City Location Choice and Household Productivity

Tan Sein Jone

1 Model of Production

$$Y_{ck}(\nu) = Z_{ck}(\nu) \tag{1}$$

Consider a closed economy consisting of N cities and a continnium of households $\nu \in [0,1]$. Each city c employs hosuehólds of type ν in sector k to produce output $Y_{ck}(\nu)$. The output of sector k in city c is given by the productivity of the household in that sector and city. The productivity of a household in sector k in city c is given by $Z_{ck}(\nu)$.

We assume no trade costs between cities, and that the price of the good produced by sector k is freely traded and priced under perfect competition. The price index of city c is Φ_c. The wage of a worker in sector k in city c is $w_{ck}(\nu) = \Phi_c Z_{ck}(\nu)$. As in EK, productivity is a random variable drawn from a max stable multivariate Fréchet distribution and is dependent on both the city and occupation. Output by a specifc occupation is assumed to be produced by a random sector.

2 Max Stable Multivariate Fréchet Distribution

We assume that the joint distribution of productivity across cities is given by

$$P[Z^{\star}_{1k}(\nu) < z, \ldots, Z^{\star}_{Nk}(\nu) < z] = exp[-\sum_{c}^{N}(T^{\star}_{ck}H_k(\nu)z^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k} \tag{2}$$

Where $Tck^{\star}$ is the scale parameter for city c and occupation k, this represents a city's absolute advantage for occupation k. $H_k(\nu)$ is the occupation specific productivity shifter for household ν in occupation k, embodying the household's inate productivity in that occupation. $\theta > 0$ is the shape parameter, characaterizing the tail behavior of the distribution. ρ_k is the occupation specific correlation parameter, which dictates the extent to which productivity draws are correlated across cities for occupation k.

A household's schedule of productivities is characterized by a vector of draws from the Fréchet distribution for each occupation k in each city c. The realized productivity a hosuehold of type ν has in city c is however the occupation that maximizes the producticity that particular household has in the city. This is given by the following:

$$Z_c(\nu) = \max_k \{Z^{\star}_{ck}(\nu)\} \tag{3}$$

Unlike sequential games where households might pick a city before picking an occupation, this schedule of productivity already determines the ideal occupation for a household in a city. The joint probability of all productivities being less than some value z for all cities is then given by the following:

$$P[Z_1(\nu) < z, \ldots, Z_N(\nu) < z] = exp\{-\sum_k [\sum_c^N (T^{\star}_{ck} H_k(\nu) z^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k}\} \tag{4}$$

The joint probability above is a max-stable multivariate Fréchet distribution with a cross-nested CES correlation function. 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 simillar to the GEV distribution, but with the added feature of a correlation function that allows for the dependence of extreme values across occupations.

A houehold of type ν has realised productivity that is hence characterised by the maxmimum productivity draw across all cities scaled by the inverse of that city's price index Φ_o and is gievn by the following:

$$Z(\nu) = \max_{c=1,\ldots,N} \left\{ \frac{Z_c(\nu)}{\Phi_c} \right\} \tag{5}$$

2.1 Correlation Function

With the case of a cross-nested CES function, we can approximate a correlation funciton. Assuming that productivity is distributed max-stable multivariate Fréchet, with scale parameter $T^{\star}_{ck}$, shape parameter θ, and correlation parameter ρ_k, we can assume the following correlation function:

$$G(Z_1^{-\theta}, \ldots, Z_N^{-\theta}) = \sum_k [\sum_c^N (T^{\star}_{ck} H_k Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k} \tag{6}$$

$$G_c(Z_1^{-\theta}, \ldots, Z_N^{-\theta}) = \frac{\partial G(Z_1^{-\theta}, \ldots, Z_N^{-\theta})}{\partial Z_c^{-\theta}} \tag{7}$$

$$\gamma = \Gamma(\frac{\theta - 1}{\theta}) \tag{8}$$

Where G_c is the derivative of the correlation function with respect to city c. This correlation function is a generalization of the CES correlation function to the multivariate case, and is used to model the dependence of extreme values across cities. The Γ function which takes in the ratio of $\theta - 1$ to θ gives us the substitutability between cities. Productivity is related to γ by the following funciton:

$$Z_c^{-\theta} = (\gamma \Phi_c)^{-\theta} T^{\star}_c \tag{9}$$

Where $T^{\star}_c$ is the city's absolute advantage that applies to all occupations and is given by $T^{\star}_c = \sum_k T^{\star}_{ck} H_k$.

2.2 Choice Shares

Under perfect competition, wages are equal to marginal revenue product and is a direct function of productivity. As established in equation (5), the productivity for a worker ν in city c is given by the maximum productivity draw across all occupations. All we have to do is intergrate over all houshold types to get the productivity index for city c:

$$Z_c = \max_k \left\{ \int_0^1 Z^{\star}_{ck}(\nu) d\nu \right\} \tag{10}$$

If productivity is distributed max-stable multivariate Fréchet and a continiously differentiable correlation function, then city c's choice shares is given by the following:

$$\pi_c = \frac{Z_c^{-\theta} G_c(Z_1^{-\theta}, \ldots, Z_N^{-\theta})}{G(Z_1^{-\theta}, \ldots, Z_N^{-\theta})} \tag{11}$$

First, the share of location choices for households has the same form as choice probabilities in GEV discrete choice models, with $Z_c^{-\theta}$ replacing choice specific utility. Second, as in EK, the share of location choices of city c equals the probability that a worker is most productive in city c. Finally, the location choice share in each city is determined by the ratio of the expected productivity in that city to the expected productivity in all cities.

$$\pi_c = \frac{T^{\star}_c Z_c^{-\theta}}{\sum_c^N T^{\star}_c Z_c^{-\theta}} \tag{12}$$

3 Simulations

Assume now that we have two occupations trades and services, with two corresponding household types that specialise in each occupation. We will also assume three cities, Detriot, Chicago and New York with Detriot and Chicago specializing in trades and New York specializing in services. We will also assume that the correlation parameter ρ_k is 0.5 for both occupations and the shape parameter θ is 2. This will give us the folliwing scale parameters for cities and hosuehold types:

	Detriot	Chicago	New York
Trades	7	8	3
Services	2	1	8

Table 1: City Scale Parameters $T^{\star}_{ck}$

	ν_1	ν_2
Trades	7	4
Services	3	8

Table 2: Houshold Scale Parameter $H_k(\nu)$

These shifters will determine the Frechet distributions from which city and occupation specific productivities are drawn for each household type. For context, this is the pdf for the Frechet distribution:

$$P[Z^{\star}_{ck}(\nu) < z] = exp(T^{\star}_{ck} H_k(\nu) z^{-\theta})^{\frac{1}{1-\rho_k}} \tag{13}$$

For each household, the draws will produce a $K \times N$ matrix which reflects that hosuehold's schedule of productivities across all cities and occupations. The realized productivity for each household is then the maximum productivity draw across all cities and occupations. Location choice is hence the column which contains that realised productivity.

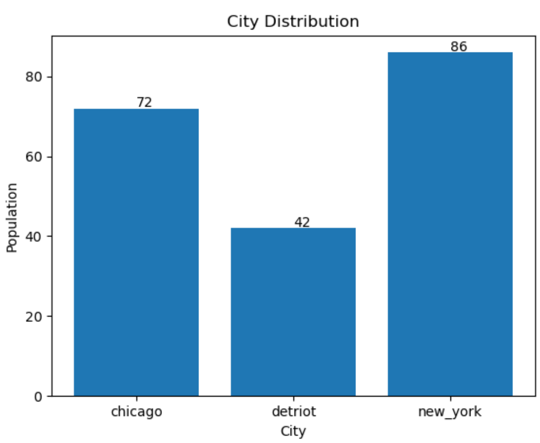

Figure 1: Initial Distribution of Populations

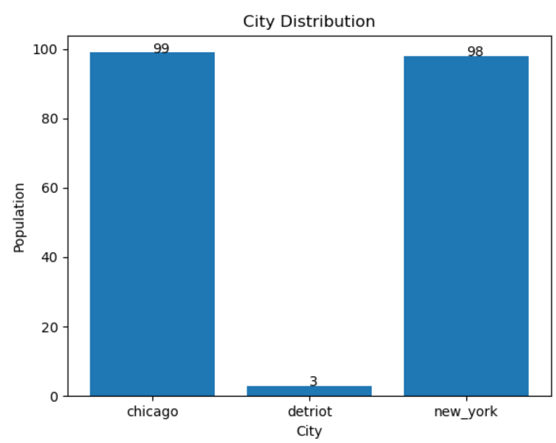

Figure 2: Post Detriot Shock

We now draw 100 distibutions of each hosuehold types for every city and occupation, simulating the dtstribution of location choices across cities. As shown in figure 1, the city with the highest population is new york, with the majority of service focused households choosing to locate in the city. Detriot and Chicago have a more even distribution of trades focused households. There is however a preference for trades focused households to locate in Chicago due to the higher specialization in trades as indicated by the scale parameter.

We then shock Detriot's productivities, reducing it by 50% and observe the change in location choices. As shown in figure 2, there is a redistibution in location choices with the majority of trades focused households now choosing to locate in Chicago. Some households have also chosen to locate in New York, due to higher productivity draws in that city.

4 Proofs

4.1 Proof of Equation (12)

$$
\begin{aligned}
G_c(Z_1^{-\theta}, \ldots, Z_N^{-\theta}) &= \frac{\partial G(Z_1^{-\theta}, \ldots, Z_N^{-\theta})}{\partial Z_c^{-\theta}} \\
&= \sum_k [(T^\star_{ck} H_k)^{\frac{1}{1-\rho_k}} (Z_c^{-\theta})^{\frac{1}{1-\rho_k}-1}]^{1-\rho_k} \\
&= Z_c^{-\theta} \sum_k [(T^\star_{ck} H_k Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k} \\
&= Z_c^{-\theta} \sum_k T^\star_{ck} H_k Z_c^{-\theta} \\
Z_c^{-\theta} G_c(Z_1^{-\theta}, \ldots, Z_N^{-\theta}) &= \sum_k T^\star_{ck} H_k Z_c^{-\theta}
\end{aligned}
$$

$$
\begin{aligned}
\frac{Z_c^{-\theta} G(Z_1^{-\theta}, \ldots, Z_N^{-\theta})}{G(Z_1^{-\theta}, \ldots, Z_N^{-\theta})} &= \frac{\sum_k T^\star_{ck} H_k Z_c^{-\theta}}{\sum_k [\sum_c^N (T^\star_{ck} H_k Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k}} \\
&= \frac{\sum_k T^\star_{ck} H_k Z_c^{-\theta}}{\sum_c^N \sum_k (T^\star_{ck} H_k Z_c^{-\theta})^{\frac{1}{1-\rho_k} \times (1-\rho_k)}} \\
&= \frac{T^\star_c H^\star Z_c^{-\theta}}{\sum_c^N T^\star_c H^\star Z_c^{-\theta}} \\
&= \frac{T^\star_c Z_c^{-\theta}}{\sum_c^N T^\star_c Z_c^{-\theta}}
\end{aligned}
$$