PAPER NOTES

Xuyuan Zhang

Update on July 30, 2024

Contents

List of Theorems

List of Definitions

LIST OF DEFINITIONS

Part I

urban economics

Chapter 1

THE ECONOMICS OF DENSITY: EVIDENCE FROM THE BERLIN WALL

1.1 THEORETICAL MODEL

The city consists of a set of discrete locations, index by $i = 1, \ldots, S$. Each block has an effective supply of floor space L_i. Floor space can be used commercially or residentially, and we denote the endogenous fractions of floor space allocated to commercial and residential use by θ_i and $1 - \theta_i$. City is populated by an endogenous measure of H workers, who are perfectly mobile within the city and the larger economy, which provides a reservation level of utility $\overline{U}$. Workers decide whether or not to move to the city before observing idiosyncratic utility shocks for each possible pair of residence and employment blocks within the city.

1.1.1 Workers

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

$$C_{ijo} = \frac{B_i z_{ijo}}{d_{ij}} (\frac{c_{ijo}}{\beta})^{\beta} (\frac{\ell_{ijo}}{1-\beta})^{1-\beta}, 0 < \beta < 1 \tag{1.1}$$

where the iceberg communiting cost $d_{ij} = e^{\kappa \tau_{ij}} \in [1, \infty)$ increases with the travel time (τ_{ij}) between blocks i and j. The parameter κ controls the size of commuting costs.

For each worker o living in block i and commuting to block j, the idiosyncratic component of utility (z_{ijo}) is drawn from an independent Frechet distribution:

$$F(z_{ijo}) = e^{-T_i E_j z_{ijo}^{-\varepsilon}}, T_i, E_j > 0, \varepsilon > 1 \tag{1.2}$$

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

The indirect utility from residing in block i and working in block j can be expressed in terms of the wage paid at this workplace (w_j), commuting costs (d_{ij}), the residential floor price (Q_i), the common component of amenities (B_i), and the idiosyncratic shock (z_{ijo}):

$$u_{ijo} = \frac{z_{ijo} B_i w_j Q_i^{\beta-1}}{d_{ij}} \tag{1.3}$$

where we have used utility maximization and the choice of the final good as numeraire. Although we model commuting costs in terms of utility, there is an isomorphic formulation in terms of a reduction in effective units of labor, because the iceberg commuting cost $d_{ij} = e^{\kappa\tau_{ij}}$ enters the indirect utility function multiplicatively. As a result, commuting costs are proportional to wages, and this specification captures changes over time in the opportunity cost of travel time.

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

$$\pi_{ij} = \frac{T_i E_j (d_{ij} Q_i^{1-\beta})^{-\varepsilon} (B_i w_j)^{\varepsilon}}{\sum_{r=1}^{S} \sum_{s=1}^{S} T_r E_s (d_{rs} Q_r^{1-\beta})^{-\varepsilon} (B_r w_s)^{\varepsilon}} \equiv \frac{\Phi_{ij}}{\Phi}. \tag{1.4}$$

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

$$\pi_{Ri} = \sum_{j=1}^{s} \pi_{ij} = \frac{\sum_{j=1}^{s} \Phi_{ij}}{\Phi}, \pi_{Mj} = \sum_{i=1}^{S} \pi_{ij} = \frac{\sum_{i=1}^{S} \Phi_{ij}}{\Phi} \tag{1.5}$$

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

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

$$\pi_{ij|i} = \frac{E_j (w_j / d_{ij})^{\varepsilon}}{\sum_{s=1}^{S} E_s (w_s / d_{is})^{\varepsilon}} \tag{1.6}$$

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

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

$$H_{M_j} = \sum_{s=1}^{S} \frac{E_j (w_j / d_{ij})^{\varepsilon}}{\sum_{s=1}^{S} E_s (w_s / d_{is})^{\varepsilon}} H_{R_i} \tag{1.7}$$

where H_{R_i} is the measure of residents in block i. Since there is a continuous measure of workers residing in each location, there is no uncertainty in the supply of workers to each employment location.

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

$$E[w_j \mid i] = \sum_{j=1}^{S} \frac{E_j(w_j/d_{ij})^{\varepsilon}}{\sum_{s=1}^{S} E_s(w_s/d_{is})^{\varepsilon}} w_j \tag{1.8}$$

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

$$E[u] = \gamma[\sum_{r=1}^{S}\sum_{s=1}^{S} T_r E_s(d_{rs}Q_r^{1-\beta})^{-\varepsilon}(B_r w_s)^{\varepsilon}]^{1/\varepsilon} = \tilde{U} \tag{1.9}$$

where E is the expectation operator and expectation is taken over the distribution for the idiosyncratic component of utility: $\gamma = \Gamma(\frac{\varepsilon-1}{\varepsilon})$ and $\Gamma(\cdot)$ is the Gamma function.

1.1.2 Production

Production of the tradable final good occurs under conditions of perfectly competition and constant returns to scale. We can assume the production technology takes the CD-form

$$y_j = A_j H_{M_j}^{\alpha} L_{M_j}^{1-\alpha} \tag{1.10}$$

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

$$H_{M_j} = (\frac{\alpha A_j}{w_j})^{1/(1-\alpha)} L_{M_j} \tag{1.11}$$

where the equilibrium wage is determined by the requirement that the demand for workers in each employment location equals the supply of workers to that location.

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

$$q_j = (1-\alpha)(\frac{\alpha}{w_j})^{\alpha/(1-\alpha)} A_j^{1/(1-\alpha)} L_{M_j}^{-1/(1-\alpha)} \tag{1.12}$$

1.1.3 Land Market Cleaning

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

$$\begin{aligned} &\theta_i = 1 \text{ if } q_i > \xi_i Q_i \\ &\theta_i \in [0,1] \text{ if } q_i = \xi_i Q_i \theta_i \quad = 0 \text{ if } q_i < \xi_i Q_i \end{aligned} \tag{1.13}$$

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

$$\mathbb{Q}_i = q_i, q_i > \xi_i Q_i, \theta_i = 1$$
$$\mathbb{Q}_i = q_i, q_i = \xi_i Q_i, \theta_i \in [0, 1] \tag{1.14}$$
$$\mathbb{Q}_i = Q_i, q_i < \xi_i Q_i, \theta_i = 0$$

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

$$L_i = \phi_i K_i^{1-\mu} \tag{1.15}$$

$$\mathbb{Q}_i = \chi \mathbb{R}_i^{1-\mu} \tag{1.16}$$

where we refer to $\phi_i = M_i^\mu$ as the density of development and χ is a constant. Residential land market clearing implies that the demand for residential floor space equals the supply of floor space allocated to residential use in each location: $(1-\theta_i)L_i$. Maximization for each worker and taking expectation over the distribution for idiosyncratic utility, the residential land market clearing condition can be expressed as:

$$\mathbb{E}[\ell_i]H_{R_i} = (1-\beta)\frac{\mathbb{E}[w_s \mid i]H_{R_i}}{Q_i} = (1-\theta_i)L_i. \tag{1.17}$$

Commercial land market cleaning requires that the demand for commercial floor space equals the supply of floor space allocated to commercial use in each location: $\theta_j L_j$. Using the FOC for profit maximization, the commercial land market clearing condition can be expressed as:

$$(\frac{(1-\alpha)A_j}{q_j})^{1/\alpha} H_{M_j} = \theta_j L_j \tag{1.18}$$

We both residential and commercial land market clearing are satisfied, total demand for floor space equals the total supply of floor space:

$$(1-\theta_i)L_i + \theta_i L_i = L_i = \phi_i K_i^{1-\mu} \tag{1.19}$$

1.1.4 General Equilibrium With Exogenous Location characteristics

Given the model's parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$, the reservation level of utility in the wider economy $\overline{U}$, and vectors of exogenous location characteristics $\{T, E, A, B, \phi, K, \xi, \tau\}$, the general equilibrium is referred with six vectors $\{\pi_M, \pi_R, Q, q, w, \theta\}$ and total city population H.

Proposition 1.1. *Assuming exogenous finite, and strictly location characteristics* ($T_i \in (0, \infty), E_i \in (0, \infty), \phi_i \in (0, \infty), K_i \in (0, \infty), \xi_i \in (0, \infty), \tau_{ij} \in (0, \infty) \times (0, \infty)$)*, and exogenous, finite and nonnegative final goods productivity* $A_i \in [0, \infty)$ *and residential amnetities* $B_i \in [0, \infty)$*, there exists a unique general equilibrium vector* $\{\pi_M, \pi_R, Q, q, w, \theta\}$

Proof. The proof can be decomposed into two parts: firstly, we show that under the assumptions that all blocks have strictly positive, finite, and exogenous location and we allow some blocks to be more attractive than others in terms of these characteristics. But workers draw idiosyncratic preferences from a Frechet distribution for pairs of residence and workplace locations, and therefore, since teh Frechet distribution is unbounded from above, any block with strictly positive characteristics has a positive measure of workers that prefer that location as a residence or workplace at a positive and finite price. Hence, all blocks with finite positive wages attract a positive measure of workers, and all blocks with finite positive floor prices attract a positive measure of residents.

Then We next show that blocks with strictly positive, finite, and exogenous location characteristics must have strictly positive and finite values of both wages and floor prices in equilibrium. □

1.1.5 Introducing Agglomenration Forces

We now introduce endogenous agglomenration forces. We allow final goods productivity to depend on production fundamentals (a_j) and production externalities (Y_j). Production fundamentals capture features of physical geography that make a location more or less productive independently of the surrounding density of economic activity. Production externalities impose structure on how the productivity of a given block is affected by the characteristics of other blocks. Specifically, we follow the standard approach in urban economics of modeling these externalities as depending on the travel-time weighted sum of workplace employment density in surrounding blocks:

$$A_j = a_j Y_j^{\lambda}, Y_j \equiv \sum_{s=1}^{S} e^{-\delta\tau_{js}}(\frac{H_{M_s}}{K_s}), \tag{1.20}$$

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

We model the externalities in workers' residential choices analygously to the externalities in firms' production choices. We allow residential amenities to depend on residential fundamentals (b_i) and residential externalities (Ω_i). Residential fundamentals capture features of physical geography that make a location a more or less attractive place to live independently of the surrounding density of economic activity. Residential externalities again impose structure on how the amenities in a given block are affected by the characteristics of other blocks. Specifically, we adopt a symmetric specification as for production externalities, and model residential externalities as depending on the travel time weighted sum of residential employment density in surrounding blocks:

$$B_i = b_i \Omega_i^{\eta}, \Omega_i \equiv \sum_{r=1}^{S} e^{-\rho\tau_{ir}}(\frac{H_{Rr}}{K_r}) \tag{1.21}$$

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

1.1.6 Recovering Location Characteristics

We now show that there is a unique mapping from the observed variables to unobserved location characteristics. Since a number of these unobserved variables enter the model isomorphically, we define the following composites denoted bya tilde:

$$\tilde{A}_i = A_i E_i^{\alpha/\varepsilon}, \tilde{a}_i = a_i E_i^{\alpha/\varepsilon}$$
$$\tilde{B}_i = B_i T_i^{1/\varepsilon} \zeta_{Ri}^{1-\beta}, \tilde{b}_i = b_i T_i^{1/\varepsilon} \zeta_{Ri}^{1-\beta}$$
$$\tilde{w}_i = w_i E_i^{1/\varepsilon}$$
$$\tilde{\phi}_i = \tilde{\phi}_i(\phi_i, E_i^{1/\varepsilon}, \xi_i)$$

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

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

Proposition 1.2. 1. *Given known values for the parameters* $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$ *and the observed data* $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \tau\}$ *there exists unique vectors of the unobserved location characteristics* $\{\tilde{\mathbf{A}}^*, \tilde{\mathbf{B}}^*, \tilde{\phi}^*\}$ *that are consistent with the data being an equilibrium of the model.*

2. *Given known values for the parameters* $\{\alpha, \beta, \mu, \varepsilon, \kappa, \lambda, \delta, \eta, \rho\}$ *and the observed data* $\{\mathbb{Q}, \mathbf{H_M}, \mathbf{H_R}, \mathbf{K}, \tau\}$ *there exists a unique vectors of the unobserved location characteristics* $\{\tilde{\mathbf{a}}^*, \tilde{\mathbf{b}}^*, \tilde{\phi}^*\}$ *that are consistent with the data being an equilibrium of the model.*

Chapter 2

THE MAKING OF THE MODERN METROPOLIS: EVIDENCE FROM LONDON

2.1 Theoretical Framework

We consider a city embedded in a wider economy (Great Britain). The economy as a whole consists of a discrete set of locations $\mathbb{M}$. Greater London is a subset of these locations $\mathbb{N} \subset \mathbb{M}$, Time is discrete and is indexed by t. The economy as a whole is populated by an exogenous continuous measure $L_{\mathbb{M}t}$ of workers, who are geographically mobile and wndowed with one unit of labor that is supplied inelastically. Workers simultaneously choose their preferred residence n and workplace i given their idiosyncratic draws. We denote the endogenous measure of workers who choose a residence-workplace pair in Greater London by $L_{\mathbb{N}t}$. 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.

2.1.1 Preferences

We assume that preferences take the CD-form, such that the indirect utility for a worker ω residing in n and working in i is:

$$U_{ni}(\omega) = \frac{B_{ni} b_{ni}(\omega) w_i}{\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}}, 0 < \alpha < 1 \tag{2.1}$$

where we suppress the time subscript from now on; P_n is the price index for consumption goods, which may include both tradeable and nontradeable consumption goods; Q_n is the price of residential floor space; w_i is the wage, κ_{ni} is an iceberg commuting cost; B_{ni} captures amenities from the bilateral commute from residence n to workplace i that are common across all workers; 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 in the city.

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

$$G(b) = e^{-b^{-\varepsilon}}, \varepsilon > 1 \tag{2.2}$$

where we normalize the Frechet scale parameter in Equation (2.2) to 1 because it enters worker choice probabilities isomorphically to common bilateral amenities B_{ni}. The Frechet shape parameter ε regulates the dispersion of idiosyncratic amenities, which controls the sensitivity of worker location decisions to economic variables. The smaller the shape parameter ε, the greater the heterogeneity in idiosyncratic amenities, and the less sensitive are worker location decisions to economic variables.

We decompose the bilateral common amenities parameter (B_{ni}) into a residence component common across all workplaces $(B_n^{\mathcal{R}})$, a workplace component common across all residences (B_i^L), and an idiosyncratic component (B_{ni}^I) specific to an individual residence-workplace pair:

$$B_{ni} = B_n^{\mathcal{R}} B_i^L B_{ni}^I, \quad B_n^{\mathcal{R}}, B_i^L, B_{ni}^I > 0 \tag{2.3}$$

We allow the levels of $B_n^{\mathcal{R}}$, B_i^I and B_{ni}^I to differ across residences n and workplace i, although when we examine the impact of the construction of railway network, we assume that B_i^L and B_{ni}^I are time-invariant. In contrast, we allow $B_n^{\mathcal{R}}$ to change over time, and for those changes to be potentially endogenous to the evolution of the surrounding concentration of economic activity through agglomeration forces.

Conditional on choosing a residence-workplace pair in Greater London, we know that the probability a worker chooses to reside in location $n \in \mathbb{N}$ and work in location $i \in \mathbb{N}$ is given by:

$$\begin{aligned} \lambda_{ni} &= \frac{L_{ni}}{L_{\mathbb{M}}} \frac{L_{\mathbb{M}}}{L_{\mathbb{N}}} = \frac{L_{ni}}{L_{\mathbb{N}}} \\ &= \frac{(B_{ni} w_i)^{\varepsilon} (\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha})^{-\varepsilon}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} (B_{k\ell} w_\ell)^{\varepsilon} (\kappa_{k\ell} P_k^{\alpha} Q_k^{1-\alpha})^{-\varepsilon}}, n, i \in \mathbb{N} \end{aligned} \tag{2.4}$$

where L_{ni} is the measure of commuters from n to i.

The probability of commuting between residence n and workplace i depends on the characteristics of that residence n, the attributes of that workplace i and bilateral commuting costs and amenities. Summing across workplaces $i \in \mathbb{N}$, we obtain the probability that a worker lives in residence $n \in \mathbb{N}$, conditional on choosing a residence-workplace pair in Greater London ($\lambda_n^R = \frac{R_n}{L_{\mathbb{N}}}$). Similarly, summing across residences $n \in \mathbb{N}$, we obtain the probability that a worker is employed in workplace $i \in \mathbb{N}$, conditional on choosing a residence-workplace pair in Greater London ($\lambda_i^L = \frac{L_i}{L_{\mathbb{N}}}$)

$$\begin{aligned} \lambda_n^R &= \frac{\sum_{i \in \mathbb{N}} (B_{ni} w_i)^{\varepsilon} (\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha})^{-\varepsilon}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} (B_{k\ell} w_\ell)^{\varepsilon} (\kappa_{k\ell} P_k^{\alpha} Q_k^{1-\alpha})^{-\varepsilon}} \\ \lambda_i^L &= \frac{\sum_{n \in \mathbb{N}} (B_{ni} w_i)^{\varepsilon} (\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha})^{-\varepsilon}}{\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} (B_{k\ell} w_\ell)^{\varepsilon} (\kappa_{k\ell} P_k^{\alpha} Q_k^{1-\alpha})^{-\varepsilon}} \end{aligned} \tag{2.5}$$

where R_n denotes employment by residence in location n and L_i denotes employment by workplace in location i. A second implication of our extreme value specification is that expected utility conditional on choosing a residence workplace pair $(\overline{U})$ is the same across all residence-workplace pairs in the economy:

$$\overline{U} = v \left[\sum_{k \in \mathbb{M}} \sum_{\ell \in \mathbb{M}} (B_{k\ell} w_{k\ell})^{\varepsilon} (\kappa_{k\ell} P_k^{\alpha} Q_k^{1-\alpha})^{-\varepsilon} \right]^{\frac{1}{\varepsilon}} \tag{2.6}$$

where the expectation is taken over the distribution for idiosyncratic amenities; $v \equiv \Gamma(\frac{\varepsilon-1}{\varepsilon})$; $\Gamma(\cdot)$ is the gamma function. Using the probability that a worker chooses a residence-workplace pair in Greater London $(\frac{L_{\mathbb{N}}}{L_{\mathbb{M}}})$, we can rewrite this probability mobility condition as:

$$\overline{U} (\frac{L_{\mathbb{N}}}{L_{\mathbb{M}}})^{\frac{1}{\varepsilon}} = v \left[\sum_{k \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} (B_{k\ell} w_{k\ell})^{\varepsilon} (\kappa_{k\ell} P_k^{\alpha} Q_k^{1-\alpha})^{-\varepsilon} \right]^{\frac{1}{\varepsilon}} \tag{2.7}$$

where only the limits of the summations differ on the right hand sides of the equations.

Intuitively, for a given common level of expected utility in the economy $(\overline{U})$, locations in Greater London must offer higher real wages adjusted for common amenities (B_{ni}) and commuting costs (κ_{ni}) to attract workers with lower idiosyncratic draws with an elasticity determined by the parameter ε.

2.1.2 Production

We assume that consumption goods are produced according to a Cobb-Douglas technology using labor, machinery capital, and commercial floor space, where commercial floor space includes both building capital and land. Cost minimization and zero profits imply that payments for labor, commercial floor space, and machinery are constant shares of revenue (X_i):

$$w_i L_i = \beta^L X_i, q_i H_i^L = \beta^H X_i, r M_i = \beta^M X_i, \beta^L + \beta^H + \beta^M = 1 \tag{2.8}$$

where q_i is the price of commercial floor space; H_i^L is commercial floor space use; M_i is machinery use; and machinery is assumed to be perfectly mobile across locations with a common price r determined in the wider economy. We allow the price of commercial floor space (q_i) to potentially depart from the price of residential floor space (Q_i) in each location i through a location-specific wedge (ξ_i):

$$q_i = \xi_i Q_i. \tag{2.9}$$

From the relationship between factor payments and revenue in equation, payments for commercial floor space are proportional to workplace income $(w_i L_i)$:

$$q_i H_i^L = \frac{\beta^H}{\beta^L} w_i L_i \tag{2.10}$$

2.1.3 Commuter Market Clearing

commuter market clearing implies that the measure of workers employed in each location (L_i) equals the measure of workers chooosing to commute to that location:

$$L_i = \sum_{n \in \mathbb{N}} \lambda_{ni|n}^R R_n \tag{2.11}$$

where $\lambda_{ni|n}^R$ is the probability of commuting to workplace i conditional on living in residence n:

$$\lambda_{ni|n}^R = \frac{\left(\frac{B_{ni} w_i}{\kappa_{ni}}\right)^{\varepsilon}}{\sum_{\ell \in \mathbb{N}} \left(\frac{B_{n\ell} w_\ell}{\kappa_{n\ell}}\right)^{\varepsilon}} \tag{2.12}$$

where all characteristics of residence n have canceled from the above equation because they do not vary across workplaces for a given residence.

Commuter market clearing also implies that per capita income by residence (v_n) is a weighted average of the wages in all locations, where the weights are these conditional commuting probabilities by residences $(\lambda_{ni|n}^R)$:

$$v_n = \sum_{i \in \mathbb{N}} \lambda_{ni|n}^R w_i. \tag{2.13}$$

2.1.4 Land Market Clearing

We assume that floor space is owned by landlords, who receive payments from the residential and commercial use of floor space and consume only consumption goods. Land market clearing implies that total income from the ownership of floor space equals the sum of payments for residential and commercial floor space use:

$$\mathbb{Q}_n = Q_n H_n^R q_n H_n^L = (1-\alpha)\left[\sum_{i \in \mathbb{N}} \lambda_{ni|n}^R w_i\right] R_n + \frac{\beta^H}{\beta^L} w_n L_n \tag{2.14}$$

where H_n^R is residential floor space use; rateable values $(\mathbb{Q}_n)$ equals the sum of prices times quantities for residential floor space $(Q_n H_n^R)$ and commercial floor space $(q_n H_n^L)$; and we have used the expression for per capita income by residence (v_n) from commuter market clearing.

From the combined land and commuter market-clearing condition, payments for residential floor space are a constant multiple of residence income $(v_n R_n)$, and payments for commercial floor space are a constant multiple of workplace income $(w_n L_n)$. Importantly, we allow the supplies of residential floor space (H_n^R) and commercial floor space (H_n^L) to be endogenous, and we allow the prices of residential and commercial floor space to potentially differ from one another through the location-specific wedge $\xi_i (q_i = \xi_i Q_i)$. In our baseline quantitative analysis below, we are not required to make assumptions about these supplies of residential and commercial floor space or this wedge between commercial and residential floor prices. The reason is that we condition on the observed rateable values in the data $(\mathbb{Q}_n)$ and the supplies and prices for residential and commercial floor space (H_n^R, H_n^L, Q_n, q_n) only after the land market-clearing condition.

2.2 Quantitative Analysis

2.2.1 Combined Land and Commuter Market Clearing

We evaluate the effect of changes in the transport network by using an "exact hat algebra" approach. In particular, we rewrite our combined land and commuter market clearing condition for another year $\tau \neq t$ in terms of the values of variables in a baseline year t and the relative changes of variables between years τ and t:

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha)\hat{v}_{nt}v_{nt}\hat{R}_{nt}R_{nt} + \frac{\beta^H}{\beta^L}\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt} \tag{2.15}$$

where $\hat{x}_{nt} = \frac{x_{n\tau}}{x_{nt}}$ for the variable x_{nt} and we now make explicit the time subscripts. The relative change in employment $(\hat{L}_{it})$ and the relative change in average per capita income by residence $(\hat{v}_{nt})$ for year τ can be expressed as:

$$\hat{L}_{it}L_{it} = \sum_{n\in\mathbb{N}} \frac{\lambda^R_{nit|n}\hat{w}^{\varepsilon}_{it}\hat{\kappa}^{-\varepsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^R_{n\ell t|n}\hat{w}^{\varepsilon}_{\ell t}\hat{\kappa}^{-\varepsilon}_{n\ell t}}\hat{R}_{nt}R_{nt} \tag{2.16}$$

$$\hat{v}_{nt}v_{nt} = \sum_{i\in\mathbb{N}} \frac{\lambda^R_{nit|n}\hat{w}_{it}}{\sum_{\ell\in\mathbb{N}}\lambda^R_{n\ell t|n}\hat{w}^{\varepsilon}_{\ell t}\hat{\kappa}^{-\varepsilon}_{n\ell t}}\hat{w}_{it}w_{it} \tag{2.17}$$

where these equations include terms in change in wages $(\hat{w}_n)$ and commuting costs $(\hat{\kappa}_{ni})$ but not in amenities, because we assume that the workplace and bilateral components of amenities are constant $(\hat{B}^L_{it} = 1$ and $\hat{B}^I_{nit} = 1)$, and changes in the residential component of amenities $(\hat{B}^R_{nt} \neq 1)$ cancel from the numerator and denominator of the fractions.

substituting the expressions to the market clearing conditions for year τ we can get the result.

Lemma 2.1. *Suppose that* $(\hat{\mathbb{Q}}_{nt}, \hat{R}_{nt}, L_{nit}, \lambda^R_{nit|n}, \mathbb{Q}_{nt}, v_{nt}, R_{nt}, w_{nt}, L_{nt})$ *are known. Given known values for model parameters* $\{\alpha, \beta^L, \beta^H, \varepsilon\}$ *and the change in bilateral commuting costs* $(\hat{\kappa}^{-\varepsilon}_{nit})$*, the combined land and commuter market clearing condition determines a unique vector of relative changes in wages* $(\hat{w}_{it})$ *in each location.*

Chapter 3

Spatial Sorting and Inequality

3.1 Change in SKill Sourting: Framework

3.1.1 Setup

On the production side, rather than modelling imperfect trade between locations, we consider an economy that is more stylized spatially, with two types of goods: (a) a homogeneous manufactured good that is freely traded across space and (b) housing, a local nontraded good.

Preferences

Consider a spatial equilibrium framework with two skill groups $\theta = U, S$, who choose where to live among locations $i \in [1, \ldots, N]$. Aggregate skill supply for each group, L^θ, is exogenously given, and each worker supplies one unit of labor for wage w_i^θ in location i. The utility of worker w, who is type θ and lives in location i, is:

$$u_i^\theta(w) = \max_{c,b} \log U^\theta(A_i, c, b) + \varepsilon_i^\theta(w), \text{ such that } c + r_i b = w_i^\theta \tag{3.1}$$

Here $\log U^\theta(\cdot)$ is the representative utility of a worker of type θ; c is the consumption of the freely traded good and is taken as the numeraire; b denote housing, with price r_i in location i; and A_i is a vector of amenities in location i. Finally, $\varepsilon_i^\theta(w)$ is a worker-specific preference shock for living in location i.

First, we make assumption of CD type preferences over traded and nontraded goods. Second, we assume that amenities are separable from consumption. We allow amenities in location i to be valued differently by the two groups, as caputured by a group-specific amenity level A_i^θ. Third, preference shocks are typically chosen to be extreme value (EV) distributed. Papers in the tradition of urban and labor economics or industrial organization tend to use logit shocks, with normalized variance $\frac{\pi^2}{6}$ shifted by a factor $\frac{1}{\kappa^\theta)}$, which together with CD utility lead to the following indirect utility of worker θ in location i:

$$v_i^\theta(w) = \log A_i^\theta + \log w_i^\theta - \alpha^\theta \log r_i + \frac{1}{\kappa^\theta} \varepsilon_i^\theta(w)$$

Equivalently, papers in the tradition of trade and economic geography typically choose Frechet shocks for $\varepsilon_i^\theta(w)$ with scale parameter $\kappa^\theta > 1$ that enter utility in a multiplicatively separable way. In that case, the indirect utility of worker θ in location i is:

$$v_i^\theta(w) = \frac{A_i^\theta w_i^\theta}{r_i^{\alpha^\theta}} \varepsilon_i^\theta(w).$$

In either case, location choices in group θ can be summarized with λ_i^θ, the share of θ workers who choose location i:

$$\lambda_i^\theta = \frac{(\frac{A_i^\theta w_i^\theta}{r_i^{\alpha^\theta}})^{\kappa^\theta}}{\sum_{j=1}^{N}(\frac{A_j^\theta w_j^\theta}{r_j^{\alpha^\theta}})^{\kappa^\theta}} \tag{3.2}$$

The parameter κ^θ captures the elasticity of population shares with respect to amenity-adjusted real wages and is therefore a measure of mobility of group θ, which we allow to be group specific. Expected utility for a worker in group θ across locations is:

$$W^\theta = \delta^\theta \left[\sum_{k=1}^{N}(\frac{A_i^\theta w_i^\theta}{r_i^{\alpha^\theta}})^{\kappa^\theta}\right]^{\frac{1}{\kappa^\theta}} \tag{3.3}$$

where $\delta^\theta = \Gamma(\frac{\kappa^\theta - 1}{\kappa^\theta})$ and $\Gamma(\cdot)$ is the gamma function in the Frechet case.

Supply of goods, amenities, and housing

We first write down the labor demand side of the economy. In location i, output is produced by perfectly competitive firms. They combine skilled and unskilled labor, who are imperfect substitutes in production:

$$Y_i = \left[(z_i^U)^{\frac{1}{\rho}}(L_i^U)^{\frac{\rho-1}{\rho}} + (z_i^S)^{\frac{1}{\rho}}(L_i^S)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} \tag{3.4}$$

In the CES production function, $\rho \geq 1$ is the elasticity of substitution between skills and z_i^θ are location- and skill specific productivity shifters. The shifters can be in part exogenous and in part endogenous, reflecting externalities. We assume that, for $\theta = \{U, S\}$ and $\forall i$,

$$z_i^\theta = z^\theta(\overline{Z}_i, L_i^U, L_i^S) \tag{3.5}$$

where $\overline{Z}_i$ is the exogenous productivity component in city i. Local productivity spillovers are allowed here to depend not just on city size or density but also on its composition (L_i^U, L_i^S). Given equation, relative labor demand in location i is:

$$\log\left(\frac{L_i^S}{L_i^U}\right) = \log\left(\frac{z_i^S}{z_i^U}\right) - \rho\log\left(\frac{w_i^S}{w_i^U}\right) \tag{3.6}$$

Furthermore, competition across cities ensures that the unit cost of production in all cities is 1, the common price of the freely traded good:

$$\sum_\theta z_i^\theta(w_i^\theta)^{1-\rho} = 1, \forall i$$

Similar to productivity, amenities A_i^θ are assumed to be driven by both exogenous differences, and endogenous differences between cities, that is,

$$A_i^\theta = A^\theta(\overline{A}_i, L_i^U, L_i^S) \tag{3.7}$$

where A_i is the exogenous amenity component of city i. Endogenous amenities capture elements of quality of life that change when the size or composition of cities changes.

Finally, we assume that housing is supplied by atomistic absentee landowners and the aggregate housing supply function in city i is:

$$H_i = \overline{H}_i r_i^{\eta_i} \tag{3.8}$$

The housing supply elasticity η_i is allowed to be city specific.

A spatial equilibrium of this economy is a set of location choices $\{\lambda_i^\theta\}_{i,\theta}$ prices $\{w_i^\theta, r_i\}_{i,\theta}$ and amenities and productivity shifters $\{z_i^\theta, A_i^\theta\}_{i,\theta}$ such that workers and firms optimize, traded good firms make no profits, and markets clear. Since amenities and productivity shifters $\{z_i^\theta, A_i^\theta\}_{i,\theta}$ typically depend on the equilibrium distribution of economic activity, these local spillovers act as feedback loops that may amplify or dampen concentration and sorting.

We now discuss conditions under which sorting arises in the equilibrium. In using the term spatial sorting, we mean the fact that the skilled and unskilled groups make different location choices, i.e., there exist locations i and j such that, denoting $\Delta X = X_i - X_j$ for any variable X,

$$\Delta \log\left(\frac{L^S}{L^U}\right) \neq 0$$

Then relative spatial supply is combined with location choices, combining labor supply and labor demand:

$$\Delta \log\left(\frac{L^S}{L^U}\right) = \underbrace{\frac{\tilde{\kappa}^S}{\rho}\Delta \log\left(\frac{z^S}{z^U}\right)}_{} \equiv \Delta z + \underbrace{\tilde{\kappa}^S \Delta \log\left(\frac{A^S}{A^U}\right)}_{} \equiv \Delta A + \underbrace{\tilde{\kappa}^S(\alpha^U - \alpha^S)\Delta \log r}_{} \equiv \Delta\alpha + \underbrace{\frac{\tilde{\kappa}^S}{\kappa^U}(1 - \frac{\kappa^U}{\kappa^S})\Delta \log L^U}_{} \equiv \Delta\kappa \tag{3.9}$$

where we denote $\tilde{\kappa}^S = \frac{\kappa^S \rho}{\kappa^S + \rho}$. Conceptually, one can therefore distinguish four sources of sourcing in this framework: We say that soring is shaped by comparative advantage in production when $\Delta \log\left(\frac{z^S}{z^U}\right) \neq 0$, by amenities when $\Delta \log\left(\frac{A^S}{A^U}\right) \neq 0$ by housing price when $\alpha^S \neq \alpha^U$, and by heterogeneous across groups when $\kappa^U \neq \kappa^S$.

A first takeaway is that, when productivity is separable between a location shifter Z_i and nationwide group productivity z^θ so that $z_i^\theta = Z_i z^\theta$, the productivity advantage of a location is skill neutral and hence does not drive sorting directly. We now assume, in contrast, that some skill group has comparative advantage in production in some location over another so that $\Delta \log\left(\frac{z^S}{z^U}\right) \neq 0$. For simplicity, we assume that local productivity depends on population, which is the most classic way to parameterize agglomeration effects, but with a different intensity γ_P^θ for different skill groups, that is,

$$z_i^\theta = \overline{z}_i^\theta (L_i^U + L_i^S)^{\gamma_P^\theta} \tag{3.10}$$

In this expression, $\overline{z}_i^\theta$ are exogenous location-group productivity shifters. Equilibrium sorting is then pinned down by:

$$\Delta \log\left(\frac{L^S}{L^U}\right) = \frac{\tilde{\kappa}^S}{\rho}\Delta \log\left(\frac{\overline{z}^S}{\overline{z}^U}\right) + \frac{\tilde{\kappa}^S}{\rho}(\gamma_P^S - \gamma_P^U)\Delta \log L + \Delta_A + \Delta_\alpha + \Delta_\kappa$$

Changes in sorting due to productivity correspond to the first two terms on the right hand side of the above equation. First, such changes may occur because of changes in exogenous comparative advantage $\Delta \log\left(\frac{\overline{z}^S}{\overline{z}^U}\right)$. Second, changes in sorting may occur because of changes in relative city sizes $\Delta \log L$ or because of changes in relative agglomeration forces $\gamma_P^S - \gamma_P^U$.

We parameterize utility derived from amenities as a CD aggregator of a vectorof amenities $\{A_{ki}\}_k$ in location i, with skill group-specific preference parameters γ_{kA}^θ. This allow both skill groups to have different preferences over each city's amenity bundle:

$$A_i^\theta = \Pi_k (A_{ki})^{\gamma_{kA}^\theta} \tag{3.11}$$

We allow a component of each amenity in the amenity bundle to be endogenous. We model the endogenous component of amenity as responding to the skill ratio $\frac{L_i^S}{L_i^U}$ of that city; that is,

$$A_{ki} = \tilde{A}_{ki}(\frac{L_i^S}{L_i^U})^{\beta_k} \tag{3.12}$$

where $\tilde{A}_{ki}$ is the exogenous component of amenity k and β_k measures how elasticity the supply of amenity k is to the skill ratio. With these formulations, equilibrium sorting is:

$$\Delta \log\left(\frac{L^S}{L^U}\right) = \frac{\tilde{\kappa}^S}{1-\tilde{\kappa}^S(\tilde{\gamma}_A^S - \tilde{\gamma}_A^U)}\Delta \log\left(\frac{\tilde{A}^S}{\tilde{A}^U}\right) + \frac{1}{1-\tilde{\kappa}^S(\tilde{\gamma}_A^S - \tilde{\gamma}_A^U)}[\Delta_z + \Delta_\alpha + \Delta_\kappa]$$

where we denote $\tilde{\gamma}_A^\theta = \sum_k[\beta_k(\gamma_{kA}^\theta)]$ and $\frac{\tilde{A}_i^S}{\tilde{A}_i^U} = \Pi_k(\tilde{A}_{ki})^{\gamma_{kA}^S - \gamma_{kA}^U}$. First, amenities are a source of sorting in themselves only to the extent that the first term is nonzero. Second, the endogenous provision of amenities ($\beta_k \neq 0$) together with their heterogeneous valuation across skills ($\gamma_{kA}^S - \gamma_{kA}^U \neq 0$) serve only as an amplifier of other sorting forces.

To make clear the specific mechanisms at play here, we shut down the other sources of sorting by making the following assumptions:

Assumption 3.1. $\Delta \log\left(\frac{z^S}{z^U}\right) = 0, \Delta \log\left(\frac{A^S}{A^U}\right) = 0$ *and* $\kappa^U = \kappa^S$.

Under assumption 1, some cities may still be more productive or have hiigher amenities than others, but in a way that is skill neutral: there exist citywide shifters Z_i and A_i and nationwide group-specific shifters z^θ, A^θ such that $z_i^\theta = Z_i z^\theta$ and $A_i^\theta = A_i A^\theta$ for all location i. If housing is a necessary, then $\alpha^U - \alpha^S > 0$ and skilled workers are overrepresented in expensive cities. Given the housing supply equation, equilibrium housing prices are implicit solution to:

$$\frac{Z_i^{\frac{1+\kappa}{\rho-1}} A_i^\kappa}{\overline{H}_i} = r_i^{\eta_i}\left[\sum_\theta w^\theta f_i(r_i) r_i^{\kappa\alpha^\theta \frac{1-\rho}{\kappa+\rho} - 1}\right]^{-1} \tag{3.13}$$

where $f_i(r_i)$ captures that, in equilibrium, wages depend on rents. This function can be shown to be equal to 1 when skills are perfect substitutes, and it is a decreasing function of r_i otherwise. Given that equation, rents r_i increase with $\frac{Z_i^{\frac{1+\kappa}{\rho-1}} A_i^\kappa}{\overline{H}_i}$: More productive cities and cities with higher amenities are more expensive in equilibrium. We denote $\mathcal{R}_i(\cdot)$ as the corresponding solution to that equation. Turning to the implication of housing rents for skill sorting, we obtain:

$$\Delta \log\left(\frac{L^S}{L^U}\right) = (\alpha^U - \alpha^S)\Delta \log \mathcal{R}(\frac{Z^{\frac{1+\kappa}{\rho-1}} A^\kappa}{\overline{H}}) \tag{3.14}$$

and high skill sorting into more productive and/or more attractive locations. First, when housing expenditure shares different across groups, Hicks-neutral city advantage is enough to drive sorting through its impact on housing prices. Second, the role of housing in mediating spatial sorting forces is stronger, all else equal, in locations with more inelastic housing supply (lower η). Inelastic supply directly leads to steeper $\mathcal{R}(\cdot)$, hence to a steeper response of housing price to productivity and amenities, and in turn to a steeper response of the skill ratio.

Heterogeneous migration elasticity

Finally, a last possible driver of sorting arises when $\kappa^U \neq \kappa^S$. Empirical findings find that higher-skill workers are more mobile than lower-skill workers, so, $\kappa^S > \kappa^U$.

Assumption 3.2. $\Delta \log\left(\frac{z^S}{z^U}\right) = 0, \Delta \log\left(\frac{A^S}{A^U}\right) = 0$, *and* $\alpha^U = \alpha^S$.

It is easy to see that:

$$\Delta \log\left(L^S\right) = \left[\underbrace{1 + \frac{\rho}{\kappa^S + \rho}\left(\frac{\kappa^S}{\kappa^U} - 1\right) > 0}\right] \Delta \log L^U \tag{3.15}$$

Under Assumption 2, there is no skill-biased amenity or productivity advantage of places. Still, the high-skill population increases faster than the low-skill population in attractive cities because members of the former group are more sensitive to city characteristics.

Urban skill premium and sorting

We now turn to considering how they shape the distribution of the skill wage premium in the cross section of cities, in equilibrium. Solving out for the equilibrium skill premium and its variation over space leads to

$$\Delta \log\left(\frac{w^S}{w^U}\right) = \frac{1}{\kappa^S}\Delta_z - \frac{1}{\rho}\Delta_A - \frac{1}{\rho}\Delta_\alpha - \frac{1}{\rho}\Delta_\kappa \tag{3.16}$$

Comparing this expression with the one that summarizes skill sorting we have:

$$\Delta \log\left(\frac{L^S}{L^U}\right) = \Delta_z + \Delta_A + \Delta_\alpha + \Delta_\kappa$$