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 attractive 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 developed by Redding (2016) which loads city heterogeneity onto a city amenity term in order to avoid a trivial solution of all households choosing to live in once city. This model however comes comes with the independence of irrelevant 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 draw 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 continuum of household types $\nu \in [0, 1]$. Each city c employs hosuehelds 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.

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

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 different 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.

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

Unlike sequential games where households might pick a city before picking an occupation, we assume that households pick both simultaneously. This is because the schedule of productivity already determines the ideal occupation of a household in a city. From here, we can obtain 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

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

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.

$$G(Z_1^{-\theta}, \dots, Z_N^{-\theta}) = \sum_c^N T_{ck} z^{-\theta} \tag{9}$$

$$G(Z_1^{-\theta}, \dots, Z_N^{-\theta}) = \sum_k \left[\sum_c^N (T_{ck} z^{-\theta})^{\frac{1}{1-\rho_k}} \right]^{1-\rho_k} \tag{10}$$

Where $Z_c^{-\theta} = (\gamma \Phi_c)^{-\theta}$ and $\gamma = \Gamma\left(\frac{\theta-1}{\theta}\right)$. $\Gamma(\cdot)$ is the gamma function. Expression 9 shows the standard Fréchet distribution with independent productivity in which the strength of comparative advantage is solely governed by th shape parameter θ. Once we abandon this assumption as in Lind and Ramondo (2023), we can obtain a correlation function that allows for correlated draws of productivity across cities as in expression 10. When here exists only one occupation, the correlation parameter will be 0 and expression 10 will reduce to expression 9.

3.2 Choice Shares

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

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

$$\pi_c = (Z_c^{-\theta})^{\frac{1}{1-\rho}} \left[\frac{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}{\sum_k \lambda_k^{1-\rho}} \right] \tag{12}$$

We define $\lambda_k = \sum_c^N T_{ck}^{\frac{1}{1-\rho}} (Z_c^{-\theta})^{\frac{1}{1-\rho}}$. $Z_c^{-\theta}$ simply acts as a city level shifter that scales the productivity of all hosueholds in that city. The second term in brackets is a measure of a city c's relative strength across occupations. As any individual T_{ck} increases, the share of that that city also increases. Conversely, as the strength of other cities in a given occupation λ_k increases, the share of that city decreases. This is because the city is less attractive relative to other cities in that occupation.

$$\pi_{ck} = \frac{Z_c^{-\theta} G_c^k(Z_1^{-\theta}, \dots, Z_N^{-\theta})}{G^k(Z_1^{-\theta}, \dots, Z_N^{-\theta})} \tag{13}$$

$$\pi_{ck} = \frac{(T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}}}{\sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}}} \tag{14}$$

$G^k(Z_1^{-\theta}, \dots, Z_N^{-\theta}) = [\sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k}$ is the occupation specific function that allows us to evaluate the city specific share of occupations. In this case, we can can abandon the simplifying assumption of $\rho_k = \rho$ and obtain the expression for the choice shares of cities within a gievn occupation. This share is simply teh share of productivities relative to all other city's productivities within a given occupation. Notice that as that occupation becomes more correlated across cities, the relative strength of that city starts to matter more.

3.3 Cross and Within City Elasticities

$$\epsilon_{cck} = \frac{\partial \ln \pi_{ck}}{\partial \ln T_{ck}} \approx \frac{1}{1-\rho} \phi_{ck} > 0 \tag{15}$$

$$\phi_{ck} = \frac{T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}} \tag{16}$$

We approximate this derivative by assuming that each city is small and therefore $\partial \ln \lambda_k / \partial \ln T_{ck} \approx 0$. This allows us to obtain the expression for the elasticity of a city's own price. Notice that a city's own price elasticity is positive, and increasing in the relavance of a given occupation for a given city. That is, as T_{ck} increases, the elasticity increases. If a given city is particularly productive in occupation k, then changes in T_{ck} will have an outsized effect on that city's overall choice share. Put it another way: becoming more productive in occupations for which you are already everyone's last choice will not do much to alter the overall choice share of your city. ϕ_{ck} captures the extent to which this city is relatively more productive at occupation k than other cities compared to all other cities in every occupation.

$$\epsilon_{cc'k} = \frac{\partial \ln \pi_{ck}}{\partial \ln T_{c'k}} \approx -\pi_{c'}\phi_{c'k}\left[1 + \left(\frac{\rho}{1-\rho}\right)\left(\frac{\phi_{ck}}{\omega_k}\right)\right] < 0 \tag{17}$$

$$\omega_k = \frac{\lambda_k^{1-\rho}}{\sum_k \lambda_k^{1-\rho}} \tag{18}$$

We can approximate the cross-price elasticity in a similar fashion. The first term in this derivative simply captures the "CES" element: if city c' is large and has a particularly string comparative advantage in occupation k, then technology shocks in this city-occupation pair will have large effects elsewhere. This is captured by the size $\pi_{c'k}$ and the comparative advantage $\phi_{c'k}$.

The second term captures the unbalanced substitution patterns associated with correlated choice probabilities. Notice first that if $\rho = 0$, this model is simple CES in substitution patterns. As $\rho \to 1$, the correlated choices dominate the CES element more and more. City c will be particularly effected by shocks to city c' in occupation k is city c is also productive in occupation k. In fact, it is the multiple of $\phi_{c'k}\phi_{ck}$ that dictates the strength of substitution.

But why is this substitution dampened by weight ω_k? Notice that is occupation k is very attractive and $\omega_k \to 1$, this implies that there are many attractive cities for households productive in occupation k. That is, we cannot determine the cross-price elasticity without considering the environment in which both c and c' exist, and the strength of substitution between them is strongest when they are both productive in k and no other cities are even close, implying a small ω_k.

$$\epsilon_{ckck} = \frac{\partial \ln \pi_{ck}}{\partial \ln T_{ck}} = \frac{1}{1-\rho_k}\left[1 - \pi_{ck}\right] > 0 \tag{19}$$

$$\epsilon_{ckc'k} = \frac{\partial \ln \pi_{ck}}{\partial \ln T_{c'k}} = -\frac{1}{1-\rho_k}\pi_{c'k} < 0 \tag{20}$$

Elasticities within occupations can similarly be derived, this time without any approximaitons. Notice that both the within and corss-price elasticities are scaled by the $\frac{1}{1-\rho_k}$ parameter, with both increasing as ρ_k increases. This means that as occupations become more correlated across cities, the elasticities increases. For the second term in the within price elasticity, as π_{ck} increases, the elasticity decreases. This illustrates a diminishing returns effect where cities with high levels of city-occupation choice shares will have lower returns for any increases in productivity. In other words, when a city has a high level of choice shares in a given occupation, it has less to gain from further increases in productivity in that occupation relative to teh shares it already has.

The opposite is true for cross-price elasticities, where the higher levels of $\pi_{c'k}$ will increase the elasticity. This is because the more productive a city is in a given occupation, the more it will be affect another city's choice shares within that given occupation. $\pi_{c'k}$ neatly illustrates the gravity effect where large cities have larger effects on other cities.

$$\epsilon_{cc'} = \frac{\partial \ln \pi_c}{\partial \ln T_{c'}} = -\frac{\rho}{1-\rho}\pi_{c'k}(\phi_{c'k} + \omega_k) \tag{21}$$

$$\epsilon_{ck} = \frac{\partial \ln \pi_c}{\partial \ln T_k} = \phi_{ck} + \frac{\rho}{1-\rho}\omega_k \tag{22}$$

4 Simulations

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

Table 1: City Scale Parameters T_{ck}

Assume now that we have two occupations, trades and services. We will also assume three cities, Detriot, Chicago and New York with Detriot and Chicago specializing in trades and New York specializing in services. The scale parameters for each city-occupation pair are as shown in Table 1. The correlation parameter ρ_k is set to 0.7 for services and 0.5 for trades and the shape parameter is set to 2.

For each hosuehold, they will draw a $K \times N$ matrix of productivities from specific Fréchet distributions where K is the number of occupations and N is the number of cities. The realized productivity of hosueholds is then the maximum productivity across all cities and occupations. The column index of this productivity determines the city at which the household chooses to locate, and the row index determines the occupation that the household chooses to work in. This will give us our choice shares for each city and occupation.

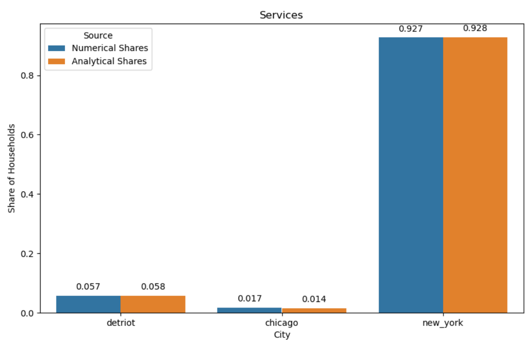

Figure 1: Service Choice Shares

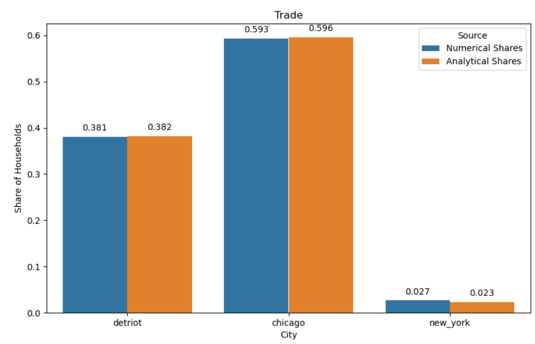

Figure 2: Trades Choice Shares

We now draw 10000 hosuheolds, simulating the distribution of location choice shares across cities and occuations. This numerical solution will now be compared to our analytical solution for choice shares as discussed in the previous section. Figure 1 and Figure 2 show the simulated occupation specific choice shares for services and trades respectively. The analytical solution is given by expression 14. The same excercise is repeated for city choice shares as shown in Figure 3. Expression 12 gives us the analytical solution for city choice shares. Consistent with our simplifying assumption, $\rho_k = \rho$ for all and it is assumed to be the average of our two ρ_k values. The numerical results closely match the analytical results, validating our analytical solution.

Now that our analytical solutions have been verified, we will now begin to examine the behaviour of our model to ensure it's consistency with intitial intuitions. Figure 4 shows the initial population distribution across cities. New York has the highest population due to it being the only city specializing in services with Detriot and Chicago sharing trades focused hosuheolds. The slight bias towards Chicago is expected due to it's higher scale parameter for trades. We will now shock Detriot's trades scale parameter by 0.5 times it's original value and observe the change in city populations. As shown in Figure 5, the population of Detriot

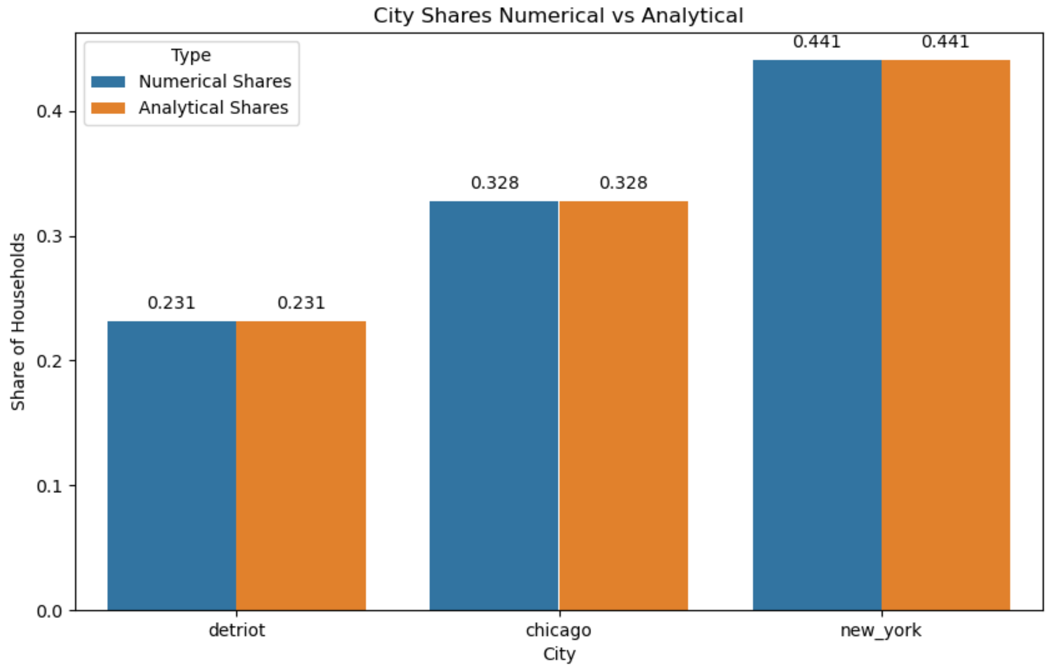

Figure 3: City Choice Shares

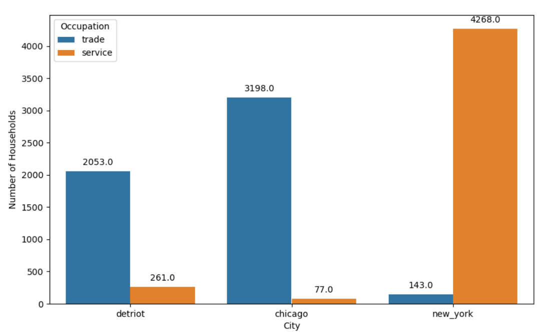

Figure 4: Initial City Population

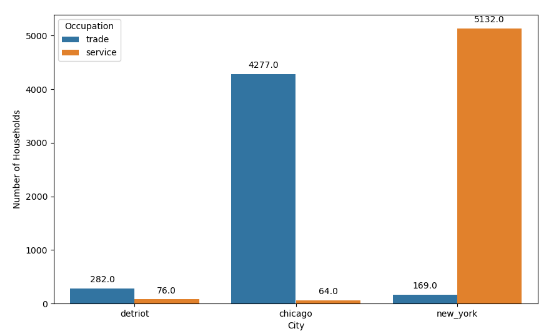

Figure 5: Post Detriot Shock City Population

decreases significantly, with Chicago and New York increasing in population. AS expected, the majority of the households working in trades have relocated to Chicago, with some moving to New York. New York meanwhile picks up the majority of households working in services from Detriot. This is consistent with our intuitions that households working in specific occupations will move to cities that specialize in those occupations.

5 Data

We now take our model to data. The data used is the 2019 American Community Survey (ACS) 1-year estimates. The data is at the metropolitan statistical area (MSA) level and consists of the number of employed individuals in each MSA by occupation. The data is aggregated into 22 occupations and 383 MSAs. Plotting the shares of employment across MSAs, we can see that there is a clear pattern of agglomeration of economic activity in certain MSAs. This is consistent with our model, where households choose to locate in cities where they are most productive. The model predicts that larger cities will tend to have higher levels of economic activity because they are able to attract more households. This is consistent with the data, where larger MSAs tend to have higher levels of employment.

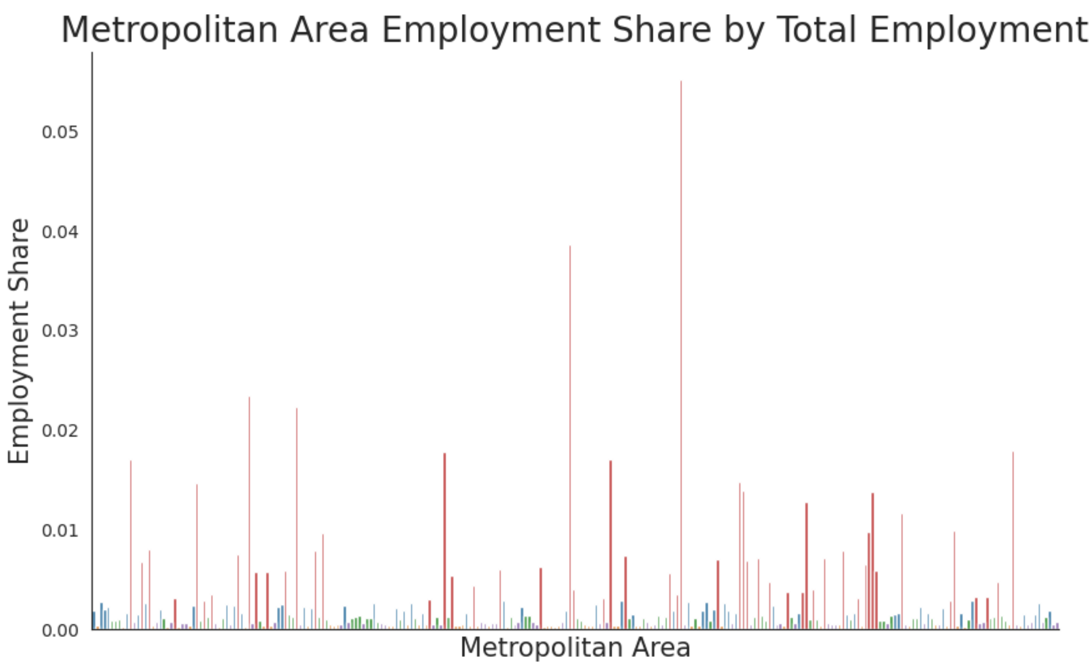

Figure 6: Metropolitan Employment Share

Zooming into the top 25 cities, Figure 7 continues to show concentration in certain cities for specific occupations, The notable example being the New York-New Jersey Metropolitan area which has a high concentration of occupation almost across the board. This illustrates the effect of the T_c shifter which makes a city more attractive overall regardless of occupation. Figure 8 meanwhile shows the composition of occupations in the top 25 cities. In other words, for any given city, which occupation does a person choose to work? We found that for the top 25 cities, the composition of occupations look very similar. This may be counterintuitive

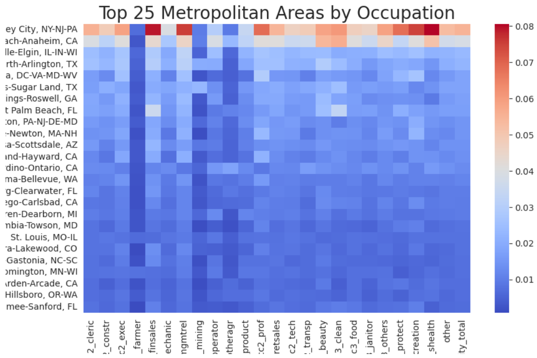

Figure 7: City Shares by Occupations

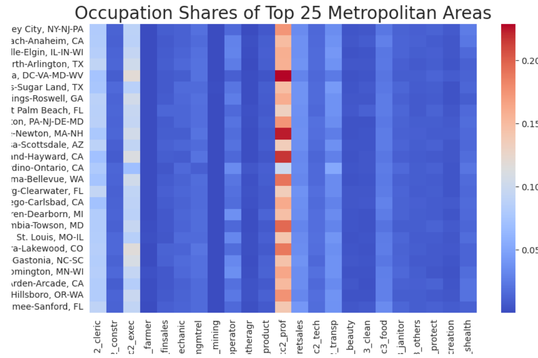

Figure 8: Occupation Shares by Cities

as cities should be specializing in different occupations. However, this may be explained by agglomoration effects where certain sectors tend to have a higher share of households overall, which leads to cities which specialize in those sectors having a higher overall share of hosueholds.

6 Proofs

6.1 City Choice Shares (Expression 12)

Begin with the following definitions:

$$G({Z_1}^{-\theta}, ..., {Z_N}^{-\theta}) = \sum_k \Big[\sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}} \Big]^{1-\rho_k}$$

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

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

Since we are summing across occupations k, we can move the derivative inside the summation:

$$\frac{\partial G({Z_1}^{-\theta}, ..., {Z_N}^{-\theta})}{\partial Z_c^{-\theta}} = \sum_k \frac{\partial}{\partial Z_c^{-\theta}} \Big[\sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho_k}} \Big]^{1-\rho_k}$$

Now we define λ_k as the following:

$$\begin{aligned} \lambda_k &= \sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho}} \\ &= \sum_c^N T_{ck}^{\frac{1}{1-\rho}} (Z_c^{-\theta})^{\frac{1}{1-\rho}} \end{aligned}$$

We can therefore simplify to the following:

$$\frac{\partial G({Z_1}^{-\theta},...,{Z_N}^{-\theta})}{\partial Z_c^{-\theta}} = \sum_k \left[(1-\rho)\lambda_k^{-\rho}\frac{\partial \lambda_k}{\partial Z_c^{-\theta}}\right]$$

$$= \sum_k \left[(1-\rho)\lambda_k^{-\rho}(\frac{1}{1-\rho})T_{ck}^{\frac{1}{1-\rho}}(Z_c^{-\theta})^{\frac{\rho}{1-\rho}}\right]$$

$$G_c({Z_1}^{-\theta},...,{Z_N}^{-\theta}) = (Z_c^{-\theta})^{\frac{\rho}{1-\rho}}\sum_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}$$

$$\pi_c = (Z_c^{-\theta})^{\frac{1}{1-\rho}}\left[\frac{\sum\limits_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}}{\sum\limits_k \lambda_k^{1-\rho}}\right]$$

6.2 City Occupation Choice Shares (Expression 14)

$$G^k(Z_1^{-\theta},\ldots,Z_N^{-\theta}) = [\sum_c^N (T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k}$$

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

$$= (1-\rho_k)[\sum_c^N (T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{-\rho_k}(\frac{1}{1-\rho_k}(T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}-1}T_{ck})$$

$$= (1-\rho_k)\frac{1}{1-\rho_k}\frac{T_{ck}}{T_{ck}Z_c^{-\theta}}(T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}[\sum_c^N (T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{-\rho_k}$$

$$Z_c^{-\theta}G_c^k(Z_1^{-\theta},\ldots,Z_N^{-\theta}) = \frac{(T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}}{\sum_c^N (T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}}[\sum_c^N (T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}]^{1-\rho_k}$$

$$\frac{Z_c^{-\theta}G_c^k(Z_1^{-\theta},\ldots,Z_N^{-\theta})}{G^k(Z_1^{-\theta},\ldots,Z_N^{-\theta})} = \frac{(T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}}{\sum_c^N (T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}}$$

$$\pi_{ck} = \frac{(T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}}{\sum_c^N (T_{ck}Z_c^{-\theta})^{\frac{1}{1-\rho_k}}}$$

6.3 Within City Elasticity (Expression 15)

$$\ln \pi_c = \frac{1}{1-\rho}\ln(Z_c^{-\theta}) + \ln[\sum_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}] - \ln[\sum_k \lambda_k^{1-\rho}]$$

We asusme that each city is small and therefore $\partial \ln \lambda_k / \partial \ln T_{ck} \approx 0$. This allows us to obtain the expression for the elasticity of a city's own price:

$$\frac{\partial \ln \pi_c}{\partial \ln T_{ck}} \approx \frac{\partial \ln[\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}]}{\partial \ln T_{ck}} = \left(\frac{\partial[\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}]}{\partial T_{ck}}\right)\left(\frac{T_{ck}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}\right)$$

$$\frac{\partial \ln \pi_c}{\partial \ln T_{ck}} \approx \left(\frac{1}{1-\rho}\right)\left(\frac{T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}{T_{ck}}\right)\left(\frac{T_{ck}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}\right)$$

$$\frac{\partial \ln \pi_c}{\partial \ln T_{ck}} \approx \Big(\frac{1}{1-\rho}\Big)\left[\frac{T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}\right] = \frac{\phi_{ck}}{1-\rho}$$

Where $\phi_{ck} = \frac{T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}$.

6.4 Cross City Elasticity (Expression 17)

$$\frac{\partial \ln \pi_c}{\partial \ln T_{c'k}} \approx \frac{\partial \ln[\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}]}{\partial \ln T_{c'k}} - \frac{\partial \ln[\sum_k \lambda_k^{1-\rho}]}{\partial \ln T_{c'k}}$$

Consider the derivative of λ_k with respect to $T_{c'k}$:

$$\frac{\partial \lambda_k}{\partial T_{c'k}} = \Big(\frac{1}{1-\rho}\Big)\Big(\frac{1}{T_{c'k}}\Big)[T_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}$$

We now decompose log-derivatives into level derivatives:

$$\frac{\partial \ln \pi_c}{\partial \ln T_{c'k}} \approx T_{ck}^{\frac{1}{1-\rho}} \Big(\frac{\partial \lambda_k^{-\rho}}{\partial T_{c'k}}\Big)\Big(\frac{T_{c'k}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}}\Big) - \Big(\frac{\partial \lambda_k^{1-\rho}}{\partial T_{c'k}}\Big)\Big(\frac{T_{c'k}}{\sum_k \lambda_k^{1-\rho}}\Big)$$

Now consider the first term:

$$
\begin{aligned}
T_{ck}^{\frac{1}{1-\rho}}\Big(\frac{\partial \lambda_k^{-\rho}}{\partial T_{c'k}}\Big)\Big(\frac{T_{c'k}}{\sum\limits_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}}\Big) &= T_{ck}^{\frac{1}{1-\rho}}(-\rho)(\lambda_k^{-\rho-1})\Big(\frac{1}{1-\rho}\Big)\Big(\frac{1}{T_{c'k}}\Big)[T_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}\Big(\frac{T_{c'k}}{\sum\limits_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}}\Big) \\
&= -\left(\frac{\rho}{1-\rho}\right)\left(\frac{[T_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}}{\lambda_k}\right)\left(\frac{T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}}{\sum\limits_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}}\right) \\
&= -\left(\frac{\rho}{1-\rho}\right)\left(\frac{[T_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}}{\lambda_k}\right)\phi_{ck}
\end{aligned}
$$

The second term can be derived in the following way:

$$
\begin{aligned}
-\Big(\frac{\partial \lambda_k^{1-\rho}}{\partial T_{c'k}}\Big)\Big(\frac{T_{c'k}}{\sum\limits_k \lambda_k^{1-\rho}}\Big) &= -(1-\rho)(\lambda_k^{-\rho})\Big(\frac{1}{1-\rho}\Big)\Big(\frac{1}{T_{c'k}}\Big)[T_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}\Big(\frac{T_{c'k}}{\sum\limits_k \lambda_k^{1-\rho}}\Big) \\
&= -(\lambda_k^{-\rho})[T_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}\Big(\frac{1}{\sum\limits_k \lambda_k^{1-\rho}}\Big)
\end{aligned}
$$

Putting these together, we can derive teh following expression for the cross-city elasticity:

$$
\begin{aligned}
\frac{\partial \ln \pi_c}{\partial \ln T_{c'k}} &\approx -(Z_{c'}^{-\theta})^{\frac{1}{1-\rho}}\left(\frac{T_{c'k}^{\frac{1}{1-\rho}}}{\lambda_k}\right)\left[\frac{\lambda_k^{1-\rho}}{\sum\limits_k \lambda_k^{1-\rho}} + \frac{\rho}{1-\rho}\phi_{ck}\right] \\
&= -\pi_{c'}\phi_{c'k}\left[1 + \left(\frac{\sum\limits_k \lambda_k^{1-\rho}}{\lambda_k^{1-\rho}}\right)\left(\frac{\rho}{1-\rho}\right)\phi_{ck}\right]
\end{aligned}
$$

$$
\frac{\partial \ln \pi_c}{\partial \ln T_{c'k}} \approx -\pi_{c'}\phi_{c'k}\Big[1 + \Big(\frac{\rho}{1-\rho}\Big)\Big(\frac{\phi_{ck}}{\omega_k}\Big)\Big]
$$

Where $\omega_k = \frac{\lambda_k^{1-\rho}}{\sum_k \lambda_k^{1-\rho}}$.

6.5 Proof for City Elasticity (Expression 21)

$$\ln \pi_c = \frac{1}{1-\rho} \ln Z_c^{-\theta} + \ln \left(\sum_k (T_c T_k t_{ck})^{\frac{1}{1-\rho}} \lambda_k^{-\rho} \right) - \ln \left(\sum_k \lambda_k^{1-\rho} \right)$$

$$\frac{\partial \ln \pi_c}{\partial \ln T_{c'}} = \frac{\partial \left(\sum_k (T_c T_k t_{ck})^{\frac{1}{1-\rho}} \lambda_k^{-\rho} \right)}{\partial T_{c'}} \frac{T_{c'}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}} - \frac{\partial \left(\sum_k \lambda_k^{1-\rho} \right)}{\partial T_{c'}} \frac{T_{c'}}{\sum_k \lambda_k^{1-\rho}}$$

Consider the derivative of λ_k with respect to $T_{c'}$:

$$\begin{aligned} \frac{\partial \lambda_k}{\partial T_{c'}} &= T_k^{\frac{1}{1-\rho}} \left(\frac{1}{1-\rho} \right) T_{c'}^{\frac{1}{1-\rho}-1} t_{c'k}^{\frac{1}{1-\rho}} (Z_c^{-\theta})^{\frac{\rho}{1-\rho}} \\ &= \frac{1}{1-\rho} \frac{1}{T_{c'}} (T_{c'k} Z_c^{-\theta})^{\frac{1}{1-\rho}} \end{aligned}$$

Consider the first term in the derivative:

$$\begin{aligned} \frac{\partial \left(\sum_k (T_c T_k t_{ck})^{\frac{1}{1-\rho}} \lambda_k^{-\rho} \right)}{\partial T_{c'}} \frac{T_{c'}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}} &= \frac{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho-1} \left(\frac{1}{1-\rho} \right) \frac{1}{T_{c'}} T_{c'k}^{\frac{1}{1-\rho}} (Z_c^{-\theta})^{\frac{1}{1-\rho}}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}} \times T_{c'} \\ &= -\frac{\rho}{1-\rho} \frac{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho} \pi_{c'k}}{\sum_k T_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}} \end{aligned}$$

Now consider the second term in the derivative:

$$
\begin{aligned}
-\frac{\partial\left(\sum_k \lambda_k^{1-\rho}\right)}{\partial T_{c'}}\frac{T_{c'}}{\sum_k \lambda_k^{1-\rho}} &= -\frac{-\sum_k \rho\lambda_k^{-\rho-1}\left(\frac{1}{1-\rho}\right)\frac{1}{T_{c'}}(T_{c'k}Z_c^{-\theta})^{\frac{1}{1-\rho}}}{\sum_k \lambda_k^{-\theta}} \times T_{c'} \\
&= \frac{\rho}{1-\rho}\frac{\sum_k \lambda_k^{-\rho}\pi_{c'k}}{\sum_k \lambda_k^{-\rho}}
\end{aligned}
$$

This will give us:

$$
\frac{\partial \ln \pi_c}{\partial \ln T_{c'}} = \frac{\rho}{1-\rho}\left[\frac{\sum_k \lambda_k^{-\rho}\pi_{c'k}}{\sum_k \lambda_k^{-\rho}} - \frac{\sum_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}\pi_{c'k}}{\sum_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}}\right]
$$

6.6 Proof for Occupation Elasticity (Expression 22)

$$
\frac{\partial \ln \pi_c}{\partial \ln T_k} = \frac{\partial\left(\sum_k (T_c T_k t_{ck})^{\frac{1}{1-\rho}}\lambda_k^{-\rho}\right)}{\partial \ln T_k}\frac{T_k}{\sum_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}} - \frac{\partial\left(\sum_k \lambda_k^{1-\rho}\right)}{\partial \ln T_k}\frac{T_k}{\sum_k \lambda_k^{1-\rho}}
$$

Consider the derivative of λ_k with respect to T_k:

$$
\begin{aligned}
\frac{\partial \lambda_k}{\partial T_k} &= \frac{1}{1-\rho}T_k^{\frac{1}{1-\rho}-1}\sum_k (T_c t_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho}} \\
&= \frac{1}{1-\rho}\frac{1}{T_k}\lambda_k
\end{aligned}
$$

Consider the first term in the derivative:

$$
\begin{aligned}
\frac{\partial\left(\sum_k (T_c T_k t_{ck})^{\frac{1}{1-\rho}}\lambda_k^{-\rho}\right)}{\partial \ln T_k}\frac{T_k}{\sum_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}} &= \frac{T_k^{\frac{1}{1-\rho}}(T_c t_{ck})^{\frac{1}{1-\rho}}(-\rho)\lambda_k^{-\rho-1}\left(\frac{1}{1-\rho}\right)\frac{1}{T_k}\lambda_k + \lambda_k^{-\rho}\left(\frac{1}{1-\rho}\right)\frac{1}{T_k}T_{ck}^{\frac{1}{1-\rho}}}{\sum_k T_{ck}^{\frac{1}{1-\rho}}\lambda_k^{-\rho}} \\
&= \phi_{ck}
\end{aligned}
$$

Consider the second term in the derivative:

$$
\begin{aligned}
-\frac{\partial\left(\sum_k \lambda_k^{1-\rho}\right)}{\partial \ln T_k}\frac{T_k}{\sum_k \lambda_k^{1-\rho}} &= -\frac{-\rho\lambda_k^{-\rho-1}\left(\frac{1}{1-\rho}\right)\frac{1}{T_k}\lambda_k}{\sum_k \lambda_k^{-\rho}} \times T_k \\
&= \frac{\rho}{1-\rho}\omega_k
\end{aligned}
$$

This will give us:

$$
\frac{\partial \ln \pi_c}{\partial \ln T_k} = \phi_{ck} + \frac{\rho}{1-\rho}\omega_k
$$