; and this corresponds to the expression in equation (5) in the main text.

Similarly, summing across residences n in equation (B.12), we obtain the probability that a worker chooses workplace i (λ_i^L):

$$\lambda_i^L = \frac{L_i}{L_{\mathbb{N}}} = \frac{\sum_{n \in \mathbb{N}} \left(B_n w_i\right)^\epsilon \left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B_k w_\ell\right)^\epsilon \left(\kappa_{k\ell} P_k^\alpha Q_k^{1-\alpha}\right)^{-\epsilon}}, \tag{B.14}$$

where L_i denotes employment by workplace in location i; and this corresponds to the expression in equation (6) in the paper.

For the measure of workers in location i (L_i), we can evaluate the conditional probability that they commute from location n (conditional on having chosen to work in location i):

$$\begin{aligned}\lambda^L_{ni|i} &= \Pr\left[u_{ni} \geq \max\{u_{ri}\}; \forall r\right], \\ &= \int_0^\infty \prod_{r \neq n} G_{ri}(u) g_{ni}(u) du, \\ &= \int_0^\infty e^{-\Psi^L_i u^{-\epsilon}} \epsilon \Psi_{ni} u^{-(\epsilon+1)} du.\end{aligned} \tag{B.15}$$

where

$$\Psi^L_i \equiv \sum_{k \in \mathbb{N}} (B_k w_i)^\epsilon \left(\kappa_{ki} P^\alpha_k Q^{1-\alpha}_k\right)^{-\epsilon}. \tag{B.16}$$

Using the result (B.11) to evaluate the integral in equation (B.15), the probability that a worker commutes from residence n to workplace i conditional on having chosen to work in location i is:

$$\lambda^L_{ni|i} = \frac{\lambda_{ni}}{\lambda^L_i} = \frac{(B_n w_i)^\epsilon \left(\kappa_{ni} P^\alpha_n Q^{1-\alpha}_n\right)^{-\epsilon}}{\sum_{k \in \mathbb{N}} (B_k w_i)^\epsilon \left(\kappa_{ki} P^\alpha_k Q^{1-\alpha}_k\right)^{-\epsilon}}, \tag{B.17}$$

which further simplifies to:

$$\lambda^L_{ni|i} = \frac{B^\epsilon_n \left(\kappa_{ni} P^\alpha_n Q^{1-\alpha}_n\right)^{-\epsilon}}{\sum_{k \in \mathbb{N}} B^\epsilon_k \left(\kappa_{ki} P^\alpha_k Q^{1-\alpha}_k\right)^{-\epsilon}}. \tag{B.18}$$

For the measure of residents of location n (R_n), we can evaluate the conditional probability that they commute to location i (conditional on having chosen to live in location n):

$$\begin{aligned}\lambda^R_{ni|n} &= \Pr\left[u_{ni} \geq \max\{u_{n\ell}\}; \forall \ell\right], \\ &= \int_0^\infty \prod_{\ell \neq i} G_{n\ell}(u) g_{ni}(u) du, \\ &= \int_0^\infty e^{-\Psi^R_n u^{-\epsilon}} \epsilon \Psi_{ni} u^{-(\epsilon+1)} du,\end{aligned} \tag{B.19}$$

where

$$\Psi^R_n \equiv \sum_{\ell \in \mathbb{N}} (B_n w_\ell)^\epsilon \left(\kappa_{n\ell} P^\alpha_n Q^{1-\alpha}_n\right)^{-\epsilon}. \tag{B.20}$$

Using the result (B.11) to evaluate the integral in equation (B.19), the probability that a worker commutes to location i conditional on having chosen to live in location n is:

$$\lambda^R_{ni|n} = \frac{\lambda_{ni}}{\lambda^R_n} = \frac{(B_n w_i)^\epsilon \left(\kappa_{ni} P^\alpha_n Q^{1-\alpha}_n\right)^{-\epsilon}}{\sum_{\ell \in \mathbb{M}} (B_n w_\ell)^\epsilon \left(\kappa_{n\ell} P^\alpha_n Q^{1-\alpha}_n\right)^{-\epsilon}}, \tag{B.21}$$

which further simplifies to:

$$\lambda^R_{ni|n} = \frac{(w_i/\kappa_{ni})^\epsilon}{\sum_{\ell \in \mathbb{N}} (w_\ell/\kappa_{n\ell})^\epsilon}. \tag{B.22}$$

Commuter market clearing requires that the measure of workers employed in each location i (L_i) equals the sum across all locations n of their measures of residents (R_n) times their conditional probabilities of commuting to i ($\lambda^R_{ni|n}$):

$$L_i = \sum_{n\in\mathbb{N}} \lambda^R_{ni|n} R_n \tag{B.23}$$
$$= \sum_{n\in\mathbb{N}} \frac{\left(w_i/\kappa_{ni}\right)^{\epsilon}}{\sum_{\ell\in\mathbb{N}}\left(w_\ell/\kappa_{n\ell}\right)^{\epsilon}} R_n,$$

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 location n equals the wages in all possible workplace locations weighted by the probabilities of commuting to those locations conditional on living in n:

$$v_n = \mathbb{E}\left[w|n\right] \tag{B.24}$$
$$= \sum_{i\in\mathbb{N}} \lambda^R_{ni|n} w_i,$$
$$= \sum_{i\in\mathbb{N}} \frac{\left(w_i/\kappa_{ni}\right)^{\epsilon}}{\sum_{\ell\in\mathbb{N}}\left(w_\ell/\kappa_{n\ell}\right)^{\epsilon}} w_i,$$

where $\mathbb{E}$ denotes the expectations operator and the expectation is taken over the distribution for idiosyncratic preferences. Intuitively, expected worker income is high in locations that have low commuting costs (low κ_{ni}) to high-wage employment locations.

Another implication of the Fréchet distribution of utility is that the distribution of utility conditional on residing in location n and commuting to location i is the same across all bilateral pairs of locations with positive residents and employment, and is equal to the distribution of utility for the economy as a whole. To establish this result, note that the distribution of utility conditional on residing in location n and commuting to location i is:

$$= \frac{1}{\lambda_{ni}} \int_0^u \prod_{s\neq i} G_{ns}(u) \left[\prod_{k\neq n}\prod_{\ell\in\mathbb{N}} G_{k\ell}(u)\right] g_{ni}(u)du, \tag{B.25}$$

$$= \frac{1}{\lambda_{ni}} \int_0^u \left[\prod_{k\in\mathbb{N}}\prod_{\ell\in\mathbb{N}} e^{-\Psi_{k\ell}u^{-\epsilon}}\right] \epsilon\Psi_{ni}u^{-(\epsilon+1)}du,$$
$$= \frac{\Psi_{\mathbb{N}}}{\Psi_{ni}} \int_0^u e^{-\Psi_{\mathbb{N}}u^{-\epsilon}} \epsilon\Psi_{ni}u^{-(\epsilon+1)}du,$$
$$= e^{-\Psi_{\mathbb{N}}u^{\epsilon}}.$$

On the one hand, lower land prices in location n or a higher wage in location i raise the utility of a worker with a given realization of idiosyncratic preferences b, and hence increase the expected utility of residing in n and working in i. On the other hand, lower land prices or a higher wage induce workers with lower realizations of idiosyncratic preferences b to reside in n and work in i, which reduces the expected utility of residing in n and working in i. With a Fréchet distribution of utility, these two effects exactly offset one another. Pairs of residence and employment locations with more attractive characteristics attract more commuters on the extensive margin until expected utility is the same across all pairs of residence and employment locations within the economy.

An implication of this result is that expected utility conditional on choosing a residence n and workplace i is the same across all residence-workplace pairs and equal to expected utility in the economy as a whole in equation (B.9):

$$U = \vartheta \Psi_{\mathbb{N}}^{1/\epsilon} = \vartheta \left[\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \left(B_k w_\ell \right)^{\epsilon} \left(\kappa_{k\ell} P_k^{\alpha} Q_k^{1-\alpha} \right)^{-\epsilon} \right]^{1/\epsilon}. \tag{B.26}$$

B.2 Existence and Uniqueness

We consider a closed-city specification with an exogenous total city population ($L_{\mathbb{N}}$), and segmented markets for floor space, with exogenous supplies for residential floor space (H_n^R) and commercial floor space (H_n^L).

The general equilibrium spatial distribution of economic activity within the city is determined by the model parameters (α, β, κ, ϵ, η_B, δ_B, η_A, δ_A) and the following exogenous location characteristics: residential fundamentals ($\overline{B}_n$), production fundamentals ($\overline{A}_n$), the supplies of residential and commercial floor space (H_n^R, H_n^L), and the transport network (τ_{ni}). Given these parameters and exogenous location characteristics, the closed-city general equilibrium of the model is referenced by residents (R_n), employment (L_n), wages (w_n), average residential income (v_n), the prices of residential and commercial floor space (Q_r, q_r), and expected utility (U), given exogenous total city population ($L_{\mathbb{N}}$). The open-city general equilibrium is analogous, except that total city population is endogenously determined by population mobility with the wider economy and its exogenous reservation level of utility ($\overline{U}$). Given these equilibrium objects, all the other endogenous variables of the model can be determined.

We now provide a sufficient condition for the existence of a unique equilibrium. We combine the general equilibrium conditions of the model to obtain a system of equations that takes the required form to apply Theorem 1 from Allen, Arkolakis and Li (2024):

$$x_{nh} = f_{nih}(x_i) = \sum_{i \in \mathbb{N}} \mathcal{K}_{nih} \prod_{h' \in \mathbb{H}} x_{ih'}^{\gamma_{nhh'}}. \tag{B.27}$$

where $n, i \in \mathbb{N}$ denote locations and $h \in \mathbb{H}$ denote economic interactions, which here include residents, employment, and the prices of residential and commercial floor space. We begin by rewriting each of the general equilibrium conditions in the form required to apply this theorem.

Population Mobility As a preliminary step, note that expected utility (7) can be re-written as:

$$\left(\frac{U}{\vartheta}\right)^{\epsilon^o} = \left[\sum_{n\in\mathbb{N}}\sum_{i\in\mathbb{N}} \left(B_n w_i\right)^{\epsilon} \left(\kappa_{ni} Q_n^{1-\alpha}\right)^{-\epsilon}\right]^{\frac{1}{\epsilon}}, \tag{B.28}$$

where we have used our choice of numeraire ($P_n = 1$). We now use this expression for expected utility (U) to rewrite the other general equilibrium conditions of the model.

Residential Choice Probabilities Using expected utility (7), the residential choice probabilities (5) for each location can be re-written as:

$$R_n = \xi \left(B_n \left(\Phi_n^R\right)^{\frac{1}{\epsilon}} Q_n^{\alpha-1}\right)^{\epsilon}, \tag{B.29}$$

where ξ is an endogenous scalar:

$$\xi \equiv L_{\mathbb{N}} \left(\frac{U}{\vartheta}\right)^{-\epsilon}, \tag{B.30}$$

and Φ_n^R our measure of residential commuting market access that is a commuting cost weighted average of wages in each workplace:

$$\Phi_n^R \equiv \sum_{i\in\mathbb{N}} \kappa_{ni}^{-\epsilon} w_i^{\epsilon}. \tag{B.31}$$

Workplace Choice Probabilities Using expected utility (7), the workplace choice probabilities (6) for each location can be re-written as:

$$L_i = \xi \left(w_i \left(\Phi_t^L\right)^{\frac{1}{\epsilon}}\right)^{\epsilon}, \tag{B.32}$$

where we defined the endogenous scalar ξ in equation (B.30) above and Φ_i^L is a measure of workplace commuting market access that is a commuting cost weighted average and amenities and the price of residential floor space in each residence:

$$\Phi_i^L \equiv \sum_{n\in\mathbb{N}} \kappa_{ni}^{-\epsilon} \left(B_n^o Q_n^{\alpha-1}\right)^{\epsilon}. \tag{B.33}$$

Residential Commuting Market Access From the workplace choice probabilities (B.32), we have the following relationship:

$$(w_i)^\epsilon = \frac{1}{\xi}\frac{L_i}{\Phi_i^L}.$$

Using this relationship, we can re-write residential commuting market access (Φ_n^R) in equation (B.31) as follows:

$$\Phi_n^R = \frac{1}{\xi}\sum_{i\in\mathbb{N}}\kappa_{ni}^{-\epsilon}\frac{L_i}{\Phi_i^L}. \tag{B.34}$$

Workplace Commuting Market Access From the residential choice probabilities (B.29), we have the following relationship:

$$\left(B_n Q_{nt}^{\alpha-1}\right)^\epsilon = \frac{1}{\xi}\frac{R_n}{\Phi_n^R}$$

Using this relationship, we can re-write workplace commuting market access (Φ_i^E) in equation (B.33) as follows:

$$\Phi_i^L = \frac{1}{\xi}\sum_{n\in\mathbb{N}}\kappa_{ni}^{-\epsilon}\frac{R_n}{\Phi_n^R}. \tag{B.35}$$

Output From the Cobb-Douglas production technology, output of the final good implies (Y_i) is:

$$Y_i = A_i\left(\frac{L_i}{\beta}\right)^\beta\left(\frac{H_i^L}{1-\beta}\right)^{1-\beta}. \tag{B.36}$$

Residential Income Using the definition of residential commuting market access from equation (B.31) in equation (B.24), average residential income (v_n) can be written as:

$$\Phi_n^R v_n = \sum_{i\in\mathbb{N}}\kappa_{ni}^{-\epsilon}\left(w_i\right)^{\epsilon+1}. \tag{B.37}$$

Wages From the zero-profit condition for the final good to be produced:

$$w_n = A_n^{\frac{1}{\beta}} q_n^{-\left(\frac{1-\beta}{\beta}\right)}, \tag{B.38}$$

where we have again used our choice of numeraire ($P_n = 1$).

Residential and Commercial Floor Space Prices The prices of residential (Q_n) and commercial (q_n) floor space are determined by the market clearing conditions for residential and commercial floor space in equations (12) and (13), respectively, given the suppliers of residential and commercial floor space (H_n^R, H_n^L).

Amenities and Productivity Amenities (B_n) and productivity (A_n) are determined by fundamentals and externalities according to equations (2) and (9), respectively.

System of General Equilibrium Conditions Using equations (B.31), (B.33), (B.34), (B.35), (B.36), (B.37), (B.38), (12), (13), (2) and (9), the system of general equilibrium conditions of the model can be written in the form of equation (B.27) as follows:

$$R_n = \xi \left(B_n \left(\Phi_n^R \right)^{\frac{1}{\epsilon}} Q_n^{\alpha-1} \right)^{\epsilon}, \tag{B.39}$$

$$L_n = \xi \left(w_n \left(\Phi_n^L \right)^{\frac{1}{\epsilon}} \right)^{\epsilon}, \tag{B.40}$$

$$\Phi_n^R = \frac{1}{\xi} \sum_{i \in \mathbb{N}} e^{-\kappa\epsilon\tau_{ni}} L_i \left(\Phi_i^L \right)^{-1}, \tag{B.41}$$

$$\Phi_n^L = \frac{1}{\xi} \sum_{i \in \mathbb{N}} e^{-\kappa\epsilon\tau_{ni}} R_i \left(\Phi_i^R \right)^{-1}, \tag{B.42}$$

$$w_n = A_n^{\frac{1}{\beta}} q_n^{-\left(\frac{1-\beta}{\beta}\right)}, \tag{B.43}$$

$$\Phi_n^R v_n = \sum_{i \in \mathbb{N}} e^{-\kappa\epsilon\tau_{ni}} \left(w_i \right)^{\epsilon+1}, \tag{B.44}$$

$$Q_n = (1-\alpha) v_n R_n \left(H_n^R \right)^{-1}, \tag{B.45}$$

$$q_n = A_n \left(\frac{L_n}{\beta} \right)^{\beta} \left(\frac{H_n^L}{1-\beta} \right)^{-\beta}, \tag{B.46}$$

$$B_n = \overline{B}_n \left(\sum_{i \in \mathbb{N}} e^{-\delta_B \tau_{ni}} R_i \right)^{\eta_B}, \tag{B.47}$$

$$A_n = \overline{A}_n \left(\sum_{i \in \mathbb{N}} e^{-\delta_A \tau_{ni}} L_i \right)^{\eta_A}, \tag{B.48}$$

where we have used our parameterization of commuting costs as $\kappa_{ni}^{-\epsilon} = e^{-\kappa\epsilon\tau_{ni}}$; and the supplies of residential floor space (H_{nt}^R) and commercial floor space (H_{nt}^L) are exogenous.

The exponents on the variables on the left-hand side of this system of equations can be represented as the following matrix:

$$\mathbf{\Lambda} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \dots & \Lambda_{1H} \\ \Lambda_{21} & \Lambda_{22} & \dots & \Lambda_{2H} \\ \vdots & \vdots & \vdots & \vdots \\ \Lambda_{H1} & \Lambda_{H2} & \dots & \Lambda_{HH} \end{bmatrix}.$$

The exponents on the variables on the right-hand side of this system of equations can be represented as the following matrix:

$$\mathbf{\Gamma} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \dots & \Gamma_{1H} \\ \Gamma_{21} & \Gamma_{22} & \dots & \Gamma_{2H} \\ \vdots & \vdots & \vdots & \vdots \\ \Gamma_{H1} & \Gamma_{H2} & \dots & \Gamma_{HH} \end{bmatrix}$$

Let $\mathbf{\Upsilon} \equiv |\mathbf{\Gamma}\mathbf{\Lambda}^{-1}|$ and denote the spectral radius (eigenvalue with the largest absolute value) of this matrix by $\rho(\mathbf{\Upsilon})$. From Theorem 1 in Allen, Arkolakis and Li (2024), a sufficient condition for the existence of a unique equilibrium (up to scale) is $\rho(\mathbf{\Upsilon}) \leq 1$.

Special Case One special case that is particularly tractable is the case with no residential floor space use ($\alpha = 1$) and no commercial floor space use ($\beta = 1$), as analyzed in Allen, Arkolakis and Li (2024). In this special case, the commuting probability (4) can be written as:

$$L_{ni} = \xi \left(\frac{B_n w_i}{\kappa_{ni}} \right)^{\epsilon}, \tag{B.49}$$

where $\xi \equiv L_{\mathbb{N}} \left(U/\vartheta\right)^{-\epsilon}$; $U = \left[\sum_{k\in\mathbb{N}} \sum_{\ell\in\mathbb{N}} \left(\frac{B_k w_\ell}{\kappa_{k\ell}} \right)^{\epsilon} \right]^{1/\epsilon}$; and the wage is now pinned down by productivity alone:

$$w_i = A_i. \tag{B.50}$$

Commuter market clearing requires:

$$L_n = \sum_{i\in\mathbb{N}} L_{in}, \tag{B.51}$$

$$R_n = \sum_{i\in\mathbb{N}} L_{ni}. \tag{B.52}$$

Substituting the commuting probability (B.49) into these commuter market clearing conditions, and using equations (2) and (9) for amenities (B_n) and productivity (A_n), the system of general equilibrium conditions can be written as:

$$\begin{aligned} L_n A_n^{-\epsilon} &= \xi \sum_{i\in\mathbb{N}} e^{-\kappa\epsilon\tau_{ni}} B_i^{\epsilon}, \\ R_n B_n^{-\epsilon} &= \xi \sum_{i\in\mathbb{N}} e^{-\kappa\epsilon\tau_{ni}} A_i^{\epsilon}, \\ A_n^{\frac{1}{\eta_A}} &= \overline{A}_n^{\frac{1}{\eta_A}} \sum_{i\in\mathbb{N}} e^{-\delta_A \tau_{ni}} L_i, \\ B^{\frac{1}{\eta_B}} &= \overline{B}_n^{\frac{1}{\eta_B}} \sum_{i\in\mathbb{N}} e^{-\delta_B \tau_{ni}} R_i, \end{aligned}$$

where we have used our parameterization of commuting costs as $\kappa_{ni}^{-\epsilon} = e^{-\kappa\epsilon\tau_{ni}}$. The exponents on the variables on the left-hand side of this system of equations can be represented as the following matrix:

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 & -\epsilon & 0 \\ 0 & 1 & 0 & -\epsilon \\ 0 & 0 & \frac{1}{\eta_A} & 0 \\ 0 & 0 & 0 & \frac{1}{\eta_B} \end{bmatrix}.$$

The exponents on the variables on the right-hand side of this system of equations can be represented as the following matrix:

$$\mathbf{\Gamma} = \begin{bmatrix} 0 & 0 & 0 & \epsilon \\ 0 & 0 & \epsilon & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

We thus have:

$$\mathbf{\Upsilon} \equiv \left|\mathbf{\Gamma}\mathbf{\Lambda}^{-1}\right| = \begin{bmatrix} 0 & 0 & 0 & |\eta_B\epsilon| \\ 0 & 0 & |\eta_A\epsilon| & 0 \\ 1 & 0 & |\eta_A\epsilon| & 0 \\ 0 & 1 & 0 & |\eta_B\epsilon| \end{bmatrix}.$$

From the Collatz-Wielandt Formula, a sufficient condition for uniqueness of the equilibrium is hence $|\eta_A|\,\epsilon \leq \frac{1}{2}$ and $|\eta_B|\,\epsilon \leq \frac{1}{2}$, which corresponds to the case in which both the agglomeration forces in the model (as parameterized by production externalities (η_A) and residential externalities (η_B)) are sufficiently weak relative to the dispersion forces from the heterogeneity in idiosyncratic amenities (as parameterized by (ϵ)).

B.3 Counterfactuals

In our numerical example of a city, we consider a counterfactual for the construction of a railway network that reduces bilateral commuting costs (κ_{ni}) between some pairs of locations by more than between other pairs of locations. We consider a closed-city with constant total population ($L_{\mathbb{N}t} = L_{\mathbb{N}}$) and an integrated market for floor space.

In our baseline specification, we assume no agglomeration forces ($\eta_A = \eta_B = \delta_A = \delta_B = 0$), such that productivity and amenities are determined solely by exogenous location characteristics ($A_n = \overline{A}_n$ and $B_n = \overline{B}_n$). In this case, there exists a unique equilibrium, and our counterfactuals yield determinate predictions for the impact of the construction of the railway line on the spatial distribution of economic activity. We also consider an augmented specification with agglomeration forces ($A_n = \overline{A}_n \mathbb{A}_n^{\eta_A}$ and $B_n = \overline{B}_n \mathbb{B}_n^{\eta_B}$). In this case, for sufficiently strong agglomeration and dispersion forces, there is the potential for multiple equilibrium in the model. We solve for a counterfactual equilibrium using starting values for the model's endogenous variables from the initial equilibrium before the construction of the railway network.

In our numerical example, all locations are incompletely specialized ($0 < \theta_n < 1$). The reason is that we assume positive and finite values for production and residential fundamentals for all locations ($\overline{A}_n > 0$ and $\overline{B}_n > 0$). In our specification without agglomeration economies, this assumption ensures that the amenity-adjusted real income is positive and finite for each residence-workplace pair. Since the support of the Fréchet distribution for idiosyncratic worker preferences is unbounded from above, this property ensures positive commuters for each residence-workplace pair.

In our specification with agglomeration economies, we assume positive spillovers of production and residential externalities across locations ($0 < \delta_A < \infty$ and $0 < \delta_B < \infty$). Therefore, even if employment and residents are zero in a location, that location continues to have positive and finite values of productivity and amenities ($A_n > 0$ and $B_n > 0$), because of the spillovers of production and residential externalities across locations. Since the support of the Fréchet distribution for idiosyncratic worker preferences is unbounded from above, this property again ensures positive commuters for each residence-workplace pair.

Finally, in our numerical example, we abstract from land use regulations ($\xi_n = 1$). This assumption, together with incomplete specialization ($0 < \theta_n < 1$), ensures that the prices of residential and commercial floor space are equalized for all locations ($Q_n = q_n$).

B.3.1 No Agglomeration Forces

We begin by considering our baseline specification with no agglomeration forces ($\eta_A = \eta_B = \delta_A = \delta_B = 0$) and exogenous productivity and amenities ($A_n = \overline{A}_n$ and $B_n = \overline{B}_n$). Given model parameters $\{\alpha, \beta, \kappa, \epsilon\}$, known values of location characteristics $\{H_n, \overline{A}_n, \overline{B}_n, \tau_{ni}\}$ and exogenous total city population ($L_{\mathbb{N}}$), we use the following solution algorithm to solve for equilibrium. We assume starting values for the price of floor space, wages and the fraction of floor space used commercially $\{Q_n^0, w_n^0, \theta_n^0\}$. Given these starting values, we use the equilibrium conditions of the model to solve for new predicted values for these endogenous variables $\{Q_n^1, w_n^1, \theta_n^1\}$:

$$\lambda_{ni}^1 = \frac{\left(\kappa_{ni}\left(Q_n^0\right)^{1-\alpha}\right)^{-\epsilon}\left(B_n w_i^0\right)^{\epsilon}}{\sum_{k\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\left(\kappa_{k\ell}\left(Q_k^0\right)^{1-\alpha}\right)^{-\epsilon}\left(B_k w_\ell^0\right)^{\epsilon}}, \tag{B.53}$$

$$\kappa_{ni} = \exp\left(\kappa\tau_{ni}\right), \tag{B.54}$$

$$\lambda_{ni|n}^{R1} = \frac{\left(w_i^0/\kappa_{ni}\right)^{\epsilon}}{\sum_{\ell\in\mathbb{N}}\left(w_\ell^0/\kappa_{n\ell}\right)^{\epsilon}}, \tag{B.55}$$

$$R_n^1 = \sum_{\ell\in\mathbb{N}}\lambda_{n\ell}^1\overline{L}, \tag{B.56}$$

$$L_i^1 = \sum_{k\in\mathbb{N}}\lambda_{ki}^1\overline{L}, \tag{B.57}$$

$$Y_i^1 = A_i \left(L_i^1\right)^{\beta} \left(\theta_i^0 H_i\right)^{1-\beta}, \tag{B.58}$$

$$w_i^1 = \frac{\beta Y_i^1}{L_i^1}, \tag{B.59}$$

$$v_n^1 = \sum_{\ell \in \mathbb{N}} \lambda_{n\ell|n}^{R1} w_\ell^0, \tag{B.60}$$

$$Q_i^1 = \frac{(1-\beta)Y_i^1}{H_i} + \frac{(1-\alpha)v_i^1 R_i^1}{H_i}, \tag{B.61}$$

$$\theta_i^1 = \frac{(1-\beta)Y_i^1}{Q_i H_i}. \tag{B.62}$$

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

$$\{Q_i^1, w_i^1, \theta_i^1\} = \{Q_i^0, w_i^0, \theta_i^0\},$$

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:

$$Q_i^2 = \varsigma Q_i^0 + (1-\varsigma) Q_i^1,$$
$$w_i^2 = \varsigma w_i^0 + (1-\varsigma) w_i^1,$$
$$\theta_i^2 = \varsigma \theta_i^0 + (1-\varsigma) \theta_i^1,$$

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 equilibrium.

B.3.2 Agglomeration Forces

We next consider our augmented specification with agglomeration forces and endogenous productivity and amenities. Given model parameters {α, β, ϕ, ϵ, η_A, η_B, δ_A, δ_B}, known values of location characteristics {H_n, $\overline{A}_n$, $\overline{B}_n$, τ_{ni}} and exogenous total city population ($L_{\mathbb{N}}$) , we use an analogous solution algorithm to solve for equilibrium. We assume starting values for the price of floor space, wages and the fraction of floor space used commercially {Q_n^0, w_n^0, θ_n^0}. Given these starting values, we use the equilibrium conditions of the model to solve for new predicted values for these endogenous variables {Q_n^1, w_n^1, θ_n^1}. We expand the equilibrium conditions of the model to include the endogenous determination of productivity and amenities as a function of production and residential externalities:

$$\mathbb{A}_n^1 = \sum_{i \in \mathbb{N}} e^{-\delta_A \tau_{ni}} L_i^1, \tag{B.63}$$

$$A_n^1 = \overline{A}_n \left(\mathbb{A}_i^1\right)^{\eta_A}, \tag{B.64}$$

$$\mathbb{B}_n^1 = \sum_{i \in \mathbb{N}} e^{-\delta_B \tau_{ni}} R_i^1, \tag{B.65}$$

$$B_n^1 = \overline{B}_n \left(\mathbb{B}_i^1\right)^{\eta_B}. \tag{B.66}$$

We continue to solve the system of equations for the equilibrium conditions of the model until the endogenous variables converge to equilibrium.

C Market Access Representation

In this section of the Online Appendix, we derive the reduced-form representation of our quantitative urban model in terms of residence market access (Φ_n^R) and workplace market access (Φ_n^L) from Section 4 of the paper. We consider a baseline specification for a closed-city, with segmented floor space markets and exogenous supplies of residential floor space (H_n^R) and commercial floor space (H_n^L), and no agglomeration forces ($\eta_A = \eta_B = \delta_A = \delta_B = 0$).

We start with the conditions for general equilibrium in the model. Substituting the wage (B.43) into employment (B.40), we obtain:

$$L_n = \xi A_n^{\frac{\epsilon}{\beta}} q_n^{-\epsilon\left(\frac{1-\beta}{\beta}\right)} \Phi_n^L.$$

Substituting the commercial floor price (B.46) into the above equation, we get:

$$L_n = \xi A_n^{\epsilon} \left(\frac{L_n}{\beta}\right)^{-\epsilon(1-\beta)} \left(\frac{H_n^L}{1-\beta}\right)^{\epsilon(1-\beta)} \Phi_n^L,$$

and hence:

$$L_n = \xi^{\frac{1}{1+\epsilon(1-\beta)}} \beta^{\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} (1-\beta)^{-\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} A_n^{\frac{\epsilon}{1+\epsilon(1-\beta)}} \left(H_n^L\right)^{\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} \left(\Phi_n^L\right)^{\frac{1}{1+\epsilon(1-\beta)}}, \tag{C.1}$$

$$\frac{L_n}{\beta} = \xi^{\frac{1}{1+\epsilon(1-\beta)}} \beta^{-\frac{1}{1+\epsilon(1-\beta)}} (1-\beta)^{-\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} A_n^{\frac{\epsilon}{1+\epsilon(1-\beta)}} \left(H_n^L\right)^{\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} \left(\Phi_n^L\right)^{\frac{1}{1+\epsilon(1-\beta)}}.$$

Substituting this result into the commercial floor price (B.46):

$$q_n = A_n \xi^{\frac{\beta}{1+\epsilon(1-\beta)}} \beta^{-\frac{\beta}{1+\epsilon(1-\beta)}} A_n^{\frac{\beta\epsilon}{1+\epsilon(1-\beta)}} \left(\frac{H_n^L}{1-\beta}\right)^{\frac{\beta\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} \left(\Phi_n^L\right)^{\frac{\beta}{1+\epsilon(1-\beta)}} \left(\frac{H_n^L}{1-\beta}\right)^{-\beta},$$

and hence:

$$q_n = \xi^{\frac{\beta}{1+\epsilon(1-\beta)}} \beta^{-\frac{\beta}{1+\epsilon(1-\beta)}} A_n^{\frac{1+\epsilon}{1+\epsilon(1-\beta)}} \left(\frac{H_n^L}{1-\beta}\right)^{-\frac{\beta}{1+\epsilon(1-\beta)}} \left(\Phi_n^L\right)^{\frac{\beta}{1+\epsilon(1-\beta)}}. \tag{C.2}$$

Substituting the residential floor price (B.45) into residents (B.39), we obtain:

$$R_n = \xi B_n^{\epsilon} (1-\alpha)^{\epsilon(\alpha-1)} v_n^{\epsilon(\alpha-1)} R_n^{\epsilon(\alpha-1)} \left(H_n^R\right)^{-\epsilon(\alpha-1)} \Phi_n^R,$$

and hence:

$$R_n = \xi^{\frac{1}{1+\epsilon(1-\alpha)}} B_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} (1-\alpha)^{\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} v_n^{\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(H_n^R\right)^{-\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(\Phi_n^R\right)^{\frac{1}{1+\epsilon(1-\alpha)}}.$$

In an equilibrium in which commuting costs between locations are prohibitive ($\kappa_{ni}^{-\epsilon} \approx 0$ for $n \neq i$), $v_n \approx w_n \approx \left(\Phi_n^R\right)^{\frac{1}{\epsilon}}$ and we can further rewrite this equation as:

$$R_n \approx \xi^{\frac{1}{1+\epsilon(1-\alpha)}} B_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} (1-\alpha)^{\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(H_n^R\right)^{-\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(\Phi_n^R\right)^{\frac{\alpha}{1+\epsilon(1-\alpha)}}. \tag{C.3}$$

Substituting this result into the residential floor price (B.45), we have:

$$Q_n \approx (1-\alpha) v_n \xi^{\frac{1}{1+\epsilon(1-\alpha)}} B_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} (1-\alpha)^{\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(H_n^R\right)^{-\frac{1+2\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(\Phi_n^R\right)^{\frac{\alpha}{1+\epsilon(1-\alpha)}}.$$

Again approximating around an equilibrium in which commuting costs between locations are prohibitive ($\kappa_{ni}^{-\epsilon} \approx 0$ for $n \neq i$), $v_n = w_n = \left(\Phi_n^R\right)^{\frac{1}{\epsilon}}$ and we can further rewrite this equation as:

$$Q_n \approx (1-\alpha) \xi^{\frac{1}{1+\epsilon(1-\alpha)}} B_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} (1-\alpha)^{\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(H_n^R\right)^{-\frac{1+2\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(\Phi_n^R\right)^{\frac{1+\epsilon}{\epsilon(1+\epsilon(1-\alpha))}}. \tag{C.4}$$

Collecting together equations (C.1), (C.2), (C.3) and (C.4), we have obtain the following system of equations for employment (L_n), the price of commercial floor space (q_n), residents (R_n) and the price of residential floor space (Q_n) in terms of commuting market access (Φ_n^L, Φ_n^R), the structural residuals (A_n, B_n, H_n^R, H_n^L), and parameters:

$$L_n = \xi^{\frac{1}{1+\epsilon(1-\beta)}} \beta^{\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} (1-\beta)^{-\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} A_n^{\frac{\epsilon}{1+\epsilon(1-\beta)}} \left(H_n^L\right)^{\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} \left(\Phi_n^L\right)^{\frac{1}{1+\epsilon(1-\beta)}},$$

$$q_n = \xi^{\frac{\beta}{1+\epsilon(1-\beta)}} \beta^{-\frac{\beta}{1+\epsilon(1-\beta)}} A_n^{\frac{1+\epsilon}{1+\epsilon(1-\beta)}} \left(\frac{H_n^L}{1-\beta}\right)^{-\frac{\beta}{1+\epsilon(1-\beta)}} \left(\Phi_n^L\right)^{\frac{\beta}{1+\epsilon(1-\beta)}},$$

$$R_n \approx \xi^{\frac{1}{1+\epsilon(1-\alpha)}} B_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} (1-\alpha)^{\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(H_n^R\right)^{-\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(\Phi_n^R\right)^{\frac{\alpha}{1+\epsilon(1-\alpha)}},$$

$$Q_n \approx (1-\alpha) \xi^{\frac{1}{1+\epsilon(1-\alpha)}} B_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} (1-\alpha)^{\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(H_n^R\right)^{-\frac{1+2\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \left(\Phi_n^R\right)^{\frac{1+\epsilon}{\epsilon(1+\epsilon(1-\alpha))}}.$$

Taking log differences between a counterfactual and an initial equilibrium, we can represent relative changes in employment ($\widehat{L}_n$), the price of commercial floor space ($\widehat{q}_n$), residents ($\widehat{R}_n$) and the price of residential floor space ($\widehat{Q}_n$) as the following system of equations:

$$\log \widehat{L}_n = \frac{1}{1+\epsilon(1-\beta)} \log \widehat{\Phi}_n^L + \log \left(\widehat{A}_n^{\frac{\epsilon}{1+\epsilon(1-\beta)}} \left(\widehat{H}_n^L\right)^{\frac{\epsilon(1-\beta)}{1+\epsilon(1-\beta)}} \right),$$

$$\log \widehat{q}_n = \frac{\beta}{1+\epsilon(1-\beta)} \log \widehat{\Phi}_n^L + \log \left(\widehat{A}_n^{\frac{1+\epsilon}{1+\epsilon(1-\beta)}} \left(\widehat{H}_n^L\right)^{-\frac{\beta}{1+\epsilon(1-\beta)}} \right),$$

$$\log \widehat{R}_n \approx \frac{\alpha}{1+\epsilon(1-\alpha)} \log \widehat{\Phi}_n^R + \log \left(\widehat{B}_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} \left(\widehat{H}_n^R\right)^{-\frac{\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \right),$$

$$\log \widehat{Q}_n \approx \frac{1+\epsilon}{\epsilon\left(1+\epsilon\left(1-\alpha\right)\right)} \log \widehat{\Phi}_n^R + \log \left(\widehat{B}_n^{\frac{\epsilon}{1+\epsilon(1-\alpha)}} \left(\widehat{H}_n^R \right)^{-\frac{1+2\epsilon(\alpha-1)}{1+\epsilon(1-\alpha)}} \right),$$

where recall that a hat above a variable denotes a relative change between the counterfactual (prime) and initial (no prime) equilibria ($\widehat{x} = x'/x$); and the last two equations involve an approximation around an equilibrium with prohibitive commuting costs between locations (such that $v_n \approx w_n \approx \left(\Phi_n^R\right)^{1/\epsilon}$).

D Quantitative Illustration

In Section 5 of the paper, we illustrate the use of our quantitative urban model for the evaluation of the impact of a transport improvement using a numerical example of a city. By focusing on a numerical example, we consider a setting in which we know the true data generating process (DGP) and model parameters. Therefore, the data are generated according to the model, and we can examine the success of alternative approaches in approximating the true impact of the transport infrastructure improvement. Our numerical example is motivated by the construction of London's 19th-century railway network.

In Section 5 of the paper, we report results for a closed-city and our baseline specification with no agglomeration forces (such that $A_n = \overline{A}_n$ and $B_n = \overline{B}_n$). In this section of the Online Appendix, we demonstrate the robustness of our quantitative results in our augmented specification with agglomeration forces (such that $A_n = \overline{A}_n \mathbb{A}_n^{\eta_A}$ and $B_n = \overline{B}_n \mathbb{B}_n^{\eta_B}$).

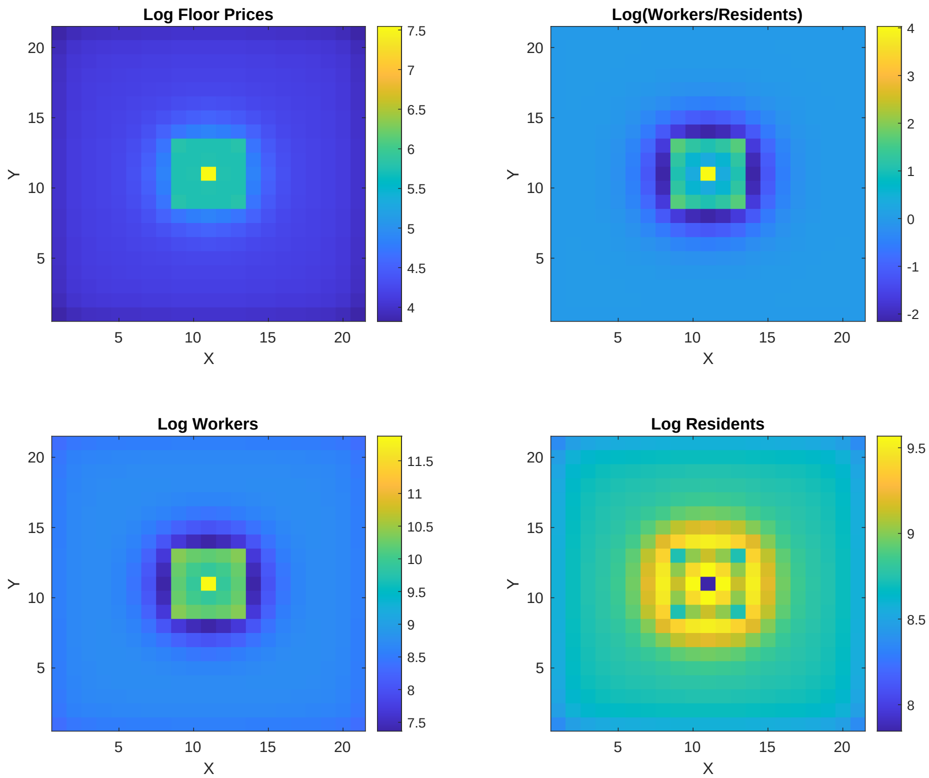
D.1 Initial Equilibrium

Figure D.1 shows the initial equilibrium distribution of economic activity before the construction of the railway network for our augmented specification with agglomeration forces. We indicate levels of economic activity in each location using a heatmap, in which higher values are denoted with lighter colors (more yellow), and lower values are denoted with darker colors (more blue).

We find a similar pattern of results as for our baseline specification in the paper. We find that the central location has the highest floor prices (top left), highest ratio of workers to residents (top right), the largest concentration of workers (bottom left), and the smallest concentration of residents (bottom right). Around the central location, we find a non-monotonic pattern of concentric rings of specialization, in which intermediate locations just beyond the boundaries of the inner city have the lowest ratios of workers to residents (top right), the smallest concentrations of workers (bottom left), and the largest concentrations of residents (bottom right).

This pattern again reflects the interaction between the comparative advantage of locations as workplaces or residences and commuting costs. In this augmented specification, the concentration of employment in the center leads to an endogenous increase in productivity through

Figure D.1: Initial Equilibrium Before the Railway Network (With Agglomeration Forces)



Note: Figure shows the initial equilibrium before the railway network in our augmented specification with agglomeration forces ($\eta_A = \eta_B = 0.10$, $\delta_A = \delta_B = 0.08$). City composed of 21×21 grid points one unit apart (interpreted as 1 kilometer). Central node (11,11) has production fundamentals (a_i) equal to 4. Surrounding nodes (9:13,9:13) have $a_i = 2$. All other locations have $a_i = 1$. All locations have residential fundamentals (b_i) equal to 1. All locations have area (K_i) equal to 100, a density of development (φ_i) equal to 1, and a supply of floor space equal to $H_i = \varphi_i K_i$. In the initial equilibrium, walking is the only mode of transport, with a travel cost of 1 per unit distance. Top-left panel shows the log price of floor space ($\log Q_i = \log q_i$). Top-right panel shows the log ratio of workers to residents ($\log(L_i/R_i)$). Bottom-left panel shows log residents ($\log R_i$). Bottom-right panel shows log workers ($\log L_i$).

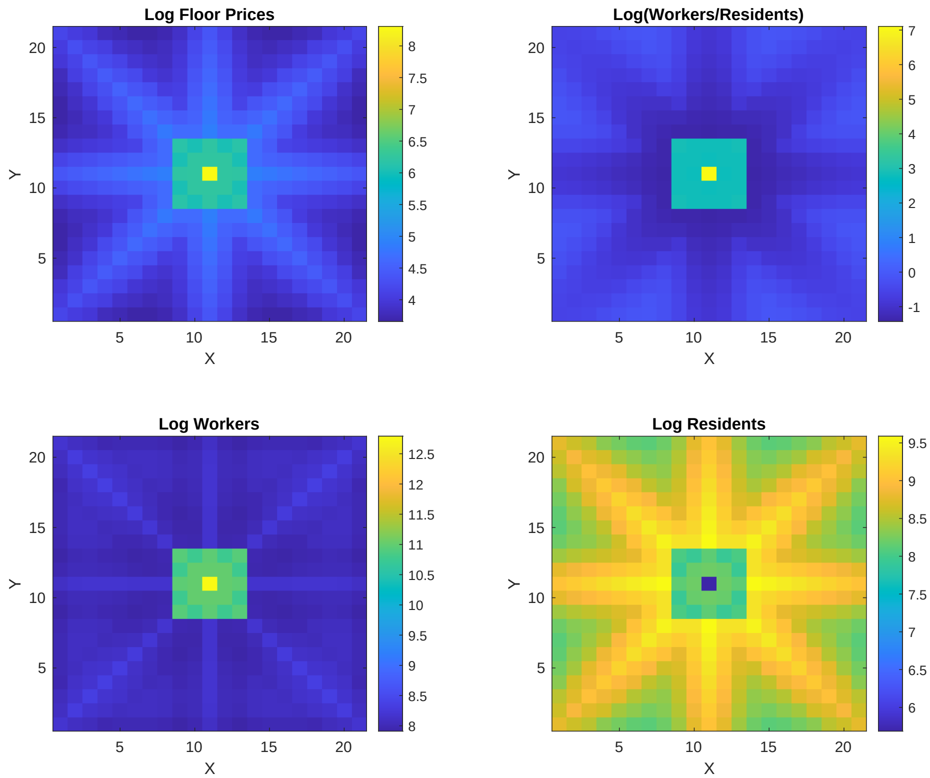
agglomeration forces, thereby magnifying the initial differences in productivity across locations. Similarly, the concentration of residents at intermediate locations just beyond the boundaries of the the inner city leads to an endogenous increase in amenities through agglomeration forces, thereby magnifying the initial differences in amenities across locations.

D.2 Counterfactual Equilibrium

Figure D.2 shows the counterfactual equilibrium distribution of economic activity after the construction of the railway network, for our augmented specification with agglomeration forces. Again we indicate levels of economic activity in each location using a heatmap, in which higher values are denoted with lighter colors (more yellow), and lower values are denoted with darker colors (more blue).

We find that the construction of the railway network leads to a similar reorganization of eco-

Figure D.2: Counterfactual Equilibrium after The Railway Network (With Agglomeration Forces)



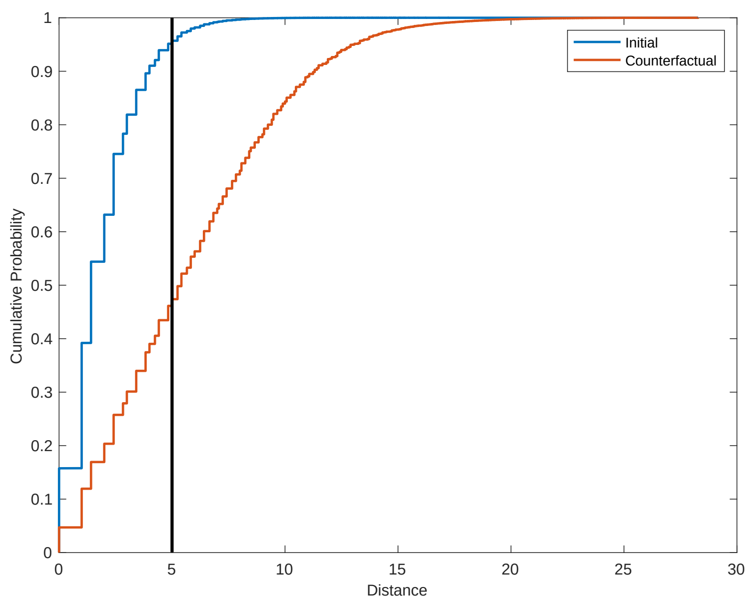
Note: Figure shows the counterfactual equilibrium after the railway network in our augmented specification with agglomeration forces ($\eta_A = \eta_B = 0.10$, $\delta_A = \delta_B = 0.08$). City composed of 21×21 grid points one unit apart (interpreted as 1 kilometer). Central node (11,11) has production fundamentals (a_i) equal to 4. Surrounding nodes (9:13,9:13) have $a_i = 2$. All other locations have $a_i = 1$. All locations have residential fundamentals (b_i) equal to 1. All locations have area (K_i) equal to 100, a density of development (φ_i) equal to 1, and a supply of floor space equal to $H_i = \varphi_i K_i$. In the counterfactual, vertical horizontal, diagonal and inverse diagonal railway lines are constructed (such that the railway network forms a "union jack." Railway lines reduce the cost of travel from 1 to $1/\gamma$ per unit distance (where $\gamma > 1$). Top-left panel shows the log price of floor space ($\log Q_i = \log q_i$). Top-right panel shows the log ratio of workers to residents ($\log{(L_i/R_i)}$). Bottom-left panel shows log residents ($\log R_i$). Bottom-right panel shows log workers ($\log L_i$).

nomic activity as in our baseline specification in the paper without agglomeration forces. Following the reduction in commuting costs from the construction of the railway network, the center of the city increasingly specializes as a workplace rather than a residence, which leads to an increase in the price of floor space. In contrast, outlying areas close to railway lines increasingly specialize as a residence, and also experience an increase in the price of floor space, as the residents of these outlying areas reallocate away from locations far from railway lines towards those close to railway lines. Again these reallocations of employment and residents across locations lead to endogenous changes in productivity and amenities through agglomeration forces.

D.3 Commuting Distributions

Figure D.3 shows the cumulative distribution of commuting distances in the initial equilibrium (blue) and counterfactual equilibrium (red) for our augmented specification with agglomeration forces. We construct this cumulative distribution in the same way as for our baseline specification in the paper. For each residence-workplace pair, we observe the equilibrium number of commuters and the bilateral distance travelled. We sort residence-workplace pairs by the distance travelled, and compute the cumulative sum of commuters for each distance travelled, divided by the total number of commuters in the city, which yields the share of workers who commute less than each distance. As for our baseline specification in the paper without agglomeration forces, we find that the construction of the railway networks leads to a substantial increase in the proportion of workers commuting over longer distance. The fraction of workers living within 5km of their residence falls from over 90 percent to around 50 percent.

Figure D.3: Cumulative Commuting Distance Distributions (With Agglomeration Forces)

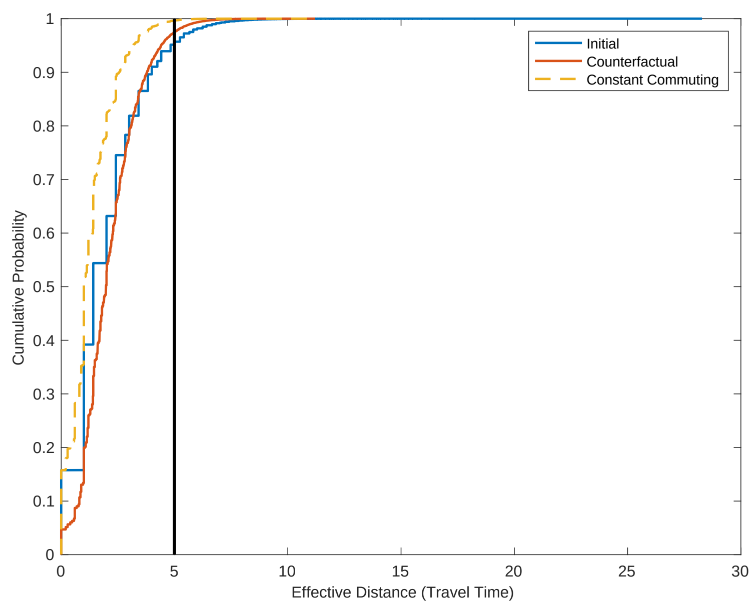


Note: Figure shows the cumulative distribution for the share of workers who commute less than each distance in the initial equilibrium (blue) and the counterfactual equilibrium (red) in the specification with agglomeration forces ($\eta_A = \eta_B = 0.10$, $\delta_A = \delta_B = 0.08$). Vertical black line shows 5 kilometers distance. In the initial equilibrium, walking is the only mode of transport, with a travel cost of 1 per unit distance. In the counterfactual, vertical horizontal, diagonal and inverse diagonal railway lines are constructed (such that the railway network forms a "union jack." Railway lines reduce the cost of travel from 1 to $1/\gamma$ per unit distance (where $\gamma > 1$).

Figure D.4 shows the cumulative distribution of effective distances, where effective distance adjusts for the different travel costs of walking and the railway, and has an interpretation as travel time. We construct this cumulative distribution in the same way as for the previous figure, but use effective distance (travel times) instead of distance. We show this cumulative distribution using initial commuting probabilities and initial travel times (blue), using initial commuting probabilities and counterfactual travel times (orange dashed line), and using counterfactual commuting probabilities and counterfactual travel times (red line).

Again we find a similar pattern of results as for our baseline specification in the paper without agglomeration forces. Although the construction of the railway network reduces travel times for the initial distribution of residents and workers (movement from the blue to the orange dashed line), it leads firms and workers to change their location decisions (movement from the orange dashed line to the red line), such that there is little change in the distribution of travel times in the new equilibrium. Therefore, the substantial increase in the number of workers who commute longer distances in Figure 5 goes hand-in-hand with little change in the distribution of travel times in Figure 6. The direct effect of the new transport technology is to reduce the time taken to travel a given distance. But workers respond to this reduction in travel costs by increasing the distance between their workplace and residence, such that there is relatively little change in equilibrium travel times.

Figure D.4: Cumulative Commuting Effective Distance (Travel Time) Distributions (With Agglomeration Forces)

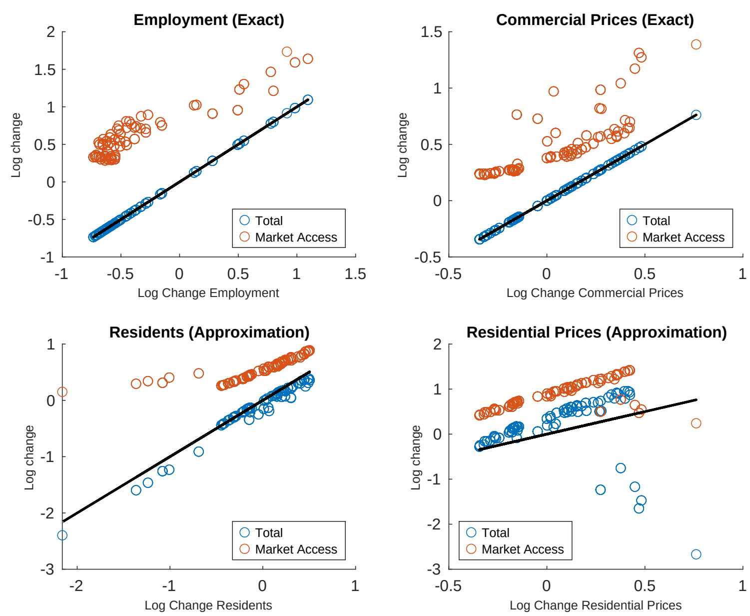


Note: Figure shows the cumulative distribution for the share of workers who commute less than each effective distance (travel time) in our augmented specification with agglomeration forces ($\eta_A = \eta_B = 0.10$, $\delta_A = \delta_B = 0.08$); blue line shows this cumulative distribution in the initial equilibrium (using initial commuting probabilities and travel times); orange dashed line shows this distribution using initial commuting probabilities and counterfactual travel times; red line shows this cumulative distribution in the counterfactual equilibrium (using counterfactual commuting probabilities and travel times). Vertical black line shows 5 kilometers distance. In the initial equilibrium, walking is the only mode of transport, with a travel cost of 1 per unit distance. In the counterfactual, vertical horizontal, diagonal and inverse diagonal railway lines are constructed (such that the railway network forms a "union jack"). Railway lines reduce the cost of travel from 1 to $1/\gamma$ per unit distance (where $\gamma > 1$). Effective distance adjusts for γ and has an interpretation as travel time.

D.4 Comparison with Market Access Predictions

Figure D.5 shows log relative changes in employment (top left), commercial floor space prices (top right), residents (bottom left), and residential floor space prices (bottom right) as a result of the construction of the railway network, for our augmented specification with agglomeration forces. On the horizontal axis of each panel, we show the true log changes in each variable from solving for the counterfactual equilibrium in the full non-linear model.

Figure D.5: Counterfactual Versus Market Access Predictions (With Agglomeration Forces)



Note: Figure shows predicted log changes in each variable between the counterfactual and initial equilibria from model counterfactuals (horizontal axis) versus market access predictions (vertical axis) in the specification with agglomeration forces ($\eta_A = \eta_B = 0.10$, $\delta_A = \delta_B = 0.08$). Red circles show predictions based on market access alone. Blue circles show predictions based on market access and a residual. For employment ($\log \widehat{L}_i$) and commercial floor space prices ($\log \widehat{q}_i$), model counterfactuals are exactly equal to the predictions based on market access and the residual (blue circles on the 45 degree line). For residents ($\log \widehat{R}_i$) and residential floor pries ($\log \widehat{Q}_i$), model counterfactual are approximately equal to the predictions based on market access and the residual (blue circles away from the 45 degree line), where the approximation is taken around an equilibrium with prohibitive travel costs.

On the vertical axis of each panel, we show the predicted log changes in each variable based on changes in market access ($\widehat{\Phi}_n^R$, $\widehat{\Phi}_n^L$) from the reduced-form system of equations (19). The blue circles (labelled total in the legend) show the overall predictions taking into account both changes in market access ($\widehat{\Phi}_n^R$, $\widehat{\Phi}_n^L$) and changes in the residuals ($\widehat{e}_n^L$, $\widehat{e}_n^q$, $\widehat{e}_n^R$, $\widehat{e}_n^Q$).[1] The red circles (labelled market access in the legend) show the predictions from changes in market access alone ($\widehat{\Phi}_n^R$, $\widehat{\Phi}_n^L$), where the changes in the residuals ($\widehat{e}_n^L$, $\widehat{e}_n^q$, $\widehat{e}_n^R$, $\widehat{e}_n^Q$) are set equal to zero.

[1] With no-arbitrage between alternative uses of floor space, the true log changes in the price of commercial and residential floor space are equal to one another: $\log(\widehat{q}_n) = \log\left(\widehat{Q}_n\right)$. Nevertheless, the overall predictions for these variables based on changes in market access and the residuals need not equal one another, because the reduced-form system of equations (19) only holds as an approximation for residential floor space prices.

Table D.1: Market Access Predictions (With Agglomeration Forces)

	Regression Slope	R-squared
Employment ($\log \widehat{L}_n$)	1.216^{***}	0.816
Commercial Floor Prices ($\log \widehat{q}_n$)	0.785^{***}	0.649
Residents ($\log \widehat{R}_n$)	1.784^{***}	0.721
Residential Floor Prices ($\log \widehat{Q}_n$)	0.608^{***}	0.546

Note: Results of regressions of the log change in each variable between the counterfactual and initial equilibria from the solution of the non-linear model on the predicted change based on market access alone from equation (19) in the specification with agglomeration forces ($\eta_A = \eta_B = 0.10$, $\delta_A = \delta_B = 0.08$). For employment ($\log \widehat{L}_i$) and commercial floor space prices ($\log \widehat{q}_i$), counterfactuals log changes in the model are exactly equal to the predictions based on market access and a residual. For residents ($\log \widehat{R}_i$) and residential floor pries ($\log \widehat{Q}_i$), counterfactual log changes in the model are approximately equal to the predictions based on market access and the residual, where the approximation is taken around an equilibrium with prohibitive travel costs.

Again we find a similar pattern of results as for our baseline specification in the paper without agglomeration forces. For employment (top left) and commercial floor space prices (top right), the overall predictions based on changes in market access and changes in the residuals are necessarily equal to the true counterfactual changes, as reflected in the blue circles lieing along the 45-degree line. This pattern of results reflects the fact that the reduced-form relationships (19) hold exactly for employment and commercial floor space prices.

For residents (bottom left) and residential floor space prices (bottom right), these overall predictions can diverge from the true counterfactual changes, as reflected in the blue circles departing from the 45-degree line. This pattern of results reflects the fact that the reduced-form relationships (19) are only approximations around an equilibrium with prohibitive commuting costs for residents and residential floor space prices. Although the gap from the 45-degree line varies across locations, the magnitude of this variation is limited, except for a relatively small number of locations for residential floor space prices. In part, these results reflect the fact that we start from an initial equilibrium in which there is relatively little commuting, with more than 90 percent of workers living within 5km of their workplace. Therefore, the approximation around an initial equilibrium with prohibitive commuting cost is relatively good.

In contrast, we find that the predictions based on market access alone can diverge substantially from the true counterfactual changes, as reflected in the red circles departing substantially from the 45-degree line for all four variables. The extent of the error is not constant, but instead differs substantially across locations. The magnitude of the departure from the 45-degree line is greater for employment and commercial floor space prices than for residents and residential floor space prices, perhaps in part because commercial economic activity is more spatially concentrated across locations than residential economic activity.

Table D.1 documents a similar pattern of results using the regression specification from equation (21). We use regression slope coefficients and R-squared to summarize the explanatory power of market access predictions. Therefore, we find that predictions based on market access can be quite misleading for the reorganization of economic activity across locations in response to a transport improvement. In our augmented specification with agglomeration forces, the residuals include both changes in productivity and amenities ($\widehat{A}_n \neq 1$ and $\widehat{B}_n \neq 1$) and changes in the allocation of floor space ($\widehat{H}_n^L \neq 1$ and $\widehat{H}_n^R \neq 1$).

References

Ahlfeldt, Gabriel, Stephen J. Redding, Daniel M. Sturm and Nikolaus Wolf (2015) "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 83(6), 2127-2189.

Allen, Treb, Costas Arkolakis, and Xiangliang Li (2024) "On the Equilibrium Properties of Spatial Models," *American Economic Review: Insights*, 6(4), 472-89.