Web Appendix for "The Making of the Modern Metropolis: Evidence from London" (Not for Publication)

Stephan Heblich

University of Bristol

Stephen J. Redding

Princeton University, NBER and CEPR

Daniel M. Sturm

London School of Economics and CEPR

A Introduction

This technical appendix contains additional supplementary material for the paper. Section B proves Lemma 1 in the paper. Section C proves Lemma 2 in the paper. Section D establishes a number of isomorphisms, in which we show that our quantitative predictions for the impact of the reversal of the railway network on historical workplace employment and commuting patterns hold in an entire class of quantitative urban models. Section E provides further details on our counterfactuals, including the proofs of Propositions E.1 and E.2 discussed in the paper. Section F reports additional empirical results discussed in the paper. Section G reports further information about the data sources and definitions.

B Proof of Lemma 1

Proof. Note that the commuter and land market clearing condition (15) in the paper can be re-written as follows:

$$\mathbb{Q}_t = \left[(1-\alpha)\mathbf{L_t} + \frac{1-\beta}{\beta}\boldsymbol{\Upsilon_t}\right]\mathbf{w_t}, \tag{B.1}$$

where $\mathbb{Q}_t$ is a vector of rateable values for each residence; $\mathbf{w_t}$ is a vector of wages for each workplace; $\mathbf{L_t}$ is the matrix of bilateral commuting flows containing the elements L_{nit} with residences n for rows and workplaces i for columns:

$$\mathbf{L_t} = \begin{pmatrix} L_{11t} & L_{12t} & \dots & L_{1Nt} \\ L_{21t} & L_{22t} & \dots & L_{2Nt} \\ \vdots & \vdots & \ddots & \vdots \\ L_{N1t} & L_{N2t} & \dots & L_{NNt} \end{pmatrix}; \tag{B.2}$$

and $\boldsymbol{\Upsilon_t}$ is a diagonal matrix with workplace employment for the diagonal and zero for the off-diagonal terms:

$$\boldsymbol{\Upsilon} = \begin{pmatrix} L_{1t} & 0 & \dots & 0 \\ 0 & L_{2t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_{Nt} \end{pmatrix}. \tag{B.3}$$

The first term in $\mathbf{L_t}$ in equation (B.1) captures payments for residential floor space use and the second term in $\boldsymbol{\Upsilon_t}$ in equation (B.1) captures payments for commercial floor space use. We assume that $\{\mathbb{Q}_{nt}, L_{nit}, L_{nt}\}$ are observed, and the parameters α and β are known, which in turn implies that the matrices $\mathbf{L_t}$ and $\boldsymbol{\Upsilon_t}$ are known. Given these known matrices, equation (B.1) is a system of linear equations in the unknown wage for each location (w_{nt}). Therefore, assuming that the rows of the matrix of commuting flows $\mathbf{L_t}$ are linearly independent, there exists a unique solution for the wage vector ($\mathbf{w_t}$) to this system of linear equations. ■

C Proof of Lemma 2

Proof. We first determine the unique vector of relative changes in wages ($\hat{\mathbf{w}}_{\mathbf{t}}$) and then recover the unique vector of relative changes in employment ($\hat{\mathbf{L}}_{\mathbf{t}}$). From equation (19) in the paper, the combined land and commuter market clearing condition for an earlier year $\tau < t$ can be written as:

$$\hat{\mathbb{Q}}_t \mathbb{Q}_t = T(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}}), \tag{C.1}$$

where $\hat{\mathbb{Q}}_t$ is the observed vector of relative changes in rateable values; $\mathbb{Q}_t$ is the observed vector of rateable values in our baseline year $t = 1921$; $\mathbf{X}_{\mathbf{t}} = \left[\hat{\mathbf{R}}_{\mathbf{t}}\ \lambda_{\mathbf{t}}^{\mathbf{C}}\ \mathbf{v}_{\mathbf{t}}\ \mathbf{R}_{\mathbf{t}}\ \mathbf{w}_{\mathbf{t}}\ \mathbf{L}_{\mathbf{t}}\right]$ is a known matrix of relative changes in variables between years τ and t and values for variables in our baseline year t; $\hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}$ is the matrix of estimated changes in commuting costs; $\hat{\mathbf{w}}_{\mathbf{t}}$ is the vector of relative changes in wages to be determined; and $T(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}})$ is an operator that is defined as:

$$\begin{aligned} T(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t,\tau}}; \mathbf{X}_{\mathbf{t}}) &= (1-\alpha)\left[\sum_{i\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{w}_{it}w_{it}\right]\hat{R}_{nt}R_{nt} \\ &+\left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)\hat{w}_{nt}w_{nt}\left[\sum_{i\in\mathbb{N}} \frac{\lambda^C_{int|i}\hat{w}^{\epsilon}_{nt}\hat{\kappa}^{-\epsilon}_{int}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{i\ell t|i}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{i\ell t}}\hat{R}_{it}R_{it}\right]. \end{aligned} \tag{C.2}$$

We now establish the following properties of equation (C.1) and the operator $T(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}})$.

Property (i): $\hat{\mathbb{Q}}_t \mathbb{Q}_t > 0$ such that $\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} > 0$ for all $n \in \mathbb{N}$

Property (ii): $T(0; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t,\tau}}; \mathbf{X}_{\mathbf{t}}) = 0$.

Property (iii): $T(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}})$ is monotonic in the vector of relative changes in wages ($\hat{\mathbf{w}}_{\mathbf{t}}$), since:

$$\begin{aligned} \frac{dT(\cdot)}{d\hat{\mathbf{w}}_{\mathbf{t}}}d\hat{\mathbf{w}}_{\mathbf{t}} &= (1-\alpha)\left[\sum_{i\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\frac{d\hat{w}_{it}}{\hat{w}_{it}}\hat{w}_{it}w_{it}\right]\hat{R}_{nt}R_{nt} \\ &\epsilon(1-\alpha)\left[\sum_{i\in\mathbb{N}}\left(1-\frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\right)\frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\frac{d\hat{w}_{it}}{\hat{w}_{it}}\hat{w}_{it}w_{it}\right]\hat{R}_{nt}R_{nt} \\ &+\left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)\frac{d\hat{w}_{nt}}{\hat{w}_{nt}}\hat{w}_{nt}w_{nt}\left[\sum_{i\in\mathbb{N}} \frac{\lambda^C_{int|i}\hat{w}^{\epsilon}_{nt}\hat{\kappa}^{-\epsilon}_{int}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{i\ell t|i}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{i\ell t}}\hat{R}_{it}R_{it}\right] \\ &+\epsilon\left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)\hat{w}_{nt}w_{nt}\left[\sum_{i\in\mathbb{N}} \frac{\lambda^C_{int|i}\hat{w}^{\epsilon}_{nt}\hat{\kappa}^{-\epsilon}_{int}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{i\ell t|i}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{i\ell t}}\frac{d\hat{w}_{nt}}{\hat{w}_{nt}}\hat{R}_{it}R_{it}\right], \end{aligned} \tag{C.3}$$

where $\frac{dT(\cdot)}{d\hat{\mathbf{w}}_{\mathbf{t}}}d\hat{\mathbf{w}}_{\mathbf{t}} > 0$ for $d\hat{\mathbf{w}}_{\mathbf{t}} > 0$.

Property (iv): $T(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}})$ is homogeneous of degree one in the vector of relative changes in wages ($\hat{\mathbf{w}}_{\mathbf{t}}$) such that $T(\boldsymbol{\xi}\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}}) = \boldsymbol{\xi}T(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}})$ for any positive scalar ξ.

From properties **(i)**-**(iv)**, starting from $\hat{w}_{nt} = 0$ for all locations n, and increasing $\hat{w}_{nt}$ for each location n, there exists a unique value for $\hat{w}_{nt}$ for which $\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = T_n(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\boldsymbol{\kappa}}_{\boldsymbol{t}}; \mathbf{X}_{\mathbf{t}})$ and equation (C.1) is satisfied.

Using these unique solutions for $\hat{\mathbf{w}}_t$, the unique vector of relative changes in employment ($\hat{\mathbf{L}}_t$) can be recovered from the commuter market clearing condition in equation (18) in the paper, as reproduced below.

$$\hat{L}_{it}L_{it} = \sum_{n\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{R}_{nt}R_{nt}, \tag{C.4}$$

where we have solved for $\hat{w}^{\epsilon}_{it}$; we estimate $\hat{\kappa}^{-\epsilon}_{nit}$; and we observe ($\hat{R}_{nt}$, L_{nt}, R_{nt}, $\lambda^C_{nit|n}$). Finally, the unique relative change in commuting flows ($\hat{L}_{nit}$) can be recovered from the conditional commuting probabilities in equation (20) in the paper, as reproduced below:

$$\hat{L}_{nit}L_{nit} = \frac{\lambda^C_{int|i}\hat{w}^{\epsilon}_{nt}\hat{\kappa}^{-\epsilon}_{int}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{i\ell t|i}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{i\ell t}}\hat{R}_{nt}R_{nt}, \tag{C.5}$$

where we have solved for $\hat{w}^{\epsilon}_{it}$; we estimate $\hat{\kappa}^{-\epsilon}_{nit}$; and we observe ($\hat{R}_{nt}$, L_{nt}, R_{nt}, $\lambda^C_{nit|n}$). ■

D Isomorphisms

In this section of the web appendix, we show that our quantitative predictions for the impact of the reversal of the railway network on workplace employment and commuting patterns hold in an entire class of quantitative urban models that satisfy the following three properties: (i) a gravity equation for bilateral commuting flows; (ii) land market clearing, such that income from the ownership of floor space equals the sum of payments for residential and commercial floor space use; (iii) Cobb-Douglas preferences and production technologies. When these three properties are satisfied, workplace income (the total income of all workers) is proportional to revenue and a sufficient statistic for payments for commercial floor space; residential income (the total income of all residents) is a sufficient statistic for payments for residential floor space; and commuting costs regulate the difference between workplace income and residential income.

An implication of this result is our baseline quantitative analysis does *not* require us to make assumptions about (a) whether productivity and amenities are exogenous or endogenous; (b) the underlying determinants of productivity and amenities including agglomeration forces; (c) whether the supply of floor space is exogenous or endogenous; (d) the underlying determinants of the supply of floor space; (e) the extent to which railways affect the cost of trading consumption goods; (f) the reservation level of utility in the wider economy. Regardless of the assumptions made about these other components of the model, we obtain the same predictions for the impact of the reversal of the railway network on workplace employment. The reason is that these predictions depend solely on the properties (i)-(iii) introduced immediately above, which hold regardless of these assumptions.

In Section D.1, we derive these predictions from the canonical urban model with a single final good that is produced under conditions of perfect competition and constant returns to scale and costlessly traded between locations. In Section D.2, we show that these predictions also hold in a new economic geography model with monopolistic competition, increasing returns to scale and costs of trading goods between locations, as in Helpman (1998), Redding and Sturm (2008) and Monte, Redding, and Rossi-Hansberg (2018). In Section D.3, we show that this new economic geography model is isomorphic to a Ricardian spatial model with perfect competition, constant returns to scale and costs of trading goods between locations, as in Eaton and Kortum (2002) and Redding (2016). In Section D.4, we demonstrate an analgous isomorphism to an Armington spatial model with neoclassical production and costs of trading goods between locations, as in Armington (1969) and Allen, Arkolakis, and Li (2017). In Section D.5, we show that it is straightforward to extend the analysis to incorporate non-traded services in additional to consumption goods. While the canonical urban model in Section D.1 takes the reservation level of utility in the wider economy as given, the new economic geography, Ricardian and Armington spatial models in Sections D.2-D.4 endogenize the level of expected utility. All of these different model structures satisfy a gravity equation for bilateral commuting flows; land market clearing; the requirement that payments for residential floor space are proportional to residential income; and the requirement that payments for commercial floor space are proportional to workplace income.

D.1 Canonical Urban Model

We start by deriving our predictions for the impact of reversing the construction of the railway network on workplace employment in the canonical urban model following Lucas and Rossi-Hansberg (2002) and Ahlfeldt, Redding, Sturm, and Wolf (2015). We consider a city (Greater London) embedded within a wider economy (the United Kingdom).

The city consists of a discrete set of locations $\mathbb{N}$ (the boroughs observed in our data). Workers are assumed to be geographically mobile and choose between the city and the wider economy. Population mobility implies that the expected utility from living and working in the city equals the reservation level of utility in the wider economy $\bar{U}_t$. If a worker chooses the city, she choose a residence n and a workplace i from the set of locations $n, i \in \mathbb{N}$ to maximize her utility. We allow locations to differ from one another in terms of their attractiveness for production and residence, as determined by productivity, amenities, the supply of floor space, and transport connections, where each of these location characteristics can evolve over time.

D.1.1 Preferences

Worker preferences are defined over consumption of a composite final good and residential floor space. The indirect utility function is assumed to take the Cobb-Douglas form such that utility for a worker ω residing in n and working in i is given by:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega) w_i}{\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}}, \qquad 0 < \alpha < 1, \tag{D.1}$$

where P_n is the price of the final good, Q_n is the price of floor space, w_i is the wage, κ_{ni} is an iceberg commuting cost, and $b_{ni}(\omega)$ is an idiosyncratic amenity draw that captures all the idiosyncratic factors that can cause an individual to live and work in particular locations within the city.

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

$$G_{ni}(b) = e^{-B_n b^{-\epsilon}}, \qquad B_n > 0,\ \epsilon > 1, \tag{D.2}$$

where the scale parameter B_n controls the average desirability of location n as a residence. The shape parameter ϵ determines the dispersion of idiosyncratic amenities, which controls the sensitivity of worker location decisions to economic variables (e.g. wages and the cost of living). The smaller is ϵ, the greater is the heterogeneity in idiosyncratic amenities, and the less sensitive are worker location decisions to economic variables.

D.1.2 Production

The final good is produced under conditions of perfect competition and constant returns to scale using labor and floor space. The production technology is assumed to take the Cobb-Douglas form with unit cost:

$$P_i = \frac{1}{A_i} w_i^{\beta} Q_i^{1-\beta}, \qquad 0 < \beta < 1, \tag{D.3}$$

where A_i is the productivity of final goods production in location i. The final good is assumed to be costlessly traded between locations within Greater London such that:

$$P_i = P, \qquad \forall i \in \mathbb{N}. \tag{D.4}$$

From profit maximization and zero profits, we obtain the results in equations (8) and (9) in the paper that payments to labor and floor space are constant shares of revenue:

$$w_i L_i = \beta Y_i, \qquad Q_i H_i^Y = (1-\beta) Y_i, \tag{D.5}$$

where L_n is workplace employment; Y_i is revenue; and H_i^Y denotes commercial floor space use. Therefore, payments for commercial floor space are proportional to workplace income:

$$Q_i H_i^Y = \frac{1-\beta}{\beta} w_i L_i. \tag{D.6}$$

Re-arranging equation (D.3), we obtain another key implication of profit maximization and zero profits for each location with positive production:

$$w_i = (P A_i)^{1/\beta} Q_i^{-(1-\beta)/\beta}. \tag{D.7}$$

Intuitively, the maximum wage (w_i) that a location can afford to pay workers is increasing in the location's productivity (A_i) and the common price of the final good (P) and decreasing in the price of floor space (Q_i).

D.1.3 Market Clearing

Land market clearing implies that the total income received by landlords as owners of floor space (which equals rateable value ($\mathbb{Q}_n$) in our data) equals the sum of payments for the use of residential and commercial floor space:

$$\mathbb{Q}_n = Q_n H_n = (1-\alpha)\, v_n R_n + \left(\frac{1-\beta}{\beta}\right) w_n L_n. \tag{D.8}$$

where H_n is the quantity of floor space; v_n is the per capita income of location n's residents, as determined below as a function of commuting patterns; and R_n is the measure of these residents.

Finally, the supply of floor space (H_n) depends on both geographical land area (K_n) and the density of development as measured by the ratio of floor space to land area (h_n). Following Saiz (2010), we allow the supply of floor space to respond endogenously to changes in its price:

$$H_n = h_n K_n, \qquad h_n = h Q_n^{\mu}, \tag{D.9}$$

where h is a constant; $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.

D.1.4 Workplace and Residence Choices

Using indirect utility (D.1) and the Fréchet distribution of idiosyncratic amenities (D.2), this canonical urban model exhibits a gravity equation for commuting flows. The unconditional probability that a worker chooses to live in location n and work in location i is given by:

$$\lambda_{ni} = \frac{B_n \left(\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon} w_i^{\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}}, \tag{D.10}$$

which is identical to equation (5) in the paper.

Summing across workplaces, we obtain the probability that an individual lives in each location ($\lambda_n^R = R_n/\bar{L}$), while summing across residences, we have the probability that an individual works in each location ($\lambda_n^L = L_i/\bar{L}$):

$$\lambda_n^R = \frac{\sum_{\ell \in \mathbb{N}} B_n w_\ell^{\epsilon} \left(\kappa_{n\ell} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}, \quad \lambda_i^L = \frac{\sum_{r \in \mathbb{N}} B_r w_i^{\epsilon} \left(\kappa_{ri} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}. \tag{D.11}$$

Both expressions take the same form as in equation (6) in the paper and labor market clearing corresponds to $\sum_{n \in \mathbb{N}} \lambda_n^R = \sum_{i \in \mathbb{N}} \lambda_i^L = 1$.

Using the unconditional commuting probability (λ_{ni}) from equation (D.10) and the residence probability (λ_n^R) from equation (D.11), the conditional commuting probability that a worker commutes to location i conditional on residing in location n takes the same form as in equation (11) in the paper:

$$\lambda_{ni|n}^C = \frac{(w_i/\kappa_{ni})^\epsilon}{\sum_{\ell\in\mathbb{N}} (w_\ell/\kappa_{n\ell})^\epsilon}. \tag{D.12}$$

Using this conditional commuting probability from equation (D.12), we obtain an identical expression for per capita residential income as in equation (13) in the paper:

$$v_n = \sum_{i\in\mathbb{N}} \lambda_{ni|n}^C w_i. \tag{D.13}$$

Commuter market clearing implies that employment in each location (L_i) equals the measure of workers choosing to commute to that location. Using the conditional commuting probabilities from equation (D.12), we obtain the same expression for this commuter market clearing condition as in equation (10) in the paper:

$$L_i = \sum_{n\in\mathbb{N}} \lambda_{ni|n}^C R_n. \tag{D.14}$$

Finally, population mobility requires that expected utility for each workplace-residence pair is equal to the reservation level of utility in the wider economy. Using the properties of the Fréchet distribution for idiosyncratic amenities, this population mobility condition takes the same form as in equation (7) in the paper:

$$\bar{U} = \mathbb{E}\left[U_{ni\omega}\right] = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right)\left[\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}} B_r \left(\kappa_{r\ell} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^\epsilon\right]^{\frac{1}{\epsilon}} \quad \text{all } n, i \in \mathbb{N}, \tag{D.15}$$

where $\mathbb{E}$ is the expectations operator; the expectation is taken over the distribution for idiosyncratic amenities; and $\Gamma(\cdot)$ is the Gamma function.

D.1.5 Comparative Statics for Changes in Commuting Costs

We now show that this canonical urban model yields exactly the same predictions for the impact of reversing the construction of the railway network on workplace employment and commuting patterns as in the paper, once we condition on the observed values of the endogenous variables in the initial equilibrium and the observed changes in residence employment and rateable values.

First, using equation (D.8), the land market clearing condition for any earlier year $\tau < t$ can be written in terms of the observed variables and model solutions for our baseline year of $t = 1921$ and the relative changes in the endogenous variables of the model between those two years:

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha)\hat{v}_{nt}v_{nt}\hat{R}_{nt}R_{nt} + \left(\frac{1-\beta}{\beta}\right)\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt}, \tag{D.16}$$

where recall that a hat above a variable denotes a relative change, such that $\hat{x}_t = x_\tau/x_t$.

Second, using equations (D.12) and (D.13), per capita residential income (v_{nt}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{v}_{nt}v_{nt} = \sum_{i\in\mathbb{N}} \frac{\lambda_{nit|n}^C \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell\in\mathbb{N}} \lambda_{n\ell t|n}^C \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}} \hat{w}_{it}w_{it}. \tag{D.17}$$

Third, using equations (D.12) and (D.14), workplace employment (L_{it}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{L}_{it} L_{it} = \sum_{n \in \mathbb{N}} \frac{\lambda^C_{nit|n} \hat{w}^{\epsilon}_{it} \hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{n\ell t}} \hat{R}_{nt} R_{nt}. \tag{D.18}$$

Finally, using equation (D.12), commuting flows ($\hat{L}_{nit}$) for any earlier year $\tau < t$ can be written in an analogous form as follows:

$$\hat{L}_{nit} L_{nit} = \frac{\lambda^C_{nit|n} \hat{w}^{\epsilon}_{it} \hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{n\ell t}} \hat{R}_{nt} R_{nt}. \tag{D.19}$$

Note that equations (D.16), (D.17), (D.18) and (D.19) above are identical to equations (16), (17), (18) and (20) in the paper. Therefore, given the same observed variables in the initial equilibrium (L_{nt}, R_{nt}, $\mathbb{Q}_{nt}$, w_{nt}, v_{nt}, L_{nit}), the same observed changes in residents and rateable values ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$) and the same estimated changes in commuting costs ($\hat{\kappa}^{-\epsilon}_{nit}$), this canonical urban model predicts the same changes in workplace employment ($\hat{L}_{it}$) and commuting patterns ($\hat{L}_{nit}$) as in the paper.

D.2 New Economic Geography Model

We now derive our predictions for the impact of reversing the construction of the railway network on workplace employment in a new economic geography model following Helpman (1998), Redding and Sturm (2008) and Monte, Redding, and Rossi-Hansberg (2018). We consider an economy that consists of a set of locations $n, i \in \mathbb{N}$. These locations are linked in goods markets through costly trade and in factor markets through migration and costly commuting. The economy as a whole is populated by a measure $\bar{L}$ of workers who are endowed with one unit of labor that is supplied inelastically.

D.2.1 Preferences and Endowments

Workers are geographically mobile and choose a pair of residence and workplace locations to maximize their utility, taking as given the choices of other firms and workers. The preferences of a worker ω who lives in location n and works in location i are defined over final goods consumption ($C_n(\omega)$), residential floor space use ($H^R_n(\omega)$), an idiosyncratic amenities shock for each workplace-residence pair ($b_{ni}(\omega)$) and iceberg commuting costs (κ_{ni}), according to the following Cobb-Douglas functional form:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega)}{\kappa_{ni}} \left(\frac{C_n(\omega)}{\alpha} \right)^{\alpha} \left(\frac{H^R_n(\omega)}{1-\alpha} \right)^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{D.20}$$

The idiosyncratic amenities shock for worker ω for each residence n and workplace i ($b_{ni}(\omega)$) is drawn from an independent Fréchet distribution:

$$G_{ni}(b) = e^{-B_n b^{-\epsilon}}, \qquad B_n > 0, \epsilon > 1, \tag{D.21}$$

where the scale parameter B_n controls the average desirability of location n as a residence; and the shape parameter $\epsilon > 1$ controls the dispersion of amenities. All workers ω residing in location n and working in location i receive the same wage and make the same choices for consumption and residential floor space use. Therefore, we suppress the implicit dependence on ω from now onwards, except where important.

The goods consumption index in location n takes the constant elasticity of substitution (CES) or Dixit-Stiglitz form and is defined over a continuum of varieties sourced from each location i,

$$C_n = \left[\sum_{i \in \mathbb{N}} \int_0^{M_i} c_{ni}(j)^\rho dj \right]^{\frac{1}{\rho}}, \qquad \sigma = \frac{1}{1-\rho} > 1, \tag{D.22}$$

where $c_{ni}(j)$ is consumption in location n of an individual variety j produced in location i; M_i is the mass of varieties produced in location i; and ρ is the CES parameter that determines the elasticity of substitution between varieties ($\sigma = 1/(1-\rho) > 1$).

Using the properties of CES preferences (D.22), the equilibrium consumption in location n of each variety j sourced from location i is determined by:

$$c_{ni}(j) = X_n P_n^{\sigma-1} p_{ni}(j)^{-\sigma}, \tag{D.23}$$

where $X_n = P_n C_n$ is total expenditure on consumption goods in location n; P_n is the price index dual to the consumption index (D.22); and $p_{ni}(j)$ is the "cost inclusive of freight" price of variety j produced in location i and consumed in location n.

Goods can be traded between locations subject to iceberg variable trade costs, such that $d_{ni} > 1$ units of a good must be shipped from location i in order for one unit to arrive in location n (where $d_{nn} = 1$). The "cost inclusive of freight" price of a variety in the location of consumption n ($p_{ni}(j)$) is thus a constant multiple of the "free on board" price of that variety in the location of production i ($p_i(j)$), with that multiple determined by the iceberg trade costs:

$$p_{ni}(j) = d_{ni} p_i(j). \tag{D.24}$$

D.2.2 Production

Production is modelled as in the new economic geography literature following Krugman (1991) and Helpman (1998). Varieties are produced under conditions of monopolistic competition. To produce a variety, a firm must incur both a fixed cost and a constant variable cost. We assume that these fixed and variable costs use labor and commercial floor space with the same factor intensity, such that the production technology is homothetic. We allow the variable cost to vary with location productivity A_i, such that the total cost of producing $x_i(j)$ units of a variety j in location i is:

$$\Gamma_i(j) = \left(F + \frac{x_i(j)}{A_i} \right) w_i^\beta Q_i^{1-\beta}, \qquad 0 < \beta < 1, \tag{D.25}$$

where w_i is the wage and Q_i is the price of floor space in location i. Profit maximization implies that the equilibrium price of each variety is a constant mark-up over marginal cost, namely:

$$p_{ni}(j) = p_{ni} = \left(\frac{\sigma}{\sigma - 1} \right) \frac{d_{ni} w_i^\beta Q_i^{1-\beta}}{A_i}. \tag{D.26}$$

Profit maximization and zero profits imply that the equilibrium output of each variety is the same for all varieties produced in location i:

$$x_i(j) = \bar{x}_i = A_i F (\sigma - 1). \tag{D.27}$$

Using the equilibrium pricing rule (D.26), free on board revenue ($y_i(j) = p_i(j) x_i(j)$) for each variety j in location i can be written as:

$$p_i(j) x_i(j) = y_i(j) = \bar{y}_i = \sigma w_i^\beta Q_i^{1-\beta} F, \tag{D.28}$$

and the common equilibrium wage bill for each variety j in location i is given by:

$$w_i l_i(j) = w_i \bar{l}_i = \beta \bar{y}_i, \tag{D.29}$$

where $l_i(j) = \bar{l}_i$ is workplace employment for variety j in location i.

Aggregating across all varieties produced within location i, profit maximization and zero profits imply that payments for labor and commercial floor space are constant shares of revenue, as in equations (8) and (9) in the paper:

$$w_i L_i = \beta Y_i, \qquad Q_i H_i^Y = (1-\beta) Y_i, \tag{D.30}$$

where L_i is total workplace employment; $Y_i = M_i \bar{y}_i$ is aggregate revenue; and H_i^Y denotes total commercial use of floor space. Therefore, payments for commercial floor space are proportional to workplace income:

$$Q_i H_i^Y = \frac{1-\beta}{\beta} w_i L_i. \tag{D.31}$$

D.2.3 Trade and Market Clearing

We assume that floor space is owned by landlords, who receive income from residents' and firms' expenditure on floor space, and consume only consumption goods where they live. Total expenditure on consumption goods equals the fraction α of the total income of residents plus the entire income of landlords. This income of landlords equals $(1-\alpha)$ times the total income of residents plus $(1-\beta)$ times revenue (which equals $(1-\beta)/\beta$ times the total income of workers). Therefore, total expenditure on consumption goods is:

$$X_n = P_n C_n = \alpha v_n R_n + (1-\alpha) v_n R_n + \frac{1-\beta}{\beta} w_n L_n = v_n R_n + \frac{1-\beta}{\beta} w_n L_n,$$

where v_n is the per capita income of location n's residents, as determined below as a function of commuting patterns, and R_n is the measure of these residents.

This new economic geography model implies a gravity equation for bilateral trade in goods between locations. Using CES demand in equation (D.23), and the fact that all varieties supplied from location i to location n charge the same price in equation (D.26), the share of location n's expenditure on goods produced in location i can be written as:

$$\pi_{ni} = \frac{M_i p_{ni}^{1-\sigma}}{\sum_{k \in \mathbb{N}} M_k p_{nk}^{1-\sigma}} = \frac{M_i \left(d_{ni} w_i / A_i\right)^{1-\sigma}}{\sum_{k \in \mathbb{N}} M_k \left(d_{nk} w_k / A_k\right)^{1-\sigma}}. \tag{D.32}$$

Therefore trade between locations n and i depends on bilateral trade costs (d_{ni}) in the numerator ("bilateral resistance") and on trade costs to all possible sources of supply k in the denominator ("multilateral resistance"). Goods market clearing and zero profits imply that payments to workers plus payments for commercial floor space use in each location equal expenditure on goods produced in that location:

$$w_i L_i + Q_i H_i^Y = \sum_{n \in \mathbb{N}} \pi_{ni} X_n. \tag{D.33}$$

Using equilibrium prices (D.26), the price index dual to the consumption index (D.22) can be rewritten as:

$$P_n = \left[\sum_{i \in \mathbb{N}} M_i \left(\frac{\sigma}{\sigma - 1} \frac{d_{ni} w_i}{A_i} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} = \left(\frac{M_n}{\pi_{nn}} \right)^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma - 1} \frac{w_n}{A_n}, \tag{D.34}$$

where the second equation uses the domestic trade share (π_{nn}) from equation (D.32) and $d_{nn} = 1$.

Labor market clearing implies that total payments to labor in each location equal the mass of varieties times labor payments for each variety. Using this relationship and the Cobb-Douglas production technology, the mass of varieties (M_i) in each location can be written as a function of total labor payments ($w_i L_i$) and firm revenue ($\bar{y}_i$) in that location:

$$M_i = \frac{w_i L_i}{w_i \bar{l}_i} = \frac{w_i L_i}{\beta \bar{y}_i}, \tag{D.35}$$

where L_i is total employment.

Land market clearing implies that the total income received by landlords as owners of floor space (which equals rateable value ($\mathbb{Q}_n$) in our data) equals the sum of payments for the use of residential and commercial floor space:

$$\mathbb{Q}_n = Q_n H_n = (1-\alpha)\, v_n R_n + \left(\frac{1-\beta}{\beta}\right) w_n L_n, \tag{D.36}$$

where H_n is the quantity of floor space.

Finally, the supply of floor space (H_n) depends on both geographical land area (K_n) and the density of development as measured by the ratio of floor space to land area (h_n). Following Saiz (2010), we allow the supply of floor space to respond endogenously to changes in its price:

$$H_n = h_n K_n, \qquad\qquad h_n = h Q_n^{\mu}, \tag{D.37}$$

where h is a constant; $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.

D.2.4 Workplace and Residence Choices

Given the direct utility function (D.20), the corresponding indirect utility function for a worker ω residing in location n and working in location i is:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega) w_i}{\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}}, \tag{D.38}$$

which takes exactly the same form as equation (3) in the paper and equation (D.1) in the canonical urban model in Section D.1 of this web appendix. The only difference from the canonical urban model is in the underlying determinants of the price index for goods consumption (P_n), as now specified in equation (D.34).

Using indirect utility (D.38) and the Fréchet distribution of idiosyncratic amenities (D.21), this new economic geography model exhibits the same gravity equation predictions for commuting flows as in the paper and in the canonical urban model in Section D.1 of this web appendix. The unconditional probability that a worker chooses to live in location n and work in location i is given by:

$$\lambda_{ni} = \frac{B_n \left(\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon} w_i^{\epsilon}}{\sum_{r\in\mathbb{N}} \sum_{\ell\in\mathbb{N}} B_r \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}}, \tag{D.39}$$

which is identical to equation (5) in the paper, except that the price index for goods consumption (P_n) is now determined by equation (D.34).

Summing across workplaces, we obtain the probability that an individual lives in each location ($\lambda_n^R = R_n/\bar{L}$), while summing across residences, we have the probability that an individual works in each location ($\lambda_n^L = L_i/\bar{L}$):

$$\lambda_n^R = \frac{\sum_{\ell\in\mathbb{N}} B_n w_\ell^{\epsilon} \left(\kappa_{n\ell} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{r\in\mathbb{N}} \sum_{\ell\in\mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}, \quad \lambda_i^L = \frac{\sum_{r\in\mathbb{N}} B_r w_i^{\epsilon} \left(\kappa_{ri} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}{\sum_{r\in\mathbb{N}} \sum_{\ell\in\mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}. \tag{D.40}$$

Both expressions take the same form as in equation (6) in the paper and labor market clearing corresponds to $\sum_{n \in \mathbb{N}} \lambda_n^R = \sum_{i \in \mathbb{N}} \lambda_i^L = 1$.

Using the unconditional commuting probability (λ_{ni}) from equation (D.39) and the residence probability (λ_n^R) from equation (D.40), the conditional commuting probability that a worker commutes to location i conditional on residing in location n takes the same form as in equation (11) in the paper:

$$\lambda_{ni|n}^C = \frac{(w_i/\kappa_{ni})^\epsilon}{\sum_{\ell \in \mathbb{N}} (w_\ell/\kappa_{n\ell})^\epsilon}. \tag{D.41}$$

Using this conditional commuting probability from equation (D.41), we obtain an identical expression for per capita income conditional on living in location n as in equation (13) in the paper:

$$v_n = \sum_{i \in \mathbb{N}} \lambda_{ni|n}^C w_i. \tag{D.42}$$

Commuter market clearing again implies that employment in each location (L_i) equals the measure of workers choosing to commute to that location. Using the conditional commuting probabilities from equation (D.41), we obtain the same expression for this commuter market clearing condition as in equation (10) in the paper:

$$L_i = \sum_{n \in \mathbb{N}} \lambda_{ni|n}^C R_n. \tag{D.43}$$

Finally, population mobility and the Fréchet distribution for idiosyncratic amenities imply that expected utility is equalized across all workplace-residence pairs and takes the same form as in equation (7) in the paper:

$$\bar{U} = \mathbb{E}\left[U_{ni\omega}\right] = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right)\left[\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r \left(\kappa_{r\ell} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^\epsilon\right]^{\frac{1}{\epsilon}} \quad \text{all } n, i \in \mathbb{N}, \tag{D.44}$$

where $\mathbb{E}$ is the expectations operator; the expectation is taken over the distribution for the idiosyncratic component of utility; and $\Gamma(\cdot)$ is the Gamma function.

D.2.5 Comparative Statics for Changes in Commuting Costs

We now show that this new economic geography model yields exactly the same predictions for the impact of reversing the construction of the railway network on workplace employment and commuting as in the paper and the canonical urban model in Section D.1 of this web appendix, once we condition on the observed values of the endogenous variables in the initial equilibrium and the observed changes in residence employment and rateable values.

First, using equation (D.36), the land market clearing condition for any earlier year $\tau < t$ can be written in terms of the observed variables and model solutions for our baseline year of $t = 1921$ and the relative changes in the endogenous variables of the model between those two years:

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha)\hat{v}_{nt}v_{nt}\hat{R}_{nt}R_{nt} + \left(\frac{1-\beta}{\beta}\right)\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt}, \tag{D.45}$$

where recall that a hat above a variable denotes a relative change, such that $\hat{x}_t = x_\tau/x_t$.

Second, using equations (D.41) and (D.42), expected residential income (v_{nt}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{v}_{nt}v_{nt} = \sum_{i \in \mathbb{N}} \frac{\lambda_{nit|n}^C \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in \mathbb{N}} \lambda_{n\ell t|n}^C \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}} \hat{w}_{it} w_{it}. \tag{D.46}$$

Third, using equations (D.41) and (D.43), workplace employment (L_{it}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{L}_{it} L_{it} = \sum_{n \in \mathbb{N}} \frac{\lambda^C_{nit|n} \hat{w}^{\epsilon}_{it} \hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{n\ell t}} \hat{R}_{nt} R_{nt}. \tag{D.47}$$

Finally, using equation (D.41), commuting flows ($\hat{L}_{nit}$) for any earlier year $\tau < t$ can be written in an analogous form as follows:

$$\hat{L}_{nit} L_{nit} = \frac{\lambda^C_{nit|n} \hat{w}^{\epsilon}_{it} \hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{n\ell t}} \hat{R}_{nt} R_{nt}. \tag{D.48}$$

Note that equations (D.45), (D.46), (D.47) and (D.48) above are identical to equations (16), (17), (18) and (20) in the paper. Therefore, given the same observed variables in the initial equilibrium (L_{nt}, R_{nt}, $\mathbb{Q}_{nt}$, w_{nt}, v_{nt}, L_{nit}), the same observed changes in residents and rateable values ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$) and the same estimated changes in commuting costs ($\hat{\kappa}_{nit}$), this new economic geography model predicts the same changes in workplace employment ($\hat{L}_{it}$) and commuting patterns ($\hat{L}_{nit}$) as in the paper and in the canonical urban model in Section D.1 of this web appendix.

D.3 Ricardian Spatial Model

We next derive our predictions for the impact of reversing the construction of the railway network on workplace employment and commuting in a Ricardian spatial model following Eaton and Kortum (2002) and Redding (2016). We again consider an economy that consists of a set of locations $n, i \in \mathbb{N}$. These locations are linked in goods markets through costly trade and in factor markets through migration and costly commuting. The economy as a whole is populated by a measure $\bar{L}$ of workers who are endowed with one unit of labor that is supplied inelastically.

D.3.1 Preferences and Endowments

Workers are geographically mobile and choose a pair of residence and workplace locations to maximize their utility, taking as given the choices of other firms and workers. The preferences of a worker ω who lives in location n and works in location i are defined over final goods consumption ($C_n(\omega)$), residential floor space use ($H^R_n(\omega)$), an idiosyncratic amenities shock for each workplace-residence pair ($b_{ni}(\omega)$) and iceberg commuting costs (κ_{ni}), according to the following Cobb-Douglas functional form:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega)}{\kappa_{ni}} \left(\frac{C_n(\omega)}{\alpha} \right)^{\alpha} \left(\frac{H^R_n(\omega)}{1-\alpha} \right)^{1-\alpha}, \qquad 0 < \alpha < 1, \tag{D.49}$$

The idiosyncratic amenities shock for worker ω for each residence n and workplace i ($b_{ni}(\omega)$) is drawn from an independent Fréchet distribution:

$$G_{ni}(b) = e^{-B_n b^{-\epsilon}}, \qquad B_n > 0, \epsilon > 1, \tag{D.50}$$

where the scale parameter B_n controls the average desirability of location n as a residence; and the shape parameter $\epsilon > 1$ controls the dispersion of amenities. All workers ω residing in location n and working in location i receive the same wage and make the same choices for consumption and residential floor space use. Therefore, we suppress the implicit dependence on ω from now onwards, except where important.

The goods consumption index for location n takes the constant elasticity of substitution (CES) form and is defined over a fixed continuum of goods $j \in [0, 1]$:

$$C_n = \left[\int_0^1 c_n(j)^{\rho} dj \right]^{\frac{1}{\rho}}, \tag{D.51}$$

where $c_n(j)$ is consumption of good j in country n; the CES parameter (ρ) determines the elasticity of substitution between goods ($\sigma = 1/(1-\rho) > 1$). The corresponding dual price index for goods consumption (P_n) is:

$$P_n = \left[\int_0^1 p_n(j)^{1-\sigma} dj\right]^{\frac{1}{1-\sigma}}, \qquad \sigma = \frac{1}{1-\rho} > 1, \tag{D.52}$$

where $p_n(j)$ is the price of good j in country n.

D.3.2 Production

Each good j can be produced in each location i with labor and commercial floor space under conditions of perfect competition and with a Cobb-Douglas production technology. If a good is produced by a location, the requirement of zero profits implies that the good's "free on board" price must equal its constant unit cost of production:

$$p_i(j) = \frac{w_i^{\beta} Q_i^{1-\beta}}{z_i(j)}, \qquad 0 < \beta < 1, \tag{D.53}$$

where w_i denotes the wage; Q_i is the price of floor space in location i; $z_i(j)$ is productivity; and to focus on Ricardian reasons for trade, we assume that factor intensity is the same for all goods, as controlled by β.

Each location i draws an idiosyncratic productivity $z_i(j)$ for each good j from an independent Fréchet distribution:

$$F_i(z) = e^{-A_i z^{-\theta}}, \qquad A_i > 0, \quad \theta > 1, \tag{D.54}$$

where the scale parameter A_i determines average productivity for location i and the shape parameter θ controls the dispersion of productivity across goods.

Goods can be traded between locations subject to iceberg variable trade costs, such that $d_{ni} > 1$ units of a good must be shipped from location i in order for one unit to arrive in location n (where $d_{nn} = 1$). The "cost inclusive of freight" price of a good in the location of consumption n ($p_{ni}(j)$) is thus a constant multiple of the "free on board" price of that good in the location of production i ($p_i(j)$) with that multiple determined by the iceberg trade costs:

$$p_{ni}(j) = d_{ni} p_i(j). \tag{D.55}$$

Combining equations (D.53) and (D.55), the cost to a consumer in location n of purchasing one unit of good j from location i is given by:

$$p_{ni}(j) = \frac{d_{ni} w_i^{\beta} Q_i^{1-\beta}}{z_i(j)}. \tag{D.56}$$

From profit maximization and zero profits, we obtain the results in equations (8) and (9) in the paper that payments to labor and floor space are constant shares of revenue:

$$w_i L_i = \beta Y_i, \qquad Q_i H_i^Y = (1-\beta) Y_i, \tag{D.57}$$

where L_i is workplace employment; Y_i is revenue; and H_i^Y denotes commercial use of floor space. Therefore, payments for commercial floor space are proportional to workplace income:

$$Q_i H_i^Y = \frac{1-\beta}{\beta} w_i L_i. \tag{D.58}$$

D.3.3 Trade and Market Clearing

We assume that floor space is owned by landlords, who receive income from residents' expenditure on floor space, and consume only consumption goods where they live. Total expenditure on consumption goods equals the fraction α of the total income of residents plus the entire income of landlords. This income of landlords equals $(1-\alpha)$ times the total income of residents plus $(1-\beta)$ times revenue (which equals $(1-\beta)/\beta$ times the total income of workers). Therefore total expenditure on consumption goods is:

$$X_n = P_n C_n = \alpha v_n R_n + (1-\alpha) v_n R_n + \frac{1-\beta}{\beta} w_n L_n = v_n R_n + \frac{1-\beta}{\beta} w_n L_n,$$

where v_n is the average income of location n's residents, as determined below as a function of commuting patterns, and R_n is the measure of these residents.

This Ricardian spatial model also implies a gravity equation for bilateral trade in goods between locations. Goods are homogeneous in the sense that one unit of a given good is the same as any other unit of that good. Therefore, the representative consumer in a given location sources each good from the lowest-cost supplier to that location. Using equilibrium prices (D.56) and the properties of the Fréchet distribution following Eaton and Kortum (2002), the share of the expenditure of location n on goods produced by location i is:

$$\pi_{ni} = \frac{A_i \left(d_{ni} w_i^{\beta} Q_i^{1-\beta}\right)^{-\theta}}{\sum_{s \in \mathbb{N}} A_s \left(d_{ns} w_s^{\beta} Q_s^{1-\beta}\right)^{-\theta}}, \tag{D.59}$$

where the elasticity of trade flows to trade costs is determined by the Fréchet shape parameter for productivity θ.

Goods market clearing and zero profits imply that payments to workers plus payments for commercial floor space use in each location equal expenditure on goods produced in that location:

$$w_i L_i + Q_i H_i^Y = \sum_{n \in \mathbb{N}} \pi_{ni} X_n. \tag{D.60}$$

Using equilibrium prices (D.26) and the properties of the Fréchet distribution, the consumption goods price index in equation (D.52) can be rewritten as:

$$P_n = \gamma \left[\sum_{i \in \mathbb{N}} A_i \left(d_{ni} w_i^{\beta} Q_i^{1-\beta}\right)^{-\theta}\right]^{-\frac{1}{\theta}}, \tag{D.61}$$

where $\gamma \equiv \left[\Gamma\left(\frac{\theta-(\sigma-1)}{\theta}\right)\right]^{\frac{1}{1-\sigma}}$; $\Gamma(\cdot)$ denotes the Gamma function; and we require $\theta > \sigma - 1$ to ensure a finite value for the price index.

Using the trade share (D.59), and noting that $d_{nn} = 1$, the consumption goods price index in equation (D.61) can be further rewritten solely in terms of the domestic trade share (π_{nn}), wages, the price of floor space, and parameters:

$$P_n = \gamma \left(\frac{A_n}{\pi_{nn}}\right)^{-\frac{1}{\theta}} \left(w_n^{\beta} Q_n^{1-\beta}\right). \tag{D.62}$$

Land market clearing implies that the total income received by landlords as owners of floor space (which equals rateable value ($\mathbb{Q}_n$) in our data) equals the sum of payments for the use of residential and commercial floor space:

$$\mathbb{Q}_n = Q_n H_n = (1-\alpha) v_n R_n + \left(\frac{1-\beta}{\beta}\right) w_n L_n, \tag{D.63}$$

where H_n is the quantity of floor space.

Finally, the supply of floor space (H_n) depends on both geographical land area (K_n) and the density of development as determined by the ratio of floor space to land area (h_n). Following Saiz (2010), we allow the supply of floor space to respond endogenously to changes in its price:

$$H_n = h_n K_n, \qquad\qquad h_n = h Q_n^{\mu}, \tag{D.64}$$

where h is a constant; $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.

D.3.4 Workplace and Residence Choices

Given the direct utility function (D.49), the corresponding indirect utility function for a worker ω residing in location n and working in location i is:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega) w_i}{\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}}, \tag{D.65}$$

which takes exactly the same form as equation (3) in the paper and equation (D.1) in the canonical urban model in Section D.1 of this web appendix. The only difference from the canonical urban model is in the underlying determinants of the price index for goods consumption (P_n), as now specified in equation (D.61).

Using indirect utility (D.65) and the Fréchet distribution of idiosyncratic amenities (D.50), this Ricardian spatial model exhibits the same gravity equation predictions for commuting flows as in the paper and in the canonical urban model in Section D.1 of this web appendix. The unconditional probability that a worker chooses to live in location n and work in location i is given by:

$$\lambda_{ni} = \frac{B_n \left(\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon} w_i^{\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}}, \tag{D.66}$$

which is identical to equation (5) in the paper, except that the price index for goods consumption (P_n) is now determined by equation (D.61).

Summing across workplaces, we obtain the probability that an individual lives in each location ($\lambda_n^R = R_n/\bar{L}$), while summing across residences, we have the probability that an individual works in each location ($\lambda_n^L = L_i/\bar{L}$):

$$\lambda_n^R = \frac{\sum_{\ell \in \mathbb{N}} B_n w_\ell^{\epsilon} \left(\kappa_{n\ell} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}, \quad \lambda_i^L = \frac{\sum_{r \in \mathbb{N}} B_r w_i^{\epsilon} \left(\kappa_{ri} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}. \tag{D.67}$$

Both expressions take the same form as in equation (6) in the paper and labor market clearing corresponds to $\sum_{n \in \mathbb{N}} \lambda_n^R = \sum_{i \in \mathbb{N}} \lambda_i^L = 1$.

Using the unconditional commuting probability (λ_{ni}) from equation (D.66) and the residence probability (λ_n^R) from equation (D.67), the conditional commuting probability that a worker commutes to location i conditional on residing in location n takes the same form as in equation (11) in the paper:

$$\lambda_{ni|n}^C = \frac{\left(w_i/\kappa_{ni}\right)^{\epsilon}}{\sum_{\ell \in \mathbb{N}} \left(w_\ell/\kappa_{n\ell}\right)^{\epsilon}}. \tag{D.68}$$

Using this conditional commuting probability from equation (D.68), we obtain an identical expression for per capita income conditional on living in location n as in equation (13) in the paper:

$$v_n = \sum_{i \in \mathbb{N}} \lambda_{ni|n}^C w_i. \tag{D.69}$$

Commuter market clearing again implies that employment in each location (L_i) equals the measure of workers choosing to commute to that location. Using the conditional commuting probabilities from equation (D.68), we obtain the same expression for this commuter market clearing condition as in equation (10) in the paper:

$$L_i = \sum_{n \in \mathbb{N}} \lambda^C_{ni|n} R_n. \tag{D.70}$$

Finally, population mobility and the Fréchet distribution for idiosyncratic amenities imply that expected utility is equalized across all workplace-residence pairs and takes the same form as in equation (7) in the paper:

$$\bar{U} = \mathbb{E}\left[U_{ni\omega}\right] = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right)\left[\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}\right]^{\frac{1}{\epsilon}} \quad \text{all } n, i \in \mathbb{N}, \tag{D.71}$$

where $\mathbb{E}$ is the expectations operator; the expectation is taken over the distribution for the idiosyncratic component of utility; and $\Gamma(\cdot)$ is the Gamma function.

D.3.5 Comparative Statics for Changes in Commuting Costs

We now show that this Ricardian spatial model yields exactly the same predictions for the impact of reversing the construction of the railway network on workplace employment and commuting as in the paper and the canonical urban model in Section D.1 of this web appendix, once we condition on the observed values of the endogenous variables in the initial equilibrium and the observed changes in residence employment and rateable values.

First, using equation (D.63), the land market clearing condition for any earlier year $\tau < t$ can be written in terms of observed variables and model solutions for our baseline year of $t = 1921$ and the relative changes in the endogenous variables of the model between those two years:

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha)\hat{v}_{nt}v_{nt}\hat{R}_{nt}R_{nt} + \left(\frac{1-\beta}{\beta}\right)\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt}, \tag{D.72}$$

where recall that a hat above a variable denotes a relative change, such that $\hat{x}_t = x_\tau / x_t$.

Second, using equations (D.68) and (D.69), expected residential income (v_{nt}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{v}_{nt}v_{nt} = \sum_{i \in \mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}_{it}^{\epsilon}\hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{n\ell t}^{-\epsilon}}\hat{w}_{it}w_{it}. \tag{D.73}$$

Third, using equations (D.68) and (D.70), workplace employment (L_{it}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{L}_{it}L_{it} = \sum_{n \in \mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}_{it}^{\epsilon}\hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{n\ell t}^{-\epsilon}}\hat{R}_{nt}R_{nt}. \tag{D.74}$$

Finally, using equation (D.68), commuting flows ($\hat{L}_{nit}$) for any earlier year $\tau < t$ can be written in an analogous form as follows:

$$\hat{L}_{nit}L_{nit} = \frac{\lambda^C_{nit|n}\hat{w}_{it}^{\epsilon}\hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{n\ell t}^{-\epsilon}}\hat{R}_{nt}R_{nt}. \tag{D.75}$$

Note that equations (D.72), (D.73), (D.74) and (D.75) are identical to equations (16), (17), (18) and (20) in the paper. Therefore, given the same observed variables in the initial equilibrium (L_{nt}, R_{nt}, $\mathbb{Q}_{nt}$, w_{nt}, v_{nt}, L_{nit}), the same observed changes in residents and rateable values ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$) and the same estimated changes in commuting costs ($\hat{\kappa}_{nit}^{-\epsilon}$), this Ricardian spatial model predicts the same changes in workplace employment ($\hat{L}_{it}$) and commuting patterns ($\hat{L}_{nit}$) as in the paper and the canonical urban model in Section D.1 of this web appendix.

D.4 Armington Model

Finally, we derive our predictions for the impact of reversing the construction of the railway network on workplace employment and commuting in an Armington spatial model following Armington (1969) and Allen, Arkolakis, and Li (2017). We again consider an economy that consists of a set of locations $n, i \in \mathbb{N}$. These locations are linked in goods markets through costly trade and in factor markets through migration and costly commuting. The economy as a whole is populated by a measure $\bar{L}$ of workers who are endowed with one unit of labor that is supplied inelastically.

D.4.1 Preferences and Endowments

Workers are geographically mobile and choose a pair of residence and workplace locations to maximize their utility, taking as given the choices of other firms and workers. The preferences of a worker ω who lives in location n and works in location i are defined over final goods consumption ($C_n(\omega)$), residential floor space use ($H_n^R(\omega)$), an idiosyncratic amenities shock for each workplace-residence pair ($b_{ni}(\omega)$) and iceberg commuting costs (κ_{ni}), according to the following Cobb-Douglas function form:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega)}{\kappa_{ni}} \left(\frac{C_n(\omega)}{\alpha} \right)^{\alpha} \left(\frac{H_n^R(\omega)}{1-\alpha} \right)^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{D.76}$$

The idiosyncratic amenities shock for worker ω for each residence n and workplace i ($b_{ni}(\omega)$) is drawn from an independent Fréchet distribution:

$$G_{ni}(b) = e^{-B_n b^{-\epsilon}}, \qquad B_n > 0, \epsilon > 1, \tag{D.77}$$

where the scale parameter B_n controls the average desirability of location n as a residence; and the shape parameter $\epsilon > 1$ controls the dispersion of amenities. All workers ω residing in location n and working in location i receive the same wage and make the same choices for consumption and residential floor space use. Therefore, we suppress the implicit dependence on ω from now onwards, except where important.

Consumption goods are assumed to be differentiated by location of origin according to the constant elasticity of substitution (CES) functional form. Therefore the consumption index in location n is:

$$C_n = \left[\sum_{i \in \mathbb{N}} c_{ni}^{\rho} \right]^{\frac{1}{\rho}}, \tag{D.78}$$

where c_{ni} denotes consumption in location n of the good produced by location i; and the CES parameter (ρ) determines the elasticity of substitution between the goods produced by each location ($\sigma = 1/(1-\rho) > 1$).

In this specification with differentiation by location of origin, the CES functional form implies that the marginal utility of consuming a location's good approaches infinity as consumption of that good converges to zero. Therefore, in equilibrium, each location consumes the goods produced by all locations. Using the properties of the CES functional form, the corresponding dual price index for goods consumption (P_n) is:

$$P_n = \left[\sum_{i \in \mathbb{N}} p_{ni}^{1-\sigma} \right]^{\frac{1}{1-\sigma}}, \qquad \sigma = \frac{1}{1-\rho} > 1, \tag{D.79}$$

where p_{ni} denotes the price in country n of the good produced by country i.

D.4.2 Production

Goods from each location of origin are produced under conditions of perfect competition using labor and commercial floor space. We assume that the production technology takes the Cobb-Douglas form. Using zero profits and the fact that the goods of all locations are consumed and produced in equilibrium, the "free on board" price of each location's good equals its constant unit cost of production:

$$p_i = w_i^{\beta} Q_i^{1-\beta} / A_i, \qquad 0 < \beta < 1, \tag{D.80}$$

where A_i denotes productivity; w_i is the wage; and Q_i corresponds to the price of floor space in location i.

Goods can be traded between locations subject to iceberg variable trade costs, such that $d_{ni} > 1$ units of a good must be shipped from location i in order for one unit to arrive in location n (where $d_{nn} = 1$). The "cost inclusive of freight" price of a good in the location of consumption n (p_{ni}) is thus a constant multiple of the "free on board" price of that good in the location of production i (p_i) with that multiple determined by the iceberg trade costs:

$$p_{ni} = d_{ni} p_i. \tag{D.81}$$

Combining equations (D.80) and (D.81), the cost to the consumer in location n of purchasing the good produced by location i is:

$$p_{ni} = d_{ni} w_i^{\beta} Q_i^{1-\beta} / A_i. \tag{D.82}$$

From profit maximization and zero profits, we obtain the results in equations (8) and (9) in the paper that payments to labor and commercial floor space are constant shares of revenue:

$$w_i L_i = \beta Y_i, \qquad Q_i H_i^Y = (1-\beta) Y_i, \tag{D.83}$$

where L_i is employment; Y_i is revenue and H_i^Y denotes commercial use of floor space. Therefore, payments for commercial floor space are proportional to revenue:

$$Q_i H_i^Y = \frac{1-\beta}{\beta} w_i L_i. \tag{D.84}$$

D.4.3 Trade and Market Clearing

We assume that floor space is owned by landlords, who receive income from residents' expenditure on floor space, and consume only consumption goods where they live. Total expenditure on consumption goods equals the fraction α of the total income of residents plus the entire income of landlords. This income of landlords equals $(1-\alpha)$ times the total income of residents plus $(1-\beta)$ times revenue (which equals $(1-\beta)/\beta$ times the total income of workers). Therefore, total expenditure on consumption goods is:

$$X_n = P_n C_n = \alpha v_n R_n + (1-\alpha) v_n R_n + \frac{1-\beta}{\beta} w_n L_n = v_n R_n + \frac{1-\beta}{\beta} w_n L_n,$$

where v_n is the average income of location n's residents, as determined below as a function of commuting patterns, and R_n is the measure of these residents.

This Armington model also implies a gravity equation for bilateral trade in goods between locations. Using again the properties of CES preferences, the share of expenditure of location n on goods produced by location i is:

$$\pi_{ni} = \frac{\left(d_{ni} w_i^{\beta} Q_i^{1-\beta} / A_i \right)^{1-\sigma}}{\sum_{s \in \mathbb{N}} \left(d_{ns} w_s^{\beta} Q_s^{1-\beta} / A_s \right)^{1-\sigma}}, \tag{D.85}$$

where the elasticity of trade to trade costs $(1-\sigma)$ is now determined by the elasticity of substitution between the goods produced by each location.

Goods market clearing and zero profits imply that payments to workers plus payments for commercial floor space use in each location equal expenditure on goods produced in that location:

$$w_i L_i + Q_i H_i^Y = \sum_{n \in \mathbb{N}} \pi_{ni} X_n. \tag{D.86}$$

We now use the expression for the equilibrium price of each location's good in equation (D.82) to rewrite the consumption goods price index in equation (D.79) as follows:

$$P_n = \left[\sum_{i \in \mathbb{N}} \left(d_{ni} w_i^{\beta} Q_i^{1-\beta} / A_i \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{D.87}$$

Using the trade share (D.85), and noting that $d_{nn} = 1$, the consumption goods price index in equation (D.87) can be further rewritten solely in terms of the domestic trade share (π_{nn}), wages, the price of floor space, and parameters:

$$P_n = \left(\frac{1}{\pi_{nn}} \right)^{\frac{1}{1-\sigma}} \left(\frac{w_n^{\beta} Q_n^{1-\beta}}{A_n} \right). \tag{D.88}$$

Land market clearing implies that the total income received by landlords as owners of floor space (which equals rateable value ($\mathbb{Q}_n$) in our data) equals the sum of payments for the use of residential and commercial floor space:

$$\mathbb{Q}_n = Q_n H_n = (1-\alpha) v_n R_n + \left(\frac{1-\beta}{\beta} \right) w_n L_n, \tag{D.89}$$

where H_n is the quantity of floor space.

Finally, the supply of floor space (H_n) depends on both geographical land area (K_n) and the density of development as determined by the ratio of floor space to land area (h_n). Following Saiz (2010), we allow the supply of floor space to respond endogenously to changes in its price:

$$H_n = h_n K_n, \qquad h_n = h Q_n^{\mu}, \tag{D.90}$$

where h is a constant; $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.

D.4.4 Workplace and Residence Choices

Given the direct utility function (D.76), the corresponding indirect utility function for a worker ω residing in location n and working in location i is:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega) w_i}{\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}}, \tag{D.91}$$

which takes exactly the same form as equation (3) in the paper and equation (D.1) in the canonical urban model in Section D.1 of this web appendix. The only difference from the canonical urban model is in the underlying determinants of the price index for goods consumption (P_n), as now determined by equation (D.87).

Using indirect utility (D.91) and the Fréchet distribution of idiosyncratic amenities (D.77), this Armington spatial model exhibits the same gravity equation predictions for commuting flows as in the paper and the canonical urban model in Section D.1 of this web appendix. The unconditional probability that a worker chooses to live in location n and work in location i is given by:

$$\lambda_{ni} = \frac{B_n \left(\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon} w_i^{\epsilon}}{\sum_{r\in\mathbb{N}} \sum_{\ell\in\mathbb{N}} B_r \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}}, \tag{D.92}$$

which is identical to equation (5) in the paper, except that the price index for goods consumption (P_n) is now determined by equation (D.87).

Summing across workplaces, we obtain the probability that an individual lives in each location ($\lambda_n^R = R_n/\bar{L}$), while summing across residences, we have the probability that an individual works in each location ($\lambda_n^L = L_i/\bar{L}$):

$$\lambda_n^R = \frac{\sum_{\ell\in\mathbb{N}} B_n w_\ell^{\epsilon} \left(\kappa_{n\ell} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{r\in\mathbb{N}} \sum_{\ell\in\mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}, \quad \lambda_i^L = \frac{\sum_{r\in\mathbb{N}} B_r w_i^{\epsilon} \left(\kappa_{ri} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}{\sum_{r\in\mathbb{N}} \sum_{\ell\in\mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}. \tag{D.93}$$

Both expressions take the same form as in equation (6) in the paper and labor market clearing corresponds to $\sum_{n\in\mathbb{N}} \lambda_n^R = \sum_{i\in\mathbb{N}} \lambda_i^L = 1$.

Using the unconditional commuting probability (λ_{ni}) from equation (D.92) and the residence probability (λ_n^R) from equation (D.93), the conditional commuting probability that a worker commutes to location i conditional on residing in location n takes the same form as in equation (11) in the paper:

$$\lambda_{ni|n}^C = \frac{(w_i/\kappa_{ni})^{\epsilon}}{\sum_{\ell\in\mathbb{N}} (w_\ell/\kappa_{n\ell})^{\epsilon}}. \tag{D.94}$$

Using this conditional commuting probability from equation (D.94), we obtain an identical expression for per capita income conditional on living in location n as in equation (13) in the paper:

$$v_n = \sum_{i\in\mathbb{N}} \lambda_{ni|n}^C w_i. \tag{D.95}$$

Commuter market clearing again implies that employment in each location (L_i) equals the measure of workers choosing to commute to that location. Using the conditional commuting probabilities from equation (D.94), we obtain the same expression for this commuter market clearing condition as in equation (10) in the paper:

$$L_i = \sum_{n\in\mathbb{N}} \lambda_{ni|n}^C R_n. \tag{D.96}$$

Finally, population mobility and the Fréchet distribution for idiosyncratic amenities imply that expected utility is equalized across all workplace-residence pairs and takes the same form as in equation (7) in the paper:

$$\bar{U} = \mathbb{E}\left[U_{ni\omega}\right] = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right)\left[\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}} B_r \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}\right]^{\frac{1}{\epsilon}} \quad \text{all } n, i \in \mathbb{N}, \tag{D.97}$$

where $\mathbb{E}$ is the expectations operator; the expectation is taken over the distribution for the idiosyncratic component of utility; and $\Gamma(\cdot)$ is the Gamma function.

D.4.5 Comparative Statics for Changes in Commuting Costs

We now show that this Armington model yields exactly the same predictions for the impact of reversing the construction of the railway network on workplace employment and commuting as in the paper and the canonical urban model in Section D.1 of this web appendix, once we condition on the observed values of the endogenous variables in the initial equilibrium and the observed changes in residence employment and rateable values.

First, using equation (D.89), the land market clearing condition for any earlier year $\tau < t$ can be written in terms of observed variables and model solutions for our baseline year of $t = 1921$ and the relative changes in the endogenous variables of the model between those two years:

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha)\hat{v}_{nt}v_{nt}\hat{R}_{nt}R_{nt} + \left(\frac{1-\beta}{\beta}\right)\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt}, \tag{D.98}$$

where recall that a hat above a variable denotes a relative change, such that $\hat{x}_t = x_\tau / x_t$.

Second, using equations (D.94) and (D.95), expected residential income (v_{nt}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{v}_{nt}v_{nt} = \sum_{i\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{w}_{it}w_{it}. \tag{D.99}$$

Third, using equations (D.94) and (D.96), workplace employment (L_{it}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{L}_{it}L_{it} = \sum_{n\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{R}_{nt}R_{nt}. \tag{D.100}$$

Finally, using equation (D.94), commuting flows ($\hat{L}_{nit}$) for any earlier year $\tau < t$ can be written in an analogous form as follows:

$$\hat{L}_{nit}L_{nit} = \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{R}_{nt}R_{nt}. \tag{D.101}$$

Note that equations (D.98), (D.99), (D.100) and (D.101) are identical to equations (16), (17), (18) and (20) in the paper. Therefore, given the same observed variables in the initial equilibrium (L_{nt}, R_{nt}, $\mathbb{Q}_{nt}$, w_{nt}, v_{nt}, L_{nit}), the same observed changes in residents and rateable values ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$) and the same estimated changes in commuting costs ($\hat{\kappa}^{-\epsilon}_{nit}$), this Armington spatial model predicts the same changes in workplace employment ($\hat{L}_{it}$) and commuting patterns ($\hat{L}_{nit}$) as in the paper and the canonical urban model in Section D.1 of this web appendix.

D.5 Non-traded Services

In this section of the appendix, we show that it is straightforward to extend our analysis to incorporate non-traded services in addition to traded goods. For simplicity, we demonstrate this extension for the canonical urban model in Section D.1 of this web appendix, but the analysis is directly analogous for the New Economic Geography model in Section D.2 of this web appendix, the Ricardian spatial model in Section D.3 of this web appendix, and the Armington model in Section D.4 of this web appendix.

D.5.1 Preferences and Production

The preferences of a worker ω who resides in location n and works in location i are again defined over a final consumption good and residential floor space, as in equation (D.1) in Section D.1 of this web appendix. This final good is assumed to be costlessly traded between locations within Greater London, such that $P_i = P$ for all i. We assume that this final good is produced using labor, residential floor space and non-traded services according to a constant returns to scale technology under conditions of perfect competition.[1] For simplicity, we assume that this production

[1] London had substantial employment in both industry and services during our sample period. It was one of the main industrial centers in the United Kingdom, with manufacturing accounting for over 25 percent of employment in Greater London in the population census of 1911. In this extension, we interpret employment in services as a non-traded input into the production of a final consumption good.

technology takes the Cobb-Douglas form. Using profit maximization and zero profits, the common final goods price is equal to its unit cost of production for all locations with positive production:

$$P = \frac{1}{A_i^F} w_i^{\beta} p_i^{\gamma} Q_i^{1-\beta-\gamma}, \qquad 0 < \beta, \gamma < 1, \qquad 0 < \beta + \gamma < 1, \tag{D.102}$$

where A_i^F is final goods productivity; w_i is the wage; p_i is the price of non-traded services in location i; and Q_i is the price of floor space.

We assume that non-traded services are produced using labor and floor space according to a constant returns to scale technology under conditions of perfect competition. For simplicity, we again assume that this production technology takes the Cobb-Douglas form. Using profit maximization and zero profits, the price for these non-traded services is equal to their unit cost of production for all locations with positive production:

$$p_i = \frac{1}{A_i^I} w_i^{\mu} Q_i^{1-\mu}, \qquad 0 < \mu < 1, \tag{D.103}$$

where A_i^I is non-traded services productivity in location i.

Using this zero-profit condition for non-traded services (D.103), the corresponding zero-profit condition for the final good (D.102) can be re-written as follows:

$$P = \frac{1}{\tilde{A}_i} w_i^{\tilde{\beta}} Q_i^{1-\tilde{\beta}}, \qquad\qquad A_i \equiv A_i^F \left(A_i^I\right)^{\gamma}, \tag{D.104}$$

$$\tilde{\beta} \equiv \beta + \gamma\mu, \qquad 1 - \tilde{\beta} = (1 - \beta - \gamma) + \gamma(1 - \mu) = 1 - (\beta + \gamma\mu), \qquad 0 < \tilde{\beta} < 1,$$

where $\tilde{\beta}$ is a composite measure of labor intensity and $\tilde{A}_i$ is a composite measure of productivity for the final goods and non-traded services sectors as a whole. Note that this composite zero-profit condition (D.104) takes exactly the same form in equation (D.3) for the canonical urban model, with $\tilde{\beta}$ and $\tilde{A}_i$ replacing β and A_i respectively.

From profit maximization and zero profits, we obtain analogous results to those in equations (8) and (9) in the paper, with payments to labor and floor space equal to constant shares of revenue:

$$w_i L_i = \tilde{\beta} Y_i, \qquad\qquad Q_i H_i^Y = (1 - \tilde{\beta}) Y_i. \tag{D.105}$$

where L_i is employment; Y_i is revenue; and H_i^Y denotes commercial floor space use. Therefore payments for commercial floor space are proportional to workplace income:

$$Q_i H_i^Y = \frac{1 - \tilde{\beta}}{\tilde{\beta}} w_i L_i, \tag{D.106}$$

where again $\tilde{\beta}$ replaces β.

Total demand for floor space equals the sum of demand for floor space for residential use, for commercial use for the final good, and for commercial use for non-traded services. Land market clearing implies that the total income received by landlords as owners of floor space (which equals rateable value ($\mathbb{Q}_n$) in our data) equals the sum of payments for the use of residential and commercial floor space:

$$\mathbb{Q}_n = Q_n H_n = (1 - \alpha)\, v_n R_n + \left(\frac{1 - \tilde{\beta}}{\tilde{\beta}}\right) w_n L_n, \tag{D.107}$$

where H_n is the quantity of floor space. Equation (D.107) again takes exactly the same form as in equation (12) in the paper, with $\tilde{\beta}$ replacing β.

Finally, the supply of floor space (H_n) depends on both geographical land area (K_n) and the density of development as determined by the ratio of floor space to land area (h_n). Following Saiz (2010), we allow the supply of floor space to respond endogenously to changes in its price:

$$H_n = h_n K_n, \qquad h_n = h Q_n^{\mu}, \tag{D.108}$$

where h is a constant; $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.

D.5.2 Workplace and Residence Choices

The indirect utility function and specification for idiosyncratic amenities remain the same as for the canonical urban model in equations (D.1) and (D.2) in Section D.1 of this web appendix. Therefore this extension of the canonical urban model to incorporate non-traded services exhibits the same gravity equation predictions for commuting flows as in the paper and in the canonical urban model in Section D.1 of this web appendix. The unconditional probability that a worker chooses to live in location n and work in location i is given by:

$$\lambda_{ni} = \frac{B_n \left(\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon} w_i^{\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}}, \tag{D.109}$$

which is identical to equation (5) in the paper.

Summing across workplaces, we obtain the probability that an individual lives in each location ($\lambda_n^R = R_n/\bar{L}$), while summing across residences, we have the probability that an individual works in each location ($\lambda_n^L = L_i/\bar{L}$):

$$\lambda_n^R = \frac{\sum_{\ell \in \mathbb{N}} B_n w_\ell^{\epsilon} \left(\kappa_{n\ell} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}, \quad \lambda_i^L = \frac{\sum_{r \in \mathbb{N}} B_r w_i^{\epsilon} \left(\kappa_{ri} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}. \tag{D.110}$$

Both expressions take the same form as in equation (6) in the paper and labor market clearing corresponds to $\sum_{n \in \mathbb{N}} \lambda_n^R = \sum_{i \in \mathbb{N}} \lambda_i^L = 1$.

Using the unconditional commuting probability (λ_{ni}) from equation (D.109) and the residence probability (λ_n^R) from equation (D.110), the conditional commuting probability that a worker commutes to location i conditional on residing in location n takes the same form as in equation (11) in the paper:

$$\lambda_{ni|n}^C = \frac{\left(w_i/\kappa_{ni}\right)^{\epsilon}}{\sum_{\ell \in \mathbb{N}} \left(w_\ell/\kappa_{n\ell}\right)^{\epsilon}}. \tag{D.111}$$

Using this conditional commuting probability from equation (D.111), we obtain an identical expression for per capita income conditional on living in location n as in equation (13) in the paper:

$$v_n = \sum_{i \in \mathbb{N}} \lambda_{ni|n}^C w_i. \tag{D.112}$$

Total employment in each location equals the sum of workers employed in producing the final good and non-traded services. Commuter market clearing again implies that this total employment in each location (L_i) equals the measure of workers choosing to commute to that location. Using the conditional commuting probabilities from equation (D.111), we obtain the same expression for this commuter market clearing condition as in equation (10) in the paper:

$$L_i = \sum_{n \in \mathbb{N}} \lambda_{ni|n}^C R_n. \tag{D.113}$$

Finally, population mobility and the Fréchet distribution for idiosyncratic amenities imply that expected utility is equalized across all workplace-residence pairs and takes the same form as in equation (7) in the paper:

$$\bar{U} = \mathbb{E}\left[U_{ni\omega}\right] = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right)\left[\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}} B_r\left(\kappa_{r\ell}P_r^{\alpha}Q_r^{1-\alpha}\right)^{-\epsilon} w_\ell^{\epsilon}\right]^{\frac{1}{\epsilon}} \quad \text{all } n, i \in \mathbb{N}, \tag{D.114}$$

where $\mathbb{E}$ is the expectations operator; the expectation is taken over the distribution for the idiosyncratic component of utility; and $\Gamma(\cdot)$ is the Gamma function.

D.5.3 Comparative Statics for Changes in Commuting Costs

We now show that this extension of the canonical urban model to incorporate non-traded services yields exactly the same predictions for the impact of reversing the construction of the railway network on workplace employment and commuting as in the paper, once we condition on the observed values of the endogenous variables in the initial equilibrium and the observed changes in residence employment and rateable values.

First, using equation (D.107), the land market clearing condition for any earlier year $\tau < t$ can be written in terms of observed variables and model solutions for our baseline year of $t = 1921$ and the relative changes in the endogenous variables of the model between those two years:

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha)\hat{v}_{nt}v_{nt}\hat{R}_{nt}R_{nt} + \left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt}, \tag{D.115}$$

where recall that a hat above a variable denotes a relative change, such that $\hat{x}_t = x_\tau / x_t$.

Second, using equations (D.111) and (D.112), expected residential income (v_{nt}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{v}_{nt}v_{nt} = \sum_{i\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{w}_{it}w_{it}. \tag{D.116}$$

Third, using equations (D.111) and (D.113), workplace employment (L_{it}) for any earlier year $\tau < t$ can be written in a similar form as:

$$\hat{L}_{it}L_{it} = \sum_{n\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{R}_{nt}R_{nt}. \tag{D.117}$$

Finally, using equation (D.111), commuting flows ($\hat{L}_{nit}$) for any earlier year $\tau < t$ can be written in an analogous form as follows:

$$\hat{L}_{nit}L_{nit} = \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{R}_{nt}R_{nt}. \tag{D.118}$$

Note that equations (D.115), (D.116), (D.117) and (D.118) take exactly the same form as equations (16), (17), (18) and (20) in the paper, with $\tilde{\beta}$ replacing β. Therefore, given the same observed variables in the initial equilibrium (L_{nt}, R_{nt}, $\mathbb{Q}_{nt}$, w_{nt}, v_{nt}, L_{nit}), the same observed changes in residents and rateable values ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$), the same estimated changes in commuting costs ($\hat{\kappa}^{-\epsilon}_{nit}$), and assuming the same value for $\tilde{\beta}$ as for β, this extension of the canonical urban model to incorporate non-traded services predicts the same changes in workplace employment ($\hat{L}_{it}$) and commuting patterns ($\hat{L}_{nit}$) as in the paper and in the canonical urban model in Section D.1 of this web appendix.

E Counterfactuals

In this section of the web appendix, we report further details on the counterfactuals from Section 7 of the paper. In Section E.1, we show that the model can be inverted to recover unique values for the unobserved relative changes in production fundamentals ($\hat{a}_{nt}$), residential fundamentals ($\hat{b}_{nt}$) and the supply of floor space ($\hat{H}_{nt}$) across locations (Proposition E.1). In Section E.2, we show that the model has a unique equilibrium in the special case in which productivity, amenities and the supply of floor space are exogenous, where $\psi = \eta = \mu = 0$ (Proposition E.2). In Section E.3, we report the counterfactuals for the "closed-city" robustness test discussed in Section 7 of the paper.

E.1 Model Inversion

Proposition E.1 *Given known values of the model parameters ($\alpha, \beta, \epsilon, \phi, \mu, \psi, \eta$) and an estimated change in commuting costs ($\hat{\kappa}_{nit}$), there is a one-to-one mapping from the observed variables in the initial equilibrium (L_{nt}, R_{nt}, $\mathbb{Q}_{nt}$, λ_{nit}) and the observed changes in residents ($\hat{R}_{nt}$) and rateable values ($\hat{\mathbb{Q}}_{nt}$) to the unobserved relative changes in production fundamentals ($\hat{a}_{nt}$), residential fundamentals ($\hat{b}_{nt}$) and the supply of floor space ($\hat{H}_{nt}$) across locations.*

Proof. We observe the following variables in the initial equilibrium in our baseline year of $t = 1921$: employment (L_{nt}), residents (R_{nt}), rateable values ($\mathbb{Q}_{nt}$), and the commuting probabilities (λ_{nit}, $\lambda^{C}_{nit|n}$). From Lemma 1, we can use the combined commuter market clearing condition in equation (15) in the paper to solve for unique values for wages (w_{nt}) and per capital residential income (v_{nt}) in the initial equilibrium in our baseline year of $t = 1921$.

We also observe changes in residents ($\hat{R}_{nt}$) and rateable values ($\hat{\mathbb{Q}}_{nt}$). From Lemma 2, we can use the combined commuter market clearing condition in equation (19) in the paper to solve for unique changes in wages ($\hat{w}_{nt}$), per capita residential income ($\hat{v}_{nt}$), and employment ($\hat{L}_{nt}$) between year $\tau < t$ and our baseline year t.

We now solve for changes in the price (Q_{nt}) and quantity of floor space (H_{nt}) from our observed rateable values ($\mathbb{Q}_{nt}$). From the floor space supply function (14) in the paper, the change in the price of floor space ($\hat{Q}_{nt}$) is:

$$\hat{Q}_{nt} = \hat{\mathbb{Q}}_{nt}^{\frac{1}{1+\mu}}, \tag{E.1}$$

where recall that $\hat{Q}_{nt} = Q_{n\tau}/Q_{nt}$ and we observe the change in rateable values ($\hat{\mathbb{Q}}_{nt}$). Dividing equation (E.1) by its geometric mean across locations, we recover the unique relative change in the price of floor space across locations:

$$\frac{\hat{Q}_{nt}}{\tilde{\hat{Q}}_{t}} = \left(\frac{\hat{\mathbb{Q}}_{nt}}{\tilde{\hat{\mathbb{Q}}}_{t}}\right)^{\frac{1}{1+\mu}}, \tag{E.2}$$

where a tilde above a variable denotes a geometric mean such that:

$$\tilde{\hat{Q}}_{t} = \left(\prod_{i \in \mathbb{N}} \hat{Q}_{it}\right)^{\frac{1}{N}}. \tag{E.3}$$

Similarly, from the floor space supply function (14) in the paper, the change in the quantity of floor space ($\hat{H}_{nt}$) is:

$$\hat{H}_{nt} = \hat{\mathbb{Q}}_{nt}^{\frac{\mu}{1+\mu}}, \tag{E.4}$$

where recall that we observe the change rateable values ($\hat{\mathbb{Q}}_{nt}$). Dividing equation (E.4) by its geometric mean across locations, we obtain unique solutions for the relative change in the supply of floor space across locations:

$$\frac{\hat{H}_t}{\tilde{\hat{H}}_t} = \left(\frac{\hat{\mathbb{Q}}_{nt}}{\tilde{\hat{\mathbb{Q}}}_{nt}} \right)^{\frac{\mu}{1+\mu}}, \tag{E.5}$$

where recall that $\hat{\mathbb{Q}}_{nt}$ is observed.

We next solve for changes in productivity and production fundamentals. From the zero-profit condition in equation (27) in the paper for years τ and t, the change in productivity ($\hat{A}_{nt}$) and the common change in the price of the final good ($\hat{P}_t$) must satisfy:

$$\hat{P}_t \hat{A}_{nt} = \hat{w}_{nt}^{\beta} \hat{Q}_{nt}^{1-\beta}, \tag{E.6}$$

where we have solved for the change in the wage ($\hat{w}_{nt}$) and the price of floor space ($\hat{Q}_{nt}$) above. Dividing equation (E.6) by its geometric mean across locations, the common change in the price of the final good ($\hat{P}_t$) cancels, and we obtain unique solutions for the relative change across locations in productivity:

$$\frac{\hat{A}_{nt}}{\tilde{\hat{A}}_t} = \left(\frac{\hat{w}_{nt}}{\tilde{\hat{w}}_t} \right)^{\beta} \left(\frac{\hat{Q}_{nt}}{\tilde{\hat{Q}}_t} \right)^{1-\beta}. \tag{E.7}$$

From the specification for productivity in equation (28) in the paper for years τ and t, the changes in productivity ($\hat{A}_{nt}$) and production fundamentals ($\hat{a}_{nt}$) must satisfy:

$$\hat{a}_{nt} = \hat{A}_{nt} \hat{L}_{nt}^{-\psi}. \tag{E.8}$$

Dividing equation (E.7) by its geometric mean across locations, we obtain:

$$\frac{\hat{a}_{nt}}{\tilde{\hat{a}}_{nt}} = \frac{\hat{A}_{nt}}{\tilde{\hat{A}}_t} \left(\frac{\hat{L}_{nt}}{\tilde{\hat{L}}_{nt}} \right)^{-\psi}. \tag{E.9}$$

Using equation (E.7) to substitute for the relative change across locations in productivity ($\hat{A}_{nt}/\tilde{\hat{A}}_t$) in equation (E.9), we obtain unique solutions for the relative change across locations in production fundamentals:

$$\frac{\hat{a}_{nt}}{\tilde{\hat{a}}_{nt}} = \left(\frac{\hat{w}_{nt}}{\tilde{\hat{w}}_t} \right)^{\beta} \left(\frac{\hat{Q}_{nt}}{\tilde{\hat{Q}}_t} \right)^{1-\beta} \left(\frac{\hat{L}_{nt}}{\tilde{\hat{L}}_{nt}} \right)^{-\psi}, \tag{E.10}$$

where ($\hat{w}_{nt}$, $\hat{Q}_{nt}$, $\hat{L}_{nt}$) are observed. All boroughs have positive changes in wages ($\hat{w}_{nt}$), the price of floor space ($\hat{Q}_{nt}$), and workplace employment ($\hat{L}_{nt}$), which in turn implies that all boroughs have positive values for the changes in production fundamentals ($\hat{a}_{nt}$).

Finally, we solve for changes in residential amenities and residential fundamentals. From the residential choice probability (λ_{nt}^{R}) in equation (6) in the paper, the change in residential amenities ($\hat{B}_{nt}$) must satisfy:

$$\hat{\lambda}_{nt}^{R} \lambda_{nt}^{R} = \frac{\sum_{\ell \in \mathbb{N}} \lambda_{n\ell t} \hat{B}_{nt} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{n\ell t}^{-\epsilon} \hat{P}_t^{-\epsilon\alpha} \hat{Q}_{nt}^{-\epsilon(1-\alpha)}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \lambda_{r\ell t} \hat{B}_{rt} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{r\ell t}^{-\epsilon} \hat{P}_t^{-\epsilon\alpha} \hat{Q}_{rt}^{-\epsilon(1-\alpha)}}, \tag{E.11}$$

From the population mobility condition in equation (7) in the paper, we also have:

$$\hat{\bar{U}}_t \bar{U}_t = \delta \left[\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \lambda_{r\ell t} \hat{B}_{rt} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{r\ell t}^{-\epsilon} \hat{P}_t^{-\epsilon\alpha} \hat{Q}_{rt}^{-\epsilon(1-\alpha)} \right]^{\frac{1}{\epsilon}}, \tag{E.12}$$

where under our open-city assumption $\hat{\tilde{U}}_t = 1$. Using this population mobility condition (E.12) in the residential choice probability (E.11), we have:

$$\hat{\lambda}_{nt}^R \lambda_{nt}^R = \left(\frac{\bar{U}}{\delta}\right)^{\epsilon} \sum_{\ell \in \mathbb{N}} \lambda_{n\ell t} \hat{B}_{nt} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{n\ell t}^{-\epsilon} \hat{P}_t^{-\epsilon\alpha} \hat{Q}_{nt}^{-\epsilon(1-\alpha)}, \tag{E.13}$$

where we have used $\hat{\tilde{U}}_t = 1$. Dividing equation (E.13) by its geometric mean across locations, we obtain:

$$\frac{\hat{\lambda}_{nt}^R}{\tilde{\hat{\lambda}}_{nt}^R} \frac{\lambda_{nt}^R}{\tilde{\lambda}_{nt}^R} = \frac{\hat{B}_{nt}}{\tilde{\hat{B}}_{nt}} \left(\frac{\hat{Q}_{nt}}{\tilde{\hat{Q}}_{nt}}\right)^{-\epsilon(1-\alpha)} \left[\frac{\sum_{\ell \in \mathbb{N}} \lambda_{n\ell t} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{n\ell t}^{-\epsilon} \hat{P}_t^{-\epsilon\alpha}}{\prod_{n \in \mathbb{N}} \left[\sum_{\ell \in \mathbb{N}} \lambda_{n\ell t} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{n\ell t}^{-\epsilon} \hat{P}_t^{-\epsilon\alpha}\right]^{\frac{1}{N}}}\right], \tag{E.14}$$

where the term $\left(\bar{U}/\delta\right)^{\epsilon}$ has cancelled from the right-hand side. Cancelling the common price term in P_t from the numerator and the denominator on the right-hand side, equation (E.14) can be re-written as:

$$\frac{\hat{\lambda}_{nt}^R}{\tilde{\hat{\lambda}}_{nt}^R} \frac{\lambda_{nt}^R}{\tilde{\lambda}_{nt}^R} = \frac{\hat{B}_{nt}}{\tilde{\hat{B}}_{nt}} \left(\frac{\hat{Q}_{nt}}{\tilde{\hat{Q}}_{nt}}\right)^{-\epsilon(1-\alpha)} \left(\frac{\widehat{RMA}_{nt}}{\widetilde{\widehat{RMA}}_t}\right)^{\epsilon}, \tag{E.15}$$

where $\widehat{RMA}_{nt}$ is a measure of the change in residents' commuting market access such that

$$\frac{\widehat{RMA}_{nt}}{\widetilde{\widehat{RMA}}_t} = \left(\frac{\sum_{\ell \in \mathbb{N}} \lambda_{n\ell t} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{n\ell t}^{-\epsilon}}{\prod_{n \in \mathbb{N}} \left[\sum_{\ell \in \mathbb{N}} \lambda_{n\ell t} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{n\ell t}^{-\epsilon}\right]^{\frac{1}{N}}}\right)^{\frac{1}{\epsilon}}. \tag{E.16}$$

Re-arranging equation (E.15), we obtain unique solutions for the relative change across locations in residential amenities:

$$\frac{\hat{B}_{nt}}{\tilde{\hat{B}}_{nt}} = \frac{\hat{\lambda}_{nt}^R}{\tilde{\hat{\lambda}}_{nt}^R} \frac{\lambda_{nt}^R}{\tilde{\lambda}_{nt}^R} \left(\frac{\hat{Q}_{nt}}{\tilde{\hat{Q}}_{nt}}\right)^{\epsilon(1-\alpha)} \left(\frac{\widehat{RMA}_{nt}}{\widetilde{\widehat{RMA}}_t}\right)^{-\epsilon}. \tag{E.17}$$

From the specification for amenities in equation (26) in the paper for years τ and t, the changes in amenities ($\hat{B}_{nt}$) and residential fundamentals ($\hat{b}_{nt}$) must satisfy:

$$\hat{b}_{nt} = \hat{B}_{nt} \hat{R}_{nt}^{-\eta} \tag{E.18}$$

Dividing equation (E.18) by its geometric mean across locations, we obtain:

$$\frac{\hat{b}_{nt}}{\tilde{\hat{b}}_{nt}} = \frac{\hat{B}_{nt}}{\tilde{\hat{B}}_t} \left(\frac{\hat{R}_{nt}}{\tilde{\hat{R}}_{nt}}\right)^{-\eta}. \tag{E.19}$$

Using equation (E.17) to substitute for the relative change across locations in amenities ($\hat{B}_{nt}/\tilde{\hat{B}}_t$) in equation (E.19), we obtain unique solutions for the relative change across locations in residential fundamentals:

$$\frac{\hat{b}_{nt}}{\tilde{\hat{b}}_{nt}} = \frac{\hat{\lambda}_{nt}^R}{\tilde{\hat{\lambda}}_{nt}^R} \left(\frac{\lambda_{nt}^R}{\tilde{\lambda}_{nt}^R}\right)^{1-\eta} \left(\frac{\hat{Q}_{nt}}{\tilde{\hat{Q}}_{nt}}\right)^{\epsilon(1-\alpha)} \left(\frac{\widehat{RMA}_{nt}}{\widetilde{\widehat{RMA}}_t}\right)^{-\epsilon}, \tag{E.20}$$

where ($\hat{\lambda}_{nt}^R$, λ_{nt}^R, $\hat{Q}_{nt}$) are observed; we also observe λ_{nit}, solve for $\hat{w}_{nt}$ and estimate $\hat{\kappa}_{nit}^{-\epsilon}$, such that $\widehat{RMA}_{nt}$ is known. All boroughs have positive changes in residence employment ($\hat{R}_{nt}$), the price of floor space ($\hat{Q}_{nt}$) and wages ($\hat{w}_{nt}$) in all years, which in turn implies that all boroughs have positive changes in residential fundamentals ($\hat{b}_{nt}$).

We have thus determined unique relative values across locations for the change in the supply of floor space ($\hat{H}_{nt}/\tilde{\hat{H}}_{nt}$), the change in production fundamentals ($\hat{a}_{nt}/\tilde{\hat{a}}_{nt}$), and the change in residential fundamentals ($\hat{b}_{nt}/\tilde{\hat{b}}_{nt}$). ■

E.2 Existence and Uniqueness

Proposition E.2 *Assume exogenous, finite and strictly positive location characteristics ($P_t \in (0,\infty)$, $A_{nt} \in (0,\infty)$, $B_{nt} \in (0,\infty)$, $\kappa_{nit} \in (0,\infty) \times (0,\infty)$, $H_{nt} = H_n \in (0,\infty)$), which corresponds to $\psi = \eta = \mu = 0$. Under these assumptions, there exists a unique general equilibrium vector (λ^L_{nt}, λ^R_{nt}, Q_{nt}, w_{nt}, $\bar{L}$).*

Proof. Assume exogenous, finite and strictly positive location characteristics ($P_t \in (0,\infty)$, $A_{nt} \in (0,\infty)$, $B_{nt} \in (0,\infty)$, $\kappa_{nit} \in (0,\infty) \times (0,\infty)$, $H_{nt} = H_n = hK_n \in (0,\infty)$), which corresponds to $\psi = \eta = \mu = 0$. Under these assumptions, all locations are incompletely specialized as both workplaces and residences, because the support of the Fréchet distribution for idiosyncratic amenities is unbounded from above. Using the probability of residing in a location (equation (6) in the paper for λ^R_{nt}), the probability of working in a location (equation (6) in the paper for λ^L_{nt}), the zero-profit condition in equation (27) in the paper, and the population mobility condition between the city and the larger economy in equation (7) in the paper, the fraction of workers residing in location n can be written as:

$$\lambda^R_{nt} = \frac{R_{nt}}{\bar{L}_t} = \left(\frac{\gamma}{\bar{U}_t}\right)^{\epsilon} \sum_{\ell \in \mathbb{N}} B_{nt} A_{\ell t}^{\epsilon/\beta} \kappa_{n\ell t}^{-\epsilon} Q_{\ell t}^{-\epsilon(1-\beta)/\beta} Q_{nt}^{-\epsilon(1-\alpha)},$$

while the fraction of workers employed in location n can be written as:

$$\lambda^L_{nt} = \frac{L_{nt}}{\bar{L}_t} = \left(\frac{\gamma}{\bar{U}_t}\right)^{\epsilon} \sum_{r \in \mathbb{N}} B_{rt} A_{nt}^{\epsilon/\beta} \kappa_{rit}^{-\epsilon} Q_{it}^{-\epsilon(1-\beta)/\beta} Q_{rt}^{-\epsilon(1-\alpha)},$$

and expected worker income conditional on residing in block i from equation (13) in the paper can be written as:

$$v_{nt} = \sum_{i \in \mathbb{N}} \frac{A_{it}^{\epsilon/\beta} \kappa_{nit}^{-\epsilon} Q_{it}^{-\epsilon(1-\beta)/\beta}}{\sum_{\ell \in \mathbb{N}} A_{\ell t}^{\epsilon/\beta} \kappa_{n\ell t}^{-\epsilon} Q_{\ell t}^{-\epsilon(1-\beta)/\beta}} \left[A_{it}^{1/\beta} Q_{it}^{-(1-\beta)/\beta} \right],$$

and the land market clearing condition from equation (12) in the paper can be written as:

$$\left(\frac{1-\beta}{\beta}\right) \frac{w_{nt}\lambda^L_{nt}}{Q_{nt}} + (1-\alpha) \frac{v_{nt}\lambda^R_{nt}}{Q_{nt}} = \frac{hK_n}{\bar{L}_t}.$$

Combining the above relationships, this land market clearing condition can be re-expressed as:

$$\begin{aligned} D_{nt}(\boldsymbol{Q_t}) = \tfrac{1-\beta}{\beta} \left[\frac{A_{nt}^{1/\beta}}{Q_{nt}^{1+(1-\beta)/\beta}} \right] \left[\textstyle\sum_{r \in \mathbb{N}} \frac{B_{rt} A_{it}^{\epsilon/\beta} \kappa_{rit}^{-\epsilon}}{Q_{it}^{\epsilon(1-\beta)/\beta} Q_{rt}^{\epsilon(1-\alpha)}} \right] \\ + \tfrac{1-\alpha}{Q_{nt}} \left[\textstyle\sum_{i \in \mathbb{N}} \left(\frac{A_{it}^{\epsilon/\beta} \kappa_{nit}^{-\epsilon} Q_{it}^{-\epsilon(1-\beta)/\beta}}{\sum_{\ell \in \mathbb{N}} A_{\ell t}^{\epsilon/\beta} \kappa_{n\ell t}^{-\epsilon} Q_{\ell t}^{-\epsilon(1-\beta)/\beta}} \right) \frac{A_{it}^{1/\beta}}{Q_{it}^{(1-\beta)/\beta}} \right] \left[\textstyle\sum_{\ell \in \mathbb{N}} \frac{B_{nt} A_{\ell t}^{\epsilon/\beta} \kappa_{n\ell t}^{-\epsilon}}{Q_{\ell t}^{\epsilon(1-\beta)/\beta} Q_{nt}^{\epsilon(1-\alpha)}} \right] - hK_n = 0, \end{aligned}$$

for all $n \in R$, where we have chosen units in which to measure utility such that $(\bar{U}_t/\gamma)^{\epsilon}/\bar{L}_t = 1$ for a given year t. The above land market clearing condition provides a system of equations for the N boroughs in terms of the N unknown floor space prices Q_{nt}, which has the following properties:

$$\lim_{Q_{nt} \to 0} D_{nt}(\boldsymbol{Q_t}) = \infty > hK_n, \qquad \lim_{Q_{nt} \to \infty} D_{nt}(\boldsymbol{Q_t}) = 0 < hK_n,$$

$$\frac{dD_{nt}(\boldsymbol{Q_t})}{dQ_{nt}} < 0, \qquad \frac{dD_{nt}(\boldsymbol{Q_t})}{dQ_{it}} < 0, \qquad \left| \frac{dD_{nt}(\boldsymbol{Q_t})}{dQ_{nt}} \right| > \left| \frac{dD_{nt}(\boldsymbol{Q_t})}{dQ_{it}} \right|.$$

It follows that there exists a unique vector of floor space prices $\boldsymbol{Q_t}$ that solves this system of land market clearing conditions. Having solved for the vector of floor space prices ($\boldsymbol{Q_t}$), the vector of wages $\boldsymbol{w_t}$ follows immediately from the zero-profit condition for production in equation (27) in the paper. Given floor space prices ($\boldsymbol{Q_t}$) and wages ($\boldsymbol{w_t}$), the probability of residing in a location ($\boldsymbol{\lambda^R_t}$) follows immediately from equation (6) in the paper, and the probability of working in a location ($\boldsymbol{\lambda_t^L}$) follows immediately from equation (6) in the paper. Having solved for ($\boldsymbol{\lambda_t^L}$, $\boldsymbol{\lambda_t^R}$, $\boldsymbol{Q_t}$, $\boldsymbol{w_t}$), the total measure of workers residing in the city can be recovered from our choice of units in which to measure utility $(\bar{U}_t/\gamma)^\epsilon/\bar{L}_t = 1)$ for our given year t, which together with the population mobility condition in equation (7) in the paper implies:

$$\bar{L}_t = \left[\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}} B_{rt} w_{\ell t}^{\epsilon}\left(\kappa_{r\ell t} Q_{rt}^{1-\alpha}\right)^{-\epsilon}\right].$$

We therefore obtain $\boldsymbol{L_t} = \boldsymbol{\lambda_t^L}\bar{L}_t$ and $\boldsymbol{R_t} = \boldsymbol{\lambda_t^R}\bar{L}_t$. This completes the determination of the equilibrium vector ($\boldsymbol{\lambda_t^L}$, $\boldsymbol{\lambda_t^R}$, $\bar{L}_t$, $\boldsymbol{Q_t}$, $\boldsymbol{w_t}$). ■

E.3 Closed-City Counterfactuals

In the paper, we consider an open-city specification, in which population mobility between Greater London and the wider economy ensures that expected utility in Greater London is constant and equal to the reservation level of utility in the wider economy ($\hat{\bar{U}}_t = 1$). In this subsection of the web appendix, we report counterfactuals for the removal of the railway network in a "closed-city" specification in which the total population of Greater London is held constant ($\hat{\bar{L}}_t = 1$), and the change in commuting costs leads to a change in expected utility in Greater London ($\hat{\bar{U}}_t \neq 1$). Again we rewrite the conditions for general equilibrium in the counterfactual equilibrium in terms of relative changes, such that $\hat{\lambda}_{nt}^L = \lambda_{nt}^{L\prime}/\lambda_{nt}^L$. The general equilibrium vector of counterfactual changes ($\hat{\lambda}_{nt}^L$, $\hat{\lambda}_{nt}^R$, $\hat{Q}_{nt}$, $\hat{w}_{nt}$, $\hat{\bar{U}}_t$) in response to the change in commuting costs ($\hat{\kappa}_{nit}$) solves the system of five equations for land market clearing (equation (30) in the paper), zero-profits in production (equation (31) in the paper), workplace choices (equation (32) in the paper), residential choices (equation (33) in the paper) and population mobility (equation (34) in the paper):

$$\hat{Q}_{nt}^{1+\mu}\mathbb{Q}_{nt} = \left\{(1-\alpha)\left[\sum_{i\in\mathbb{N}}\frac{\lambda_{nit|n}^{C}\hat{w}_{it}^{\epsilon}\hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell\in\mathbb{N}}\lambda_{n\ell t|n}^{C}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{n\ell t}^{-\epsilon}}\hat{w}_{it}w_{it}\right]\hat{\lambda}_{nt}^{R}\lambda_{nt}^{R} + \left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)\hat{w}_{nt}w_{nt}\hat{\lambda}_{nt}^{L}\lambda_{nt}^{L}\right\}\bar{L}_t, \tag{E.21}$$

$$\hat{Q}_{nt} = \hat{A}_{nt}^{1/(1-\beta)}\hat{w}_{nt}^{-\beta/(1-\beta)}, \tag{E.22}$$

$$\hat{\lambda}_{nt}^{L}\lambda_{nt}^{L} = \frac{\sum_{n\in\mathbb{N}}\lambda_{n\ell t}\hat{B}_{nt}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{n\ell t}^{-\epsilon}\hat{Q}_{nt}^{-\epsilon(1-\alpha)}}{\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\lambda_{r\ell t}\hat{B}_{rt}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{r\ell t}^{-\epsilon}\hat{Q}_{rt}^{-\epsilon(1-\alpha)}}, \tag{E.23}$$

$$\hat{\lambda}_{nt}^{R}\lambda_{nt}^{R} = \frac{\sum_{\ell\in\mathbb{N}}\lambda_{n\ell t}\hat{B}_{nt}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{n\ell t}^{-\epsilon}\hat{Q}_{nt}^{-\epsilon(1-\alpha)}}{\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\lambda_{r\ell t}\hat{B}_{rt}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{r\ell t}^{-\epsilon}\hat{Q}_{rt}^{-\epsilon(1-\alpha)}}, \tag{E.24}$$

$$\hat{\bar{U}}_t = \left[\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\lambda_{r\ell t}\hat{B}_{nt}\hat{w}_{\ell t}^{\epsilon}\hat{\kappa}_{r\ell t}^{-\epsilon}\hat{Q}_{rt}^{-\epsilon(1-\alpha)}\right]^{\frac{1}{\epsilon}}. \tag{E.25}$$

where our specifications of productivity (equation (28) in the paper) and amenities (equation (26) in the paper) imply:

$$\hat{A}_{nt} = \hat{L}_t^{\psi}, \tag{E.26}$$

$$\hat{B}_{nt} = \hat{R}_t^{\eta}; \tag{E.27}$$

where recall that $\hat{\lambda}_{nt}^L = \lambda_{nt}^{L\prime}/\lambda_{nt}^L$; we observe or have solved for the initial equilibrium values of the endogenous variables ($\mathbb{Q}_{nt}$, λ_{nt}^L, λ_{nt}^R, $\bar{L}_t$, λ_{nit}, $\lambda_{nit|n}^C$, w_{it}); we assume that production and residential fundamentals, the price of the final good, and total population in Greater London remain constant ($\hat{a}_{nt} = 1$, $\hat{b}_{nt} = 1$, $\hat{P}_t$ and $\hat{\bar{L}}_t = 1$); we have used $\hat{\mathbb{Q}}_n = \hat{Q}_n^{1+\mu}$; and we have chosen units in which to measure the initial level of expected utility such that $\bar{U}_t = 1$.

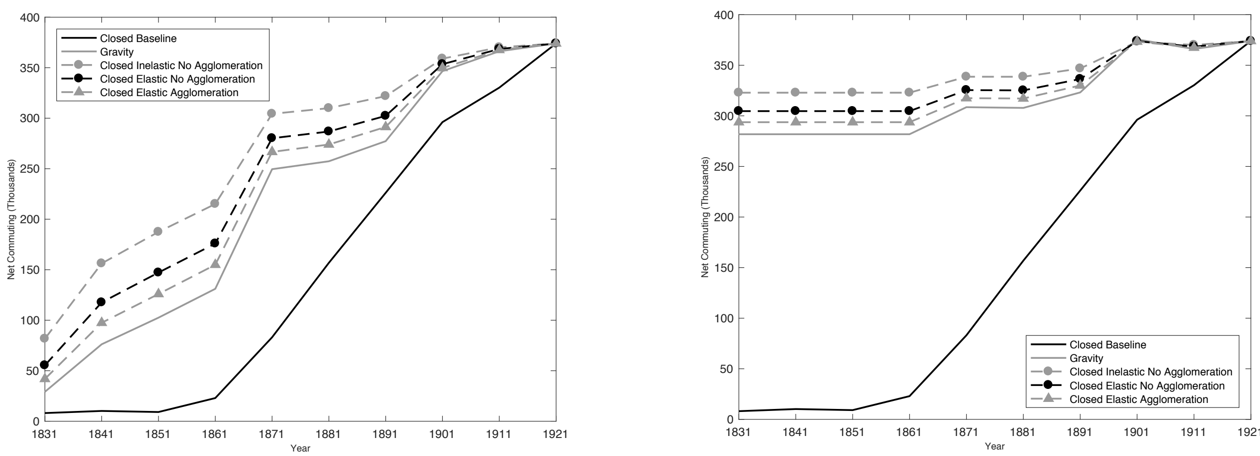
In Figure E.1, we display net commuting into the City of London in our closed-city counterfactuals for removing the entire railway network (left panel) and eliminating only the underground network (right panel). To provide a point of comparison, the solid black line (no markers and labelled "baseline") shows net commuting into the City of London in our baseline quantitative analysis of the model in Section 6 of the paper. As another point of comparison, the solid gray line (no markers and labelled "Constant Workplace/Residence") displays the results of our mechanical predictions based on the gravity equation alone from Section 7.1 of the paper.

The gray dashed line with circle markers (labelled "Inelastic No Agglomeration") shows our counterfactual equilibrium predictions for the special case of the model with no agglomeration forces ($\psi = \eta = 0$) and a perfectly inelastic supply of land ($\mu = 0$). We find that removing the entire railway network back to 1831 reduces net commuting into the City of London to around 82,000, while removing the underground network back to 1861 reduces these net commuting flows to just under 323,000. Unsurprisingly, these effects with a constant total population of Greater London are somewhat smaller than in the open-city specification in the paper, because the decline in the total population of Greater London in the open-city specification reduces commuting flows between all locations. Nevertheless, comparing Figure E.1 for the closed-city specification with the corresponding Figure 13 for the open-city specification in the paper, it is striking just how similar are the model's counterfactual predictions for net commuting into the City of London across all these quite different specifications. This robustness is consistent with our findings much of the increased separation of workplace and residence in the City of London in the late 19th-century is driven by the direct effect of the change in commuting costs, which is the same across all these quite different specifications.

The black dashed line with circle markers (labelled "Elastic No Agglomeration") shows our closed-city counterfactuals with our calibrated floor space supply elasticity ($\mu = 2.86$) and no agglomeration forces ($\psi = \eta = 0$). The gray dashed line with triangle markers (labelled "Elastic Agglomeration") displays our closed-city counterfactuals with our calibrated floor space supply elasticity ($\mu = 2.86$) and values for production and residential externalities in line with the range of estimates reviewed in Rosenthal and Strange (2004) ($\psi = \eta = 0.05$). Again we find a similar pattern of results for net commuting into the City of London as in the open-city specification in the paper. Both an elastic supply of floor space and agglomeration forces magnify the reduction in these net commuting flows into the City of London. In our counterfactual for removing the entire railway network back to 1831, the net flow of workers commuting into the City of London falls to 55,000 with an elastic supply of floor space and exogenous productivity and amenities, and falls to 42,000 with an elastic supply of floor space and agglomeration forces. Again the volume of net commuting into the City of London in these closed-city specifications is somewhat higher than in the corresponding open-city specifications in the paper, because the fall in the total population of Greater London in the open-city specifications reduces commuting flows between all locations.

Therefore, across the wide range of specifications reported in the paper and this section of the web appendix, we continue to find that much of the increased separation of workplace and residence in the City of London in the late-19th century is driven by the new transport technology of the steam railway. In these closed-city specifications, we also find substantial effects of the change in commuting costs on the expected utility of workers in Greater London. Across all of the closed-city specifications reported in this section of the web appendix, we find that the removal of the entire railway network reduces expected utility by around 18 percent, while the removal of the underground network alone reduces expected utility by about 5 percent.

Figure E.1: Counterfactual Net Commuting into the City of London 1831-1921 (Closed-City Specification)



(a) All Rail Network

(b) Underground Rail Network

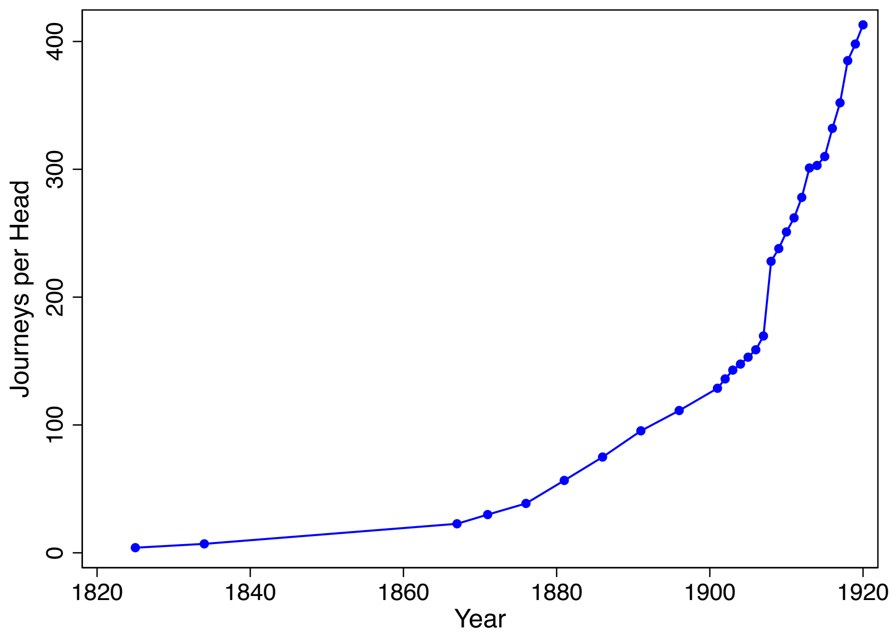
Note: "Baseline model prediction" shows net commuting from our baseline quantitative analysis from Section 6 of the paper; "Gravity" shows net commuting in our gravity-based counterfactuals from Section 7.1 of the paper; "Closed Inelastic No Agglomeration" shows net commuting in our model-based counterfactuals with a perfectly inelastic supply of floor space ($\mu = 0$) and exogenous productivity and amenities ($\psi = \eta = 0$); "Closed Inelastic No Agglomeration" shows net commuting in our model-based counterfactuals with a positive floor space supply elasticity (2.86) and exogenous productivity and amenities ($\psi = \eta = 0$); "Closed Elastic Agglomeration" shows net commuting in our model-based counterfactuals with a positive floor space supply elasticity (2.86) and positive production and residential externalities ($\psi = \eta = 0.05$). All these counterfactuals assume a closed-city specification with a fixed total population in Greater London.

F Additional Empirical Results

In this section of the web appendix, we report additional empirical results that are discussed in the paper. In Figure F.1, we provide further evidence on a change in transport use by graphing passenger journeys using public transport per head of population in the County of London in each year (see also Barker 1980). Public transport includes underground and overground railways, horse and electric trams, short-stage coaches, and horse and motor omnibuses. As discussed in Section 4.1 of the paper, the increasing specialization of locations as workplace or residence from the mid-19th century onwards is reflected in a sharp increase in the intensity of public transport use, with journeys per head of population increasing from around 7 in 1834 to just under 400 in 1921.

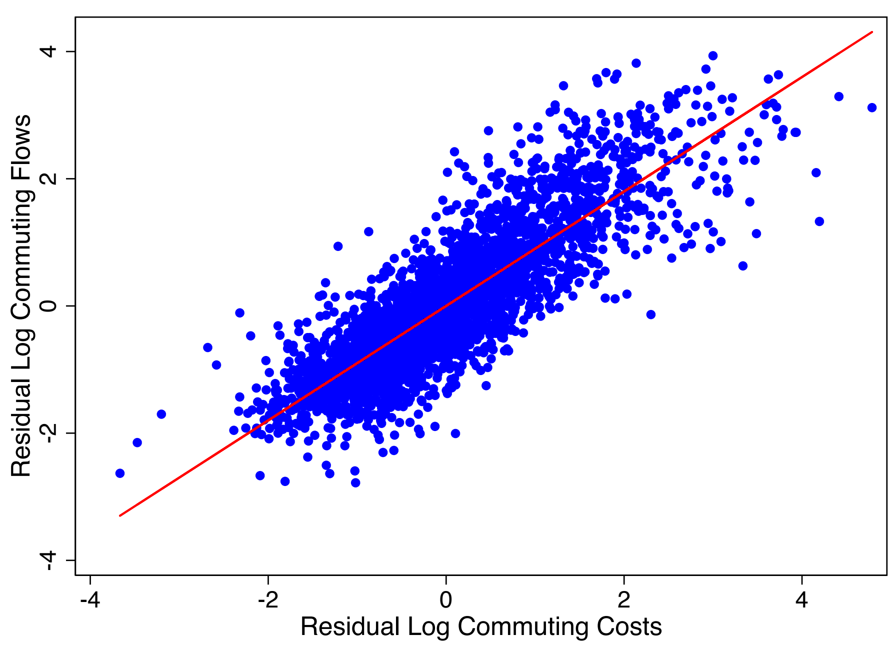
In Figure F.2, we examine the within-sample fit of our estimated cost commuting parameters using the 1921 data on bilateral commuting flows. As discussed in Section 6.4 of the paper, we show the conditional correlation between the log unconditional commuting probabilities to other boroughs $n \neq i$ and our estimates of bilateral commuting costs, after removing workplace and residence fixed effects. Consistent with our model's predictions, we find a tight and approximately log linear relationship between bilateral commuting probabilities and our estimates of bilateral commuting costs, with a conditional correlation of over 0.7. While our parametrization of commuting costs necessarily abstracts from many idiosyncratic factors that could affect commuting costs for individual workplace-residence pairs, these results suggest that it provides a good approximation to observed cross-section bilateral commuting flows.

Figure F.1: Public Transport Passenger Journeys per Head in the County of London



Note: Journeys per head measured as millions of passengers carried per year on public transport divided by the population of the County of London. Public transport includes underground and overground railways, horse and electric trams, short-stage coaches, and horse and motor omnibuses. Sources: Barker (1980) and London Statistics, 1921.

Figure F.2: Conditional Correlation Between Bilateral Commuting Flows and Estimated Commuting Costs in 1921



Note: Conditional correlation after removing workplace and residence fixed effects between log bilateral unconditional commuting probabilities (equation (5) in the paper) and log estimated bilateral commuting costs to other boroughs in 1921.

G Data Appendix

This section of the web appendix reports additional information about the data sources and definitions, supplementing the discussion in Section 3 of the paper. Section G.1 discusses the population data from the population censuses of England and Wales from 1801-1921. Section G.2 summarizes the rateable value data from 1815-1921. Section G.3 outlines the historical data on day population from the City of London Day Censuses. Section G.4 explains the construction of our geographical information systems (GIS) shapefiles of the overground and underground railway networks over time. Section G.5 reports analogous information for our GIS shapefiles of the omnibus and tram network over time. Section G.7 discusses our bilateral commuting data for 1921 from the population census of England and Wales. Section G.8 contains further details on our historical data on commuting patterns from the personnel ledgers of Henry Poole Bespoke Tailors.

G.1 Population Data

Population data from the population censuses of England and Wales from 1801-1891 was provided by the *Cambridge Group for the History of Population and Social Structure* (Cambridge Group), as documented in Wrigley (2011). The original sources for the population data are as follows:

- 1801 Census Report, Abstract of answers and returns, PP 1801, VI
- 1811 Census Report, Abstract of answers and returns, PP 1812, XI
- 1821 Census Report, Abstract of answers and returns, PP 1822, XV
- 1831 Census Report, Abstract of the Population Returns of Great Britain, PP 1833, XXXVI to XXXVII
- 1841 Census Report, Enumeration Abstract, PP 1843, XXII
- 1851 Census Report: Population Tables, part II, vols. I to II, PP 1852-3, LXXXVIII, parts I to II
- 1861 Census Report: Population tables, vol. II, PP 1863, LIII, parts I to II
- 1871 Census Report: vol. III, Population abstracts: ages, civil condition, occupations and birthplaces of people, PP 1873, LXXI, part I
- 1891 Census Report: vol. II, Area, Houses and Population: registration areas and sanitary districts, PP 1893-4, CV [which also includes the 1881 data, as used in our analysis]

The smallest unit of observation and the lowest tier of local government are civil parishes, which we refer to simply as *parishes*. The boundaries of these parishes can change across the population censuses. To create a consistent panel of mappable spatial units over time, the Cambridge Group has developed a two-stage procedure. First, they spatially match parish level polygons and geographical units from each census to derive all spatial units that existed in any period between 1801–1891. They refer to this as *CGKO* (Cambridge Group Kain Oliver) map. This dataset includes 456 polygons for the Greater London Authority (GLA). Next, they employ a *Transitive Closure Algorithm* from graph theory (see for example Cormen, Leiserson, Rivest and Stein 2009) to determine the lowest common unit between parish polygons in different years, which defines the *mappable units*. This procedure implies that these mappable units do not necessarily represent real parishes, but for simplicity we continue to refer to them as *parishes*. After applying the transitive closure algorithm, we obtain 283 mappable units for the GLA. The average mappable unit has a size of 5.60 square kilometers, with 4,042 inhabitants in 1801 and 19,686 inhabitants in 1891. Within the GLA, we further distinguish London City Council (LCC) which encloses 183 mappable units with an average size of 1.64 square kilometers and an average number of inhabitants of 5,432 (22,890) in 1801 (1891) respectively; and the City of London (COL) with 111 mappable units of an average size of 0.02 square kilometers and an average number of inhabitants of 1,219 (1801) and 348 (1891) respectively.

Population data from 1901-1921 stem from the Integrated Census Microdata Project (I-CeM). The majority of parishes did not experience any change in boundaries from 1891-1901. Therefore, we can simply extend the parish panel for 1801-1891 discussed above to 1901. However, from 1911 onwards, there are a number of major changes in parish boundaries. Most notably, the COL consisted of more than 100 parishes in the censuses for 1801-1901, which were amalgamated into a single parish in 1907. To avoid having to make assumptions in order to disaggregate the 1911 and 1921 population data for the COL, we end our parish-level panel dataset in 1901.

The next smallest unit of observation is referred to as either metropolitan borough, urban district or rural district in the population census, depending on the level of urbanization of that location. For simplicity, we refer to these units as *boroughs*. We use the boundaries of these boroughs from the 1921 population census to construct consistent panel data on the population of boroughs from 1801-1921. There are 99 of these boroughs in the GLA, 29 in the LCC, and the COL is its own borough. The average borough has an area of 16 square kilometers in the GLA (10.81 in the LCC and 2.98 in the COL). For 1921, we obtain borough population data directly from the population census for that year. For the years before 1921, we overlay the 1921 boroughs and the mappable units discussed above, and allocate the population of the mappable units to the 1921 boroughs, by weighting the values for each mappable unit by its share of the geographical area of the 1921 boroughs. Given that mappable units have a much smaller geographical area than boroughs, most of them lie within a single borough.

G.2 Rateable Value Data

We measure the value of floor space using rateable values, which correspond to the annual flow of rent for the use of land and buildings, and equal the price times the quantity of floor space in the model. In particular, these rateable values correspond to "The annual rent which a tenant might reasonably be expected, taking one year with one another, to pay for a hereditament, if the tenant undertook to pay all usual tenant's rates and taxes ... after deducting the probable annual average cost of the repairs, insurance and other expenses" (see London County Council 1907).

These rateable values cover all categories of property, including public services (such as tramways, electricity works etc), government property (such as courts, parliaments etc), private property (including factories, warehouses, wharves, offices, shops, theaters, music halls, clubs, and all residential dwellings), and other property (including colleges and halls in universities, hospitals and other charity properties, public schools, and almshouses). As discussed in Stamp (1922), there are three categories of exemptions: (1) Crown property occupied by the Crown (Crown properties leased to other tenants are included); (2) Places for divine worship (church properties leased to other tenants are included); (3) Concerns listed under No. III Schedule A, namely: (i) Mines of coal, tin, lead, copper, mundic, iron, and other mines; (ii) Quarries of stone, slate, limestone, or chalk; ironworks, gasworks, salt springs or works, alum mines or works, waterworks, streams of water, canals, inland navigations, docks, drains and levels, fishings, rights of markets and fairs, tolls, railways and other ways, bridges, ferries, and cemeteries. Rateable values were assessed at the parish level approximately every five years during our sample period. All of the above categories of properties are included, regardless of whether or not their owners are liable for income tax.

These rateable values have a long history in England and Wales, dating back to the 1601 Poor Relief Act, and were originally used to raise revenue for local public goods. Different types of rateable values can be distinguished, depending on the use of the revenue raised: Schedule A Income Taxation, Local Authority Rates, and Poor Law Rates. Where available, we use the Schedule A rateable values, since Schedule A is the section of the national income tax concerned with income from property and land, and these rateable values are widely regarded as corresponding most closely to market valuations. For example, Stamp (1922) argues that "It is generally acknowledged that the income tax, Schedule A, assessments are the best approach to the true values." (page 25) After the Metropolis Act of 1869, all rateable values for the County of London are computed on the basis of Schedule A Income Taxation. Where these Schedule A rateable values are not available, we use the Local Authority rateable values, Poor Law rateable values, or property valuations for income tax. For years for which more than one of these measures is available, we find that they are highly correlated with one another across parishes.

The original sources for the rateable values data used for each year are as follows:

- **1815**: Property valuations for income tax. Return to an address of the Honourable the House of Commons, dated 21 February 1854; House of Commons Papers, vol. LVI.1, paper no: 509.
- **1843**: Property valuations for income tax. Return to an address of the Honourable the House of Commons, dated 21 February 1854; House of Commons Papers, vol. LVI.1, paper no: 509.
- **1847**: Poor Law Rateable Values. Return to an order of the Honourable the House of Commons, dated 31 August 1848; House of Commons Papers, vol. LIII.11, paper no: 735.
- **1852**: Property valuations for income tax. Return to an address of the Honourable the House of Commons, dated 21 February 1854; House of Commons Papers, vol. LVI.1, paper no: 509.
- **1860**: Property valuations for income tax. Return to an order of the Honourable the House of Commons, dated 13 August 1860; House of Commons Papers, vol. XXXIX, paper no: 546.
- **1881**: Poor Law Rateable Values. A Statement of the Names of the Several Unions And Poor Law Parishes In England And Wales; And of the Population, Area, And Rateable Value Thereof in 1881. London: Her Majesty's Stationery Office, 1887.
- **1896**: Schedule A Rateable Values. Agricultural Rates Act, 1896. Reports separate data on the rateable value of agricultural land and the rateable value of other land and buildings. Return to an order of the Honourable the House of Commons, dated 27 July 1897; House of Commons Papers, paper no: 368; 1897.
- **1905**: Schedule A Rateable Values. Local taxation returns (England and Wales). The annual local taxation returns. Year 1904-05. Part I. House of Commons Papers, vol. CI.1, paper no: 311, 387; 1906.
- **1911**: Schedule A Rateable Values. Local taxation returns (England and Wales). The annual local taxation returns. Year 1910-11. Part I. House of Commons Papers, vol. LXXII.1, Paper no: 264, 268, 364, 282; 1912.

- **1921**: Local Authority Rateable Values, Ministry of Health. Statement showing, for each borough and other urban district in England and Wales, and for 100 typical rural parishes, the amount of the local rates per pound of assessable value, for the financial years 1920-21 and 1921-22, and the assessable values in force at the commencement of the year 1921-22. House of Commons Papers, vol. XVII. 625, Paper no: 1633; 1922.

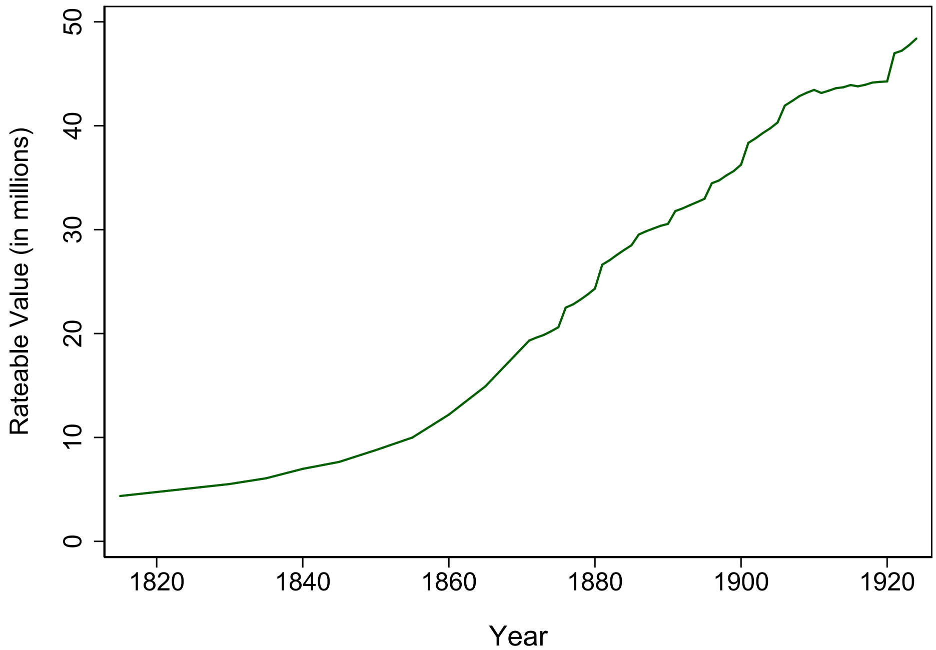
To create consistent spatial units over time, we manually match parishes with the spatial units provided by the *CGKO* (Cambridge Group Kain Oliver) map, as discussed for the population data in Section G.1 of this web appendix. We then use area weights to create the same *mappable units* as for the population data. This procedure gives us a parish-level panel for the years 1815, 1843, 1848, 1852, 1860, 1881 and 1896. Finally, we aggregate these parish data to the 1921 boroughs using area weights, as discussed for the population data in Section G.1 above.

For the years 1905, 1911 and 1921, we use rateable values at the borough level. For the year 1921, we observe rateable values for all 1921 boroughs, including metropolitan boroughs, urban districts and rural districts. However, some boroughs were created for the first time in 1921, when a previously-existing borough was sub-divided into separate urban and rural districts. Therefore, for the years 1905 and 1911, we are missing data for these newly-created sub-divisions: Croydon (created in 1912), Orset (created in 1912), Watford (created in 1906), Hitchin (created in 1919), Dartford (created in 1920), Bexley (created in 1920), Uxbridge (created in 1920), Chertsey (created in 1909) and Hambeldon (not separately reported in 1905 and 1911). To deal with these sub-divisions, we allocate the data for the larger 1905 and 1911 spatial units across their 1921 sub-divisions using area weights.

As discussed above, Schedule A rateable values, local authority rateable values, poor law rateable values and property valuations for income tax are highly correlated across parishes. Nonetheless, the level of the property valuations for income tax is somewhat lower than the rateable values, which is consistent with the fact that rateable values include all properties, regardless of whether their owners are liable for income tax, whereas the property valuations for income tax are based on income tax liability. To address this difference, we use a consistent time-series on Schedule A rateable values for the County of London. This consistent time-series was constructed for an aggregate of 28 boroughs in the County of London by London County Council for the years 1830, 1835, 1840, 1845, 1850, 1855, 1860, 1865, 1871, 1876, 1881 and 1891. We extend this time-series forward to 1921, using the reported Schedule A rateable values for these 28 boroughs reported in *London Statistics*. We also extend this time-series back to 1815, using the annual population growth rate for these 28 boroughs from 1815-1830. In Figure G.1, we display the resulting time-series for this aggregate of 28 boroughs in the County of London.

For each year, we first construct an adjustment factor, which is equal to the ratio of our property valuation to the consistent Schedule A rateable value for the aggregate of 28 boroughs in the County of London. We next adjust our property valuation upwards or downwards for all boroughs using this adjustment factor. The adjustment factors for each year are 0.72 (1815), 0.58 (1843), 0.97 (1847), 0.96 (1852), 0.57 (1860), 1.01 (1881), 1.03 (1896), 1.00 (1905), 0.99 (1911) and 1 (1921). These adjustment factors are all close to one in years for which we use rateable values, which is consistent with the idea that Schedule A rateable values, Local Authority rateable values and Poor Law rateable values are all highly correlated with one another. These adjustment factors are less than one for years in which we use property valuations for income tax, which is consistent with the fact that the Schedule A rateable values include all properties, regardless of whether the owners of those properties are liable for income tax, whereas the property valuations for income tax are based on income tax liability.

Figure G.1: Schedule A Rateable Value for an Aggregate of 28 Boroughs in the County of London from 1815-1921



Notes: Current price millions of pounds, London County Council (LCC) and authors' calculations.

Finally, we use linear interpolation in between the above years to construct a time-series on rateable values for each borough and for each census decade, from immediately before the arrival of the railway in 1831, to the end of our sample period in 1921.

G.3 Employment and Day Population Data

For 1921, we observe the bilateral matrix of flows of commuters from each residence borough (rows) to each workplace borough (columns) from the population census for England and Wales. Summing across columns in this matrix, we obtain employment by residence for 1921 (which we refer to as "residence employment") for each borough. Summing across rows in this matrix, we obtain employment by workplace for 1921 (which we refer to as "workplace employment"). We also construct an employment participation rate for each borough in 1921 by dividing residence employment by population.

For years prior to 1921, we construct residence employment using our population data from the population censuses for England and Wales. Assuming that the ratio of residence employment to population is stable for a given borough over time, we use the 1921 value of this ratio and the historical population data to construct residence employment for each borough for each decade from 1801-1911. Consistent with a stable employment participation rate, we find relatively little variation in the ratio of residence employment to population across boroughs in 1921.

Data on workplace employment are not available prior to 1921. Therefore, in our structural estimation of the model, we use our bilateral commuting data for 1921, together with our data on residence employment and rateable values for earlier years, to generate model predictions for workplace employment for earlier years. In overidentification checks, we compare these model predictions to the data on the day population from the City of London Day Censuses. In the face of the increased commuting from the mid-19th century onwards, the City of London Corporation recognized that the measure of population from the population census of England and Wales, which is based on where one slept on census night ("night population"), could be a misleading indicator of the population present during the daytime ("day population"). Therefore, the City of London Corporation undertook Day Censuses in 1866, 1881, 1891 and 1911 to record "... every person, male or female, of all ages, residing, engaged, occupied, or employed in each and every house, warehouse, shop, manufactory, workshop, counting house, office, chambers, stable, wharf, etc, and to include all persons, of both sexes and all ages, on the premises during the working hours of the day, whether they sleep or do not sleep there ..." (Salmon 1891, page 97). Therefore, the "day population" includes both those employed in the City of London and those resident in the City of London and present during the data (e.g. because they are economically inactive). We generate an analogous measure of day population in the model, which equals our model's prediction for workplace employment plus observed economically-inactive residents (observed population minus observed residence employment).

The original sources for the City of London day census data are as follows:

- Day Census, City of London, Report, Local Government and Taxation Committee, 13th December, 1866.
- Report on the City Day Census, 1881, By the Local Government and Taxation Committee of the Corporation of London, Second Edition, London: Longmans, Green and Company.
- Ten Years' Growth of the City of London, Report, Local Government and Taxation Committee of the Corporation, by James Salmon, London: Simpkin, Marshall, Hamilton, Kent and Company, 1891.
- City of London Day Census, 1911, Report, County Purposes Committee of the Corporation, by Henry Percival Monckton, London: Simpkin, Marshall, Hamilton, Kent and Company.

G.4 Overground and Underground Railway Network

We have geographical information systems (GIS) information on the location of all railway lines and stations opened for the public carriage of passengers and/or goods and the year in which they opened. This GIS dataset was provided by the *Cambridge Group for the History of Population and Social Structure*, which based its digitization on Cobb (2003). Using these data, we construct separate networks for overground and underground railways in each census year. In Table G.1 below, we summarize the opening years and respective lengths of the London Underground lines. In Figures G.2 to G.10 below, we show the decennial evolution of the underground and overground railway network across the Greater London Authority (GLA), where 1841 is the first census year in which an overground railway exists, and 1871 in the first census year in which an underground railway exists.

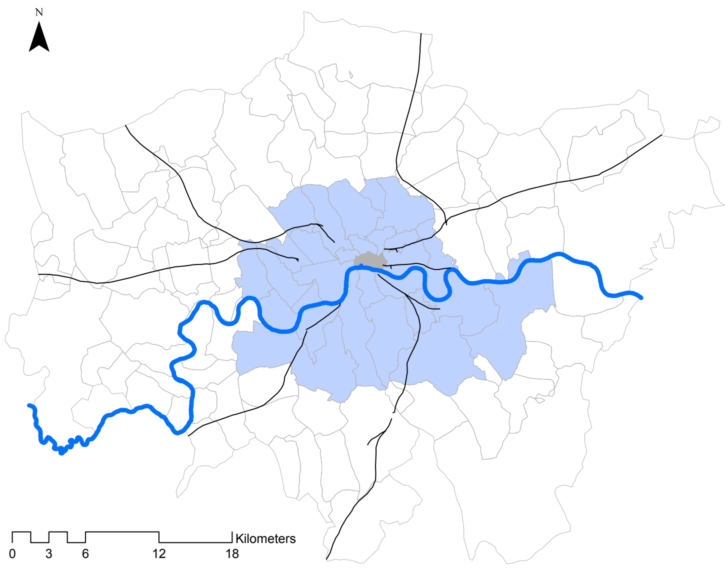
In Table G.1, we use the modern names of underground lines, which do not always correspond to their names in 1921. We further exclude parts of the London Underground that did not exist in 1921. The *Victoria Line* and the *Jubilee Line* did not open until 1968 and 1979 respectively. We do not list the *Circle Line* and the *Hammersmith & City Line*, because they were both part of the network of the District and Metropolitan line in 1921. We also exclude the *Waterloo and City Line*, because it was not classified as part of the London Underground when it opened in 1898, even though its tracks run underground from Waterloo station underneath the River Thames to Bank station in the City of London with no intermediate stops. We follow this convention and classify this line, which is only 1.59 miles long, as an overground railway. Although this line was formally owned by the Waterloo and City Railway Company, it was operated from the start by an overground railway company: the London and South Western Railway (LSWR). In 1907, the LSWR formally absorbed the Waterloo and City Railway Company. From that year onwards, the Waterloo and City Line continued to be operated by overground railway companies, and was only officially taken over by the London Underground system in 1994. Finally, we classify the *East London Line* as part of the London Underground, because it was initially operated by a consortium that included the District and Metropolitan lines. After 1933, this line became known as the East London Part of the Metropolitan Line. Today, it is part of the London overground railway network. This line is mostly above ground, but it uses the Thames tunnel built by Isambard Kingdom Brunel between 1825 and 1843 for horse-drawn carriages.

In measuring construction costs for underground railways, we distinguish shallow lines built using "cut-and-cover" techniques and deep lines built using "bored tubes," as discussed in Section G.10 of this web appendix. In central London, the *District Line* and the *Metropolitan Line* are cut-and-cover lines, while the remaining lines use bored tubes. Outside central London, parts of both types of line are above ground. We also measure the length of each type of line that is below and above ground, and take this into account in our measures of construction costs.

Table G.1: Opening Years and Lengths of London Underground Lines up to 1921

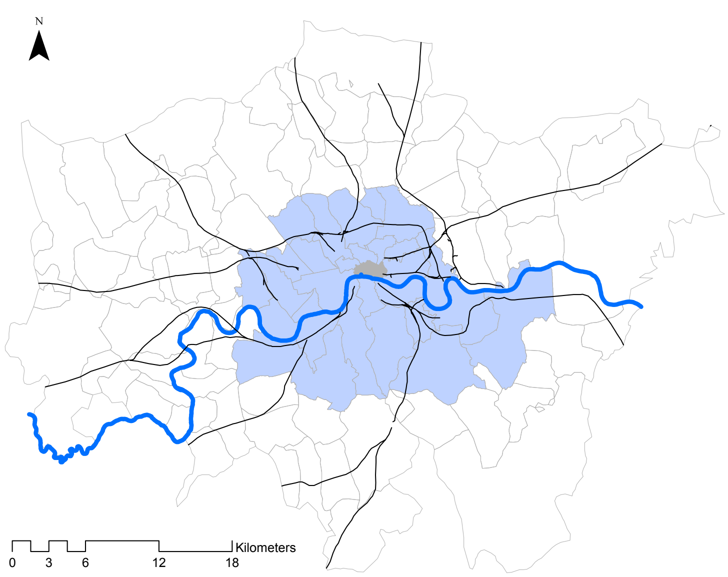
	All tracks		Tracks Underneath Ground	
Line	Opening Date	Length (in km)	Opening Date	Length (in km)
Bakerloo Line	1906	6.04	1906	6.04
Bakerloo Line	1907	1.04	1907	1.04
Bakerloo Line	1913	0.76	1913	0.76
Bakerloo Line	1915	3.32	1915	3.32
Bakerloo Line	1916	1.25		
Bakerloo Line	1917	11.29		
		23.70		11.15
Central Line	1900	9.30	1900	9.30
Central Line	1908	0.76	1908	0.76
Central Line	1912	0.63	1912	0.63
Central Line	1920	6.92		
		17.62		10.69
District Line	1868	4.49	1868	4.48
District Line	1869	2.71	1869	1.03
District Line	1871	2.84	1871	2.84
District Line	1874	2.34		
District Line	1877	7.65		
District Line	1879	3.79		
District Line	1880	2.17	1880	0.75
District Line	1883	8.85		
District Line	1884	3.66	1884	3.01
District Line	1902	2.59	1902	2.59
District Line	1905	0.69		
		41.79		14.70
East London Line	1869	3.45	1869	0.40
East London Line	1871	2.23		
East London Line	1876	2.67		
East London Line	1880	1.77		
		10.12		0.40
Metropolitan Line	1863	6.46	1863	6.46
Metropolitan Line	1865	0.71	1865	0.71
Metropolitan Line	1868	7.16	1868	7.16
Metropolitan Line	1875	0.52	1875	0.52
Metropolitan Line	1876	0.44	1876	0.44
Metropolitan Line	1879	3.48	1879	0.60
Metropolitan Line	1880	8.87		
Metropolitan Line	1884	0.35		
Metropolitan Line	1885	3.48		
Metropolitan Line	1887	5.93		
Metropolitan	1904	14.76	1904	4.70
		52.16		20.60
Northern Line	1890	3.87	1890	3.87
Northern Line	1900	3.77	1900	3.77
Northern Line	1901	2.30	1901	2.30
Northern Line	1907	14.70	1907	14.70
		24.63		24.63
Piccadilly Line	1903	8.12		
Piccadilly Line	1906	14.59	1906	14.59
Piccadilly Line	1907	0.71	1907	0.71
Piccadilly Line	1910	1.70		
		25.11		15.30

Figure G.2: Overground Railway Network in Greater London 1841



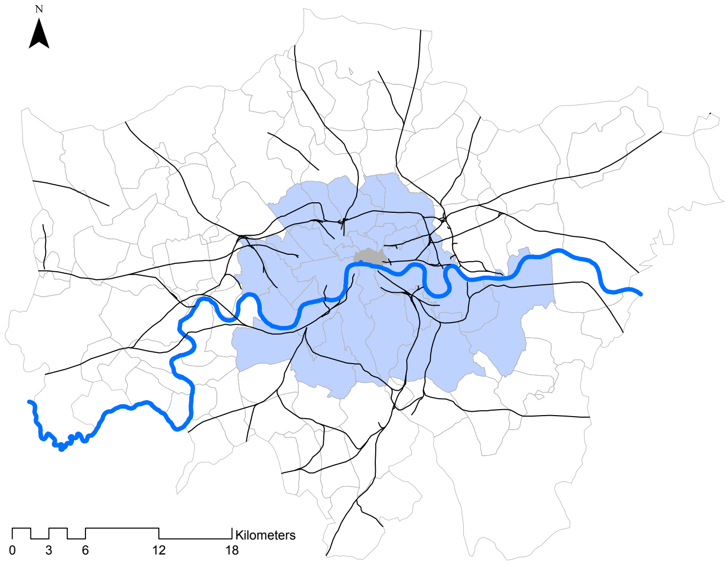
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black.

Figure G.3: Overground Railway Network in Greater London 1851



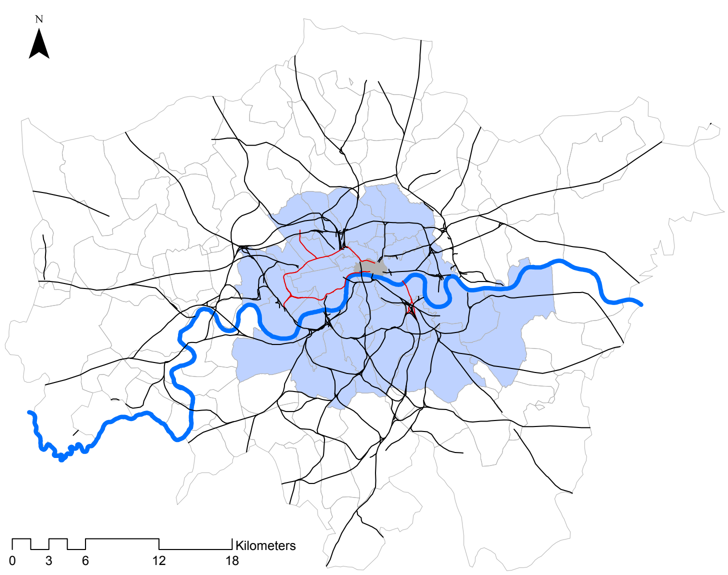
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black.

Figure G.4: Overground Railway Network in Greater London 1861



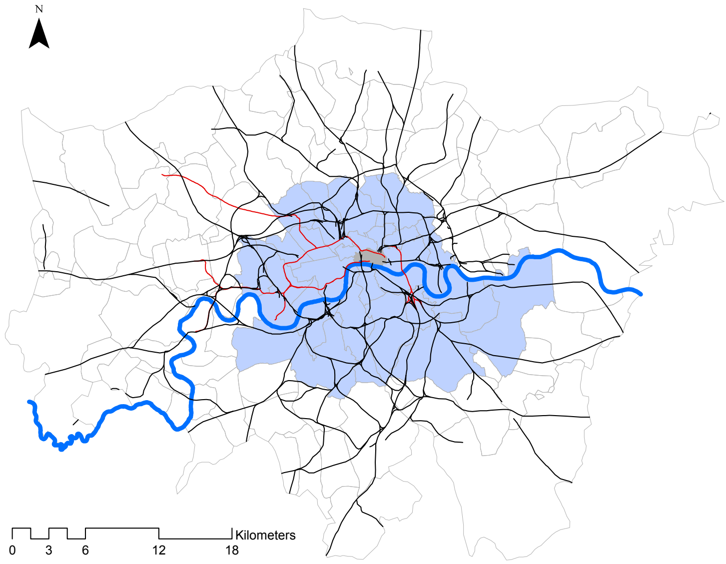
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black.

Figure G.5: Overground and Underground Railway Network in Greater London 1871



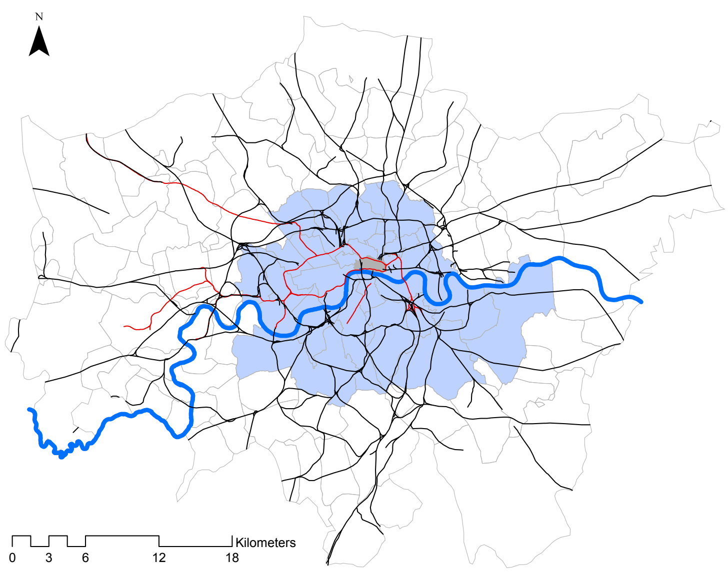
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

Figure G.6: Overground and Underground Railway Network in Greater London 1881



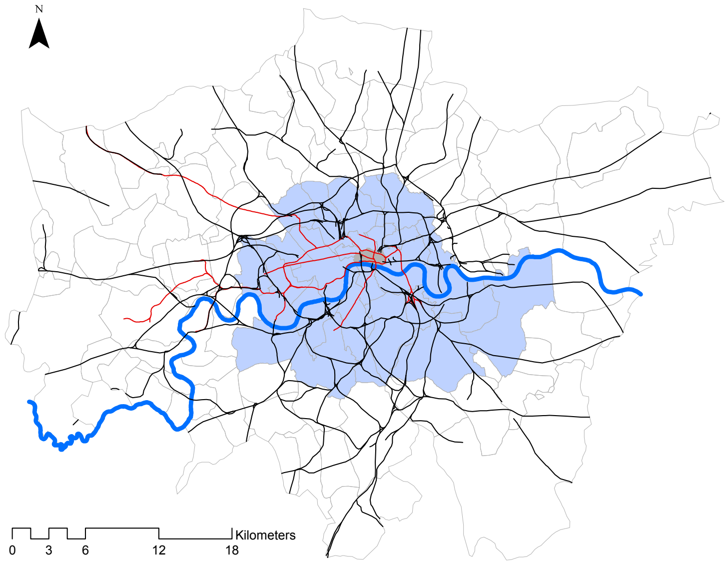
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

Figure G.7: Overground and Underground Railway Network in Greater London 1891



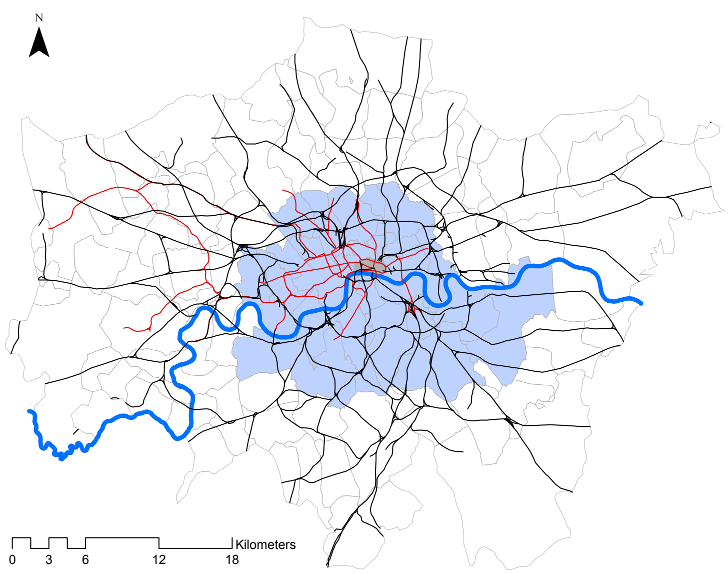
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

Figure G.8: Overground and Underground Railway Network in Greater London 1901



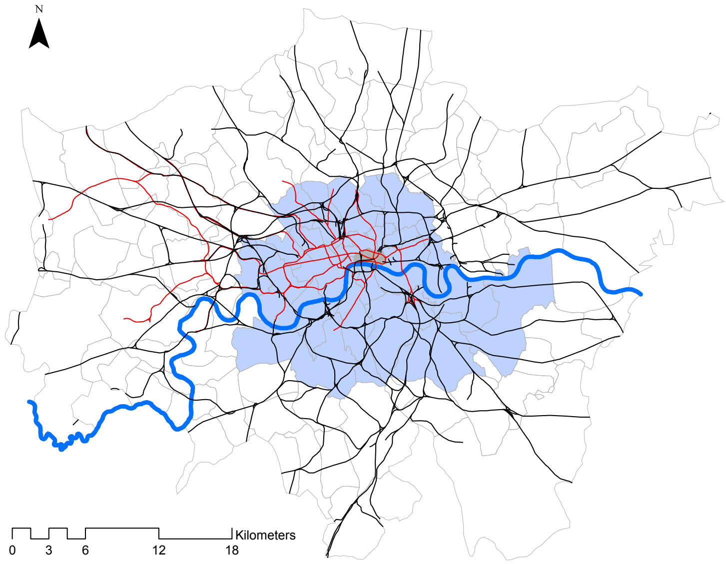
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

Figure G.9: Overground and Underground Railway Network in Greater London 1911



Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

Figure G.10: Overground and Underground Railway Network in Greater London 1921



Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

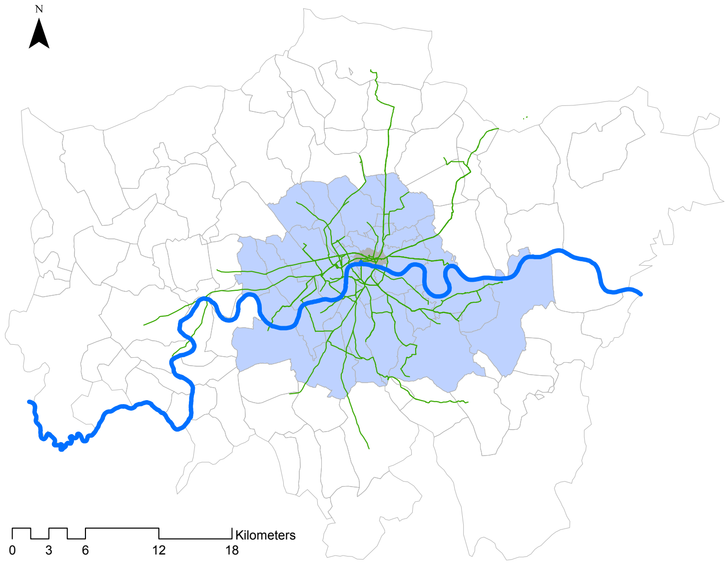
G.5 Omnibus and Tram Network

In addition to our geographic information systems (GIS) shapefiles of the railway network for each year, we also constructed an analogous shapefile for each year for the combined horse omnibus, horse tram, motor omnibus, and electric tram networks in each year. Horse omnibuses were first introduced from Paris to London in the 1820s; the first horse trams in London appeared in 1860; the first motor omnibus ran on the streets of London in 1898; and the first fully-operational electric tram service started in 1901. As discussed in Section G.6 of this web appendix, the average reported travel speeds for horse omnibuses and horse trams are around 6 miles per hour (mph), which are close to those reported for motor omnibuses and electric trams of around 7-8 mph. Therefore, we construct a single shapefile for each year that contains the combined networks of all forms of omnibus and tram in that year, assuming a single average travel speed of 6 mph.

To construct our GIS shapefiles for all years except 1831, 1841 and 1851, we start with georeferenced original maps of horse omnibus, horse tram, motor omnibus, and electric tram networks. Using these georeferenced maps, we construct line shapefiles for the combined omnibus and tram network in each year. For the years 1831, 1841 and 1851, we use reported route information (origin, intermediate stops and destination) to construct line shapefiles for these networks in each year. The original sources for the data used for each of our census decades are as follows:

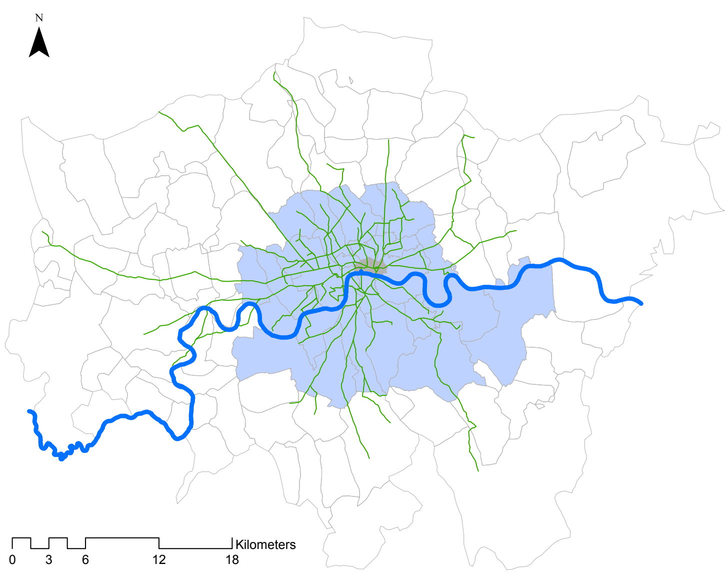
- **1831 and 1841**: Appendix 2 of Barker and Robbins (1963) reports bilateral routes for the Board of Stamps list of omnibuses and short-stage coaches licensed to operate in the London area in 1838-9.
- **1851 and 1861**: The Illustrated Omnibus Guide, with a Complete Guide to London, 1851, Simpkin and Company, Stationers' Court, W. H. Smith and Son, 136 Strand, Reprinted for Railwayana Ltd. by Oxford Publishing Company, 1971.
- **1871**: London Horse Omnibus Routes in 1871, map compiled by J. C. Gillham, chiefly from John Murray's Guide to London of 1871 and Adam and Charles Black's Guide.
- **1881**: London Horse Bus and Tram Routes in 1879, map compiled by J. C. Gillham, chiefly from John Murray's, Herbert Fry's and Lambert's Golden Guide Books to London 1879.
- **1891**: London Omnibus Routes early in 1895, map compiled by J. C. Gillham from a list published in 1895 by the London County Council statistical department.
- **1901**: London Omnibus Routes at the end of 1902, map compiled by J. C. Gillham from a list published in 1902 by the London County Council statistical department.
- **1911**: London Omnibus Routes in July 1911, map compiled by J. C. Gillham, chiefly from the London Traffic Report of the Board of Trade, 1911.
- **1921**: London County Council Tramway Map, 1916, London General Omnibus Company Motor Bus Map, 1921.

Figure G.11: Omnibus Network in Greater London in 1839



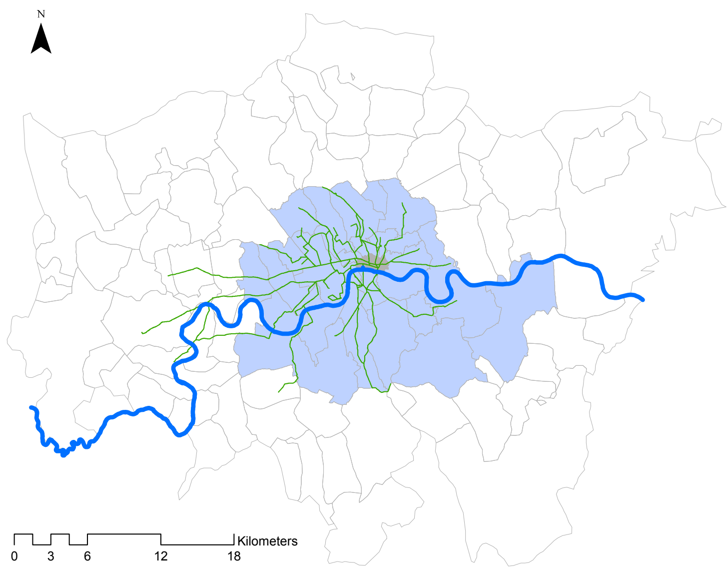
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus lines are shown in green; we use this omnibus network in 1839 for census years 1831 and 1841.

Figure G.12: Omnibus Network in Greater London in 1851



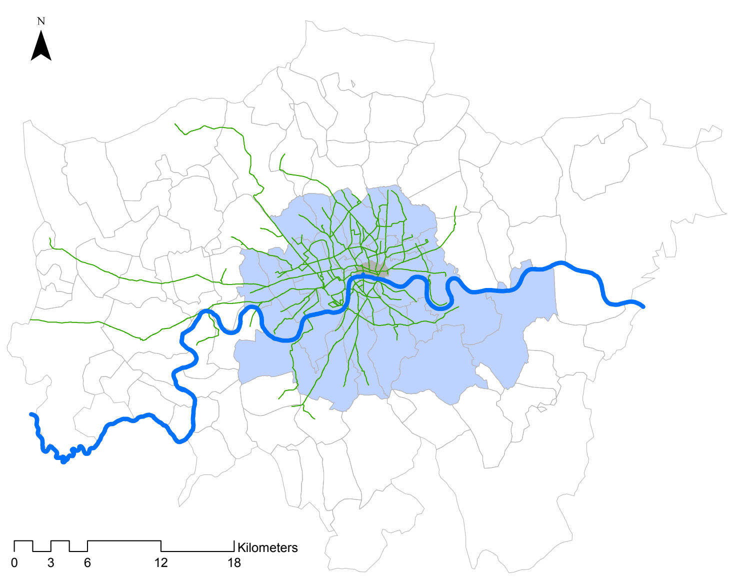
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus lines are shown in green; we use this omnibus network in 1851 for census years 1851 and 1861.

Figure G.13: Omnibus and Tram Network in Greater London in 1871



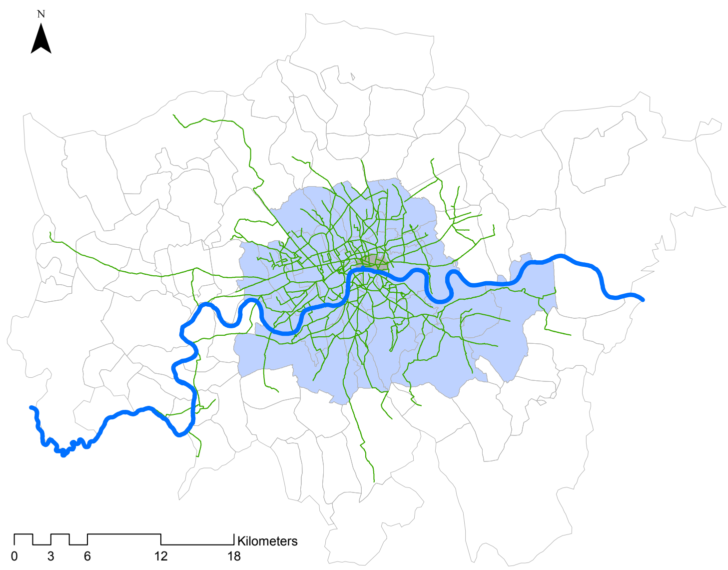
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus and tram lines are shown in green.

Figure G.14: Omnibus and Tram Network in Greater London in 1881



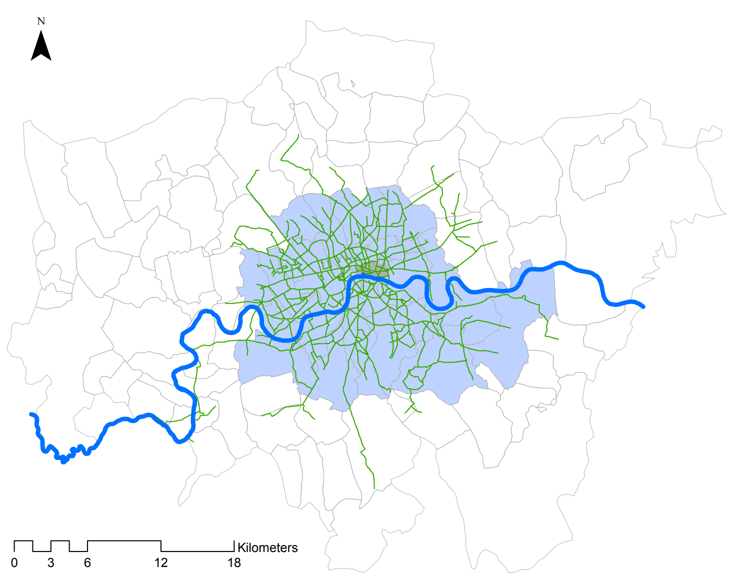
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus and tram lines are shown in green.

Figure G.15: Omnibus and Tram Network in Greater London in 1891



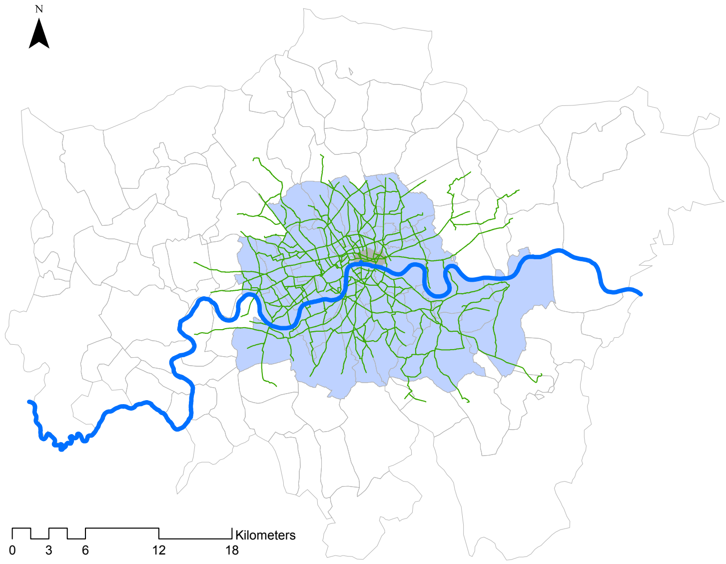
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus and tram lines are shown in green.

Figure G.16: Omnibus and Tram Network in Greater London in 1901



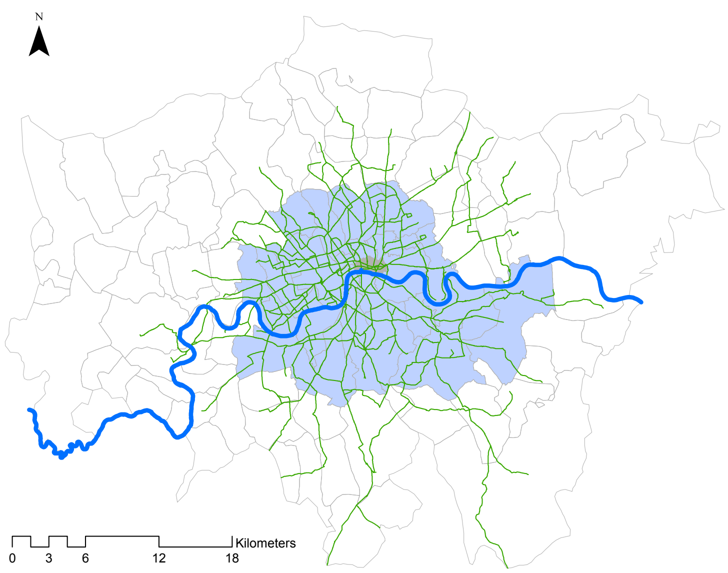
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus and tram lines are shown in green.

Figure G.17: Omnibus and Tram Network in Greater London in 1911



Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus and tram lines are shown in green.

Figure G.18: Omnibus and Tram Network in Greater London 1921



Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; omnibus and tram lines are shown in green.

G.6 Data on the Travel Speeds using Alternative Transport Modes

As discussed in the paper, at the beginning of the 19th-century, walking was the most common mode of transport, with average travel speeds in good road conditions of 3 miles per hour (mph). The state of the art technology for long distance travel was the stage coach, but it was expensive because of the multiple changes in teams of horses required over long distances, and hence was relatively infrequently used. Even with this elite mode of transport, poor road conditions limited average long distance travel speeds to around 5 mph (see for example Gerhold 2005).

With the growth of urban populations, attempts to improve existing modes of transport led to the introduction of the horse omnibus from Paris to London in the 1820s, as discussed in Barker and Robbins (1963). The main innovation of the horse omnibus relative to the stage coach was increased passenger capacity for short-distance travel. However, the limitations of horse power and road conditions ensured that average travel speeds remained low. As reported in Table G.2, average travel speeds using horse omnibuses remained around 6 mph for routes through both central and outlying areas. A further innovation along the same lines was the horse tram, which was introduced in London in 1860. However, average travel speeds again remained low, in part because of road congestion. As reported in Table G.3, average travel speeds during both rush and slack hours using horse trams were little different from those using horse omnibuses. A later innovation was the replacement of the horse tram with the electric tram, with the first fully-operational services starting in 1901. As reported in Table G.3, although electric trams brought some improvement in travel speeds relative to horse trams, during the rush hours when most commuting occurred, average travel speeds remained 5.5-7 mph.

In contrast, overground and underground railways brought a substantial improvement in average travel speeds, transforming the relationship between distance travelled and time taken. The world's first overground railway to be built specifically for passengers was the London and Greenwich railway, which connected what was then the village of Greenwich to Central London, and opened in 1836. As reported in Table G.4, average travel speeds using overground railways and steam locomotives were around 21 mph, with some variation depending on the track layout and number of intermediate stops. The world's first underground railway was the Metropolitan and District Railway, which connected the London termini of Paddington, Euston and Kings Cross with Farringdon Street in the City of London, and opened in 1863. When first opened, the Metropolitan and District Railway used steam locomotives, like its overground counterparts. The City and South London Railway was the first underground railway to use electric traction from its opening in 1890 onwards. As reported in Table G.5, average travel speeds using underground railways were slightly slower than overground railways at around 15 mph, reflecting both differences in the engineering conditions and frequency of intermediate stops. In London today, average travel speeds using overground and underground railways are little different from those reported in Tables G.4 and G.5.[2]

We assume relative weights for different modes of transport based on these average travel speeds in our quantitative analysis of the model. Normalizing the weight for overground railways to 1, we assume the following weights for the other modes of transport: walking 7 (21/3); horse and motor omnibuses and horse and electric trams 3.5 (21/6); and underground railways 1.4 (21/15).

[2] See "Commuter Journeys Slower than Before the War," David Millward, *Daily Telegraph*, 18th August 2015.

Table G.2: Travel Speeds for Horse and Motor Omnibuses for 1907

	Route or Section of Route		Horse Omnibuses		Motor Omnibuses	
		Approximate Distance (Miles)	Time for Journey (minutes)	Speed (Miles Per Hour)	Time for Journey (minutes)	Speed (Miles Per Hour)
	Routes Through Central Areas					
1.	Liverpool Street to Wormwood Scrubs	7.2	70	6.2	56	7.7
	Wormwood Scrubs to Liverpool Street	7.3	74	5.9	61.5	7.1
2.	Bank to Shepherd's Bush	5.7	61	5.6	57.5	6
	Shepherd's Bush to Bank	5.7	62	5.5	45	7.6
3.	Oxford Circus to Kilburn	3.5	37	5.7	27.5	7.7
	Kilburn to Oxford Circus	3.5	34	6.2	26.5	8.1
4.	Bank to Putney	7	67	6.3	55	7.7
	Putney to Bank	7	75	5.6	61	6.8
5.	Bank to Hammersmith	6.5	61	6.4	49	8.0
	Hammersmith to Bank	6.4	61	6.3	56.5	6.8
	Average Speed			6.0		7.3
	Routes Through Outlying Areas					
6.	Clapham ("Plough") to Putney	3.9	37	6.3	27	8.6
	Putney to Clapham ("Plough")	3.9	40	5.8	26	8.9
7.	Shepherd's-Bush to Putney	3.2	31.5	6.1	20	9.6
	Putney to Shepherd's-Bush	3.2	29.5	6.4	20.5	9.4
	Average Speed			6.1		9.1
	Average Speed for all the Above Journeys			6.0		7.5

Source: *London Statistics*, 1907.

Table G.3: Travel Speeds for Trams 1904

Type of Tram	Speed During Rush Hours (Miles Per Hour)	Speed During Slack Hours (Miles Per Hour)
Horse	2-5	5.5-8
Electric	5.5-7	8.5-11.5

Source: *Royal Commission on London Traffic*, 1904.

Table G.4: Travel Speeds for Overground Railways for 1907 from 8-9am

Railway and Terminus	Number of inward suburban trains 8-9am	Average speed of inward suburban trains 8-9am
Great Central (Marylebone)	5	29.9
London and North Western (Euston)	7	28.7
South-Eastern and Chatham		
(London-Bridge)	3	28.5
(Holborn)	8	20.3
(Victoria)	9	19.0
(Charing Cross)	12	18.0
(Cannon Street)	10	17.1
(Ludgate Hill, St. Paul's, Moorgate-Street and Farringdon-Street)	16	16.4
Great Western (Paddington)	10	27.2
Midland		
(St. Pancras)	5	26.2
(Moorgate Street)	6	15.3
Great Northern		
(King's Cross)	16	23.3
(Moorgate-Street)	7	15.6
London, Tilbury and Southend (Fenchurch Street)	6	21.3
London and South-Western (Waterloo)	27	21.2
Great Eastern		
(Liverpool Street)	38	20.3
(Fenchurch Street)	14	16.0
London, Brighton and South Coast		
(London-Bridge)	28	20.2
(Victoria)	15	18.3
North London (Broad Street)	28	18.1

Source: *London Statistics*, 1907.

Table G.5: Travel Speeds for Underground Railways for 1907

Railway	Total Length (miles)	Number of Intermediate Stations	Time for Journey (minutes)	Approximate Speed per Hour (miles)
Great Northern, Piccadilly and Brompton	8.9	19	32	16.7
Baker Street and Waterloo	4.25	9	16	15.9
Great Northern and City	3.44	4	13	15.9
Central London	5.75	11	23	15.0
Metropolitan and District				
(Inner Circle)	13	26	50	15.6
(Ealing-Whitechaptel)	12.74	24	51	15.0
City and South London	7.45	13	30	14.9
Charing Cross, Euston and Hampstead				
(Charing Cross - Golder's Green)	5.93	10	24	14.8
(Charing Cross - Highgate)	4.28	10	18	14.3

Source: *London Statistics*, 1907.

G.7 Bilateral Commuting Data for 1921

The 1921 population census of England and Wales reports bilateral flows of commuters from each residence borough to each workplace borough. The data for London and its surrounding Home Counties of Essex, Hertfordshire, Kent, Middlesex and Surrey are reported in Census of England and Wales, 1921, County of London, Tables, Part III (Supplementary), Workplaces in London and Five Home Counties. Residence is measured based on the population present within the area on Census night whether as permanent residence or as temporary visitors. Workplace is measured based on usual place of work. Bilateral flows of less than 20 people are not reported for confidentiality reasons and are omitted. Summing these reported bilateral flows, the resulting sums of workplace employment and residence employment are close to the totals for workplace employment and residence employment (including flows of less than 20 people) that are separately reported in the population census.

G.8 Commuting Data form Henry Poole Tailors

Prior to the 1921 census, there is no comprehensive data on commuting flows in London. To provide some evidence on commuting in the years before 1921, we consulted a number of company archives. While many archives contain lists of employees, most companies do not seem to have recorded the home addresses of their employees. An exception is Henry Poole Tailors, a high-end bespoke tailoring firm, which was founded in 1802, and has been located on London's Savile Row since 1828. Savile Row and its immediate surroundings have been a concentration of bespoke tailoring firms in London for several hundred years. The archives of Henry Poole contain ledgers with the names and home addresses of their employees going back to 1857.

This data has previously been used by the historian David Green in Green (1988). He kindly made available to us his raw data with his transcriptions of the ledgers surviving in the archives of Henry Poole. The data has a simple format. It lists the name of the employee and the year in which he or she joined the company. It also records a history of home addresses for each person. Unfortunately, for most address changes, no year is recorded. We therefore follow Green (1988) and only use the first reported home address and assume that it is the address at which the person lived in the year in which he or she joined Henry Poole Tailors.

The data is organised in two ledgers. The first ledger contains employees that joined the company between 1857 and 1877. For these employees, David Green did not transcribe the exact year in which they joined the company. The second ledger mainly contains employees who were hired after 1893 and David Green transcribed the data in this ledger up to 1914. However, this second ledger also contains a number of employees who in some cases were hired decades before 1893. For employees in this second ledger who were hired before the 1890s, it is uncertain whether the first recorded address is their home address at the time they were hired, or their home address at the time this second ledger was started. Therefore, we ignore these observations, and only use observations from the first ledger to compute commuting distances for the period from 1857 to 1877.[3] Finally, for a small number of observations in the second ledger, the year hired is missing.

David Green geo-located the home addresses of the Henry Poole employees in these two ledgers by hand on printed maps and then worked out the distances to the workshop of Henry Poole on Savile Row. The straight-line distances that he measured have unfortunately been lost. We therefore geolocated each address using Google Maps.

[3] There are a number of employees who appear in both ledgers, as discussed further below.

In a large number of cases, these addresses do not exist any more today, because of street name changes or changes in London's street layout. To track street name changes, we used *Bruce's List of London Street Name Changes*, which contains a record of all street name changes in London between 1857 and 1966 collected by Bruce Hunt. A copy of this book is available at www.maps.thehunthouse.com. If addresses were located on streets that were renamed or no longer exist today for other reasons, the location of the address was traced with the help of the geo-referenced historical maps of London provided on www.oldmapsonline.org.

Of the employees contained in the first ledger covering the period 1857 to 1877, David Green geolocated the home addresses of 162 workers. We managed to geolocate 156 of these addresses.[4] For some addresses, only a street name and house number is provided, with no further information, and a given street name can appear multiple times in London. Therefore, for each address, we have documented how confident we are that we have correctly geolocated this address, by assigning a confidence level of low, low to medium, medium, medium to high, or high, depending on the further information provided in the ledger (such as the name of a suburb) and the frequency of use of the street name in London. In this first ledger, there are 135 addresses for which our confidence that the geolocation is correct is medium to high or high. We use these 135 addresses for our analysis, for which the median and 95th percentile commuting distances are 1.9 and 4.8 kilometers respectively.

Of the employees contained in the second ledger, David Green concentrated on those who were employed by Henry Poole Tailors up to and including 1899. We instead considered all 190 employees whose addresses David Green transcribed from the second ledger. Of these employees, the year in which they were hired by Henry Poole was missing for 10 names. We drop a further 18 names for employees who were hired before 1878 according to the second ledger and also appear in the first ledger. The first address reported for these workers differs across the two ledgers, suggesting that they were added to the second ledger at the time this ledger was created, and instead of transferring across their entire address history, only their address at the time they were added to the second ledger was transcribed.[5] There were a further 10 names for which we could not geolocate their address. Of the remaining 150 workers, 111 were hired between 1890 and 1914, and we were able to geolocate 95 of these addresses with a medium-to-high or high confidence. We use these 95 addresses for our analysis, for which the median and 95th percentile commuting distances are 5.2 and 16.2 kilometers respectively.

G.9 Floor Space Supply Elasticity

In this section, we discuss our calibration of the floor space supply elasticity. As discussed in the paper, we assume that the supply of floor space (H_n) depends on both geographical land area (K_n) and the density of development as measured by the ratio of floor space to land area (h_n). Following Saiz (2010), we allow the supply of floor space to respond endogenously to changes in its price, as in equation (14) in the paper, which is reproduced below:

$$H_n = h_n K_n, \qquad h_n = hQ_n^{\mu} \tag{G.1}$$

where h is a constant; $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.

[4] Of these 156 workers, Mr S. Codling reports living in Newcastle, and is probably a representative or salesmen of Henry Poole Tailors. Therefore, we drop him from the sample, as he cannot have commuted to London on a daily basis.

[5] We also drop John Blanchard Ash, who lives in Dover when he is hired by Henry Poole Tailors, and whose second recorded address is in Brussels. While we are not certain of his role, he is likely to have been a salesman or foreign representative who did not work in London. Finally, we also drop Rene Jones, who lives in Paris, and is described as an assistant to the Paris branch.

To calibrate the floor space supply elasticity, we use data on the evolution of office rents and the supply of floor space for the City of London during the second half of the 19th century. Using data from the rent rolls, deed books and lease registers of five property investment companies whose archives are held at the Guildhall Library in London, Devaney (2010) estimates that office rents in the City of London grew at an average annual rate of 0.6 percent over the period 1867-1913. This average annual growth rate implies a cumulative growth in office rents (the price of floor space) between our census years of 1871 and 1911 of 27 percent:

$$\frac{Q_{1911}}{Q_{1871}} = (1 + 0.006)^{1911-1871} = 1.2703, \tag{G.2}$$

Over the same period, the cumulative growth of rateable values in the City of London in our data is 125 percent:

$$\frac{\mathbb{Q}_{1911}}{\mathbb{Q}_{1871}} = 2.2530. \tag{G.3}$$

Using the fact that rateable values equal the price times the quantity of floor space, we can recover the cumulative growth in the quantity of floor space from these estimates, which equals 77 percent:

$$\frac{H_{1911}}{H_{1871}} = \frac{\mathbb{Q}_{1911}/\mathbb{Q}_{1871}}{Q_{1911}/Q_{1871}} = \frac{2.2530}{1.2703} = 1.7736. \tag{G.4}$$

Assuming the stable supply function for floor space in equation (G.1), we estimate the floor space supply elasticity as the percentage change in the quantity of floor space divided by the percentage change in the price of floor space:

$$\epsilon_H = \frac{\left(\frac{H_{1911}}{H_{1871}} - 1\right)}{\left(\frac{Q_{1911}}{Q_{1871}} - 1\right)} = \frac{0.7736}{0.2703} = 2.8620. \tag{G.5}$$

As an independent check on this calibration, the implied 77 percent growth in the supply of floorspace in the City of London over this period is close to the estimate of at least 50 percent in Turvey (1998): "An interesting guesstimate is that of all the buildings that existed in 1855, about four-fifths had been rebuilt by 1905, and that while the average street block in 1840 probably carried an amount of floorspace equal to twice its gross area, the blocks largely occupied by late Victorian buildings carried twice as much. Thus it is likely that during Victorian times, City floorspace increased by at least one-half." (page 57). Additionally, our finding of a substantial positive estimated floor space supply elasticity is consistent with London's rapid 19th-century growth predating the planning and zoning regulations that were introduced in the aftermath of the Second World War following the Barlow Commission Report of 1940, as discussed in Foley (1963). Finally, our estimate for the floor space supply elasticity of 2.86 is also in line with the values reported for US metropolitan areas with relatively light planning regulations in Saiz (2010).

G.10 Construction Costs Estimates of the Railway and Underground Lines

In this section of the web appendix, we estimate separate construction costs per mile of line for overground and underground railways. For underground railways we further distinguish between the different construction costs of shallow "cut-and-cover" and deep "bored-tube" underground railways.

G.10.1 Underground Railways

The earliest London Underground lines (such as the Metropolitan District Railway that was opened in 1863) used shallow "cut-and-cover" construction methods, with the lines frequently running underneath existing streets. These "cut-and-cover" methods involve excavating a trench for the underground line and constructing a roof overhead to bear the load of whatever is above. Following improvements to existing tunneling shields by James Greathead in the 1870s and 1880s, the first London Underground railway that was built using deep "bored-tube" techniques was the City and South London railway, which opened in 1890. Greathead's shield consisted of an iron cylinder, which was inched forward as the working face was excavated, while behind it a permanent tunnel lining of cast iron segments was fitted into place.[6]

Before the Second World War, both overground and underground lines in the United Kingdom were constructed by private companies, who raised capital to finance construction and paid dividends on this capital from their ticket revenue. Construction of each railway line or extension of an existing line had to be approved by Parliament. A simple measure of the construction costs of overground and underground railway lines is therefore the amount of capital that these private companies raised per mile of line constructed. This measure should capture the full costs of constructing railway lines, including not only the cost of construction work (tunnels, rails, stations and other buildings and structures) but also the cost of purchasing land, rolling stock and fees for operating permissions.[7]

For a 1901 parliamentary report ("Report From The Joint Select Committee of The House of Lords on London Underground Railways") Henry L. Cripps compiled an overview of extensions constructed and proposed by London Underground companies. His data is contained in Appendix B of the report. Cripps reports the length of the extensions and the amount of capital authorized by parliament per mile of underground line. From Cripps's data we extract, for each London Underground company, the years in which the extensions were authorized, the total length in miles of line authorized, and the average authorized capital per mile as summarized in Table G.6.[8]

Table G.6: Authorized Capital per Mile for London Underground Companies

Underground Railway Company	Years Extensions Authorized	Miles of Line Authorized	Authorized Capital per mile in pounds (current prices)	Authorized Capital per mile in pounds (1921 prices)
City and South London	1884-1898	6.88	319,709	374,060
Central London	1891-1892	6.79	559,852	649,428
Great Northern and City	1892-1897	3.48	598,561	682,360
Baker Street and Waterloo	1893-1900	5.25	605,523	653,965
Charing Cross, Euston and Hampstead	1893-1899	6.10	388,196	438,661
Waterloo and City	1893	1.59	453,543	512,504
Brompton and Picadilly Circus	1897-1899	2.41	552,538	574,640
Metropolitan District	1897	4.87	328,205	354,461

Note: Taken from Appendix B of the "Report From The Joint Select Committee of The House of Lords on London Underground Railways" (1901), compiled by Henry L. Cripps. Current year prices converted to constant 1921 prices using a price deflator based on the ratio of overground construction costs in those years from the Railway Returns.

We classify each company in Table G.6 as either a "bored-tube" or "cut-and-cover" tube operator, based on the classification provided by Croome and Jackson (1993), in order to generate separate cost estimates for these two construction types. During the period surveyed by Cripps, the extensions to the underground network were pre-

[6]See the discussion in chapter 4 of Barker and Robbins (1963) for a description of the construction of the Metropolitan line tunnels by "cut-and-cover". Croome and Jackson (1993) provide a detailed history of the construction of the London Underground system from the early "cut-and-cover" construction to the later "bored-tube" lines that could be constructed under all parts of Central London.

[7]Kellet (1969) examines the capital accounts of 26 railway companies in the UK and finds that on average across these companies, as a share of the total cost of construction, the cost of land represented 25 percent and Parliamentary expenses represented a further 6 percent.

[8]Cripps's survey also includes a few projects that were authorized but were never actually built. We exclude such extensions that were proposed but never built from Table G.6.

dominantly for "bored-tube" lines. With the exception of the Metropolitan District extension authorized in 1897, all other projects listed in Table G.6 are "bored-tube" line extensions. For our baseline measure, we use the unweighted average of the authorized capital per mile for these two groups of companies, which yields £330,000 and £500,000 for "cut-and-cover" and "bored-tube" lines, respectively. These averages are reported to the nearest £10,000. We convert these figures in current year prices into common 1921 prices using a price deflator based on the ratio of overground railway construction costs in those years from the "Railway Returns: Returns of the Capital, Traffic, Receipts, and Working Expenditure of the Railway Companies of Great Britain." We thus obtain authorized capital per mile in 1921 prices for "cut-and-cover" and "bored-tube" lines of £355,000 and £555,000. Using a weighted average of authorized capital per mile for of the "bored-tube" underground companies, where the weights are the miles of line authorized, yields a similar figure in 1921 prices of £544,000.

By 1921 many of the London underground lines had added sections to their network that ran overground rather than in tunnels. Typically, as tube lines reached the boundary of the densely built-up area of London the line continued above ground to avoid the heavy costs of tunneling. Based on the work of the *Cambridge Group for the History of Population and Social Structure*, discussed further in Section G.4 of this web appendix, we are able to classify the parts of each underground line that run in tunnels and those that are above ground. From this data, it is clear that the extensions to the underground network considered by Cripps were all parts of the network running in tunnels, and hence his authorized capital per mile is for the construction of lines in tunnels. In estimating the overall costs of constructing the underground network, we make the natural assumption that the construction costs per mile for the parts of each underground line that are above ground are the same as those for overground railways, as discussed in the next subsection.

G.10.2 Overground Railways

We estimate the construction costs for overground railway lines using the 1921 edition of the "Railway Returns: Returns of the Capital, Traffic, Receipts, and Working Expenditure of the Railway Companies of Great Britain", which compiled large volumes of data on the railways of the United Kingdom. In particular, the Railway Returns report both the total line length of the UK railway network as well as both the authorized capital and paid-up capital of all railway companies. For our basic estimate of construction costs we divide the capital values by the length of line to calculate the average capital per mile.[9] For our baseline measure, we use the authorized capital per mile of railway line, which which rounded to the nearest one thousand pounds is £60,000 in 1921 prices. Using instead the average capital paid-up per mile would result in a similar figure of £57,000 in 1921 prices.[10]

G.10.3 Robustness

A potential concern with our estimate of overground construction costs is that UK averages may not be representative of construction costs in London and its surroundings. To examine this possibility, we use data from the 1921 edition of the Railway Returns for individual railway companies. This data reports the total authorized capital and the length of

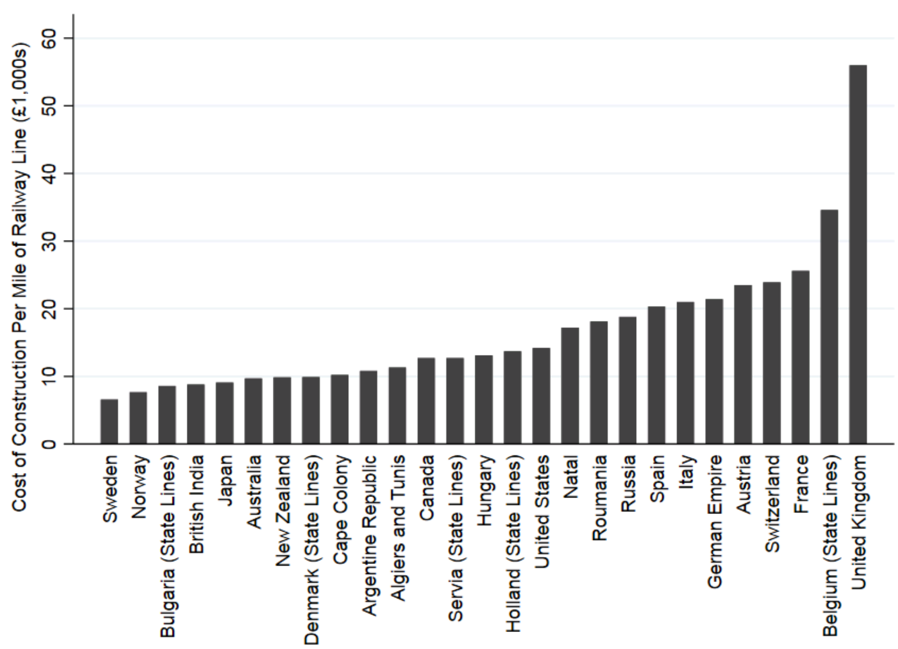
[9]Line length is defined treating each line as a single-track line, regardless of whether or not there are multiple tracks that run in parallel.

[10]The total length of railway lines in the UK and also the aggregate capital stocks in the summary table of the Railway Returns implicitly include the London Underground system. As underground railways are more costly to construct than overground railways, this will bias our estimate of construction costs for overground railways upwards, and implies that our estimates of these construction costs for overground railways are upper bounds. In practice, this bias is likely to be small, because the London Underground network is a tiny fraction of the 23,724 miles of railway line that were open in the United Kingdom in 1921.

the operated network for each major railway company in the UK. From this data, we selected the railway companies that operate services to and from London and computed their average authorized capital per mile of line, which results in a figure of £60,000 to the nearest one thousand pounds.[11] Therefore, London-based companies do not appear to have faced substantially different construction costs from the UK average, and we assume a construction cost of £60,000 per mile for overground railways in Greater London.

As another robustness check, we compared our historical estimates of construction costs for overground railways with those from other sources. The 1911 edition of the "New Dictionary of Statistics" provides a comparison of railway construction costs per mile across 27 countries over the years 1905-1908. Figure G.19 displays the data from the Dictionary of Statistics. The notes provided in the Dictionary of Statistics show that their estimate of the construction costs for the UK in 1907 of £56,000 is also based on the data from the Railway Returns and simply divides total authorized capital by the length of the UK railway network. This estimate for 1907 therefore uses the same data and approach as our estimate of overground construction costs in 1921. The figure shows that construction costs in the UK are, if anything, at the upper bound of construction costs in other countries.

Figure G.19: Construction Costs of Railway Lines Across Countries



Note: Taken from pp. 512-513 of the 1911 edition of the "The New Dictionary of Statistics" compiled by Augustus D. Webb. The data complied by Webb comes from the years 1905 to 1908 depending on data availability for different countries.

There could be a number of reasons why construction costs in the UK were higher than that of other countries. First, the cost of purchasing land for railway construction in the UK might have been higher than those in other countries. Sir Josiah Stamp, who was the chairman of the London, Midland and Scottish Railway Company, for example, argued that "in international comparisons of costs Great Britain notoriously suffers through the heavy initial outlay in lands and parliamentary expenses to acquire them against bitter opposition." Second, the Dictionary of

[11]The main railway companies that have a terminus in London in 1912 are the Great Eastern, Great Northern, Great Western, London and North Western, London and South Western, London, Brighton and South Coast, North Eastern and South Eastern and Chatham company.

Statistics notes that the UK probably had a higher share of multiple track railway lines compared to other countries, in particular the US. As a single track line is less expensive to build than multiple track lines, this could contribute to the higher costs per mile of railway line built in the UK. Finally, the UK was one of the leading industrial nations before the First World War. The resulting higher wages compared to other countries are likely to have made construction work, which was labor intensive at this time, more expensive.

As a final robustness check, we examined the evolution of overground railway construction costs over time using the Railway Returns. For each year between 1871 and 1912 and also 1920 and 1921, we can compute the authorized capital per mile of line. A similar time series can be generated for paid-up capital per mile. We find that both capital per mile measures steadily increase over time. In 1871, authorized capital per mile was approximately £40,000. By 1900, this number reaches £60,000, after which it plateaus and remains relatively constant until 1921. These reported changes in authorized capital per mile reflect both inflation and changes in the real costs of railway construction. Nevertheless, they suggest that the early parts of the UK network could have been constructed at lower costs, implying that our estimate of £60,000 based on 1921 is again an upper bound on construction costs.

References

Allen, Treb, Costas Arkolakis and Xiangliang Li (2015) "Optimal City Structure," Yale University, mimeograph.

Armington, Paul S. (1969) "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers*, 16(1), 159-178.

Barker, T. (1980) "Towards an Historical Classification of Urban Transport Development Since the Later Eighteenth Century," *Journal of Transport History*, 1(1), 75-90.

Barker, T. C., and M. Robbins (1963) *A History of London Transport: The Nineteenth Century*, vol. 1, London: George Allen and Unwin.

Cobb, Michael H. (2003) *The Railways of Great Britain: a Historical Atlas*, Patrick S. Cobb.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest and Clifford Stein (2009) *Introduction to Algorithms*, Third Edition, Cambridge MA: MIT Press.

Croome, Desmond F. and Alan A. Jackson (1993) *Rails Through the Clay: a History of London's Tube Railways*, London: Capital Transport Publishing.

Devaney, Steven (2010) "Trends in Office Rents in the City of London: 1867-1959," *Explorations in Economic History*, 47, 198-212.

Eaton, Jonathan and Samuel S. Kortum (2002) "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741-1779.

Foley, D. L. (1963) *Controlling London's Growth: Planning the Great Wen*, Berkeley: University of California Press.

Gerhold, D. (2005) *Carriers and Coachmasters: Trade and Travel Before the Turnpikes*, Bodmin, Cornwall: Phillimore.

Green, David R. (1988) "Distance to Work in Victorian London: A Case Study of Henry Poole, Bespoke Tailors," *Business History*, 30(2), 179-194.

Helpman, Elhanan (1998) "The Size of Regions," in *Topics in Public Economics: Theoretical and Applied Analysis*, (ed.) D. Pines, E. Sadka, and I. Zilcha. Cambridge University Press, Cambridge.

Kellet, John R. (1969) *The Impact of Railways on Victorian Cities*, London: Routlege and Kegan Paul.

Krugman, Paul (1991) "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483-499.

Monte, Ferdinando, Stephen J. Redding and Esteban Rossi-Hansberg (2018) "Commuting, Migration and Local Employment Elasticities," *American Economic Review*, forthcoming.

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

Redding, Stephen J. (2016) "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 101, 148-167.

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

Saiz, Albert (2010) "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3), 1253-1296.

Salmon, James (1891) *Ten Years' Growth of the City of London*, Report, Local Government and Taxation Committee of the Corporation, London: Simpkin, Marshall, Hamilton, Kent and Company.

Stamp, Josiah (1922) *British Incomes and Property: The Application of Official Statistics to Economic Problems*, London: P. S. King and Son.

Turvey, Ralph (1998) "Office rents in the City of London, 1867-1910," *The London Journal*, 23, 53-67.

Wrigley, Edward Anthony (2011) *The Early English Censuses*, Oxford: Oxford University Press.