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 trend is driven by the benefits cities provide, such as access to jobs, amenities, and social networks. While these advantages are often considered universal to all households, regardless of their occupation-specific skills, the benefits of each city clearly vary for different households. For instance, an auto worker from Detroit will not view New York as an equivalent substitute for a city to live in compared to a finance worker from New York. Given the billions of dollars spent on place-based policies, it is crucial to understand this dimension of comparative advantage, where a city can provide higher levels of productivity for households working in specific occupations.

Forces of spacial inequality and agglomeration forces have been well documented in the literature (Behrens et al. (2014); Behrens and Robert-Nicoud (2015)) where large, dense cities have been shown to have higher levels of productivity, resulting in higher wages. These cities however tend to attract households with specific skills, reflected by the occupational composition of the cities. As seen in Lagakos and Waugh (2013), households will self select into occupations that they are most productive in, which in the case of large cities are occupations that disproportionately benefit from co location effects. The effect is however also seen in smaller cities which specialize in occupations such as farming, where the households who choose to locate in these cities are those who are most productive in farming. One benefit to this self selection is the co-location with households who are also productive in that occupation leading to knowledge spillovers. Davis and Dingel (2019) show that the exchange of ideas plays an important role as an agglomeration force in cities. Aside from occupations, education has also been documented to be one important factor of sorting even within a given city, where highly educated households tend to live in areas of high productivity and wages (Diamond (2016); Diamond and Gaubert (2022)).

A city's comparative advantage exists in three dimensions. First, there is a city-specific advantage that makes all occupations more productive in that city, primarily due to its size and population, which results in better amenities, more jobs, and a larger social network. Second, there is an occupation-specific advantage

that is universal across all cities; some jobs are simply more productive than others, and cities specializing in these jobs will be observed as having higher overall choice shares of households. Third, there is a city-occupation specific advantage, where a city may have specific policies that make it more attractive for a particular occupation, such as tax breaks for mechanics in Chicago. It is crucial to distinguish between occupations and sectors: sectors are broad categories of jobs (e.g., manufacturing, services, agriculture), while occupations are specific jobs (e.g., mechanic, clerk). Cities can specialize in occupations by focusing on sectors that intensively use those occupations.

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. However, both these models 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 apply this structure to the urban setting by treating occupations as independent nests, capturing substitution patterns where cities with similar occupational structures are treated as better substitutes for each other. This approach enables us to capture all three dimensions of comparative advantage.

Our model is a part of the larger class of quantitative spacial models as proposed by Redding and Rossi-Hansberg (2017), with modifications that allow us to capture the effects of trade shocks at the group level, similar to the work of Galle et al. (2022). As shown by Autor et al. (2013), trade shocks have heterogenous effects on industries with differing exposure to trade shocks, leading to an sharp movements in wages but slow adjustments in labor market outcomes (Artuç et al. (2010)). We will be able to capture these effects in counterfactual exercises once we have estimated the parameters of our model. Given our class of model, we will be able to draw from techniques and estimations which already exist in the literature such as Albert and Monras (2022) who estimate migration elasticity and Kim and Vogel (2021) who estimate the elasticity of labor market adjustments to trade shocks. We also make use of shift share instruments (Bartik (1991)), specifically shifts in total factor productivity for industries which have different shares of occupations, that's widely used in trade to estimate our parameters. This type will require us to base our identification on either exogenous shares (Goldsmith-Pinkham et al. (2020)) or exogenous shifts (Borusyak et al. (2021)).

We begin in the first section by adapting the model of Lind and Ramondo (2023) to the urban setting. Unlike

their original paper where nests are implied, we explicitly model the nests as occupations. This specification yields several implications. 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: city-specific, occupation-specific, and city-occupation specific elasticities. The last elasticity will serve as a sanity check for the direction of substitution, while 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 with a certain occupational composition will have larger effects on cities with similar compositions due to correlations in productivity draws.

In the second section, we present our estimation strategy, exploiting variation over time to recover two important parameters of our model: the correlation parameter ρ and the shape parameter θ. As we will discuss later, these parameters determine the within and across occupation migration patterns. In this process, we will also recover our city, occupation, and city-occupation specific shifters.

In the third section, we apply our estimation equations to data from the American Community Survey at the metropolitan area and occupational level. We begin by presenting stylized facts about cities and their compositions, comparing similarities among top cities. We then discuss our estimation results, focusing on the correlation parameter ρ, as well as our city, occupation, and city-occupation specific shifters. As we will demonstrate, ρ can be interpreted as the degree of within-occupation migration, while ρ represents the degree of across-occupation migration. Using this fully calibrated model, we conduct counterfactual analyses and simulate the effects of shocks at both the city and occupation level. This analysis allows us to evaluate the effectiveness of place-based policies and provide recommendations for designing such policies to maximize their impact.

2 Model

$$Y_{ck} = Z_{ck}$$

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

3

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. A household has symmetric CES preferences over cities and occupations and will choose a city-occupation pair that maximizes their utility. We will assume that households are perfectly mobile across cities and that there are no fixed costs to moving. The productivity of a household within a given city-occupation is drawn from a Fréchet distribution. In Eaton and Kortum (2002), the Fréchet distribution is used to model productivity draws in a trade setting and has two main components. The shape parameter θ which determines the dispersion of draws across the distribution and reflects the heterogeneity of cities. θ will also determine the gains from city choices, ie how much households will gain from having more choices of cities to live in, which is analogous to gains from trade. The scale parameter T_{ck} is a level shifter that determines the absolute advantage of a city-occupation pair, reflecting its attractiveness. Our specification so far assumes independent draws of productivity across cities, the joint distribution of which is given by:

$$P[Z_1 < z_1, \ldots, Z_N < z_N] = \prod_c^N P[Z_c < z_c]$$
$$= \exp\left\{-\sum_c^N T_{ck} Z_c^{-\theta}\right\}$$

This result implies that the choice shares of households for a given city is equal to the probability of households choosing that city. So far, we have stuck with the the base EK specification with the only modification being the mapping of productivity onto wages rather than prices which allows us to obtain city choice shares. We will now abandon the independence assumption and introduce a structure that allows for correlated draws across cities.

$$P[Z_1 < z_1, \ldots, Z_N < z_N] = \exp\left\{-\left(\sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho}}\right)^{1-\rho}\right\}$$

Notice with this structure that the shape of the distribution is no longer purely determined by θ but also by the correlation parameter ρ. Where the distribution of productivity within a given occupation decreases as ρ increases. Meaning that within a given occupation, the higher correlation will result in cities being more

similar with each other in terms of productivity. When $\rho = 0$, the distribution reduces down to the original joint distribution with the shape being purely determined by θ and where draws are independent. As $\rho \to 1$, the draws become perfectly correlated and the productivity is identical across all cities.

2.1 Cross Nested CES

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

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.

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.

$$Z = \max_c \left\{ \frac{Z_c}{\Phi_c} \right\}$$

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   Household Problem

$$\max_{c,k} \left\{ U(\nu) = \frac{y(\nu)}{\Phi_c} \right\} \tag{2}$$

Now assume that the economy has households of type $\nu \in [0,1]$ who earn an income $y(\nu)$. The household's utility is given by the ratio of income to the price index of the city. Each household will simultaneously choose a city occupation pair that maximizes their utility. Given that all goods are traded freely and all firms operate in perfect competition, we will make the assumption that all cities will have a common price index $\Phi_c = \Phi$. This will allow us to simplify the household's problem to the following:

$$\max_{c,k} \left\{ U(\nu) = \frac{L(\nu)w_{ck}(\nu)Z_{ck}(\nu)}{\Phi} \right\}$$

where $y(\nu) = L(\nu)w_{ck}(\nu)Z_{ck}(\nu)$, $L(\nu)$ is the portion of time dedicated to labor, $w_{ck}(\nu)$ is the wage of the household in city c working in occupation k, and $Z_{ck}(\nu)$ is the productivity of the household in city c working in occupation k. We will now assume that households spend all their time working $L(\nu) = 1$ and that the wage is the only source of income. Expanding $w_{ck} = T_c T_k t_{ck}$, we wil get the final household maximization problem.

$$U^{\star}(\nu) = \max_{c,k} \frac{1}{\Phi} \{T_c T_k t_{ck} Z_{ck}(\nu)\}$$

Here, we can see that the household maximization problem is separated into two components. The first being the expected wage which households will receive in a given city-occupation pair, determined by the scale parameters T_c, T_k, and t_{ck}. And the second is the idiosyncratic productivity draw $Z_{ck}(\nu)$ which is drawn from a Fréchet distribution and is unique to each household type resulting in deviations from the expected choice shares. The assumption of a common price index would mean that any changes in shares of

6

households across cities would not affect the overall price index. This results in a partial equilibrium model where changes resulting from a shock are assumed to be stable. While this is a limitation of the model, it is a necessary assumption in order to make the model tractable.

2.3 Equilibrium

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

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

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

Where $G_c(Z_1^{-\theta}, \ldots, Z_N^{-\theta}) = \partial G(Z_1^{-\theta}, \ldots, Z_N^{-\theta})/\partial Z_c^{-\theta}$. 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 $\rho_k = \rho$.

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

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. The following definitions will be useful in simplifying the expression for city shares.

7

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

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:

$$\phi_{ck} \equiv \frac{t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}} \tag{5}$$

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

$$\omega_k \equiv \frac{T_k \lambda_k^{1-\rho}}{\sum_k T_k \lambda_k^{1-\rho}} \tag{6}$$

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.

$$\frac{\pi_{ck}}{\pi_c} \equiv \frac{\phi_{ck}}{\omega_k} \tag{7}$$

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.

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

$$\pi_c = \sum_k^K \pi_{ck} \omega_k \tag{8}$$

By using our definitions, we arrive at the above expression for city shares, which is now the sum of all occupation specific shares weighted by the market share of each occupation.

2.4 City-Occupation Elasticities

We now turn to the city-occupation derivatives: city c's own technology elasticity and cross technology 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. We then obtain the following for the own technology shock:

$$\frac{\partial \ln \pi_{ck}}{\partial \ln t_{ck}} \approx \phi_{ck} \left(\frac{1}{1-\rho} \right) \tag{9}$$

This elasticity is positive and increasing in a city's comparative advantage in occupation k. The intuition

is that a city that is more attractive for occupation k will see a larger increase in the choice share of that occupation in response to a marginal increase in the productivity of that occupation. The more specialized the city is in that occupation, the larger the effect of a productivity shock in that occupation. As $\rho \to 1$, the elasticity approaches infinity, as within any given occupation, if they become more correlated, cities within that given occupation become perfect substitutes for each other.

$$\frac{\partial \ln \pi_c}{\partial \ln t_{c'k}} = -\pi_{c'}\phi_{ck}\left[1 + \left(\frac{\rho}{1-\rho}\right)\left(\frac{\pi_{ck}}{\pi_c}\right)\right] \tag{10}$$

Turning our attention to the cross city technology elasticity, we see that the first term captures the CES element of consumer choice. As the city c' increases in size and becomes increasingly competitive in that occupation, any shocks to this city-occupation pair will have an increasingly negative effects on all other cities.

How much the other cities will be affected is captured by the second term which captures the unbalanced substitution patterns. If a city is relatively concentrated in occupation k, then the affects of a shock in that given occupation will be felt more severely, this is shown by the ratio π_{ck}/π_c. Notice also that if $\rho = 0$, this elasticity simplified to the standard CES elasticity. As $\rho \to 1$, the elasticity approaches infinity and the correlated choices start to dominate.

2.5   Aggregate Elasticities

We now consider the elasticity of city shares with respect to aggregate city technology shocks T_c and aggregate occupation shocks T_k. Beginning with the former, if we assume again that each cit is small and that $\partial \ln \lambda_k / \partial \ln T_c \approx 0$, we get the following for a city's own technology shock:

$$\frac{\partial \ln \pi_c}{\partial \ln T_c} \approx \frac{1}{1-\rho}$$

This is simply the summation of city-occupation elasticities which was derived in expression 9. And given that by definition $\sum_{k}^{K} \phi_{ck} = 1$, we get the following derivations:

$$\frac{\partial \ln \pi_c}{\partial \ln T_c} = \sum_k \frac{\partial \ln \pi_c}{\partial \ln t_{ck}}$$

$$= \sum_k \phi_{ck} \left(\frac{1}{1-\rho} \right)$$

$$= \left(\frac{1}{1-\rho} \right) \sum_k \phi_{ck} = \frac{1}{1-\rho}$$

The result tells us that if $\rho = 0$ and there's no correlation between occupations, then the elasticity is simply 1. We get a similar story when deriving the cross city technology elasticity. Using the same assumption, we arrive at the result:

$$\frac{\partial \ln \pi_c}{\partial \ln T_{c'}} = -\pi_{c'} \left[\sum_k \phi_{c'k} \left[1 + \left(\frac{\rho}{1-\rho} \right) \left(\frac{\pi_{ck}}{\pi_c} \right) \right] \right] \tag{11}$$

Similarly, this is simply the sum over all city-occupation elasticities derived in expression 10. This is shown in the following derivation:

$$\frac{\partial \ln \pi_c}{\partial \ln T_{c'}} = \sum_k \frac{\partial \ln \pi_c}{\partial \ln t_{c'k}}$$

$$= \sum_k \left[-\pi_{c'} \phi_{ck} \left[1 + \left(\frac{\rho}{1-\rho} \right) \left(\frac{\pi_{ck}}{\pi_c} \right) \right] \right]$$

$$= -\pi_{c'} \left[\sum_k \phi_{c'k} \left[1 + \left(\frac{\rho}{1-\rho} \right) \left(\frac{\pi_{ck}}{\pi_c} \right) \right] \right]$$

The interpretation of this elasticity is the same as expression 10. The first term captures the CES element of consumer choice and the second term captures the unbalanced substitution patterns, with ρ governing the strength of this unbalanced substitution. Now we simply sum over all occupations to get the aggregate elasticities.

Lastly, we derive the elasticity of city shares with respect to aggregate occupation shocks T_k. This is fairly straightforward as wel can simply evaluate the following:

$$\frac{\partial \ln \pi_c}{\partial \ln T_k} = \left(\frac{\partial \phi_{ck}}{\partial T_k}\right)\left(\frac{T_k}{\phi_{ck}}\right) - \left(\frac{\partial \omega_k}{\partial T_k}\right)\left(\frac{T_k}{\omega_k}\right)$$

$$= \phi_{ck} - \omega_k$$

By using the identity in expression 7, we can rewrite the elasticity as the following:

$$\frac{\partial \ln \pi_c}{\partial \ln T_k} = \omega_k \left[\frac{\pi_{ck}}{\pi_c} - 1\right] \tag{12}$$

You should first notice that the second term in brackets can be either positive or negative depending on the city's relative concentration in that occupation. If the city is relatively concentrated and $\pi_{ck}/\pi_c > 1$, then the city will face a negative shock. If the city is relatively less concentrated and $\pi_{ck}/\pi_c < 1$, then the city will face a negative shock. We can think of this as certain occupations start to become more attractive, cities that are less concentrated in that occupation will face a negative shock as households start to move to cities that are more concentrated in that occupation. This second term is scaled by the market share of that occupation, ω_k, which is simply the market share of that occupation across all cities.

3   Estimation Strategy

3.1   Within-Occupation Migration

Notice that for us to estimate our shifters, we have to first recover our ρ parameter. To do so, we exploit a variation in city-occupation shares over time. Consider the following expression at the city-occupation-year level. From expression 4, we can get the following:

$$\ln \pi_{ck} = \frac{1}{1-\rho} \ln t_{ckt} - \frac{1}{1-\rho} \ln Z_{ct}^{-\theta} - \ln \lambda_{kt}$$

Now we assume the following relationship productivity and wages at the city-occupation level: $t_{ckt} = w_{ckt}\eta_{ckt}$. Where w_{ckt} are pages paid and η_{ckt} is productivity. This will give us the following model:

$$\ln \pi_{ckt} = \left(\frac{1}{1-\rho} \right) \ln w_{ckt} + \delta_{ct} + \delta_{kt} + \eta_{ckt}$$

With some abuse of notation, we have defined $\eta_{ckt} = \ln \eta_{ckt}$. We now take differences over time to eliminate level shifts.

$$\Delta_t \ln \pi_{ckt} = \left(\frac{1}{1-\rho} \right) \Delta_t \ln w_{ckt} + \delta_{ct} + \delta_{kt} + \Delta_t \eta_{ckt}$$

We can estimate this equation by regressing wages on city occupation shares and using city and occupation fixed effects at the city-occupation level for every year with δ_{ct} being city fixed effects and δ_{kt} being occupation fixed effects. Given that $\hat{\rho} \in (0,1)$, we will end up with the restriction $\hat{\beta} \in [1, +\infty)$.

With this regression, we are testing whether migration decisions are primarily within occupations rather than across occupations. If the correlation parameter is close to 1, then migration decisions are purely within occupations. Notice that as $\hat{\rho}$ increases, $\hat{\beta}$ increases. From this perspective, we can interpret ρ as governing within-occupation migration and θ as governing across-occupation migration.

$$Y_{st} = A_{st} \prod_k^K Q_{kst}^{\gamma_{ks}}$$

One concern with the initial specification is the endogeneity of wages with the error term,more explicitly $cov[w_{ckt}, \eta_{ckt}] \neq 0$. In order to instrument for the change in wages, we will use a shift share in which we make use of sectors. Here, we define a Cobb-Douglas strcture of production where the intensity of occupation k in sector s is given by γ_{ks} such that $\sum_k^K \gamma_{ks} = 1$. We will use the shift share of occupation k in sector s as an instrument for the change in wages. A_{st} will be the total factor productivity of sector s in year t.

$$w_{ckt}^{IV} = \sum_s^S \gamma_{ks} \kappa_{cs} A_{st}$$

where κ_{cs} is the share of occupation k in sector s in city c. This will give us the following first stage regression:

$$\ln w_{ckt}^{IV} = \beta^{IV} \ln w_{ckt}^{IV} + \delta_{ct} + \delta_{kt}^{IV} + \eta_{ckt}^{IV}$$

We will then use the predicted values of $\ln w_{ckt}^{IV}$ as our instrument in the second stage regression.

$$\ln \pi_{ckt} = \left(\frac{1}{1-\rho}\right) \ln w_{ckt}^{IV} + \delta_{ct} + \delta_{kt} + \eta_{ckt}$$

3.2   City Shifters

Given that we've identified ρ, we can now proceed to estimate our city scale parameter T_c. We start with occupation market shares ω_k in expression 6. By taking relative shares, we get the following:

$$\frac{\omega_{kt}}{\omega_{k't}} = \frac{T_{kt}\lambda_{kt}^{1-\rho}}{T_{k't}\lambda_{k't}^{1-\rho}}$$

Subject to normalization to some base occupation year, we can estimate $T_{kt}\lambda_{kt}^{1-\rho}\forall(k,t)$. We can then use this to recover the denominator in ω_k and get $\sum_k^K T_{kt}\lambda_{kt}^{1-\rho}\forall t$. This will be useful in recovering our city and city occupation specific scale parameters. Now consider the city-occupation shares from expression 4 taken in logs and transformed into a form with fixed effects:

$$\ln \pi_{ckt} = \frac{1}{1-\rho} \ln t_{ckt} - \frac{1}{1-\rho} \ln Z_{ct}^{-\theta} - \ln \lambda_{kt}$$
$$= \hat{\epsilon}_{ckt} + \hat{\delta}_{ct} + \hat{\delta}_{kt}$$

such that,

$$\hat{\lambda}_{kt} = \exp\left(-\hat{\delta}_{kt}\right)$$

$$\hat{t}_{ckt} = \exp\left((1-\rho)\hat{\epsilon}_{ckt}\right)$$

We can then use the estimated $\hat{\lambda}_{kt}$ to recover our occupation specific scale parameters:

$$\hat{T}_{kt} = \frac{\sum_k^K \widehat{T_{kt}\lambda_{kt}^{1-\rho}}}{\hat{\lambda}_{kt}^{1-\rho}}$$

Now we need some way for us to recover productivity. Consider the following modification to city shares using our recovered parameters:

$$\tilde{\pi}_{ct} = \pi_{ct} \frac{\sum_k^K T_{kt}\lambda_{kt}^{-\rho}}{\sum_k^K t_{ckt}^{\frac{1}{1-\rho}} T_{kt}\lambda_{kt}^{1-\rho}}$$

$$= Z_{ct}^{-\frac{\theta}{1-\rho}}$$

We now take logs and difference the equation to eliminate the level shift. Recalling our definition of $Z_c^{-\theta} = (\gamma)^{-\theta}(T_c)^\theta$, we get the following estimation equation:

$$\Delta_t \ln \hat{\pi}_{ct} = \hat{\beta}\Delta_t \ln w_{ct} + \Delta_t \hat{\eta}_{ct}$$

The differencing will eliminate γ, which is constant over time, leaving us with a regression of changes in wages and city fixed effects over time on changes in city shares. Using this regression, we recover city level shifters from fixed effects estimates.

$$\hat{T}_{ct} = exp\left((1 - \rho)\hat{\eta}_{ct}\right)$$

4 Empirical Results

4.1 Data

We will be using data from the American Community Survey (ACS) which is a survey conducted by the US Census Bureau. The ACS is a survey that collects data on the metropolitan area and occupational level. We will be using data from 2010 to 2022. The ACS provides data on the household level, detailing wages, metropolitan area, occupation and sector of employment based on NAICS codes. This will be used to construct our share data π_c π_{ck}, ω_k, ϕ_{ck} and ϕ_{cs} as well as obtain wage data w_{ckt}. We will also be using data from the Bureau of Labour Statistics to obtain total factor productivity data at the sector level A_{st}.

Figure 1: Summary Statistics of Metropolitan Statistical Areas (MSA) in 2019
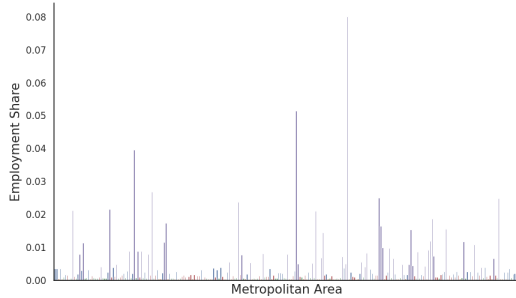
(a) Employment Share by MSA in 2019
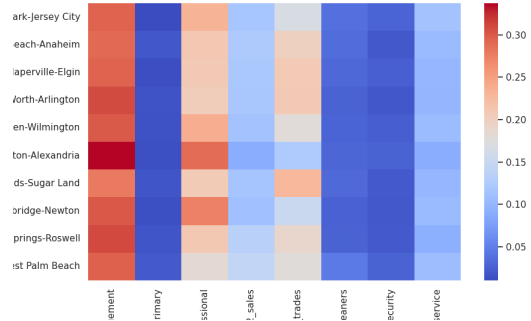
(b) Composition of Top 10 MSAs in 2019

Note: We will be referring to MSAs as cities throughout the paper. MSAs in Figure 1a are arranged in alphabeltical order on the x axis.

From the data in Figure 1a, we see that a few metropolitan areas own a disproportionate share of employment which illustrates the size effects of cities to attract employment and households. Intuitively, we should expect these cities to be very different from each other in terms of their specialization, otherwise we would expect the largest city with a particular composition to dominate in the share of households. When zooming into the compositions of the top 10 cities however, we see a different story. In Figure 1b, we see that the top 25 cities are very similar with each other in terms of their specialization. With all cities specializing in occupations like management and professional services. While there is some heterogeneity in the composition of cities,

the specialization of cities is not as pronounced as we would expect. This indicates to us that the occupation shifter will play a large role in determining the attractiveness of a city.

4.2 Results

To recover our ρ parameter, we run a two stage least squares regression with a shift share instrument. We find that this regression yields a high F-Statistic and a significantly higher estimate for $\hat{\beta}$ compared to ols[1]. This ultimately results in a fairly high estimate of $\hat{\rho} = 0.67$. The result means that draws across cities within any given occupation id highly correlated and indicates to us that people are very likely to migrate within occupations. This is consistent with the idea that people are more likely to move to cities where they can find jobs in their occupation.

Table 1: Estimations for T_c and T_k for 2019

(A) Occupation Shifters

Occupation	Occupation Shifter
Management	0.849566
Professional	0.743607
Trades	0.610380
Service	0.344887
Sales	0.318461
Cleaners	0.090246
Primary	0.066206
Security	0.066183

(B) Top City Shifters

City	City Shifter
New York-Newark-Jersey City	3.029541
Los Angeles-Long Beach-Anaheim	2.645708
Chicago-Naperville-Elgin	2.333439
Dallas-Fort Worth-Arlington	2.158189
Houston-The Woodlands-Sugar Land	2.093633
Philadelphia-Camden-Wilmington	2.028458
Washington-Arlington-Alexandria	2.007037
Atlanta-Sandy Springs-Roswell	2.001475

Note: Occupation and city shifters are not directly comparable. Recall that occupations are scaled only by ρ while occupations are scaled by ρ and θ. Comparisons should only be relative and be made within each table.

In 2019, we find that occupation shifters are generally lower compared to city shifters. However, since cities are scaled by an additional factor of theta, the resulting effect from occupation shifters remains higher than city shifters. For occupation shifters, it's unsurprising that professional and management services have the highest values, given the proportion of households employed in these occupations. Conversely, primary sector occupations have the lowest shifters. Coupled with the fact that these occupations are generally performed in cities with low shifters themselves, this results in cities specializing in primary occupations having low shares of households. It's important to note, however, that while a city-occupation pair may be generally unattractive to most households, those who specialize in these occupations will be very willing to move to cities with occupations matching their abilities.

---

[1]Refer to appendix 8.2 for full regression table

Most dense metropolitan areas tend to specialize in occupations with high shifters, explaining the large pattern of rural-urban migration to these cities in the past. However, now that most cities have developed similar occupational compositions, we see large cities competing with each other for household shares, which is why no particular city is universally most attractive to all households. While large cities tend to have similar overall compositions, they still manage to distinguish themselves by having distinct compositions in other occupations. For example, although New York may be concentrated in management and professional services, it has a higher concentration in sales compared to Los Angeles. This suggests that these cities are large because of their dominant occupations, but they don't compete purely based on scale with other cities.

Unsurprisingly, cities with the highest shifters are dense metropolitan areas such as New York and Los Angeles, reflecting the large populations they attract. Notable in our estimates is the variance in shifter values between cities, with some shifters being up to three times larger than others, indicating significant heterogeneity in city attractiveness. This makes sense as large cities tend to offer more and better amenities as well as better access to jobs, making them more attractive to households. We also find that city shifters tend to be fairly stable over time, with changes generally below 10% over the 10-year period. Given the time required for cities to build infrastructure and other amenities, in addition to needing to differentiate themselves from other cities, this stability is not surprising.

5 Quantitative Exercises

Using our parameter and shifter estimates, we can now perform two counterfactual exercises. First, we will simulate a city-specific shock to observe migration patterns across cities, focusing on the unbalanced substitution patterns that favor similar cities. Second, we will examine the migration patterns resulting from a shock to a particular occupation. This effectively simulates the effects of a trade shock that impacts some occupations more than others, such as the China shock. These exercises will provide insights into how different types of economic shocks affect urban migration and occupation choices.

In order to perform these exercises, we will assume aa parameter θ of 3 which is an estimate used by Redding (2016) that is obtained from the literature. This will give us our across occupation elasticity.

Figure 2: Relative Change in City Shares Following New York Shock ($\rho = 0$ vs $\rho = 0.64$)
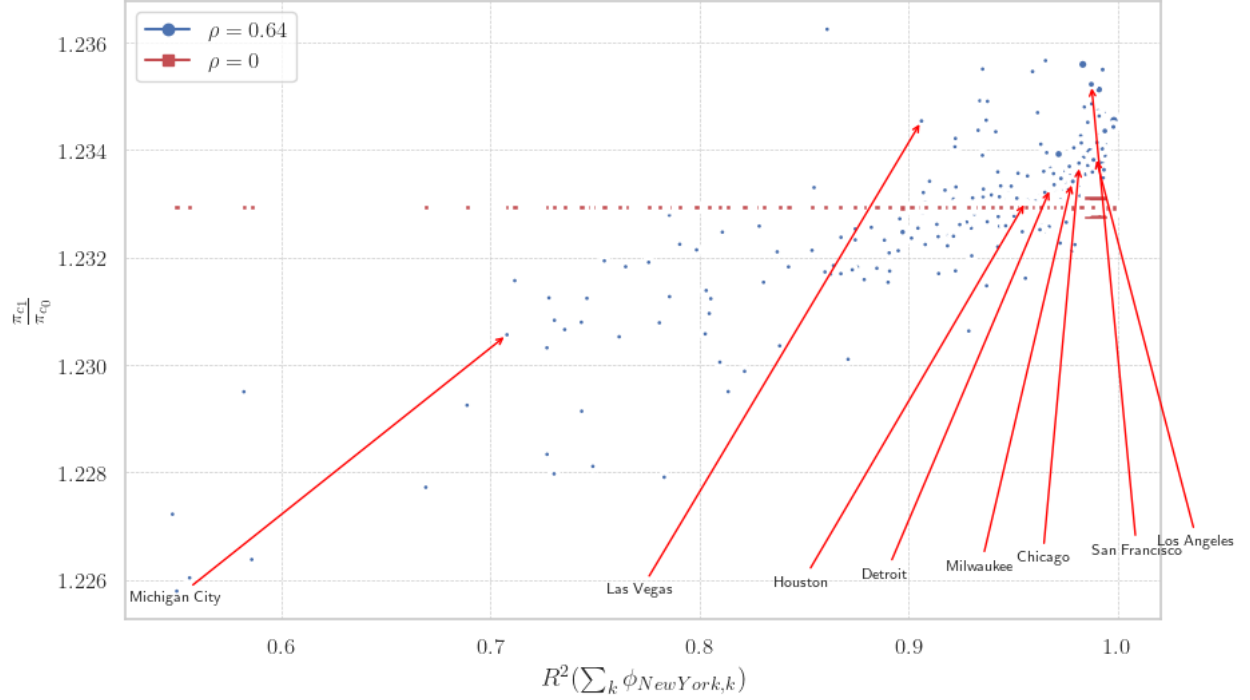
Note: *The x axis represents the correlation of city compositions of other cities relative to New York. New York is excluded from the plot. The size of each plot is relative to the size of each city before the shock. π_{c_0} is the city share before the shock, and π_{c_1} is the city share after the shock. $\phi_{c,k}$ is the share of occupation k in city c. Notable cities are pointed out with a red arrow with the name annotated.*

5.1 City Specific Shock

Our results from the previous section suggest that while we may observe some migration to large cities purely due to their size, we should expect to see more migration to cities with similar occupational makeups. For this exercise, we induce a negative shock of 5% to all occupations in New York, a city primarily focused on management and professional services occupations. The results are shown in Figure 2. Analyzing the results, we find that the cities which benefit the most from this shock are cities with occupational compositions that are highly correlated with New York, meaning that the cities which are most similar to New York benefit the most from this negative shock. This is consistent with our previous results that households are more likely to move to cities with similar occupational compositions. Correlation in city composition is however not the only determinant of household migration. The size of the city also plays a large role in determining where households move. This is why we see that at the top end of the correlation, it's large cities such as Los Angeles and Chicago that benefit the most from the New York shock. One interpretation of this is that households who are looking to relocate will want to move to cities that are similar to New York, but given

19

that all cities are very similar, the determining factor in where they move is the size of the city, which in our case is perfectly correlated with the benefits each city provides.

What's interesting is when we turn off within-occupation migration by setting $\rho = 0$. Recalling our discussion in section 2.1, when $\rho = 0$, we should expect the distribution to reduce to a standard CES model. This is exactly what we see in the results. When $\rho = 0$, we see that households are much more evenly redistributed across cities. Given that all occupations are equally attractive to all households, all cities will hence become perfect substitutes for each other[2].

This exercise perfectly illustrates the goal of the model, which is to simulate better match the substitution patterns of households across cities. The model illustrates that migration patterns of households are mainly determined by the occupational composition of each city. Meaning that households are very willing to relocate, but only to cities which offer opportunities in their field. But within the set of cities which meet this criteria, the size of the city becomes the deciding factor in where households move. This is in contrast to the standard QSM model proposed by Redding (2016) which would predict that households will simply move to the city with the highest amenity value.

5.2 Occupation Specific Shock

What happens when there's a shock to particular occupations? We've established that households move to cities with similar occupational profiles when there's a negative shock to specific cities, but how do they respond when a shock affects all cities in a particular occupation? To explore this, we induce a negative shock of 5% to management occupations across all cities.

Naturally, we see a decrease in city shares for the cities that are relatively more concentrated in management occupations. This is regardless of the value of ρ, with large cities that are the least concentrated gaining the most from this shock. In Figure 3, we can see that cities such as San Francisco and New York lose out from this shock, while cities such as Michigan tend to benefit. One additional observation is that the relative change for each city is the same regardless of the value of ρ, which we can see when looking at the 45 degree line. This is consistent with our findings in city technology elasticities.

Recall Expression 12, which tells us that the elasticity of a city with respect to an occupation shock is governed by the relative concentration of that occupation in the city. Crucially, this elasticity is not governed by ρ,

---

[2]Expression 11 tells us that the elasticity of each city is governed by ρ. When $\rho = 0$, $\sum_k^K \phi_{c'k} = 1$, all cities will share the same elasticity of $\pi_{c'}$.

Figure 3: Relative Change in City Shares Following A Shock in Management Occupations ($\rho = 0$ vs $\rho = 0.64$)
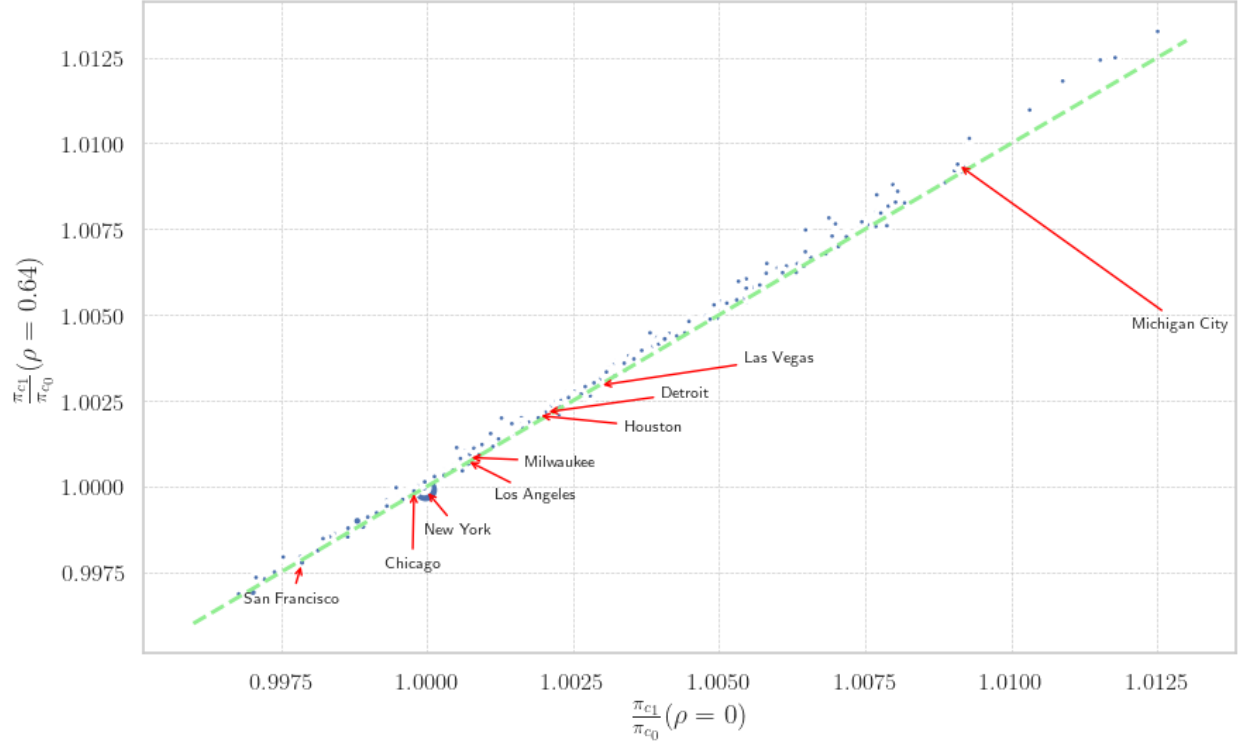
Note: *π_{c_0} is the share of households in city c before the shock, and π_{c_1} is the share of households in city c after the shock. The size of each plot represents the size of each city before the shock. The green line is a 45 degree line. The plot does not start from 1, some cities have experienced a decrease in the share of households and some have experienced an increase. Notable cities are pointed out with a red arrow with their name annotated.*

meaning that it should be consistent across different values of ρ. The interpretation is that when there is a shock in cities, households will be able to find employment in the same occupation in other cities, leading to a migration to cities of similar compositions. But when there is a shock in occupations, employment within that given occupation will be equally affected across all cities, which will lead to migration across occupations to cities that are less affected by the shock. This is exactly what we see, an even redistribution of households to all other occupations, leading to cities that are less concentrated in management occupations to benefit the most from this shock, which in our case tend to be smaller cities.

6   Discussion

Given the empirical and counterfactual results, we believe there are a couple of important policy implications to consider. Place-based policies that aim to make a city overall more productive in all occupations will

generally be ineffective in attracting migration. This is because households tend to pick cities based on their occupations rather than the city itself. A more effective place-based policy should target either occupations that the city already specializes in, or occupations that have a high productivity shifter. Importantly, to increase a city's relative competitiveness, policymakers should focus on occupations that make it less similar to other cities in order to decrease direct competition.

Policies aimed at responding to trade shocks should be focused on transitioning households to other occupations. Unlike city shocks where households can simply move to other cities for better opportunities, trade shocks will result in households being persistently unemployed due to the lack of opportunities in their given occupation specialization. To tackle this form of unemployment, governments need to facilitate the transition of these households to other occupations. This can be done through retraining programs or by providing incentives for firms to hire these workers. Such measures will not only help households find jobs, but will also help cities that are highly specialized in these affected occupations to attract more households. By adopting these targeted approaches, policymakers can more effectively address the challenges posed by both localized economic changes and broader trade shocks, ultimately fostering more resilient urban economies and labor markets.

7   Conclusion

In this paper, we have developed a model of city and occupation choice that enables us to recover city and occupation-specific shifters and substitution parameters. Our findings demonstrate that occupations play a more significant role in migration decisions compared to cities themselves, with households more likely to relocate to cities with similar occupational compositions. We have also shown that within-occupation migration is more prevalent than across-occupation migration, consistent with the notion that people are more inclined to move to cities where they can find jobs in their current occupation. In the future, we hope to extend this model to include a housing market and unemployment which will allow up to explain housing prices in relatively low productivity cities as well as persistent unemployment in certain cities.

References

ALBERT, C. AND J. MONRAS (2022): "Immigration and Spatial Equilibrium: The Role of Expenditures in the Country of Origin," *American Economic Review*, 112, 3763–3802.

ARTUÇ, E., S. CHAUDHURI, AND J. MCLAREN (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–45.

AUTOR, D. H., D. DORN, AND G. H. HANSON (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–68.

BARTIK, T. J. (1991): *Who Benefits from State and Local Economic Development Policies?*, W.E. Upjohn Institute.

BEHRENS, K., G. DURANTON, AND F. ROBERT-NICOUD (2014): "Productive Cities: Sorting, Selection, and Agglomeration," *Journal of Political Economy*, 122, 507–553.

BEHRENS, K. AND F. ROBERT-NICOUD (2015): "Chapter 4 - Agglomeration Theory with Heterogeneous Agents," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Elsevier, vol. 5 of *Handbook of Regional and Urban Economics*, 171–245.

BORUSYAK, K., P. HULL, AND X. JARAVEL (2021): "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, 89, 181–213.

DAVIS, D. R. AND J. I. DINGEL (2019): "A Spatial Knowledge Economy," *American Economic Review*, 109, 153–70.

DIAMOND, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106, 479–524.

DIAMOND, R. AND C. GAUBERT (2022): "Spatial Sorting and Inequality," *Annual review of economics*, 14, 795–819.

EATON, J. AND S. S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

GALLE, S., A. RODRÍGUEZ-CLARE, AND M. YI (2022): "Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade," *The Review of Economic Studies*, 90, 331–375.

GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2020): "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 110, 2586–2624.

KIM, R. AND J. VOGEL (2021): "Trade Shocks and Labor Market Adjustment," *American Economic Review: Insights*, 3, 115–30.

LAGAKOS, D. AND M. E. WAUGH (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103, 948–80.

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.

REDDING, S. J. AND E. ROSSI-HANSBERG (2017): "Quantitative Spatial Economics," *Annual review of economics*, 9, 21–58.

8 Appendix

8.1 Derivations

City Choice Shares π_c: As discussed in Lind and Ramondo (2023), the choice shares can be expressed as:

$$\pi_c = \frac{Z_c^{-\theta} G_c}{G()}$$

We define $G()$ as the following, given the definition of T_{ck} and λ_k provided in the main text:

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

$G_c()$ is simply the derivative of $G()$ with respect to $Z_c^{-\theta}$.

$$
\begin{aligned}
\frac{\partial G()}{\partial Z_c^{-\theta}} &= \sum_k \frac{\partial}{\partial Z_c^{-\theta}} \left[\sum_c^N (T_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho}} \right]^{1-\rho} \\
&= \sum_k \frac{\partial}{\partial Z_c^{-\theta}} \left[T_k^{\frac{1}{1-\rho}} \sum_c^N (t_{ck} Z_c^{-\theta})^{\frac{1}{1-\rho}} \right]^{1-\rho} \\
&= \sum_k T_k \left[(1-\rho) \lambda_k^{-\rho} \frac{\partial \lambda_k}{\partial Z_c^{-\theta}} \right] \\
&= \sum_k T_k \left[(1-\rho) \lambda_k^{-\rho} (\frac{1}{1-\rho}) t_{ck}^{\frac{1}{1-\rho}} (Z_c^{-\theta})^{\frac{\rho}{1-\rho}} \right] \\
&= (Z_c^{-\theta})^{\frac{\rho}{1-\rho}} \sum_k T_k t_{ck}^{\frac{1}{1-\rho}} \lambda_k^{-\rho}
\end{aligned}
$$

Putting these three terms together, we derive the choice shares:

$$
\begin{aligned}
\pi_c &= \frac{Z_c^{-\theta} G_c}{G()} \\
&= (Z_c^{-\theta})^{\frac{1}{1-\rho}} \left[\frac{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}}{\sum_k T_k \lambda_k^{1-\rho}} \right]
\end{aligned}
$$

Own Occupation-Specific Elasticity: We wish to evaluate $\partial \ln \pi_c / \partial \ln t_{ck}$, under the assumption that $\partial \ln \lambda_k / \partial \ln t_{ck} = 0$. Notice this assumption yields the following:

$$\frac{\partial \ln \pi_c}{\partial \ln t_{ck}} \approx \frac{\partial \ln \pi_{ck}}{\partial \ln t_{ck}} - \frac{\partial \ln \phi_{ck}}{\partial \ln t_{ck}}$$

given the definitions of π_{ck} and ϕ_{ck} provided in the text.

Notice that this only leaves one term to evaluate if we take the logarithm of π_c and the derivative of this logarithm with respect to $\ln t_{ck}$:

$$\frac{\partial \ln \pi_c}{\partial \ln t_{ck}} \approx \frac{\partial \ln[\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}]}{\partial \ln t_{ck}}$$

$$= \left(\frac{\partial[\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}]}{\partial t_{ck}} \right) \left(\frac{t_{ck}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}} \right)$$

$$= \left(\frac{1}{1-\rho} \right) \left(\frac{t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}}{t_{ck}} \right) \left(\frac{t_{ck}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}} \right)$$

$$= \left(\frac{1}{1-\rho} \right) \left[\frac{t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}} \right]$$

$$= \frac{\phi_{ck}}{1-\rho}$$

Cross-City Occupation-Specific Elasticity: We wish to evaluate $\partial \ln \pi_c / \partial \ln t_{c'k}$. Notice that this elasticity can be expressed as the following two terms:

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

It will be useful to first define the derivative of λ_k with respect to our variable of interest, $t_{c'k}$:

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

We can now use chain rule in order to evaluate our elasticity of interest. The first term becomes the following:

$$t_{ck}^{\frac{1}{1-\rho}} T_k \left(\frac{\partial \lambda_k^{-\rho}}{\partial t_{c'k}} \right) \left(\frac{t_{c'k}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}} \right) = t_{ck}^{\frac{1}{1-\rho}} T_k (-\rho)(\lambda_k^{-\rho-1}) \left(\frac{1}{1-\rho} \right) \left(\frac{[t_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}} \right)$$

$$= -\left(\frac{\rho}{1-\rho} \right) \left(\frac{[t_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}}{\lambda_k} \right) \left(\frac{t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}}{\sum_k t_{ck}^{\frac{1}{1-\rho}} T_k \lambda_k^{-\rho}} \right)$$

$$= -\left(\frac{\rho}{1-\rho} \right) \left(\frac{[t_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}}{\lambda_k} \right) \phi_{ck}$$

$$= -\left(\frac{\rho}{1-\rho} \right) \pi_{c'k} \phi_{ck}$$

The second term can be derived in the following way:

$$-T_k \left(\frac{\partial \lambda_k^{1-\rho}}{\partial t_{c'k}} \right) \left(\frac{t_{c'k}}{\sum_k T_k \lambda_k^{1-\rho}} \right) = -T_k (\lambda_k^{-\rho})[t_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}} \left(\frac{1}{\sum_k T_k \lambda_k^{1-\rho}} \right)$$

$$= -\left(\frac{[t_{c'k}(Z_{c'}^{-\theta})]^{\frac{1}{1-\rho}}}{\lambda_k} \right) \left(\frac{T_k \lambda_k^{1-\rho}}{\sum_k T_k \lambda_k^{1-\rho}} \right)$$

$$= -\pi_{c'k} \omega_k$$

Putting these together, we derive the cross-city occupation-specific elasticity as the following:

$$\frac{\partial \ln \pi_c}{\partial \ln t_{c'k}} = -\pi_{c'k} \left[\omega_k + \left(\frac{\rho}{1-\rho} \right) \phi_{ck} \right]$$

Notice that this is equivalent to the following, given the identity linking π_c, π_{ck}, ϕ_{ck}, and ω_k.

$$\frac{\partial \ln \pi_c}{\partial \ln t_{c'k}} = -\pi_{c'} \phi_{ck} \left[1 + \left(\frac{\rho}{1-\rho} \right) \left(\frac{\pi_{ck}}{\pi_c} \right) \right] \tag{13}$$

Cross-City Aggregate Elasticity: We wish to evaluate $\partial \ln \pi_c / \partial \ln T_{c'}$. Notice that:

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

Notice also that:

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

We can therefore solve for this cross-derivative as the following:

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

8.2 Regression Tables

8.2.1 First Stage Shift Share Regression

	Dependent variable: wage
	(1)
sim_wage	0.523***
	(0.017)
Observations	19280
R^2	0.881
Adjusted R^2	0.863
Residual Std. Error	0.198 (df=16799)
F Statistic	49.973*** (df=2480; 16799)
Note:	*p<0.1; **p<0.05; ***p<0.01

8.2.2 OLS vs IV Regression

	Dependent variable: actual_shares	
	IV	OLS
	(1)	(2)
predicted_wage	2.777***	
	(0.041)	
wage		0.291***
		(0.011)
Observations	19280	19280
R^2	0.956	0.946
Adjusted R^2	0.949	0.938
Residual Std. Error	0.255 (df=16799)	0.282 (df=16799)
F Statistic	145.837*** (df=2480; 16799)	118.423*** (df=2480; 16799)
Note:		*p<0.1; **p<0.05; ***p<0.01