City Location Choice and Household Productivity

Tan Sein Jone

1   Introduction

The location choice of households is an important determinant of the spatial distribution of economic activity. The choice of location is influenced by a variety of factors such as the availability of jobs, the cost of living, and the quality of public services. In this paper, I develop a model of city location choice that incorporates the productivity of households as an important determinant of location choice. I show that the productivity of households is an important determinant of the spatial distribution of economic activity and that it can help explain the observed patterns of urbanization.

The model is based on the idea that households are more productive in cities because of the agglomeration of economic activity. In particular, households are more productive in certain cities due to the presence of occupations that are complementary to their skills. For example, a household with skills in finance may be more productive in a city with a large financial sector. Similarly, a household with skills in technology may be more productive in a city with a large technology sector. The model shows that households choose to locate in cities where they are most productive, leading to the agglomeration of economic activity in certain cities.

This leads to two key outcomes. First, larger cities will tend to have higher levels of economic activity because they are able to attract more households. Second, cities that specialize in specific occupations will tend to be more attratcive to households with those skills. This leads to the agglomeration of economic activity in cities that specialize in specific occupations.

We achieve this by developing a model of city location choice that is based on Quantitative Spacial Models develoepd by Redding (2016) which loads city heterogeneity onto a city ammenity term in order to avoid a trivial solution of all hosueholds choosing to live in once city. This model however comes comes with the independence of irrelevent alternatives (IIA) assumption, which makes cities perfect substitutes for each other. We relax this assumption by using a nested CES structure with correlation proposed by Lind and Ramondo (2023) to allow for correlated draws of productivity between cities, making cities imperfect substitutes for each other.

Both of these models are based off the class of discrete choice models proposed by Eaton and Kortum (2002) which determines production shares of different goods by taking probabalistic draws of productivity of different countries. Rather than having these productivities determine prices, we will draaw productivities for households which determine wages and hence utility across different cities.

1.1 Literature Review

2 Model of Production

$$Y_c(\nu) = \prod_k Q_{ck}(\nu)^{\omega_{ck}} \tag{1}$$

Consider a closed economy consisting of N cities and a continuim of hosuehold types $\nu \in [0, 1]$. Each city c employs hosueholds of type ν in occupation k to produce output $Y_{ck}(\nu)$ using a Cobb-Douglas production function with technology $Z_{ck}(\nu)$. We assume no trade costs between cities, and that the price of goodproduced by occupation k is freely traded and priced under perfect competition.

The price index of city c is Φ_c. The wage of a hosuehold in occupation k in city c is $w_{ck}(\nu) = \Phi_c Z_{ck}(\nu)$. We asusme that when a hosuehold chooses to work in a particular occupation k, they will be randomly allocated to a sector within a given city. As in EK, productivity is a random variable drawn from a max stable mutivariate Fréchet distribution and is dependent on both the city and the occupation. Output produced by a specific occuaption is also assumed to be produced by a random sector.

$$U(\nu) = \max_c \left\{ \Phi_c Z_{ck}(\nu) \right\} \tag{2}$$

A hosuehold's utility is purely determined by wages and is given by the maximum wage across all cities. We assume that hosueholds are risk neutral and hence maximize expected utility. When choosing a prticular city to maximize utility, households will simultaneously choose an occupation to work in. We assume that hosueholds are perfectly mobile and can move to any city at no cost.

3 Max Stable Multivariate Fréchet Distribution

$$P[Z_{1k}(\nu) < z, \ldots, Z_{Nk}(\nu) < z] = \exp\left\{\left[-\sum_{c}^{N}(T_{ck}H_k(\nu)z^{-\theta})^{\frac{1}{1-\rho_k}}\right]^{1-\rho_k}\right\} \tag{3}$$

We assume that productivity is distributed max stable multivariate Fréchet where T_{ck} is the scale parameter for city c and occupation k. $H_k(\nu)$ is the occupation specific productivity shifter for household of type ν in occupation k. The heterogeneity of this shifter reflects the difference in human capital across different hosuehold types. $\theta > 0$ is the shape parameter, characterizing the tail behaviour of the distribution. ρ_k is the occupation specific correlation parameter, which dictates the extent to which productivity draws are correlated acorss cities for occupation k.

$$T_{ck} = T_c T_k t_{ck} \tag{4}$$

Looking at the scale parameter T_{ck}, we can separate it into three components. T_c is the city specific scale parameter which captures the attractiveness of a city, shifting the distribution for all productivity draws within that city. T_k is the occupation specific scale parameter which is common across all cities, this captures the aggregate effect that occupation k has on productivity. An example shock to this parameter is the effect new computers have on an occupation that heavily relies on computation. t_{ck} is the city-occupation specific scale paramter which captures the idiosyncratic effect of city c on occupation k. The effect on new york on finance is an example of this.

$$Z_c(\nu) = \max_{k}\{Z_{ck}(\nu)\} \tag{5}$$

A household's schedule of productivities is characterized by a vector of draws from diferent fréchet distributions for each city occupaiton pair. The realized productivity of a household of type ν has in city c is the occupation that maximizes the productivity that particular hosuehold has in the city.

$$P[Z_1(\nu) < z, \ldots, Z_N(\nu) < z] = \exp\left\{-\sum_k \left[\sum_c^N (T_c H(\nu) z^{-\theta})^{\frac{1}{1-\rho_k}}\right]^{1-\rho_k}\right\} \tag{6}$$

Unlike sequential games where households might pick a city before picking an occupation, we assume that households pick both simultaneously. This is because thhe schedule of productivity already determines the ideal occupation of a household in a city. From here, we can obatin the joint distribution of all productivities being less than some value z for all cities and occupations.

$$Z(\nu) = \max_c \left\{\frac{Z_c(\nu)}{\Phi_c}\right\} \tag{7}$$

The joint probability takes on the form of a max stable multivariate Fréchet distribution with a cross-nested CES correlation structure as proposed by Lind and Ramondo (2023). This distribution is a generalization of the Fréchet distribution to the multivariate case, and is used to model the joint distribution of extreme values. This is similar to the GEV distribution, but with the added flexibility of allowing for correlation between the draws of different cities across occupations. A household of type ν has the realized productivity that is hence characterized by the maximum of all productivity draws across all cities and occupations scaled by the inverse of that city's price index Φ_c.

3.1 Correlation Function

$$P[Z_1 < z, \ldots, Z_N < z] = \exp\left\{-\sum_k \left[\sum_c^N (T_c z^{-\theta})^{\frac{1}{1-\rho_k}}\right]^{1-\rho_k}\right\} \tag{8}$$

When we intergrate over all hosuehold types for a particular hosuehold shifter, we get the average productivity shifter for that occupation which we assume to be 1 for all occupations $\int_0^1 H_k(\nu) d\nu = 1$. With the case of a corss-nested CES function, we can approximate a correlation function. Assuming that productivity is distributed max-stable multivariate Fréchet, with scale parameter T_{ck} shape parameter θ and correlation parameters ρ_k.

$$G(Z_1^{-\theta}, \ldots, Z_N^{-\theta}) = \sum_k \left[\sum_c (T_{ck} z^{-\theta})^{\frac{1}{1-\rho_k}} \right]^{1-\rho_k} \tag{9}$$

$$G_c(Z_1^{-\theta}, \ldots, Z_N^{-\theta}) = \frac{\partial G(Z_1^{-\theta}, \ldots, Z_N^{-\theta})}{\partial Z_c^{-\theta}} \tag{10}$$

$$\gamma = \Gamma\left(\frac{\theta - 1}{\theta}\right) \tag{11}$$

$$Z_c^{-\theta} = (\gamma \Phi_c)^{-\theta} \tag{12}$$

As in Lind and Ramondo (2023), if we consider the case of one nest, ρ will be perfectly uncorrelated and be equal to 0. This means that our CNES structure will collapse down to the regular CES case.