

City Location Choice and Household Productivity

Tan Sein Jone

1 Introduction

Cities capture a disproportionate amount of household location choices in the world. In 2018, 55% of the world's population lived in urban areas, and this is expected to increase to 68% by 2050. This is because of the benefits cities provide, such as access to jobs, amenities, and social networks. These benefits are often thought of as universal to all households, regardless of their occupation specific skills. However, the benefits of every city is clearly not the same for every household. An auto worker from Detroit will not view New York as an equivalent substitute for a city to live in as a finance worker from New York. Given the billions of dollars spent on place based policies, it is important for us to separate out this dimension of comparative advantage, where a city is able to provide higher levels of productivity for households working in specific occupations.

A city's comparative advantage exists in three dimensions. First, there is a city specific advantage which makes all occupations more productive in that city. This is due to the city's size, more specifically it's population which results in better amenities, more jobs, and a larger social network. Second, there is an occupation specific advantage which is universal across all cities. Some jobs are simply more productive than others and cities that specialize in that job will be observed as having a higher overall choice shares of households. Think cities that specialize in highly attractive occupations such as professional services. Finally, there is a city-occupation specific advantage. A city with a comparative advantage in this dimension may have specific policies that make it more attractive for a specific occupation such as tax breaks for mechanics in Chicago. One crucial distinction we need to make is between occupations and sectors. Sectors are broad categories of jobs such as manufacturing, services, and agriculture. Occupations are specific jobs such as mechanic and clerk. Cities can specialize in occupations by specializing in sectors that most intensely use that occupation.

In this paper, we propose a model of city location choices based on a class of models in the quantitative trade literature first proposed by Eaton and Kortum (2002). These models treat productivity as random draws from Fréchet distributions, allowing for a non zero production of goods which will help us to account for small cities. Redding (2016) applies this model to an urban setting by loading city heterogeneity onto

amenities, capturing the first dimension of comparative advantage. Both these models however rely on the independence assumption which treats that goods and cities are perfect substitutes for each other. Lind and Ramondo (2023) break this assumption by introducing a nested CES utility structure which allows for correlated draws of productivity within nests across countries. We will apply this structure to the urban setting by treating occupations as independent nests. This will allow us to capture substitution patterns where cities that have a similar occupational structure will be treated as better substitutes for each other. By doing this, we are able to capture the second and third dimensions of comparative advantage.

We start in the first section by adapting the model of Lind and Ramondo (2023) to the urban setting. Unlike their original paper where the nests are implied, we will explicitly model the nests as occupations. A few implications will result from this specification. First, we will be able to observe both overall city shares and city shares specific to each occupation. As well as city occupation compositions which can be used as a measure of similarity between cities. Second, using this data, we will derive elasticities for all three dimensions of comparative advantage, that being the city specific, occupation specific and city-occupation specific elasticities. While the last elasticity will serve as a sanity check for the direction of substitution. City and occupation elasticities will be used to verify our claim of unbalanced substitution patterns that favor similar cities. In other words, shocks to cities of a certain occupational composition will have larger effects of cities with similar compositions due to correlations in productivity draws.

In the second section, we will use data from the American Community Survey which has on the metropolitan area and occupational level to present some stylized facts about cities. We will then be presenting our estimation equations that will be used to recover our parameters. Once the model is calibrated, we will use it to run counterfactuals and simulate the effects of a policy at both the city and occupation level. This is where it is possible for us to evaluate the effectiveness of place based policies and also where we will make statements about how these policies should be designed to maximize their effectiveness.

2 Model

Consider a closed economy consisting of N cities and K occupations. Each city c employs households in occupation k to produce output Y_{ck} . We assume no trade costs between cities and that the price of the good produced is freely traded and priced under perfect competition. The price index of city c is given by Φ_c and the wage of a household living in city c working in occupation k is given by $w_{ck} = \Phi_c Z_{ck}$ where Z_{ck} is the productivity of the household. We assume that when a household chooses to work in a particular

occupation, they will be randomly allocated to a sector within a given city. As in Eaton and Kortum (2002), productivity is a random variable drawn from a Fréchet distribution and is dependent on both the city and the occupation. Output produced by a specific occupation is also assumed to be produced by a random sector.

A household's utility is purely determined by wages and is given by the maximum wage across all cities. When households choose a city, they will simultaneously choose an occupation. We will also assume that households are perfectly mobile across cities and that there are no fixed costs to moving.

2.1 Fréchet Distribution

$$P[Z_1(\nu) < z, \dots, Z_N(\nu) < z] = \exp \left\{ - \sum_k^K \left[\sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}} \right]^{1-\rho_k} \right\} \quad (1)$$

We will assume that productivity is distributed max stable multivariate Fréchet with T_{ck} being the scale parameter for city c and occupation k . $\theta > 0$ is the shape parameter that determines the dispersion of draws across the distribution and ρ_k is the occupation specific correlation parameter. As in Lind and Ramondo (2023), we adopt a cross nested CES structure with a correlational structure within independent occupational nests. This has the characteristic of distributions not purely determined by θ but also by the correlation parameter. This is because the correlation parameter determines the degree of correlation between draws of productivity across cities. When $\rho_k = 0$, the distribution is purely determined by θ and the draws are independent. When $\rho_k = 1$, the distribution is purely determined by the correlation parameter and the draws are perfectly correlated.

In order to identify the three dimensions of comparative advantage, we will be separating out the scale parameter into three components where $T_{ck} = T_c T_k t_{ck}$. T_c is the city specific scale parameter which captures the attractiveness of a city, shifting the distribution of all productivity draws within that city upwards. T_k is the occupation specific scale parameter which is common across all cities, this captures the effect that any occupation has across all cities. t_{ck} is the city-occupation specific scale parameter which captures the effect of a city on a specific occupation. This will give us the final dimension of comparative advantage.

$$Z = \max_c \left\{ \frac{Z_c}{\Phi_c} \right\} \quad (2)$$

A household's schedule of productivity is characterized by a vector of draws from different Fréchet distributions for each city occupation pair. The realized productivity a household has in city c is the occupation that maximizes the productivity has in the city. Unlike sequential games where households might pick a city before picking an occupation, we assume that households pick both simultaneously. A way to contextualize this is to think of households as already having an ideal occupation in mind when picking a city, hence $Z_c = \max_k \{Z_{ck}\}$. This productivity is then scaled by the price index of a city to reflect how high prices reduces the purchasing power of the household within that city, hence making it less attractive.

2.2 Correlation Function

$$G(Z_1^{-\theta}, \dots, Z_N^{-\theta}) = \sum_k^K T_k \left[\sum_c^N (t_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}} \right]^{1-\rho_k} \quad (3)$$

We first define $Z_c^{-\theta} = (\gamma \Phi_c)^{-\theta} T_c$ and $\gamma = \Gamma(\frac{\theta-1}{\theta})$. Given our definition of $Z_c^{-\theta}$, T_c will be absorbed within $Z_c^{-\theta}$, giving us the correlation function $G(Z_1^{-\theta}, \dots, Z_N^{-\theta})$ as defined in equation 3.

$$\pi_c = \frac{Z_c^{-\theta} G_c(Z_1^{-\theta}, \dots, Z_N^{-\theta})}{G(Z_1^{-\theta}, \dots, Z_N^{-\theta})}$$

Where $G_c(Z_1^{-\theta}, \dots, Z_N^{-\theta}) = \partial G(Z_1^{-\theta}, \dots, Z_N^{-\theta}) / \partial Z_c^{-\theta}$. The expression above gives us the city specific choice shares of households across all occupations. In order to evaluate the expression, we make the simplifying assumption that the correlation parameter is the same across all occupations $\rho_k = \rho$.

$$\lambda_k = \sum_c^N (t_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho}}$$

To simplify the final expression of city shares, we introduce the above expression which measures the appeal of occupation k . Given that the expression does not include T_k , it does not capture the attractiveness of occupation k . Instead the expression captures the extent to which a given occupation k exhibits city-specific productivities which are correlated with aggregate city-level attractiveness.

$$\pi_c = \sum_k^K \frac{(Z_c^{-\theta} t_{ck})^{\frac{1}{1-\rho}}}{\lambda_k} \frac{T_k \lambda_k^{1-\rho}}{\sum_k^K T_k \lambda_k^{1-\rho}}$$

When we solve for city shares, we arrive at the above expression. To further simplify it, we first have to consider occupation specific choice shares.

$$\pi_{ck} = \frac{Z_c^{-\theta} G_c^k(Z_1^{-\theta}, \dots, Z_N^{-\theta})}{G^k(Z_1^{-\theta}, \dots, Z_N^{-\theta})}$$

To obtain occupation specific choice shares, we can evaluate the derivative for city shares at the occupation level. More specifically, rather than $G(Z_1^{-\theta}, \dots, Z_N^{-\theta})$, we will be evaluating $G^k(Z_1^{-\theta}, \dots, Z_N^{-\theta}) = \sum_c^N (t_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}}$. This will give us the following:

$$\pi_{ck} = \frac{(t_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho}}}{\lambda_k} \tag{4}$$

Moving forward, it will be convenient for us to provide the following definitions. First, we define ϕ_{ck} as the comparative advantage of city c in occupation k . This can be interpreted as an occupation's within city choice shares. As such ϕ_{ck} will be between 0 and 1. We define this formally as:

$$\phi_{ck} \equiv \frac{t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}}$$

Second, we define an occupation specific choice share which takes into account both the size of each occupa-

tion, as measured by productivity T_k , and the extent to which individual city-occupation productivities are correlated with overall city attractiveness. This is defined as the following expression:

$$\omega_k \equiv \frac{T_k \lambda_k^{1-\rho}}{\sum_k T_k \lambda_k^{1-\rho}}$$

From the perspective of discrete choice demand models, ω_k can be interpreted as the "market share" of occupation k . This is because it captures the extent to which the productivity of occupation k is correlated with overall city attractiveness. The expression is normalized by the sum of all ω_k to ensure that the sum of all market shares is equal to 1.

$$\frac{\pi_{ck}}{\pi_c} \equiv \frac{\phi_{ck}}{\omega_k} \quad (5)$$

The expression above tells us the extent to which a given occupation k is concentrated in city c above and beyond the predicted concentration, given the size of the city π_c is equivalent to the within-city choice share in occupation k divided by the aggregate choice share of that occupation.

$$\pi_c = \sum_k^K \pi_{ck} \omega_k \quad (6)$$

By using our definitions, we can further simplify the expression for city shares which is simply now the sum of all occupation specific choice shares weighted by the market share of each occupation.

2.3 Elasticities

We now turn to the city-occupation derivatives: city c 's "own price" elasticity and "cross price" elasticity with respect to another city c' . This derivative will be approximated by assuming that each city is small, and therefore $\partial \ln \lambda_k / \partial \ln t_{ck} \approx 0$. That is, the aggregate "price index" associated with location choice in occupation k is not affected by a marginal change in the productivity of a given city-occupation pair.

References

- EATON, J. AND S. S. KORTUM (2002): “Technology, Geography, and Trade,” *Econometrica*, 70, 1741–1779.
- LIND, N. AND N. RAMONDO (2023): “Trade with Correlation,” *American Economic Review*, 113, 317–53.
- REDDING, S. J. (2016): “Goods trade, factor mobility and welfare,” *Journal of International Economics*, 101, 148–167.