Automation, Spatial Sorting, and Job Polarization*

Jan Eeckhout[‡], Christoph Hedtrich[§], and Roberto Pinheiro[¶]

May 2019

Preliminary and Incomplete version - Please do not cite

Abstract

We present evidence showing that more expensive cities – measured by rental costs – have not only invested proportionately more in automation (measured by investment in Enterprise Resource Planning software) but also have seen a higher decrease in the share of routine abstract jobs (clerical workers and low-level white collar workers). We propose an equilibrium model of location choice by heterogeneously skilled workers where each location is a small open economy in the market for computers and software. We show that if computers are substitutes to middle skill workers – commonly known as the automation hypothesis – in equilibrium large and expensive cities invest more in computers and software, substituting middle skill workers with computers. Intuitively, in expensive cities, the relative benefit of substituting computers for routine abstract workers is higher, since workers must be compensated for the high local housing prices. Moreover, if the curvature of the production function is the same across skills, the model also delivers the thick tails in large cities' skill distributions presented by Eeckhout et al. (2014).

Keywords: Automation, Skill Distributions, City Sizes, Job and Wage Polarization.
JEL Codes: D21, J24, J31, R23.

*We would like to thank Tim Dunne, Bruce Fallick, Pawel Krolikowski, Stuart Rosenthal, Henry Siu, Murat Tasci, the participants of the University of Kentucky and Cleveland Fed Microeconomic Workshop and participants of the Conference on Housing, Urban Development, and the Macroeconomy at USC Dornsife. Any remaining errors are our responsibility. Jan Eeckhout acknowledges financial support from grant ECO2015-67655-P and ERC Advanced Grant 339186. The views in this article do not necessarily reflect those of the Federal Reserve System or the Board of Governors.

‡UPF and UCL `jan.eeckhout@upf.edu`.
§UPF, `christoph.hedtrich@upf.edu`
¶Federal Reserve Bank of Cleveland, `Roberto.Pinheiro@clev.frb.org`.

1 Introduction

The disappearance of mid-skill high-paying jobs has dominated the academic and policy discussions. The pay gap between highly educated workers and low and middle-education workers, represented by the college wage premium, increased steadily from the early 1980s until the late 2000s, exceeding 97 % (Acemoglu and Autor (2011)). Moreover, according to Goldin and Katz (2009), this is the highest level since 1915, the earliest year for which representative data are available. In many cases, the process of automation and job polarization has been pointed out as the underlying reason for the increasing inequality and "disappearing middle" of the income distribution.

The rapid diffusion of new technologies has directly substituted capital for labor in tasks previously performed by moderately skilled workers. In particular, the automation hypothesis formulates that machines are most likely to displace jobs that are intensive in routine tasks, both manual and abstract. Hence, automation reduces job opportunities in the middle of the skill distribution, including clerical, administrative, production, and operative occupations. Jobs less affected by automation would demand either non-routine abstract tasks – requiring high levels of education and commanding high compensation – or non-routine manual tasks – which tend to be low-paying manual jobs. Consequently, we have a "hollowing out" of the compensation distribution, in line with the results obtained in empirical work.

In this paper, we show that the substitution of routine jobs and tasks with machines, computers, and software has not happened evenly in space. In fact, the relative benefit of replacing routine-task jobs by computers and software depend on the cost of hiring a worker in a particular location. Consequently, living costs – in particular housing costs – play a key role. For example, let's consider two offices that are demanding for some standard accounting services that can be performed either by an accounting assistant or by an accounting software. One of these offices is located in New York city, the other in Akron - OH. In order to hire a new accounting assistant, the New York office must pay a wage that allows the new employee to live in an area close enough to the company's office in order to go to work every day. Since housing costs in the New York area are significantly higher than in Akron - OH, the New York-based firm must pay more to hire the same accounting assistant. In comparison, accounting softwares are the same price in both cities. Consequently, automation is a more attractive option to the New York firm. In equilibrium, it is more likely that the New York firm will introduce the new software, while the Akron office hires an additional accounting assistant.

Our empirical results show that large and more expensive cities – measured by population size in the Metropolitan Statistical Area (henceforth MSA) as well as rental price index – not only have invested more in Enterprise Resource Planning softwares (henceforth ERPs), but have also experienced the largest decrease in the fraction of routine abstract workers in the population of employed workers. As pointed out by Bloom et al. (2014), ERP is the generic name for software systems that integrate several data sources and processes of an organization into a unified system. These applications are used to store, retrieve, and share information on any aspect of the sales and firm organizational processes in real time.[1] Consequently, the introduction of ERPs reduces the need for clerical and low-level white collar workers. Moreover, in contrast with Personal Computers (henceforth PCs), which are general purpose technologies (Jovanovic and Rousseau (2005)), the introduction of ERPs have as its main goal the replacement of clerical work.

We propose an equilibrium model of heterogeneous workers' locations across cities that illustrate this intuition. In our model, prices play a key role in our equilibrium model of city choice. Heterogeneously skilled citizens earn a living based on a competitive wage and choose housing in a competitive housing market. Under perfect mobility, their location choice make them indifferent between consumption-housing bundles, and therefore between different wage-housing price pairs across cities. Wages are generated by firms that compete for labor and that have access to a city-specific technology summarized by that city's total factor productivity (TFP). This naturally gives rise to a price-theoretic measure of skills. Larger cities pay higher wages, and are more expensive to live in. Under worker mobility, revealed preference location choices imply that wages adjusted for housing prices are a measure of skills.

Within this framework, we introduce capital in a simple way. First, we consider that capital is produced globally and all cities are small open economies in the market for capital. Therefore, firms in all cities can rent any quantity of capital and take capital's rental rate as given. Second, we test two competing hypothesis that have been championed by the previous literature. The automation hypothesis, which consider that mid-skill workers and capital are substitutes, and the Skill Biased Technological Change hypothesis (henceforth SBTC) which consider that capital and high-skill workers are complements. While we believe that these hypothesis are not mutually exclusive, this simplification allows us to draw some stark comparisons in order to identify the driving forces behind the changes in the employment and wage distributions across cities.

We show that the automation hypothesis is able to match the empirical patterns that we find in the data,i.e, we observe an increasing substitution of routine abstract jobs with ERP and computers as the latter became more affordable. Moreover, our model shows that the automation hypothesis is also able to deliver the thick tails distribution in the skill distribution, documented by Eeckhout et al. (2014), in one of its sub-cases. In contrast, in the same set-up,

[1] This information includes not only standard metrics like production, deliveries, machine failures, orders and stocks, but also broader metrics on human resources and finance.

the SBTC hypothesis would deliver first order stochastic dominance (henceforth FOSD) in the skill distribution. In this sense, while we do not discard the possibility of SBTC, our results point to the importance of including the automation hypothesis in order to match some key patterns presented by the empirical evidence.

Our paper is closest to Autor and Dorn (2013). In their paper, they show that areas in which we have a high concentration of workers performing routine tasks, there is a push towards automation. In this sense, we could imagine an initial large sunk cost of implementing automation – particularly true for routine manual workers – which would be more profitable the more workers the new machines would substitute. Our results point towards a different dynamics, that hinges on the differences of local prices. Through our results, even though clerical workers may be a somewhat smaller fraction of the labor force in New York City than in Akron - OH, the fact that hiring a new accounting assistant is significantly more expensive in New York City makes it more attractive to New York-based firms to introduce the new software. Consequently, it is not necessarily the absolute fraction of the work force in routine tasks that induce automation, but the relative cost of introducing the new technology vs. routine task workers. Our results suit quite well the introduction of technologies that do not demand large initial sunk costs – as the introduction of ERP softwares.

Moreover, we are the first to document the effects of introducing new technologies while looking at technology investments that are not only tied to geographical locations, but also to a particular use. In this sense, we focus on softwares whose use is clearly related to the activities performed by routine abstract workers, instead of general purpose technologies, such as PCs.

Our paper is divided into 6 sections. Sections 2 and 3 present our model and theoretical results, as well as some simple numerical exercises. Section 4 and 5 describe the data and empirical results, respectively. Finally, section 6 concludes the paper. All proofs are presented in the Appendix.

2 Model

Population. Consider an economy with heterogeneously skilled workers. Workers are indexed by a skill type i. For now, let the types be discrete: $i \in \mathcal{I} = \{1, ..., I\}$. Associated with this skill order is a level of productivity x_i. Denote the country-wide measure of skills of type i by M_i. Let there be J locations (cities) $j \in \mathcal{J} = \{1, ..., J\}$. The amount of land in a city is fixed and denoted by H_j. Land is a scarce resource.

Preferences. Citizens of skill type i who live in city j have preferences over consumption c_{ij}, and the amount of land (or housing) h_{ij}. The consumption good is a tradable numeraire good with price equal to one. The price per unit of land is denoted by p_j. We think of the expenditure on housing as the flow value that compensates for the depreciation, interest on capital, etc. In a competitive rental market, the flow payment will equal the rental price.[2] A worker has consumer preferences over the quantities of goods and housing c and h that are represented by: $u(c, h) = c^{1-\alpha}h^{\alpha}$, where $\alpha \in [0, 1]$. Workers are perfectly mobile, so they can relocate instantaneously and at no cost to another city. Because workers with the same skill are identical, in equilibrium each of them should obtain the same utility level wherever they choose to locate. Therefore for any two cities j, j' it must be the case that the respective consumption bundles satisfy $u(c_{ij}, h_{ij}) = u(c_{ij'}, h_{ij'})$, for all skill types $\forall i \in \{1, ..., I\}$.

Technology. Cities differ in their total factor productivity (TFP) which is denoted by A_j. For now, we assume that TFP is exogenous. We think of it as representing a city's productive amenities, infrastructure, historical industries, persistence of investments, etc.

In each city, there is a technology operated by a representative firm that has access to a city-specific TFP A_j. Output is produced by choosing the right mix of differently skilled workers i as well as the amount of capital k. While labor markets are local and workers must live in the city in which they are employed, capital markets are global and even large cities are small open economies in the capital markets. We also consider that firms rent capital that is owned by a zero measure of absentee capitalists. For each skill i, a firm in city j chooses a level of employment m_{ij} and produces output: $A_jF(m_{1j}, ..., m_{Ij}, k_j)$. Firms pay wages w_{ij} for workers of type i. It is important to note that wages depend on the city j because citizens freely locate between cities not based on the highest wage, but, given housing price differences, based on the highest utility. Like land and capital, firms are owned by absentee capitalists (or equivalently, all citizens own an equal share in the mutual fund that owns all the land and all the firms). Finally, we consider that the rental price for capital is given by $r > 0$ which is determined in the global market and taken as given by firms in the different cities.

Market Clearing. In the country-wide market for skilled labor, markets for skills clear market by market, and for housing, there is market clearing within each city:

$$\sum_{j=1}^{J} C_j m_{ij} = M_i, \ \forall i \qquad\qquad \sum_{i=1}^{I} h_{ij} m_{ij} = H_j, \ \forall j. \tag{1}$$

where C_j denotes the number of cities with TFP A_j.

[2] We will abstract from the housing production technology; for example, we can assume that the entire housing stock is held by a zero measure of absentee landlords.

3 The Equilibrium Allocation

The Citizen's Problem. Within a given city j and given a wage schedule w_{ij}, a citizen chooses consumption bundles $\{c_{ij}, h_{ij}\}$ to maximize utility subject to the budget constraint (where the tradable consumption good is the numeraire, i.e. with price unity)

$$\begin{aligned} \max_{\{c_{ij},h_{ij}\}} u(c_{ij}, h_{ij}) &= c_{ij}^{1-\alpha} h_{ij}^{\alpha} \\ \text{s.t. } c_{ij} + p_j h_{ij} &\leq w_{ij} \end{aligned} \tag{2}$$

for all i, j. Solving for the competitive equilibrium allocation for this problem we obtain $c_{ij}^{\star} = (1-\alpha)w_{ij}$ and $h_{ij}^{\star} = \alpha \frac{w_{ij}}{p_j}$. Substituting the equilibrium values in the utility function, we can write the indirect utility for a type i as:

$$U_i = \alpha^{\alpha} (1-\alpha)^{1-\alpha} \frac{w_{ij}}{p_j^{\alpha}} \quad \Longrightarrow \quad w_{ij} = U_i p_j^{\alpha} \frac{1}{\alpha^{\alpha} (1-\alpha)^{1-\alpha}}, \tag{3}$$

where U_i is constant across cities from labor mobility. This allows us to link the wage distribution across different cities j, j'. Wages across cities relate as:

$$\frac{w_{ij}}{w_{ij'}} = \left(\frac{p_j}{p_{j'}} \right)^{\alpha}. \tag{4}$$

The Firm's Problem. All firms are price-takers and do not affect wages. Wages are determined simultaneously in each submarket i, j while capital rent is determined in the global market. Given the city production technology, a firm's problem is given by:

$$\max_{m_{ij}, \forall i} A_j F(m_{1j}, ..., m_{Ij}, k_j) - \sum_{i=1}^{I} w_{ij} m_{ij} - r k_j, \tag{5}$$

subject to the constraint that $m_{ij} \geq 0$ and $k \geq 0$. The first-order conditions are: $A_j F_{m_{ij}} (m_{ij}, k_j) = w_{ij}, \forall i$ and $A_j F_{k_j} (m_{ij}, k_j) = r$.[3]

Because there is no general solution for the equilibrium allocation in the presence of an unrestricted technology, we focus on variations of the Constant Elasticity of Substitution (CES) technology, where the elasticity is allowed to vary across skill types. As a benchmark therefore,

[3]In what follows, the non-negativity constraint on m_{ij} and k_j are dropped. This is justified whenever the technology satisfies the Inada condition that marginal product at zero tends to infinity whenever A_j is positive. This will be the case since we focus on variations of the CES technology.

we consider the following *separable* technology:

$$A_jF(m_{1j},...,m_{Ij},k_j) = A_j\left(\sum_{i=1}^{I} m_{ij}^{\gamma_i}x_i + k_jx_k\right) \tag{6}$$

with $\gamma_i < 1$, $\forall i \in \{1,...,I\}$. In this case the first-order conditions are $A_j\gamma_i m_{ij}^{\gamma_i-1}x_i = w_{ij}, \forall i$ and $A_j\gamma_i k_j^{\gamma_i-1}x_k = r$. Notice that if $\gamma_i \equiv \gamma,\ \forall i \in \{1,...,I\}$ we have a CES production function.

In an on line appendix, we solve the allocation under *separable* technology as a special case of the more general technologies presented in the paper. Even without fully solving the system of equations for the equilibrium wages, observation of the first-order condition reveals that productivity between different skills i in a given city is governed by three components: (1) the productivity x_i of the skilled labor and how fast it increases in i; (2) the measure of skills m_{ij} employed (wages decrease in the measure employed from the concavity of the technology); and (3) the degree of concavity γ_i, indicating how fast congestion builds up in a particular skill. Without loss of generality, we assume that wages are monotonic in the order i.[4] This is consistent with our price-theoretic measure of skill.

We now proceed by introducing varying degrees of complementarity/substitutability between different skills and capital, starting from the *separable* technology. In this way, we are able to address different theories in terms of the impact of technology in either boosting the productivity of some types, as presented by the literature on Skill Bias Technological Change (henceforth SBTC) or replacing workers, as in the literature about automation. For tractability, let there be two cities, $j \in \{1,2\}$ and three skill levels $i \in \{1,2,3\}$. We will also consider the degree of complementarity/substitutability by nesting a CES production function within the overall production function. Consequently, if we assume that there is a degree of complementarity between skill i and capital, while none between the remaining skills, then we consider that the technology can be written as $\left(m_{ij}^{\theta}x_i + k^{\theta}x_k\right)^{\frac{\gamma_i}{\theta}} + \sum_{l=-i} m_{lj}^{\gamma_j}x_l$. Notice that if $\gamma_i > \theta$, skill i and capital are gross complements, while if $\gamma_i < \theta$, capital and skill i are gross substitutes.

Definition 1 *Consider the following technologies:*

[4] For a given order i, wages may not be monotonic as they depend on the relative supply of skills as well as on x_i. If they are not, we can relabel skills such that the order i corresponds to the order of wages. Alternatively, we can allow for the possibility that higher skilled workers can perform lower skilled jobs. Workers will drop job type until wages are non-decreasing. Then the distribution of workers is endogenous, and given this endogenous distribution, all our results go through. For clarity of the exposition, we will assume that the distribution of skills ensures that wages are monotonic.

I. ***Automation.*** *Capital and middle skill workers are substitutes.*

$$A_j F(m_{1j}, m_{2j}, m_{3j}, k) = A_j \left\{ m_{1j}^{\gamma_1} x_1 + \left(m_{2j}^{\theta} x_2 + k_j^{\theta} x_k \right)^{\frac{\gamma_2}{\theta}} + m_{3j}^{\gamma_3} x_3 \right\} \quad \textit{where} \quad \gamma_2 < \theta \tag{7}$$

II. ***Skill-Bias Technological Change.*** *Capital and high skill workers are complements.*

$$A_j F(m_{1j}, m_{2j}, m_{3j}, k) = A_j \left\{ m_{1j}^{\gamma_1} x_1 + \left(m_{3j}^{\theta} x_3 + k_j^{\theta} x_k \right)^{\frac{\gamma_3}{\theta}} + m_{2j}^{\gamma_2} x_2 \right\} \quad \textit{where} \quad \gamma_3 > \theta \tag{8}$$

3.1 Automation

We first derive the equilibrium conditions for case I, Automation. The first-order conditions (henceforth FOCs) are for each j and all skill types i and capital, respectively:

$$\begin{aligned}
&(m_{1j}): && A_j \gamma_1 m_{1j}^{\gamma_1 - 1} x_1 = w_{1j}, \quad \forall j \in J; \\
&(m_{2j}): && A_j \tfrac{\gamma_2}{\theta} \left(m_{2j}^{\theta} x_2 + k_j^{\theta} x_k \right)^{\frac{\gamma}{\theta} - 1} \theta m_{2j}^{\theta - 1} x_2 = w_{2j}, \quad \forall j \in J; \\
&(k_j): && A_j \tfrac{\gamma_2}{\theta} \left(m_{2j}^{\theta} x_2 + k_j^{\theta} x_k \right)^{\frac{\gamma_2}{\theta} - 1} \theta k_j^{\theta - 1} x_k = r, \quad \forall j \in J; \\
&(m_{3j}): && A_j \gamma_3 m_{3j}^{\gamma_3 - 1} x_3 = w_{3j}, \quad \forall j \in J;
\end{aligned} \tag{9}$$

Using labor mobility, we can write the wage ratio in terms of the house price ratio for all i, $\frac{w_{i2}}{w_{i1}} = \left(\frac{p_2}{p_1} \right)^{\alpha}$ and equate the first-order condition in both cities for a given skill. If we then compare the results for low- and high skill workers and use both the utility equalization condition, due to labor mobility, and the housing market clearing conditions for cities 1 and 2 we have:

$$m_{11} = \frac{\left[\left(\frac{p_1}{p_2} \right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_1 - 1}} M_1}{\left\{ 1 + \left[\left(\frac{p_1}{p_2} \right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_1 - 1}} \right\}} \quad \text{and} \quad m_{31} = \frac{\left[\left(\frac{p_1}{p_2} \right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_3 - 1}} M_3}{\left\{ 1 + \left[\left(\frac{p_1}{p_2} \right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_3 - 1}} \right\}} \tag{10}$$

and likewise for city 2. Finally, using the FOCs for skill 2 and capital, jointly with utility equalization and labor market condition for skill 2 in city 1, we have:

$$m_{21} = \frac{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{\theta - 1}} \frac{k_1}{k_2}}{\left[1 + \left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{\theta - 1}} \frac{k_1}{k_2} \right]} M_2 \quad \text{and} \quad k_2 = \frac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1 - \gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1 - \theta}} \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1 - \gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}} \tag{11}$$

and likewise for city 2.

So far we have consumer optimization for consumption and housing, the location choice by the worker, and firm optimization given wages. The next step is to allow for market clearing in the housing market given land prices. The system is static and solved simultaneously, which is reported in the Appendix. In what follows, we assume $H_j = H$ for all cities j. Below, we will discuss the implications where this simplifying assumption has bite.

The Main Theoretical Results. First we establish the relationship between TFP and house prices. When cities have the same amount of land, we can establish the following result.

Proposition 1 (Automation, TFP, and Housing Prices) *Assume* $\gamma_2 < \theta$. $A_i > A_j \Rightarrow p_i > p_j, \forall j \in \{1, 2\}$

Consequently, the city with the highest TFP is also the one with the highest housing prices. We establish this result for cities with an identical supply of land. Clearly, the supply of land is important in our model since in a city with an extremely small geographical area, labor demand would drive up housing prices all else equal. This may therefore make it more expensive to live in even if the productivity is lower. Because in our empirical application we consider large metropolitan areas (NY city for example includes large parts of New Jersey and Connecticut), we believe that this assumption does not lead to much loss of generality.[5]

We now focus on the demand for capital and TFP. As proposition 2 shows, the city with higher TFP also demands more capital. The intuition is straightforward. In cities with higher TFP, housing prices are higher and workers must be compensated in order to afford living in a more expensive place. Furthermore, since firms with higher TFP hire more of all skill levels, the decreasing marginal returns are also more strong, pushing towards the increase in the use of capital in order to replace middle skills in this case. Hence, high-TFP cities demand more capital.

Proposition 2 (Automation, TFP and capital demand) *Assume* $\gamma_2 < \theta$. $A_i > A_j \Rightarrow k_i > k_j$.

Then, in theorem 1 we show that the city with the high TFP is also larger. In fact, we are able to show that, in equilibrium, the high-TFP city has more workers at all skill levels.

Theorem 1 (Automation and City Size) *Assume* $\gamma_2 < \theta$ *and* $A_1 > A_2$. *We have that* $S_1 > S_2$.

[5] In fact, the equal supply of housing condition is only sufficient for the proof, but not necessary. However, our model does not address the important issue of within-city geographical heterogeneity, as analyzed for example in Lucas and Rossi-Hansberg (2002). In our application, all heterogeneity is absorbed in the pricing index by means of the hedonic regression.

Finally, theorem 2 shows that, in the case in which $\gamma_i \equiv \gamma$ for all skills and $\gamma < \theta$, high-TFP city has proportionately more of high and low skill workers than low-TFP cities. This is true even though high-TFP cities have more of all types. Consequently, the high-TFP city is more unequal in terms of its skill distribution.

Theorem 2 (Automation and Spatial Sorting) *Assume $\gamma_i \equiv \gamma,\ \forall i \in \{1,2,3\}$ and $\gamma < \theta$. If $A_1 > A_2$ we have that city 1 has thick tails in the skill distribution.*

3.2 Skill Biased Technological Change

We now consider the case of *Skill-Bias Technological Change* (henceforth SBTC) in which capital and high-skill workers are complements. In this case, the FOCs for each city j, skill type i, and capital, respectively are:

$$\begin{aligned}
&(m_{1j}) : A_j \gamma_1 m_{1j}^{\gamma_1 - 1} x_1 = w_{1j} \\
&(m_{2j}) : A_j \gamma_2 m_{2j}^{\gamma_2 - 1} x_2 = w_{2j} \\
&(m_{3j}) : A_j \gamma_3 \left(m_{3j}^{\theta} x_3 + k_j^{\theta} x_k \right)^{\frac{\gamma_3}{\theta} - 1} m_{3j}^{\theta - 1} x_3 = w_{3j} \\
&(k_j) : A_j \gamma_3 \left(m_{3j}^{\theta} x_3 + k_j^{\theta} x_k \right)^{\frac{\gamma_3}{\theta} - 1} k_j^{\theta - 1} x_k = r
\end{aligned} \tag{12}$$

Using labor mobility, we can write the wage ratio in terms of the house price ratio for all i, $\frac{w_{i2}}{w_{i1}} = \left(\frac{p_2}{p_1}\right)^{\alpha}$ and equate the first-order condition in both cities for a given skill. If we then compare the results for low- and middle-skill workers and use both the utility equalization condition, due to labor mobility, and the housing market clearing conditions for cities 1 and 2 we have:

$$m_{11} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_1 - 1}} M_1}{\left\{ 1 + \left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_1 - 1}} \right\}} \quad \text{and} \quad m_{21} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_2 - 1}} M_2}{\left\{ 1 + \left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_2 - 1}} \right\}} \tag{13}$$

and likewise for city 2. Finally, using the FOCs for skill 3 and capital, jointly with utility equalization and labor market condition for skill 2 in city 1, we have:

$$m_{31} = \frac{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{\theta - 1}} \frac{k_1}{k_2}}{\left[1 + \left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{\theta - 1}} \frac{k_1}{k_2} \right]} M_3 \quad \text{and} \quad k_2 = \frac{M_3 x_3^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1 - \theta}} \left[\left(\frac{r}{A_1 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - x_k \right]^{\frac{1}{\theta}}} \tag{14}$$

and likewise for city 2.

So far we have consumer optimization for consumption and housing, the location choice by the worker, and firm optimization given wages. The next step is to allow for market clearing in the housing market given land prices. The system is static and solved simultaneously, which is reported in the Appendix. In what follows, we assume $H_j = H$ for all cities j. Below, we will discuss the implications where this simplifying assumption has bite.

The Main Theoretical Results. First we establish the relationship between TFP and house prices. When cities have the same amount of land, we can establish the following result.

Proposition 3 (SBTC, TFP, and Housing Prices) *Assume* $\gamma_3 > \theta$. $A_i > A_j \Rightarrow p_i > p_j$, $\forall j \in \{1, 2\}$

Consequently, the city with the highest TFP is also the one with the highest housing prices. We establish this result for cities with an identical supply of land. Clearly, the supply of land is important in our model since in a city with an extremely small geographical area, labor demand would drive up housing prices all else equal. This may therefore make it more expensive to live in even if the productivity is lower. Because in our empirical application we consider large metropolitan areas (NY city for example includes large parts of New Jersey and Connecticut), we believe that this assumption does not lead to much loss of generality.

We now focus on the demand for capital and TFP. As proposition 4 shows, the city with higher TFP also demands more capital.

Proposition 4 (SBTC, TFP, and capital demand) *Assume* $\gamma_3 > \theta$. $A_i > A_j \Rightarrow k_i > k_j$.

Corollary 1 shows that the high TFP city also attracts more high-skill workers.

Corollary 1 (SBTC and demand for high skill) *Assume* $\gamma_3 > \theta$. $A_i > A_j \Rightarrow m_{3i} > m_{3j}$.

Finally, theorem 3 shows that in the case in which $\gamma_i \equiv \gamma$ for all skills and $\gamma > \theta$, high-TFP city attracts proportionately more skilled workers. In particular, we show that the skill distribution in the high-TFP city stochastically dominates in first order the skill distribution in the low-TFP city.

Theorem 3 *Assume* $\gamma_i \equiv \gamma$, $\forall i \in \{1, 2, 3\}$ *and* $\gamma > \theta$. *If* $A_1 > A_2$, *we have that city 1's skill distribution F.O.S.D. city 2's skill distribution.*

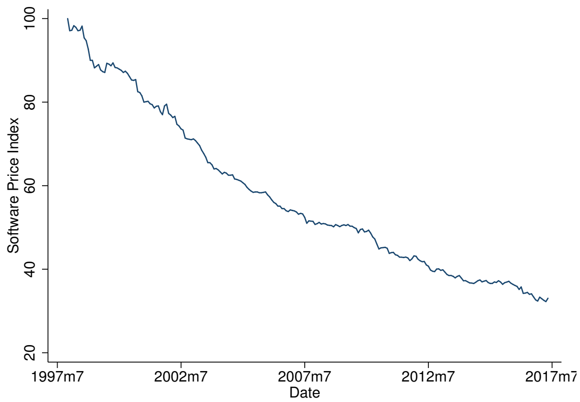
Differently from the case of Automation, **SBTC does not imply that the high-TFP city is larger**. In the appendix, we present two examples that illustrate that results can go either way, i.e., depending on the parameters we may have the high-TFP city to be either larger or smaller than the low-TFP city.

In the next section, we simulate the model in order to get a better understanding of the model's mechanisms and how changes in the parameters may affect the two regions' labor markets. We focus on two parameter changes that are related to the observed evolution of computing power prices over the last twenty years. First, the price of PCs and software went down significantly over this time period. Second, personal computers became significantly more powerful, being able to do operations that needed servers or computer networks previously. While this distinction seems subtle at first sight, it is an important difference for the model. Reductions in price, while increasing the benefit of renting more capital, do nothing to counteract the decreasing marginal contribution of capital. Differently, increases in computer power per machine, by increasing x_k, avoids the decreasing forces of marginal productivity. Moreover, we also believe it is an important distinction in reality. Increasing computer power through the use of servers or connected networks, while possible, demands a lot of coordination and knowledge by its users. These additional user costs reduce the widespread implementation of internal networks and local servers. Moreover, while prices for information technology have gone down, the wide decline in the price indexes for technology, presented in figures (a) and (c) in figure 1 are mostly due to the increase in the processing power which is factored in by the Bureau of Labor Statistics (BLS). Furthermore, even though there is some evidence that the gross investment in personal computers and peripherals has stalled in the latter period, once we control for processing power, the investment in computers has continued to go up, as we present in figures (b) and (d) in figure 1. Consequently, it is important to take into account a potential difference between quality and quantity when we are dealing with changes due to technological progress over time.

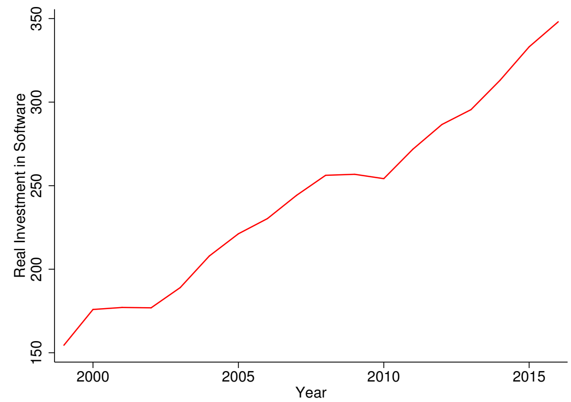
3.3 Numerical Example

3.3.1 Benchmark parametrization

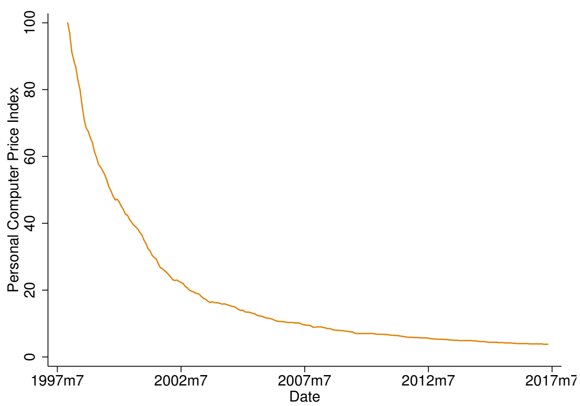
In this section, we show a simple numerical example that illustrates the results of the model. In order to be able to interpret the results more properly, we use results found in the previous literature and data in order to calibrate our parameters. We start using parameter values described by Eeckhout et al. (2014)'s table 2 in order to pin down the values for city TFP and workers' labor productivity. We consider the case that $\gamma_i \equiv \gamma, \forall i \in \{1, 2, 3\}$ and use Eeckhout et al. (2014)'s table 2 to set γ as well. Moreover, we follow Davis and Ortalo-Magné (2011) and set $\alpha = 0.24$. Finally, we must specify values for both θ and the housing stock. We will keep



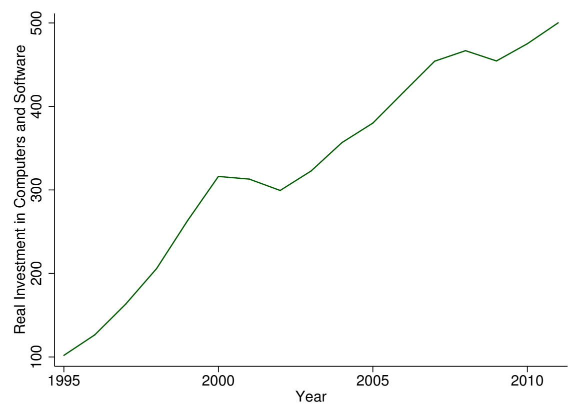
(a) Software's Price Index: 1997–2017
Source: Bureau of Labor Statistics



(b) Real Investment in Software: 1999–2017
Source: Bureau of Economic Analysis



(c) PC's Price Index: 1997–2017
Source: Bureau of Labor Statistics



(d) Real Investment in PCs: 1995–2011
Source: Bureau of Economic Analysis

Figure 1: Price Index and Real Investment in Technology

these values as given at $H_i = 62,559,000, \forall i \in \{1,2\}$ which is close to the BEA's estimate for half of the total housing units for the United States in 2005Q2, and $\theta = 0.85$. We present these parameters in table 1. We assume that these parameters are fixed over time in our numerical exercise.

Table 1: Maintained Parameters – from Eeckhout et al. (2014)

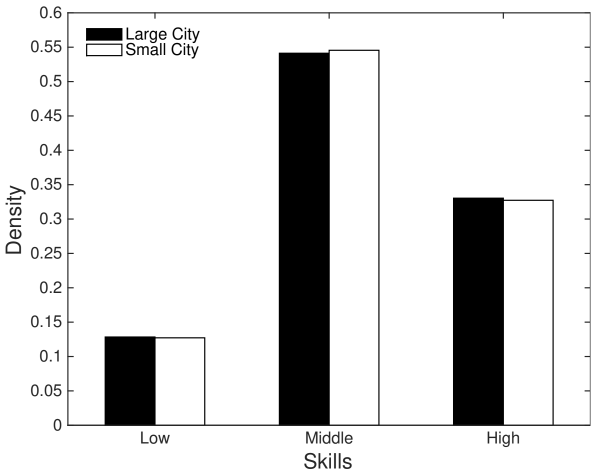
γ	θ	A_1	A_2	y_1	y_2	y_3	H_i	α
0.8	0.85	19,118	9,065	0.3189	1	1.4733	62,559,000	0.24

We then consider two periods in time: 1995 and 2015. We consider changes in the size and composition of the population – measured by the size of the labor force and the distribution across occupations. We follow the distribution of the population across routine and non-routine manual and cognitive occupations for the years 1989 and 2014 as presented by Cortes et al. (2016). We combine routine cognitive and manual occupations to form the middle-skill measure, while we consider non-routine cognitive occupations as high skill and non-routine manual as low skill. Finally, we disregard the unemployed. Similarly, we consider changes in the technology. We pin down x_k by normalizing it at 1 in 1995 and using the estimates for multi-factor productivity (MFP) growth for softwares as presented by Byrne et al. (2017)'s table 3B in order to pin down x_k in 2015. Similarly, in order to consider the changes in the price for technology, we normalize $r = 700$ in 1995 – close to the value that Eeckhout et al. (2014) implied for a middle-skill worker in the small city – and use Byrne and Corrado (2017)'s estimate of price decrease in the cost of ICT investments (Table 4 – software), in order to pin down the value for r in 2015. The calibrated values are presented in table 2.

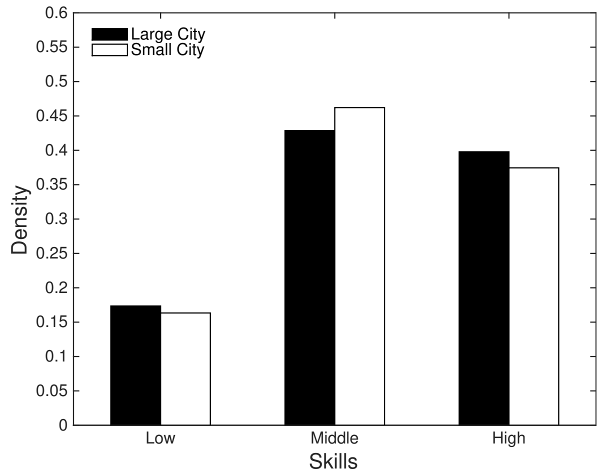
Table 2: Adjusted Parameters - Experiments

	y_k	r	M_1	M_2	M_3
1995	1	700	15,836,150	66,973,717	40,745,094
2015	1.333	635.58	26,640,565	67,576,067	61,078,368

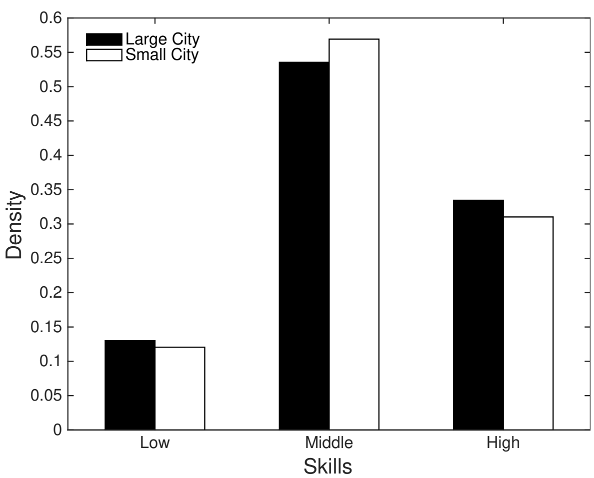
Results are presented in figure 2 and table 3. As we can see from figures 2(a) and 2(b) and table 3's panel B, between 1995 and 2015, city 1 not only became even bigger than city 2, but it also became more unequal – the proportion of mid-skilled workers went down significantly more in city 1 than in city 2. While this result is in line with the overall increase in inequality that we observed over time, jointly showing a geographical component, it does not clearly indicates the underlying reason for this increase in inequality. From our parameters in table 2, we have that many things changed between 1995 and 2015. First, not only the population has grown, but the distribution of skills across the overall population has developed fatter tails. Second, technology became cheaper as well as more productive. In order to disentangle these effects, we consider two counterfactuals. In the first counterfactual, we keep the overall population size and skill distribution at its 1995 levels and only allow technology to become cheaper and more productive, presented in figure 2(c) and in table 3 **Pop. Fixed** lines. In the second counterfactual, we keep technology at its 1995 levels of cost and productivity, while allowing population and skill distribution to adjust to its 2015 levels, presented in 2(d) and in table 3 **Tech. Fixed** lines. As we can see from the results, while changes in population may be responsible for the bulk of the change in the overall shape of the distributions between 1995 and 2015, the changes in technology cost and productivity are the leading factors behind the big cities becoming increasingly more unequal compared to smaller ones.



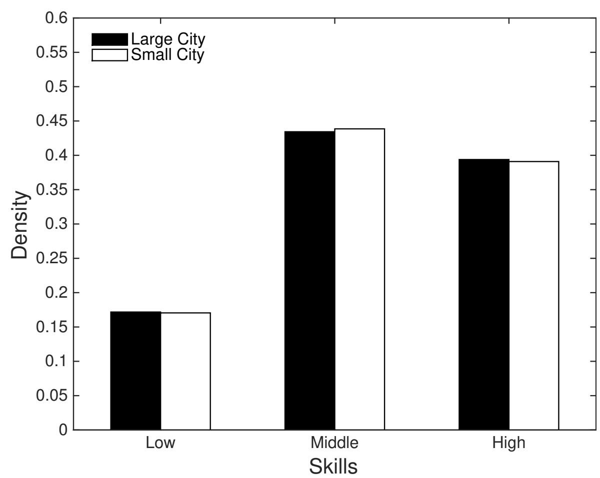
(a) Skill Distribution – 1995



(b) Skill Distribution – 2015



(c) Skill Distribution – Population Constant



(d) Distribution – Technology Constant

Figure 2: Skill Distribution across cities – 1995 vs. 2015

4 Data Sources and Measurement

Data on Workers

Our main data source is the Census Public Use Microdata. We use the 5% Samples for 1980, 1990, and 2000 and for 2013-2015 we combine the American Community Survey yearly files. From these files, we construct labor force and price information at the Metropolitan Statistical Area (MSA) level. The definition of an MSA we use is the 2000 Combined Metropolitan Statistical Areas (CMSA) by the Census for all MSAs that are part of an CMSA and otherwise the MSA itself. For simplicity, we will refer to this definition as MSA from now on. We follow the same procedure as Baum-Snow and Pavan (2013) in order to match the Census Public Use Microdata Area (PUMA) of each Census sample to the 2000 Census Metropolitan Area definitions. The

Table 3: Numerical Exercise Results

Panel A: Prices and Wages

	p_1	p_2	w_{11}	w_{12}	w_{21}	w_{22}	w_{31}	w_{32}
1995	188.38	28.193	184.73	117.10	432.80	274.36	706.45	447.82
2015	224.91	34.466	166.30	106.02	422.28	269.21	650.81	414.91
Pop. Fixed	185.38	28.572	184.48	117.77	422.85	269.95	705.52	450.4
Tech. Fixed	227.91	34.084	166.48	105.52	432.05	273.83	651.53	412.94

Panel B: City Size and Skill Distribution

	S_1	f_{11}	f_{21}	f_{31}	S_2	f_{12}	f_{22}	f_{32}
1995	99,936,000	12.84%	54.12%	33.04%	23,620,000	12.72%	54.55%	32.73%
2015	125,058,000	21.71%	42.86%	39.79%	30,237,100	16.33%	46.21%	37.45%
Pop. Fixed	99,342,000	12.93%	53.54%	33.45%	24,213,500	12.06%	56.92%	31.02%
Tech. Fixed	125,633,000	21.60%	43.43%	39.39%	29,664,100	17.05%	43.85%	39.09%

Census data restricts us to consider only MSAs which are sufficiently large, as they are otherwise not identifiable due to the minimal size of a PUMA. For each year we then construct information on the labor force in each MSA and the local price level. We focus our attention to full-time full-year workers aged 25-54. In order to obtain an estimate of the price level at the MSA level, we consider a simple price index including both consumption goods – which sell at a the same price across different locations – and housing, which is priced differently in each MSA. Based on a hedonic regression using rental data and building characteristics, we calculate the difference in housing values across cities. In large parts of our empirical analysis we focus on the occupational composition of MSAs. To do so, we aggregate the census occupations into broad groups based on their task content as in Cortes et al. (2014). Table 5 shows the classification into groups by task components and the corresponding titles of occupation groups in the Census 2010 Occupation Classification system[6].

Table 4 presents sample averages and standard deviations in the subsample of MSAs for which we have data in all years in the Census and information on technology adoption. We present descriptive statistics for the main variables used in the analysis: occupation shares, employment levels and our MSA rent index.

Technology Data

Our technology data comes from the Ci Technology Database, produced by the Aberdeen Group (formerly known as Harte-Hanks). The data has detailed hardware and software infor-

[6]See `https://www.census.gov/people/io/files/2010_OccCodeswithCrosswalkfrom2002-2011nov04.xls` for the detailed list of Census 2010 Occupations and Cortes et al. (2014) for the mapping to previous Census Occupation Classifications

Table 4: Descriptive Statistics

	1980 mean (st. dev.)	2015 mean (st. dev.)
Occupation Share		
Non-Routine Cognitive	0.32 (0.04)	0.44 (0.05)
Non-Routine Manual	0.09 (0.02)	0.14 (0.02)
Routine Cognitive	0.28 (0.02)	0.23 (0.02)
Routine Manual	0.30 (0.06)	0.19 (0.04)
Rent and Size		
log rent index	0.01 (0.13)	0.01 (0.22)
Employment in 000s	991.78 (1160.60)	1546.51 (1687.15)
Observations	253	253

Note: Averages and standard deviations are weighted by MSA employment. Subsample of MSAs for which we have complete data in all years.

Table 5: Occupation Groups by Tasks

Tasks	Census Occupations
Non-routine Cognitive	Management Business and financial operations Computer, Engineering and Science Education, Legal, Community Service, Arts and Media Occupations Healthcare Practitioners and Technical Occupations
Non-routine Manual	Service Occupations
Routine Cognitive	Sales and Related Office and Administrative Support
Routine Manual	Construction and Extraction Installation, Maintenance and Repair Production Transportation and Material Moving

mation for over 200,000 sites in 2015[7], including not only installed capacity but also expected

[7] In fact, the overall sample is significantly larger than 200,000, but we are restricting the sample to the plants and sites to which we have detailed software information.

future expenses in technology. Their data also includes detailed geographical location for the interviewed sites, as well as aggregation to the firm level. Finally, they also collect some basic information about the sites, such as detailed industry code, number of employees, and total revenue. We have available information for three years – 1990, 1996, and 2015. While the sample seems representative for all 3 available years, we focus on 1996 and 2015, not only due to a larger sample than 1990, but also due to the detailed information on software installation.

Our measure of ERP considers the fraction of plants in the MSAs that adopted ERP softwares. We consider ERPs that help managing the following areas: Accounting, Human Resources, Customer & Sales Force, Collaborative and Integration, Supply Chain Management, as well as bundle softwares like the ones produced by SAP, which are usually called Enterprise Applications. We clean the data in order to eliminate any state-run or governmental department from the sample of sites. Figure 3 shows the geographical dispersion of ERP concentration across the country in 2015. First of all geographical coverage is quite good, with only very few MSAs missing. In fact, the missing MSAs are due to the matching procedure of the Census PUMA to the 2000 Census Metropolitan Area definitions as described by Baum-Snow and Pavan (2013). Moreover, there seems to be a lot of dispersion in the ERP shares across MSAs even in 2015, when we should expect already a more widespread use of technology. These results are corroborated by the results presented in table 6. As we can see, we have at least some information on 280 MSAs across the country. Moreover, we can see that, while on average about 47% of the establishments have at least some form of ERP, there is substantial variation across the country. Some MSAs have a fraction as low as 26%, while others have more than 61% of establishments with some form of ERP. Even more, as we show in figure 4, the degree of adoption seems closely tied to the size as well as cost of living in the MSA, proxied by the rental index.

Table 6: Descriptive Statistics of technology Adoption across MSAs – 2015

	Mean	Median	St. Dev.	Min	Max	N
Share of Sites with ERP	0.47	0.48	0.06	0.26	0.61	280
Number of Sites with ERP in MSA	308.77	85.50	999.60	5.00	14128.00	280
Number of Sites in MSA	625.17	184.50	2090.36	11.00	30541.00	280

5 Empirical Evidence

In this section we describe our main evidence regarding the adoption of automation technology and the occupational composition of cities. We focus on the two main predictions of the theory: (1) locations with higher housing costs should implement automation technology at higher

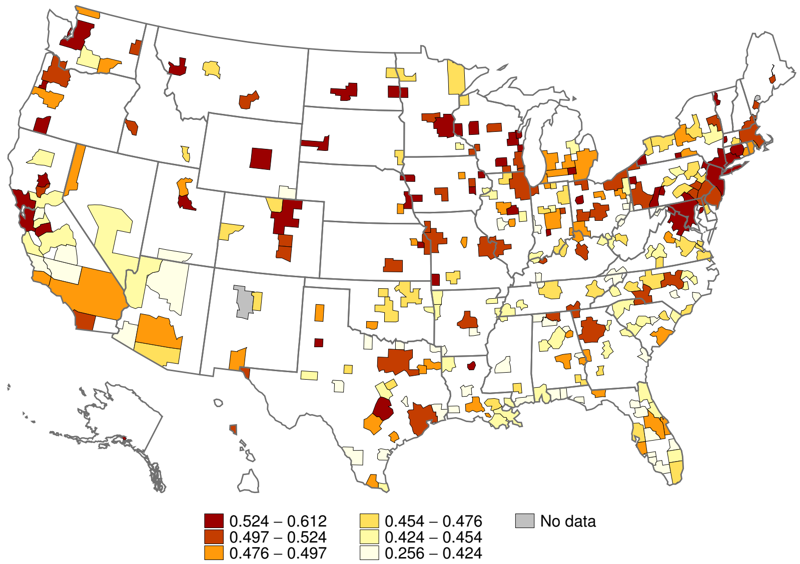


Figure 3: Geographical distribution of ERP adoption shares across MSAs – 2015

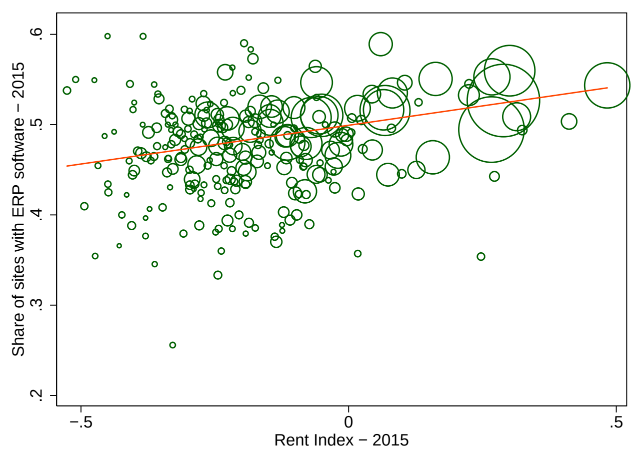


Figure 4: Entreprise Resource Planning Software vs local price level

rates and (2) locations with higher housing costs should also see decreasing shares of their workforce being employed in middle-skill occupations, whose tasks are being replaced by automation technology. To test these predictions, we focus on one specific technology: Enterprise Resource Planning (ERP) software. This type of software allows to automate many tasks usually done by clerical and administrative support staff, which generally fall in the category of routine-cognitive tasks.

Table 7 shows the results of a linear regression of the index of ERP usage on the local price index and the share of routine cognitive workers in 1980. Columns 1 and 2 indicate that, when considered separately, both the current local price level and the past share of routine-cognitive

Table 7: Enterprise Resource Planning Software

	(1) ERP 2015	(2) ERP 2015	(3) ERP 2015
log rent index	0.0858***		0.0777***
	(0.0153)		(0.0180)
routine cognitive share 1980		0.508***	0.205
		(0.133)	(0.137)
Observations	253	253	253
R^2	0.230	0.082	0.242

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Each observation (an MSA) is weighted by its employment in 2015.

workers can explain a substantial amount of variation in ERP adoption. In the first specification, a 10% higher local price index (about half a standard deviation) is associated with a 0.8% increase in the share of sites that use ERP. The most expensive places in our sample have about 8% more sites that use ERP, relative to the cheapest places. In the second specification, a 1% higher the share of routine-cognitive workers in 1980 (about half a standard deviation) is associated with a 0.05% higher share of sites with ERP. However, when considering both variables jointly the effect of the lagged routine-cognitive share conditional on the local price level shrinks substantially and turns out insignificant. Yet, the coefficient on the price level is stable across the specifications. These results indicate that the usage of ERP is more widespread in MSAs with a higher local price level, but conditional on the current price level the past concentration of routine-cognitive workers does not predict the change in concentration. This is in line with the theoretic prediction that cities with high living costs invest more in automation technology.

We turn to the second prediction of the theory: High cost locations should feature a decline in the share of workers, whose tasks can be automated after the introduction of new technology. We use 1980 as the pre-technology period and compare to the occupational composition in 2015. Our focus on such a long span of time is motivated by the fact that we compare steady state predictions of the model and ignore short-term dynamics.

Table 8 presents the results of linear regressions of the change in the routine-cognitive share of MSAs between 1980 and 2015 on its 1980 level and the local rent index in 2015. Again, columns 1 and 2 present the bivariate specifications and column 3 the model with both covariates. The first column indicates that a 10% higher local house price level is associated with a 0.7% larger drop in the routine-cognitive share over 1980-2015. Thus, the most expensive places have about a 7% larger drop in the routine-cognitive share relative to the cheapest locations, this is a substantially

Table 8: Change in routine-cognitive share, 1980-2015

	(1)	(2)	(3)
	Δ rout-cog	Δ rout-cog	Δ rout-cog
log rent index	-0.0663***		-0.0395**
	(0.0143)		(0.0143)
routine cognitive share 1980		-0.831***	-0.677***
		(0.0765)	(0.0763)
Observations	253	253	253
R^2	0.307	0.490	0.582

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Each observation (an MSA) is weighted by its employment in 2015.

difference compared to the average routine-cognitive share of 23% in 2015. Column 2 presents the bivariate specification with the share of routine-cognitive workers in 1980. A 1% higher routine-cognitive share in 1980 is associated with a 0.8% larger drop subsequently. In column 3 the results for the multivariate regression are presented. Both variables are strongly related to the decline in the routine-cognitive share of workers, even after accounting for their covariation. However, the partial effect of each is smaller. The effect of a 10% higher house price drops to 0.4% and the effect of a 1% higher initial share drops to 0.7%. This confirms the prediction, that expensive locations should see a decline in the share of workers, whose tasks are replaced by machines, after the introduction of the new technology.

5.1 Robustness Check - Weighting

In the main analysis of the data we weighted each MSA by its employment. This may raise concerns that we are relying too much on large cities to inform our estimates. Therefore, we consider a different weighting scheme to check for robustness. One may argue that we should compare our previous results to results without any weighting. However, our estimates of the ERP share are very noisy for some MSAs due to small sample size. Therefore we use the inverse of the standard error of the ERP index as weights. We calculate the standard error σ_{ERP} according to the following formula

$$\hat{\sigma}_{ERP} = \sqrt{\frac{\hat{\pi}_{ERP}(1 - \hat{\pi}_{ERP})}{N}} \tag{15}$$

where $\hat{\pi}_{ERP}$ is the empirical share of ERP sites in a location and N the number of sites.

In order to examine the differences in empirical results between this weighting scheme and

Table 9: Entreprise Resource Planning Software - weights $\hat{\sigma}_{ERP}^{-1}$

	(1) ERP 2015	(2) ERP 2015	(3) ERP 2015
log rent index	0.0769***		0.0650***
	(0.0141)		(0.0158)
routine cognitive share 1980		0.420***	0.225*
		(0.110)	(0.112)
Observations	253	253	253
R^2	0.114	0.060	0.128

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Each observation (an MSA) is weighted by the inverse of the estimated standard error of the ERP measure.

the original weighting by size table 9 and 10 replicate the corresponding estimations of the main section with the new weighting scheme. The results in table 7 and 9 show the results regarding ERP adoption with the two weighting schemes. We find that the parameter estimates are very similar.

Comparing the results regarding the change in the routine-cognitive share of employment, table 8 and 10, we find again the same pattern. Results are very similar, albeit parameter estimates on rent being slightly smaller in absolute size. From this exercise we conclude that results were not simplz driven by very large MSAs.

Table 10: Change in routine-cognitive share, 1980-2015 $\hat{\sigma}_{ERP}^{-1}$

	(1) Δ rout-cog	(2) Δ rout-cog	(3) Δ rout-cog
log rent index	-0.0700***		-0.0326***
	(0.00903)		(0.00902)
routine cognitive share 1980		-0.801***	-0.703***
		(0.0466)	(0.0426)
Observations	253	253	253
R^2	0.240	0.558	0.601

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Each observation (an MSA) is weighted by the inverse of the estimated standard error of the ERP measure.

Finally, in appendix C we replace the log rent index by log employment. As presented in section 2, the model delivers an equilibrium in which not only the more productive city is more expensive, but it is also bigger. As we can see from tables 16 and 17, results are qualitatively the same as the ones presented here.

5.2 Measures of Concentration

We now calculate measures of concentration of skills across regions. These measures allow us to test if we have observed an increase in the spatial dispersion of skills across MSAs in the last 30 years. Moreover, these measures abstract from issues of long-run trends in the composition of labor force. Consequently, we are able to focus on the correlation between the spatial dispersion of skills and MSAs characteristics – in particular size and cost of housing. We consider three simple measures: The location quotient that compares the skill distribution in the MSA against the overall skill distribution in the economy, the Ellison and Glaeser (1997) index of industry concentration, and an adjusted version of this index proposed by Oyer and Schaefer (2016). The latter two indexes attempt to measure concentration by comparing it against a distribution that would be obtained by chance (the "dartboard approach").

5.2.1 Location Quotient

As a first pass, we consider a concentration measure that compares the distribution in a given MSA against the overall economy distribution. In particular, we consider that the degree of concentration of skill i in city j (λ_{ij}) is given by:

$$\lambda_{ij} = \frac{\frac{m_{ij}}{S_j}}{\frac{M_i}{\sum_{l=1}^{N} M_l}} \tag{16}$$

Intuitively, if a MSA is more concentrated in skill level i than the economy at large, this index's value would be above 1. Moreover, this measure has two additional benefits. First, by focusing on shares, it reduces the impact of the MSA's overall size on the analysis. Second, by comparing the region against the economy-wide distribution, it takes into account the potential changes in the national labor market. Consequently, it allows us to focus on the increase/decrease of concentration across regions as well as how it correlates to these regions' characteristics.

In our analysis, we consider two time periods – 1980 and 2015. Moreover, following Cortes et al. (2016), we divide the occupations in 4 groups: non-routine manual, routine manual, routine cognitive, and non-routine cognitive. We divide the regions in two groups around the median. As a first pass, we divide MSAs in terms of the size of its labor force, i.e., large vs. small. Similar results are obtained if we use the log rent index, i.e. cheap vs. expensive, as the measure to separate the MSAs. Results are presented in table 11.

Table 11: Simple Measure of Concentration across skill and city size groups

Panel A: 1980								
	Non-Routine Manual		*Routine Manual*		*Routine Cognitive*		*Non-Routine Cognitive*	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Large City	0.99	0.95	1.08	1.05	0.97	0.97	0.95	0.95
Small City	1.05*	1.03†	1.11	1.11	0.92**	0.91††	0.93	0.90
Panel B: 2015								
	Non-Routine Manual		*Routine Manual*		*Routine Cognitive*		*Non-Routine Cognitive*	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Large City	0.99	0.96	1.07	1.05	1.02	1.01	0.96	0.97
Small City	1.02	1.02	1.21**	1.19††	1.00	0.99	0.90*	0.89††

**,* represent significant at 1 and 5 % respectively in a t-test of means with unequal variances.
††,† represent significant at 1 and 5% respectively in a Wilcoxon rank-sum test of medians.

As we can see from table 11, in 1980, small cities had on average a higher concentration in non-routine manual jobs, a lower concentration in routine cognitive jobs, and were at par in routine manual and non-routine cognitive once compared to large cities. Differently, in 2015 we see small cities being on average more concentrated in routine manual jobs, less concentrated in non-routine cognitive jobs, and at par in routine cognitive and non-routine manual jobs. Taken as a whole, table 11 shows an increase in the concentration of routine cognitive and routine manual jobs in small cities, jointly with a decrease in non-routine manual and non-routine cognitive jobs, as expected from our theory.

Finally, figure 5 presents the density distribution of the simple concentration index for small and large cities across skill groups and time. While we observe that there is significant variance in this index across CMSAs, the overall message is the same as the one presented in table 11.

5.2.2 Ellison-Glaeser (1997) Index

We now adapt the concentration index presented by Ellison and Glaeser (1997) for the skill distribution context. Denote γ_i as the EG concentration index for skill i. To define this index, we first introduce some notation. Define s_{ij} as the share of workers of skill i in city j, i.e., $s_{ij} = \frac{m_{ij}}{M_i}$. Let x_j be the share of total employment in city j, i.e., $x_j = \frac{S_j}{\sum_{l=1}^{N} M_l}$. Then, our measure of spatial concentration of skill i is given by:

$$\gamma_i = \frac{\sum_j (s_{ij} - x_j)^2}{1 - \sum_j x_j^2} \tag{17}$$

According to Ellison and Glaeser (1997), there are several advantages in using this index. First, it is easy to compute with readily available data. Second, the scale of the index allows us to make comparisons with a no-agglomeration case in which the data is generated by the simple dartboard model of random location choices (in which case $E(\gamma_i) = 0$). Finally, the index is comparable across populations of different skill sizes. Notice that in this case, we have one index per skill group per year. Consequently, we are unable compare large and small cities. However, we are able to see if skill groups became more or less concentrated across cities over time.

Table 12: Ellison-Glaeser Index

	1980	**2015**	**% Change**
Non-Routine Manual	0.00063	0.00044	-0.29659
Routine Manual	0.00080	0.00068	-0.15094
Routine Cognitive	0.00011	0.00014	0.24356
Non-Routine Cognitive	0.00026	0.00029	0.11259

Results are presented in table 12. As we can see, manual occupations have seen a decline in concentration, whereas cognitive occupations have seen a (small) increase in concentration. These results complement the results regarding the location index, by indicating how concentration of each occupation group has changed across cities. While these results are generally in line with what we should expect given our model's outcomes, we are not able to precisely link them to city characteristics. In order to do that, in the next section we follow Oyer and Schaefer (2016) and adapt the Ellison and Glaeser (1997) to create a city's skill concentration index.

5.2.3 Oyer-Schaefer (2016) Index

We now consider an adapted version of the EG concentration index that we call the Oyer-Schaefer index (henceforth OS index). Hence, denote ζ_j the OS concentration index for city j. To define this index, we first introduce some notation. Define $\tilde{x}_i$ the overall share of workers of skill i in the economy, i.e. $\tilde{x}_i = \frac{M_i}{\sum_{l=1}^{N} M_l}$. Similarly, define $\tilde{s}_{ij}$ the share of workers of skill i in city j, i.e., $\tilde{s}_{ij} = \frac{m_{ij}}{S_j}$, where S_j is city j's labor force size. Then, the OS index is define as:

$$\zeta_j = \frac{S_j}{S_j - 1} \frac{\sum_i (\tilde{s}_{ij} - \tilde{x}_i)^2}{1 - \sum_i \tilde{x}_i^2} - \frac{1}{S_j - 1} \tag{18}$$

Differently from the EG index, in the OS index we are able to compare the degree of concentration across city sizes or across cities with different housing costs. Unfortunately, we are unable to pin down the source of the increase/decrease in within-city concentration. In particular, we are unable to tie the changes in concentration to changes in the shares of each particular skill group. In this sense, EG and OS indexes, while complementing each other, both have its weaknesses and do not give a complete picture of the changes in concentration.

Table 13 presents the results for 1980 and 2015. As we can see, in both periods, small cities are consistently more concentrated than large cities, although there is also more variance of concentration across small cities. Furthermore, while both small and large cities have seen a reduction in concentration over time, the reduction has been on average larger at large cities.

Finally, we present the changes in the density distribution of the OS index in figure 6. As we can see, the distribution of the OS index became more concentrated as we move from 1980 to 2015.

Table 13: OS Index across city sizes and time

Panel A: 1980					
	Mean	Median	St. Dev.	Min	Max
Large City	0.01193	0.00551	0.01732	0.00012	0.10032
Small City	0.01879	0.00965	0.02132	0.00037	0.11660
Panel B: 2015					
	Mean	Median	St. Dev.	Min	Max
Large City	0.00896	0.00406	0.01156	0.00014	0.06074
Small City	0.01835	0.01259	0.01738	0.00003	0.10652

6 Estimation

In order to complement the descriptive evidence in the previous sections and to perform quantitative counterfactuals we estimate an extended version of the model. The extended model embeds a more realistic housing market by introducing Stone-Geary preferences and a finite supply elasticity of housing. Furthermore, the production function is generalized to allow for differential returns to scale of labor by occupation and a finite elasticity of substitution between occupation level output.

The identification of the main model parameters is relatively simple, given the static structure of the model. In the following we shortly introduce the model extensions and then discuss identification of the main model parameters. The identification arguments also deliver a practical estimation protocol.

6.1 Extended Model Setup

We extend the model to capture the key features of housing, labor and capital allocations in the data.

[!Add workers i here!] There are J locations each endowed with amenities $a_{j,i}$, TFP A_j and housing supply schedule $p_j(h) = h_j^0 h^{\epsilon_j}$.

Furthermore, we allow for Stone-Geary preferences

$$u_{j,i}(c,h) = a_{j,i} c^{1-\alpha_i} (h - \underline{h}_i)^{\alpha_i} \tag{19}$$

Thus, housing spending shares can vary with the local price level.

$$A_j F_j(\boldsymbol{m}, \boldsymbol{k}, \boldsymbol{x}, \boldsymbol{x^k},) = A_j \left\{ \sum_i \left[(m_{j,i} x_i)^{\gamma_i \phi_i} + (k_{j,i} x_i^k)^{\gamma_i} \right]^{\frac{\lambda}{\gamma_i}} \right\}^{\frac{\beta}{\lambda}} \tag{20}$$

As before factor remuneration is determined competitively, workers and capital are paid according to their marginal product. Wages are endogenously determined in equilibrium, but the rental rate of capital r is taken as given. The definition of a spatial equilibrium remains the same as before.

6.2 Identification and Estimation Protocol

In this section we present how we identify the structural parameters of the model. As the identification arguments are mostly based on linear equations they can be directly estimated.

The utility function parameters $(\alpha_i, \underline{h}_i)$ are related to the spending share on housing and its covariation with the local price level, see equation (21).

$$\frac{p_j h_{j,i}}{w_{j,i}} = \alpha_i + (1-\alpha_i)\underline{h}_i \frac{p_j}{w_{j,i}} \tag{21}$$

Local spending on housing $p_j h_{j,i}$ and wages $w_{j,i}$ are directly observable in the data. Local house prices p_j are estimated as a hedonic index. The parameters α_i and $\underline{h}_i$ can be identified as long as $J \geq 2$ under the assumption that both α_i and $\underline{h}_i$ do not vary systematically with unobserved local characteristics.

The production function parameters can be identified up to some basic normalization.

Suppose β is known. Then we are left with the labor productivity parameters x_i, capital productivity x_i^k, the elasticity of substitution between capital and labor governed by γ_i and the labor returns to scale ϕ_i.

The ratio of the marginal product of capital to labor allows us to identify ϕ_i and γ_i independently of other occupations output or local TFP. The productivity of capital and labor can only be identified in relative terms from this equation. At this point we normalize one of x_i, this is just to set the scale of the model productivity parameters.

$$\log\left(\frac{r}{w_{ij}}\right) = \gamma_i \log\left(\frac{x_i^k}{\phi_i x_i^{\phi_i}}\right) + (\gamma_i - 1)\log(k_{j,i}) - (\gamma_i\phi_i - 1)\log(m_{j,i}) \tag{22}$$

This equation can be readily estimated and separately identifies γ_i and ϕ_i. However, from this equation we can not yet separate x_i and x_i^k.

In occupations that do not use capital equation (22) does not apply. Let us denote such an occupation by i'. For occupations i' divide its wage by wage of occupation i. We are left with an equation that is that is linear in $\phi_{i'}$ and λ. All other terms are known at this point.

$$\begin{aligned}\log(w_{j,i'}) - \log(w_{j,i}) =& \lambda\phi_{i'}\log(x_{i'}) + (\phi_{i'}\lambda - 1)\log(m_{j,i'}) - \left(\frac{\lambda}{\gamma_i} - 1\right)\log\left[(m_{j,i}x_i)^{\gamma_i\phi_i} + (k_{j,i}x_i^k)^{\gamma_i}\right]\dots \\ & - \log(\lambda) - \log\left[\phi_i\,(x_i)^{\gamma_i\phi_i}\,(m_{j,i})^{\gamma_i\phi_i - 1}\right]\end{aligned} \tag{23}$$

The equations used to show how to identify the parameters can be directly estimated. Note that identification can be achieved without the knowledge of TFP and local amenities. Threats to identification would be local factor specific productivity that systematically varies with the amount of workers.

Note that the production function specifies that only local inputs are used in production.

Therefore if inputs are actually freely tradable across locations we will obtain an elasticity of substitution between occupation outputs of ∞ ($\lambda = 1$). Therefore, the potential misspecification of the trade structure does not invalidate the estimation protocol.

TFP and amenities are pinned down through residuals. TFP is the residual of wages under normalization of β, where we use the average residual of all wages in the city in order to pin down TFP conditional on (log) city size.

6.2.1 Parameter Estimates and Model Fit - Preliminary

Table 14: Estimates – Production Function Parameters

	non-routine cognitive	**non-routine manual**	**routine cognitive**	**routine manual**
γ	1.0	1.0	0.779	1.0
	$(\cdot)$	$(\cdot)$	(0.00106)	$(\cdot)$
ϕ	1.06	1.03	1.06	1.02
	(8.49e-6)	(1.16e-5)	(2.48e-5)	(1.42e-5)
x	1.66	1.48	1.0	2.11
	(0.00351)	(0.00445)	(.)	(0.00947)
x_k	.	.	3.13	.
	$(\cdot)$	$(\cdot)$	(0.0307)	$(\cdot)$

Table 15: Estimates – Stone-Geary Parameters

	$\hat{\alpha}_i$	$\hat{\underline{h}}_i$
Non-Routine Cognitive	0.086	3.842
Non-Routine Manual	0.221	2.086
Routine Cognitive	0.112	3.303
Routine Manual	0.079	3.277

6.3 Counterfactuals – TBD

1. real price of capital r and x_k
2. skill biased technological changes (x)
3. composition M

7 Conclusion

In this paper, we show that the substitution of routine jobs and tasks with machines, computers, and software has not happened evenly in space. In fact, the relative benefit of replacing middle-skill workers that perform routine tasks by computers and software depend on the cost of hiring a worker in this particular location. Consequently, living costs – in particular housing costs – play a key role. Our empirical results show that the share of routine-abstract jobs has gone down proportionately more in expensive and large cities. Moreover, these areas also have seen a larger investment in technologies directly associated with the tasks previously exercised by routine-abstract workers. In order to rationalize the observed empirical patterns, we propose an equilibrium model of location choice by heterogeneously skilled workers where each location is a small open economy in the market for computers and software. We show that if computers are substitutes to middle-skill workers – commonly known as the automation hypothesis – we have that in equilibrium large and expensive cities will invest more in automation, as they are more likely to substitute middle-skill workers with computers. Intuitively, in large and expensive cities, the relative benefit of substituting computers for routine abstract workers is higher than in cheaper and smaller places, since computers have the same price everywhere, while workers must reside locally, having to be compensated for the high local housing prices.

References

Acemoglu, D. and D. Autor (2011): "Skills, tasks and technologies: Implications for employment and earnings," *Handbook of labor economics*, 4, 1043–1171.

Autor, D. and D. Dorn (2013): "The growth of low-skill service jobs and the polarization of the US labor market," *The American Economic Review*, 103, 1553–1597.

Baum-Snow, N. and R. Pavan (2013): "Inequality and city size," *Review of Economics and Statistics*, 95, 1535–1548.

Bloom, N., L. Garicano, R. Sadun, and J. Van Reenen (2014): "The distinct effects of information technology and communication technology on firm organization," *Management Science*, 60, 2859–2885.

Byrne, D., S. Oliner, and D. Sichel (2017): "Prices of High-Tech Products, Mismeasurement, and Pace of Innovation," .

Byrne, D. M. and C. A. Corrado (2017): "ICT Prices and ICT Services: What do they tell us about Productivity and Technology?" .

Cortes, G. M., N. Jaimovich, C. J. Nekarda, and H. E. Siu (2014): "The micro and macro of disappearing routine jobs: A flows approach," Tech. rep., National Bureau of Economic Research.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2016): "Disappearing Routine Jobs: Who, How, and Why?" .

Davis, M. A. and F. Ortalo-Magné (2011): "Household expenditures, wages, rents," *Review of Economic Dynamics*, 14, 248–261.

Eeckhout, J., R. Pinheiro, and K. Schmidheiny (2014): "Spatial Sorting," *Journal of Political Economy*, 122, 554–620.

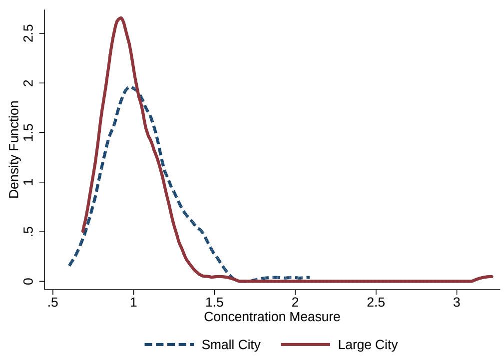
Ellison, G. and E. L. Glaeser (1997): "Geographic concentration in US manufacturing industries: a dartboard approach," *Journal of political economy*, 105, 889–927.

Goldin, C. D. and L. F. Katz (2009): *The race between education and technology*, Harvard University Press.

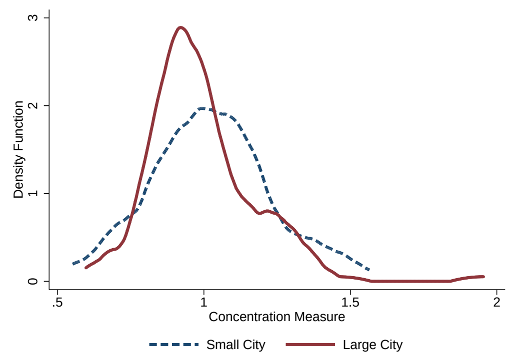
Jovanovic, B. and P. L. Rousseau (2005): "General purpose technologies," *Handbook of economic growth*, 1, 1181–1224.

Lucas, R. E. and E. Rossi-Hansberg (2002): "On the internal structure of cities," *Econometrica*, 70, 1445–1476.

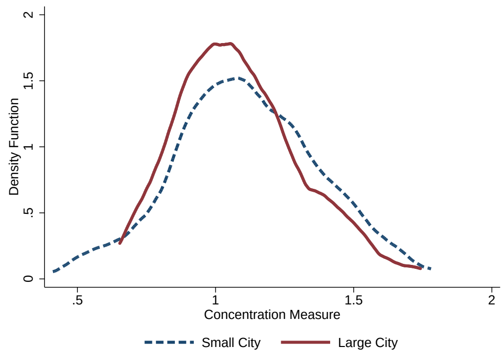
Oyer, P. and S. Schaefer (2016): "Firm/Employee Matching: An Industry Study of US Lawyers," *ILR Review*, 69, 378–404.



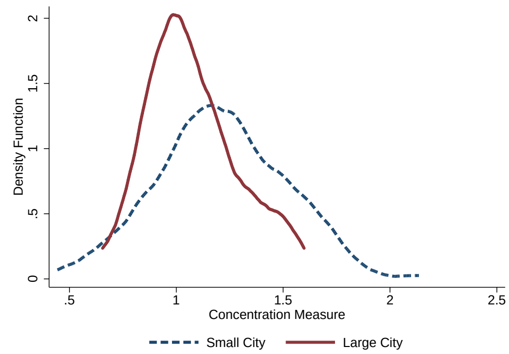
(a) Non-Routine Manual: 1980



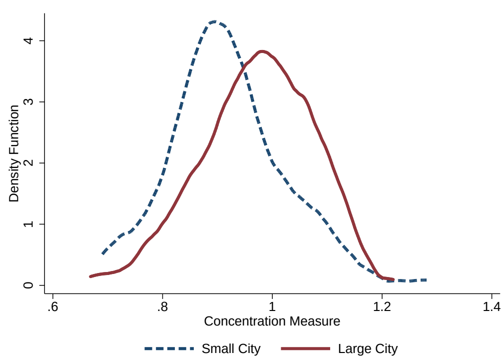
(b) Non-Routine Manual: 2015



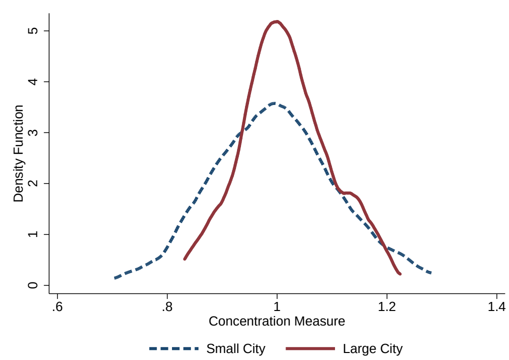
(c) Routine Manual: 1980



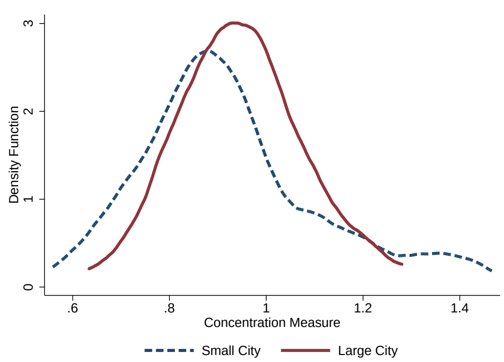
(d) Routine Manual: 2015



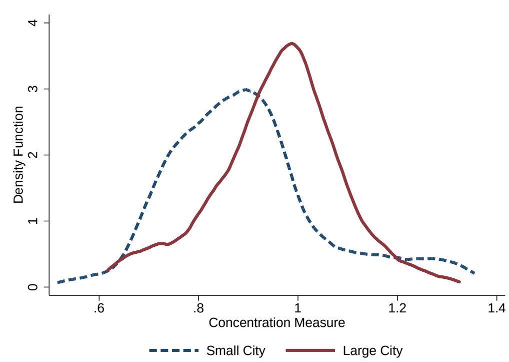
(e) Routine Cognitive: 1980



(f) Routine Cognitive: 2015

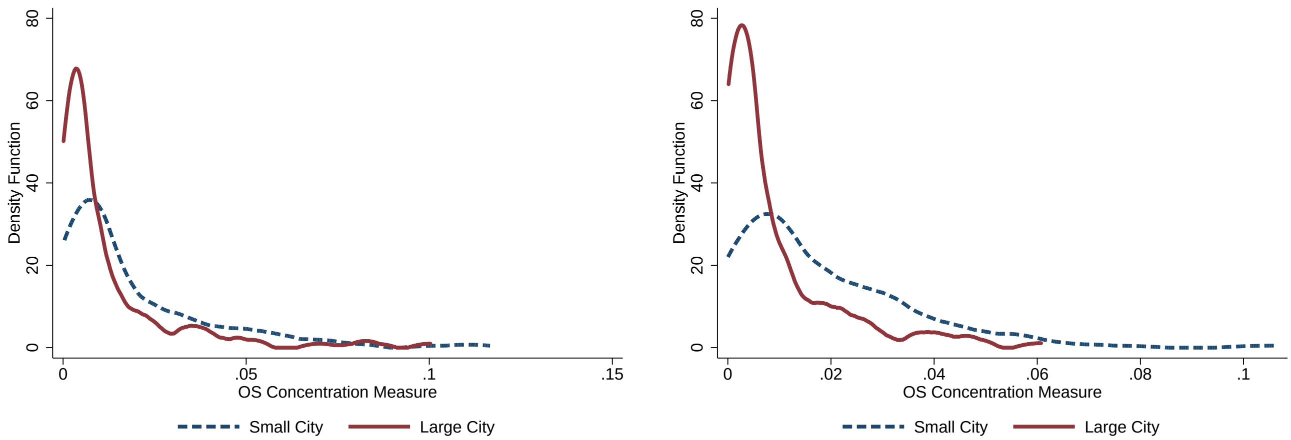


(g) Non-Routine Cognitive: 1980



(h) Non-Routine Cognitive: 2015

Figure 5: Skill Distribution across city sizes and time



(a) 1980

(b) 2015

Figure 6: Distribution of OS index across city sizes and time

A Theory - Preliminary Steps - Automation

A.1 Automation

Closing the Model

The final steps to close the model involve simplifying the model such that we have a system with only two equations and two unknowns (k_1 and $\frac{p_1}{p_2}$. Based on the calculations presented in the paper for k_2, k_1 and their respective FOCs, we obtain:

$$F_j(m_{1j}, m_{2j}, m_{3j}, k_j) = A_j \left[m_{1j}^{\gamma_1} x_1 + \left(m_{2j}^{\theta} x_2 + k_j^{\theta} x_k \right)^{\frac{\gamma_2}{\theta}} + m_{3j}^{\gamma_3} x_3 \right] \tag{24}$$

FOCs:

$$\begin{aligned}
&(m_{1j}) : A_j \gamma_1 m_{1j}^{\gamma_1 - 1} x_1 = w_{1j} \\
&(m_{2j}) : A_j \gamma_2 \left(m_{2j}^{\theta} x_2 + k_j^{\theta} x_k \right)^{\frac{\gamma_2}{\theta} - 1} m_{2j}^{\theta - 1} x_2 = w_{2j} \\
&(m_{3j}) : A_j \gamma_3 m_{3j}^{\gamma_3 - 1} x_3 = w_{3j} \\
&(k_j) : A_j \gamma_2 \left(m_{2j}^{\theta} x_2 + k_j^{\theta} x_k \right)^{\frac{\gamma_2}{\theta} - 1} k_j^{\theta - 1} x_k = r
\end{aligned}$$

Since from utility equalization, we have:

$$\frac{w_{ij}}{w_{ij'}} = \left(\frac{p_j}{p_{j'}} \right)^{\alpha}, \quad \forall i \in \{1, 2, 3\} \text{ and } \forall j \in \{1, 2\} \tag{25}$$

From (m_{11}), (m_{12}), and feasibility condition for skill 1, we have:

$$m_{11} = \frac{\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1 - 1}} M_1}{1 + \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1 - 1}}} \tag{26}$$

Similarly, for skill 3:

$$m_{31} = \frac{\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3 - 1}} M_3}{1 + \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3 - 1}}} \tag{27}$$

From (m_{21}), (k_1), (m_{22}), (k_2), labor market clearing, and the utility equalization condition, we have:

$$\left(\frac{m_{21}}{m_{22}}\right) = \left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{\theta - 1}} \frac{k_1}{k_2} \tag{28}$$

Now let's go back to the expression for (k_1). Manipulating it, we have that:

$$m_{21} = \left\{\frac{1}{x_2}\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k\right]\right\}^{\frac{1}{\theta}} k_1 \tag{29}$$

Similarly, for (k_2), we have:

$$m_{22} = \left\{\frac{1}{x_2}\left[\left(\frac{r}{A_2\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2 - \theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k\right]\right\}^{\frac{1}{\theta}} k_2 \tag{30}$$

Dividing (56) by (57)and substituting (55), we have:

$$\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta - 1}} = \left\{\frac{\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k\right]}{\left[\left(\frac{r}{A_2\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2 - \theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k\right]}\right\} \tag{31}$$

Manipulating and simplifying it, we have:

$$k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} = \left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_2 - \theta}} \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} + \left(\frac{r}{A_2\gamma_2 x_k}\right)^{\frac{\theta}{\theta - \gamma_2}} \left[1 - \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1 - \theta}}\right] x_k$$

Now, we also can use the fact that $m_{21} + m_{22} = M_2$. Then, we have that:

$$M_2 x_2^{\frac{1}{\theta}} = \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}} k_1 + \left[\left(\frac{r}{A_2 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}} k_2 \tag{32}$$

Substituting (58) and manipulating, we have:

$$k_2 = \frac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}}} \tag{33}$$

Substituting (60) into (59) and manipulating, we have:

$$\begin{gathered} \left\{ \frac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} = \\ = \left(\frac{A_2}{A_1} \right)^{\frac{\theta}{\gamma_2-\theta}} \left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} + \left(\frac{r}{A_2 \gamma_2 x_k} \right)^{\frac{\theta}{\theta-\gamma_2}} \left[1 - \left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{1-\theta}} \right] x_k \end{gathered} \tag{34}$$

which implicitly pins down k_1 as a function of $\frac{p_1}{p_2}$.

Finally, in order to pin down the equilibrium, we need to work with the housing market equilibrium conditions. Looking at the ratio of the housing market clearing conditions, we have:

$$\frac{w_{11} m_{11} + w_{21} m_{21} + w_{31} m_{31}}{w_{12} m_{12} + w_{22} m_{22} + w_{32} m_{32}} = \frac{p_1}{p_2}$$

Now substituting wages and labor demands and rearranging it, we have:

$$\begin{gathered} \left\{ \begin{array}{c} \left(m_{21}^{\theta} x_2 + k_1^{\theta} x_k \right)^{\frac{\gamma_2-\theta}{\theta}} m_{21}^{\theta} x_2 - \\ - \frac{A_2}{A_1} \frac{p_1}{p_2} \left(m_{22}^{\theta} x_2 + k_2^{\theta} x_k \right)^{\frac{\gamma_2-\theta}{\theta}} m_{22}^{\theta} x_2 \end{array} \right\} = \\ \left\{ \begin{array}{c} \left(\frac{M_1}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_1 - 1}}} \right)^{\gamma_1} x_1 \left[\frac{A_2}{A_1} \frac{p_1}{p_2} - \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{\gamma_1}{\gamma_1 - 1}} \right] \\ + \left(\frac{M_3}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_3 - 1}}} \right)^{\gamma_3} x_3 \left[\frac{A_2}{A_1} \frac{p_1}{p_2} - \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{\gamma_3}{\gamma_3 - 1}} \right] \end{array} \right\} \end{gathered} \tag{35}$$

Then, from the ratio of (m_{21}) and (m_{22}), we have:

$$\left(m_{22}^{\theta}x_2 + k_2^{\theta}x_k\right)^{\frac{\gamma_2-\theta}{\theta}} = \left(\frac{p_2}{p_1}\right)^{\alpha}\left(m_{21}^{\theta}x_2 + k_1^{\theta}x_k\right)^{\frac{\gamma_2-\theta}{\theta}} \times \left(\frac{m_{21}}{m_{22}}\right)^{\theta-1} \times \left(\frac{A_1}{A_2}\right) \tag{36}$$

Substituting (63) into (62) and rearranging, we have:

$$\begin{gathered}\left\{ \left[1 - \left(\frac{p_1}{p_2}\right)^{1-\alpha}\frac{M_2-m_{21}}{m_{21}}\right]\left(m_{21}^{\theta}x_2 + k_1^{\theta}x_k\right)^{\frac{\gamma_2-\theta}{\theta}} m_{21}^{\theta}x_2 \right\} = \\ \left\{ \begin{array}{l} \left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1 \left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right] \\ + \left(\frac{M_3}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3} x_3 \left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_3}{\gamma_3-1}}\right] \end{array} \right\}\end{gathered} \tag{37}$$

But then, from equation (56), we have that:

$$m_{21}^{\theta}x_2 = \left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\theta)}{\gamma_2-\theta}} - k_1^{\theta}x_k \tag{38}$$

Similarly, from (k_1), we have:

$$\left(m_{21}^{\theta}x_2 + k_1^{\theta}x_k\right)^{\frac{\gamma_2-\theta}{\theta}} = \left(\frac{r}{A_1\gamma_2x_k}\right)k_1^{1-\theta} \tag{39}$$

Then, from (65) and (66), we have:

$$\left(m_{21}^{\theta}x_2 + k_1^{\theta}x_k\right)^{\frac{\gamma_2-\theta}{\theta}} m_{21}^{\theta}x_2 = \left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\gamma_2}{\gamma_2-\theta}} k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2-\theta}} - \frac{r}{A_1\gamma_2}k_1 \tag{40}$$

Substituting equation (60) into (55) and manipulating, we have:

$$\frac{M_2 - m_{21}}{m_{21}} = \frac{M_2x_2^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}}{k_1\left[\left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}} \tag{41}$$

Consequently:

$$\left[1-\left(\frac{p_1}{p_2}\right)^{1-\alpha}\frac{M_2-m_{21}}{m_{21}}\right]=\frac{\left\{\begin{array}{c}\left(1+\left(\frac{p_1}{p_2}\right)^{1-\alpha}\right)k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}\\ -\left(\frac{p_1}{p_2}\right)^{1-\alpha}M_2x_2^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}} \tag{42}$$

Then, from equations (67) and (69), we have that:

$$\left[1-\left(\frac{p_1}{p_2}\right)^{1-\alpha}\frac{M_2-m_{21}}{m_{21}}\right]\left(m_{21}^{\theta}x_2+k_1^{\theta}x_k\right)^{\frac{\gamma_2-\theta}{\theta}}m_{21}^{\theta}x_2=$$
$$\frac{\left\{\begin{array}{c}\left(1+\left(\frac{p_1}{p_2}\right)^{1-\alpha}\right)k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}\\ -\left(\frac{p_1}{p_2}\right)^{1-\alpha}M_2x_2^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}}\times\left\{\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\gamma_2}{\gamma_2-\theta}}k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2-\theta}}-\frac{r}{A_1\gamma_2}k_1\right\} \tag{43}$$

Notice that the LHS of equation (70) is the same of the one of equation (64). Substituting it back, we have:

$$\frac{\left\{\begin{array}{c}\left(1+\left(\frac{p_1}{p_2}\right)^{1-\alpha}\right)k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}\\ -\left(\frac{p_1}{p_2}\right)^{1-\alpha}M_2x_2^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}}\times\left\{\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\gamma_2}{\gamma_2-\theta}}k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2-\theta}}-\frac{r}{A_1\gamma_2}k_1\right\}=$$
$$\left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}x_1\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right]\\ +\left(\frac{M_3}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3}x_3\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_3}{\gamma_3-1}}\right]\end{array}\right\} \tag{44}$$

Finally, notice that equations (71) and (61) generate a system with two equations and two unknowns (k_1 and $\frac{p_1}{p_2}$):

$$\left\{\begin{array}{l} \dfrac{\left\{\begin{array}{c} \left(1+\left(\frac{p_1}{p_2}\right)^{1-\alpha}\right) k_1 \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}} \\ -\left(\frac{p_1}{p_2}\right)^{1-\alpha} M_2 x_2^{\frac{1}{\theta}} \end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}} \times \left\{\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\gamma_2}{\gamma_2-\theta}} k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2-\theta}} - \frac{r}{A_1\gamma_2}k_1\right\} = \\ \left\{\begin{array}{l} \left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1 \left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right] \\ + \left(\frac{M_3}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3} x_3 \left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_3}{\gamma_3-1}}\right] \end{array}\right\} \qquad (F.1) \\ \left\{\dfrac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} = \qquad (F.2) \\ = \left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_2-\theta}} \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} + \left(\frac{r}{A_2\gamma_2 x_k}\right)^{\frac{\theta}{\theta-\gamma_2}} \left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right] x_k \end{array}\right.$$

Preliminary Results

In this subsection, we present some preliminary results that will help us to show the main results presented in the paper.

Lemma A.1: *The distribution of skills across cities is identical if and only if* $\frac{m_{i1}}{m_{i2}} = constant, \forall i \in \{1,2,3\}$.

Proof: ($\Rightarrow$) Consider that the distribution across cities is constant, then $pdf_{i1} = pdf_{i2}, \forall i \in \{1,2,3\}$, i.e.:

$$\frac{m_{i1}}{m_{11}+m_{21}+m_{31}} = \frac{m_{i2}}{m_{12}+m_{22}+m_{32}} \tag{45}$$

But that means that $\frac{m_{i1}}{m_{i2}} = \eta = \frac{S_1}{S_2} = \frac{m_{11}+m_{21}+m_{31}}{m_{12}+m_{22}+m_{32}}$. The other direction is trivial. ■

Lemma A.2: *Assume* $\gamma_2 < \theta$. $p_1 = p_2$ *if and only if* $A_1 = A_2$.

Proof: Towards a contradiction, let's assume that $A_1 = A_2$ and $p_1 > p_2$. From the RHS of

$(F.1)$, we have:

$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1 \left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right] \\ + \left(\frac{M_3}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3} x_3 \left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_3\alpha}{\gamma_3-1}}\right] \end{array} \right\} > 0$$

Since $p_1 > p_2$, $\gamma_1 < 1$, and $\gamma_3 < 1$. Therefore, the LHS of $(F.1)$ must also be positive in order for the equality to be satisfied. Then, from equation (67), we have:

$$\left(m_{21}^{\theta} x_2 + k_1^{\theta} x_k\right)^{\frac{\gamma_2-\theta}{\theta}} m_{21}^{\theta} x_2 = \left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\gamma_2}{\gamma_2-\theta}} k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2-\theta}} - \frac{r}{A_1\gamma_2} k_1$$

So the second term on the LHS of $(F.1)$ must be positive. Moreover, from (66), we have that:

$$k_1 \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}} = m_{21} x_2^{\frac{1}{\theta}} > 0$$

Consequently, in order to satisfy $(F.1)$, we must have:

$$\frac{M_2 x_2^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}} < k_1 \left(\frac{p_1}{p_2}\right)^{\alpha-1}$$

Dividing both sides by $\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}$, we have:

$$\frac{M_2 x_2^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}} < k_1 \left(\frac{p_2}{p_1}\right)^{\left(1+\frac{\alpha\theta}{1-\theta}\right)} \tag{46}$$

Now, from $(F.2)$, we have that, due to $p_1 > p_2$ and $\gamma_2 < \theta$:

$$\frac{M_2 x_2^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}} > \left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}} k_1 \tag{47}$$

Then, notice that:

$$1+\frac{\alpha\theta}{1-\theta}-\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}=1+\frac{\alpha\theta}{1-\theta}\left[1-\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}\right]=1+\frac{\alpha\theta}{1-\theta}\left[\frac{\gamma_2(1-\theta)}{\theta(1-\gamma_2)}\right]>0 \quad (48)$$

Therefore the exponent at $\frac{p_2}{p_1}$ is higher at the RHS of (46). Since $\frac{p_2}{p_1}\in(0,1)$, we have that:

$$k_1\left(\frac{p_2}{p_1}\right)^{\left(1+\frac{\alpha\theta}{1-\theta}\right)}<\left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}}k_1$$

consequently, equations (46) and (47) give us a contradiction.

Now, again towards a contradiction, let's assume $p_2>p_1$. In this case, from the RHS of $(F.1)$, we have:

$$\left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}x_1\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right]\\+\left(\frac{M_3}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3}x_3\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_3\alpha}{\gamma_3-1}}\right]\end{array}\right\}<0$$

Since $p_1<p_2$, $\gamma_1<1$, and $\gamma_3<1$. Therefore, the LHS of $(F.1)$ must also be negative. Since we already showed that the second term in the LHS and the denominator of the first term in the LHS must be positive, this requirement of a negative LHS implies, after dividing both sides by $\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}$:

$$\frac{M_2x_2^{\frac{1}{\theta}}-k_1\left[\left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}}>k_1\left(\frac{p_2}{p_1}\right)^{\left(1+\frac{\alpha\theta}{1-\theta}\right)} \quad (49)$$

Then, from $(F.2)$, since $p_1<p_2$, the last term on the RHS is positive. Consequently, once $\gamma_2<\theta$, we have:

$$\frac{M_2x_2^{\frac{1}{\theta}}-k_1\left[\left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2x_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-x_k\right]^{\frac{1}{\theta}}}<\left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}}k_1 \quad (50)$$

Since:

$$1+\frac{\alpha\theta}{1-\theta}-\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}=1+\frac{\alpha\theta}{1-\theta}\left[1-\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}\right]=1+\frac{\alpha\theta}{1-\theta}\left[\frac{\gamma_2(1-\theta)}{\theta(1-\gamma_2)}\right]>0$$

and $p_2 > p_1$, we have that:

$$k_1\left(\frac{p_2}{p_1}\right)^{\left(1+\frac{\alpha\theta}{1-\theta}\right)} > \left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}} k_1$$

Consequently, equations (49) and (50) give us a contradiction. Therefore, we have that $p_1 = p_2 \Leftrightarrow A_1 = A_2$. ■

A.2 SBTC

The final steps to close the model involve simplifying the model such that we have a system with only two equations and two unknowns (k_1 and $\frac{p_1}{p_2}$. Based on the calculations presented in the paper for k_2, k_1 and their respective FOCs, we obtain:

$$F_j(m_{1j}, m_{2j}, m_{3j}, k_j) = A_j\left[m_{1j}^{\gamma_1}x_1 + \left(m_{3j}^{\theta}x_3 + k_j^{\theta}x_k\right)^{\frac{\gamma_3}{\theta}} + m_{2j}^{\gamma_2}x_2\right] \tag{51}$$

FOCs:

$$\begin{aligned}
&(m_{1j}) : A_j\gamma_1 m_{1j}^{\gamma_1-1}x_1 = w_{1j}\\
&(m_{2j}) : A_j\gamma_2 m_{2j}^{\gamma_2-1}x_2 = w_{2j}\\
&(m_{3j}) : A_j\gamma_3\left(m_{3j}^{\theta}x_3 + k_j^{\theta}x_k\right)^{\frac{\gamma_3}{\theta}-1} m_{3j}^{\theta-1}x_3 = w_{3j}\\
&(k_j) : A_j\gamma_3\left(m_{3j}^{\theta}x_3 + k_j^{\theta}x_k\right)^{\frac{\gamma_3}{\theta}-1} k_j^{\theta-1}x_k = r
\end{aligned}$$

Since from utility equalization, we have:

$$\frac{w_{ij}}{w_{ij'}} = \left(\frac{p_j}{p_{j'}}\right)^{\alpha}, \quad \forall i \in \{1,2,3\} \text{ and } \forall j \in \{1,2\} \tag{52}$$

From (m_{11}), (m_{12}), and feasibility condition for skill 1, we have:

$$m_{11} = \frac{\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}} M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}} \tag{53}$$

Similarly, for skill 2:

$$m_{21} = \frac{\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}} M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}} \tag{54}$$

From (m_{31}), (k_1), (m_{32}), (k_2), labor market clearing, and the utility equalization condition, we have:

$$\left(\frac{m_{31}}{m_{32}}\right) = \left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{\theta-1}} \frac{k_1}{k_2} \tag{55}$$

Now let's go back to the expression for (k_1). Manipulating it, we have that:

$$m_{31} = \left\{ \frac{1}{x_3} \left[\left(\frac{r}{A_1 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right] \right\}^{\frac{1}{\theta}} k_1 \tag{56}$$

Similarly, for (k_2), we have:

$$m_{32} = \left\{ \frac{1}{x_3} \left[\left(\frac{r}{A_2 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_2^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right] \right\}^{\frac{1}{\theta}} k_2 \tag{57}$$

Dividing (56) by (57)and substituting (55), we have:

$$\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta-1}} = \left\{ \frac{\left[\left(\frac{r}{A_1 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right]}{\left[\left(\frac{r}{A_2 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_2^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right]} \right\} \tag{58}$$

Manipulating and simplifying it, we have:

$$k_2^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} = \left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_3-\theta}} \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} + \left(\frac{r}{A_2 \gamma_3 x_k}\right)^{\frac{\theta}{\theta-\gamma_3}} \left[1 - \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}} \right] x_k$$

Now, we also can use the fact that $m_{31} + m_{32} = M_3$. Then, we have that:

$$M_3 x_3^{\frac{1}{\theta}} = \left[\left(\frac{r}{A_1 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right]^{\frac{1}{\theta}} k_1 + \left[\left(\frac{r}{A_2 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_2^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right]^{\frac{1}{\theta}} k_2 \tag{59}$$

Substituting (58) and manipulating, we have:

$$k_2 = \frac{M_3 x_3^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_3 x_k} \right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k \right]^{\frac{1}{\theta}}} \tag{60}$$

Substituting (60) into (59) and manipulating, we have:

$$\left\{\frac{M_3 x_3^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} = \tag{61}$$
$$= \left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_3-\theta}}\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} + \left(\frac{r}{A_2\gamma_3 x_k}\right)^{\frac{\theta}{\theta-\gamma_3}}\left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right] x_k$$

which implicitly pins down k_1 as a function of $\frac{p_1}{p_2}$.

Finally, in order to pin down the equilibrium, we need to work with the housing market equilibrium conditions. Looking at the ratio of the housing market clearing conditions, we have:

$$\frac{w_{11}m_{11} + w_{21}m_{21} + w_{31}m_{31}}{w_{12}m_{12} + w_{22}m_{22} + w_{32}m_{32}} = \frac{p_1}{p_2}$$

Now substituting wages and labor demands and rearranging it, we have:

$$\left\{\begin{array}{c}\left(m_{31}^{\theta}x_3 + k_1^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}x_3 - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta}x_3 + k_2^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta}x_3\end{array}\right\} =$$
$$\left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right] \\ +\left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} x_2\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_2}{\gamma_2-1}}\right]\end{array}\right\} \tag{62}$$

Then, from the ratio of (m_{31}) and (m_{32}), we have:

$$\left(m_{32}^{\theta}x_3 + k_2^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} = \left(\frac{p_2}{p_1}\right)^{\alpha}\left(m_{31}^{\theta}x_3 + k_1^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} \times \left(\frac{m_{31}}{m_{32}}\right)^{\theta-1} \times \left(\frac{A_1}{A_2}\right) \tag{63}$$

Substituting (63) into (62) and rearranging, we have:

$$\left\{\left[1-\left(\frac{p_1}{p_2}\right)^{1-\alpha}\frac{M_3-m_{31}}{m_{31}}\right]\left(m_{31}^{\theta}x_3 + k_1^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}x_3\right\} =$$
$$\left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right] \\ +\left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} x_2\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_2}{\gamma_2-1}}\right]\end{array}\right\} \tag{64}$$

But then, from equation (56), we have that:

$$m_{31}^{\theta} x_3 = \left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\theta)}{\gamma_3-\theta}} - k_1^{\theta} x_k \tag{65}$$

Similarly, from (k_1), we have:

$$\left(m_{31}^{\theta} x_3 + k_1^{\theta} x_k\right)^{\frac{\gamma_3-\theta}{\theta}} = \left(\frac{r}{A_1\gamma_3 x_k}\right) k_1^{1-\theta} \tag{66}$$

Then, from (65) and (66), we have:

$$\left(m_{31}^{\theta} x_3 + k_1^{\theta} x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta} x_3 = \left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\gamma_3}{\gamma_3-\theta}} k_1^{\frac{\gamma_3(1-\theta)}{\gamma_3-\theta}} - \frac{r}{A_1\gamma_3} k_1 \tag{67}$$

Substituting equation (60) into (55) and manipulating, we have:

$$\frac{M_3 - m_{31}}{m_{31}} = \frac{M_3 x_3^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}}}{k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}}} \tag{68}$$

Consequently:

$$\left[1 - \left(\frac{p_1}{p_2}\right)^{1-\alpha} \frac{M_3 - m_{31}}{m_{31}}\right] = \frac{\left\{\begin{array}{c}\left(1 + \left(\frac{p_1}{p_2}\right)^{1-\alpha}\right) k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2}\right)^{1-\alpha} M_3 x_3^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}}} \tag{69}$$

Then, from equations (67) and (69), we have that:

$$\begin{gathered}\left[1 - \left(\frac{p_1}{p_2}\right)^{1-\alpha} \frac{M_3 - m_{31}}{m_{31}}\right] \left(m_{31}^{\theta} x_3 + k_1^{\theta} x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta} x_3 = \\ \frac{\left\{\begin{array}{c}\left(1 + \left(\frac{p_1}{p_2}\right)^{1-\alpha}\right) k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2}\right)^{1-\alpha} M_3 x_3^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - x_k\right]^{\frac{1}{\theta}}} \times \left\{\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\gamma_3}{\gamma_3-\theta}} k_1^{\frac{\gamma_3(1-\theta)}{\gamma_3-\theta}} - \frac{r}{A_1\gamma_3} k_1\right\}\end{gathered} \tag{70}$$

Notice that the LHS of equation (70) is the same of the one of equation (64). Substituting it back, we have:

$$\frac{\left\{\begin{array}{c}\left(1+\left(\frac{p_1}{p_2}\right)^{1-\alpha}\right)k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-x_k\right]^{\frac{1}{\theta}}\\ -\left(\frac{p_1}{p_2}\right)^{1-\alpha}M_3x_3^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-x_k\right]^{\frac{1}{\theta}}}\times\left\{\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\gamma_3}{\gamma_3-\theta}}k_1^{\frac{\gamma_3(1-\theta)}{\gamma_3-\theta}}-\frac{r}{A_1\gamma_3}k_1\right\}=$$

$$\left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}x_1\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right]\\ +\left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2}x_2\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_2}{\gamma_2-1}}\right]\end{array}\right\} \tag{71}$$

Finally, notice that equations (71) and (61) generate a system with two equations and two unknowns (k_1 and $\frac{p_1}{p_2}$):

$$\left\{\begin{array}{l}\frac{\left\{\begin{array}{c}\left(1+\left(\frac{p_1}{p_2}\right)^{1-\alpha}\right)k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-x_k\right]^{\frac{1}{\theta}}\\ -\left(\frac{p_1}{p_2}\right)^{1-\alpha}M_3x_3^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-x_k\right]^{\frac{1}{\theta}}}\times\left\{\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\gamma_3}{\gamma_3-\theta}}k_1^{\frac{\gamma_3(1-\theta)}{\gamma_3-\theta}}-\frac{r}{A_1\gamma_3}k_1\right\}= \\ \qquad \left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}x_1\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right]\\ +\left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2}x_2\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_2}{\gamma_2-1}}\right]\end{array}\right\} \qquad (F.1) \\ \left\{\frac{M_3x_3^{\frac{1}{\theta}}-\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-x_k\right]^{\frac{1}{\theta}}k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-x_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}= \qquad (F.2) \\ =\left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_3-\theta}}\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}+\left(\frac{r}{A_2\gamma_3 x_k}\right)^{\frac{\theta}{\theta-\gamma_3}}\left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right]x_k\end{array}\right.$$

Preliminary Results

In this subsection, we present some preliminary results that will help us to show the main results presented in the paper.

Lemma A.3: *The distribution of skills across cities is identical if and only if* $\frac{m_{i1}}{m_{i2}} = constant, \forall i \in \{1, 2, 3\}$.

Proof: ($\Rightarrow$) Consider that the distribution across cities is constant, then $pdf_{i1} = pdf_{i2}, \forall i \in \{1, 2, 3\}$, i.e.:

$$\frac{m_{i1}}{m_{11} + m_{21} + m_{31}} = \frac{m_{i2}}{m_{12} + m_{22} + m_{32}} \tag{72}$$

But that means that $\frac{m_{i1}}{m_{i2}} = \eta = \frac{S_1}{S_2} = \frac{m_{11}+m_{21}+m_{31}}{m_{12}+m_{22}+m_{32}}$. The other direction is trivial. ■

Lemma A.4: *Assume* $\gamma_3 > \theta$. $p_1 = p_2$ *if and only if* $A_1 = A_2$.

Proof: Towards a contradiction, let's assume that $A_1 = A_2$ and $p_1 > p_2$. Consequently, $\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta-1}} < 1$. From (58), we have $k_1 < k_2$. But then, from equation (55), we obtain $m_{31} < m_{32}$. Finally, from the RHS of (62), we have:

$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1 \left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right] \\ + \left(\frac{M_2}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} x_2 \left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_2\alpha}{\gamma_2-1}}\right] \end{array} \right\} > 0$$

Since $p_1 > p_2$, $\gamma_1 < 1$, and $\gamma_2 < 1$. However, given the results we obtained from (58) and (55), the LHS of (62) gives us:

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta} x_3 + k_1^{\theta} x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta} x_3 - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta} x_3 + k_2^{\theta} x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta} x_3 \end{array} \right\} < 0$$

which is a contradiction.

Similarly, again towards a contradiction, let's consider $A_1 = A_2$ and $p_1 < p_2$. Then $\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta-1}} > 1$. Again from (58), we have $k_1 > k_2$. Similarly, from (55), we obtain $m_{31} > m_{32}$. But then, from (62), we have that:

$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1 \left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right] \\ + \left(\frac{M_2}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} x_2 \left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_2\alpha}{\gamma_2-1}}\right] \end{array} \right\} < 0$$

given $p_1 < p_2$. Then RHS(62) < 0. While

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta}x_3 + k_1^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}x_3 - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta}x_3 + k_2^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta}x_3 \end{array} \right\} > 0.$$

which again gives you a contradiction. Therefore, we have that $p_1 = p_2$. Consequently, we have that $A_1 = A_2 \Rightarrow p_1 = p_2$.

Now, let's show that $p_1 = p_2 \Rightarrow A_1 = A_2$. Assume $p_1 = p_2$. Then, from (55), we have:

$$\frac{m_{31}}{m_{32}} = \frac{k_1}{k_2} \tag{73}$$

From (58) , we have

$$\frac{k_1}{k_2} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{74}$$

Combining (73) and (74), we have:

$$\frac{m_{31}}{m_{32}} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{75}$$

But then, from LHS(62), substituting (73) and (75) given $p_1 = p_2$, we have:

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta}x_3 + k_1^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}x_3 - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta}x_3 + k_2^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta}x_3 \end{array} \right\} = \left[\left(\frac{A_1}{A_2}\right)^{\frac{\gamma_3}{1-\gamma_3}} - \frac{A_2}{A_1}\right]\left(m_{32}^{\theta}x_3 + k_2^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta}x_3 \tag{76}$$

while the RHS(62) gives us:

$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_1}}}\right)^{\gamma_1} x_1 \left[\frac{A_2}{A_1} - \left(\frac{A_1}{A_2}\right)^{\frac{\gamma_1}{1-\gamma_1}}\right] \\ + \left(\frac{M_2}{1+\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_2}}}\right)^{\gamma_2} x_2 \left[\frac{A_2}{A_1} - \left(\frac{A_1}{A_2}\right)^{\frac{\gamma_2}{1-\gamma_2}}\right] \end{array} \right\} \tag{77}$$

Then, consider the case in which $A_1 > A_2$. From (76), we have that LHS(62)> 0, while (77) gives us RHS(62)< 0. Similarly, if $A_1 < A_2$, (76) gives us LHS(62)< 0 while (77) gives us RHS(62)> 0. Consequently, (62) is only satisfied if $A_1 = A_2$, concluding our proof.$\blacksquare$

B Proofs

Proof of Proposition 1

Proof. Towards a contradiction, assume that $A_2 > A_1$ and $p_1 > p_2$. Then, the RHS of $(F.1)$ is positive. Consequently, in order to satisfy $(F.1)$, $(F.1)$'s LHS must also be positive. Following the same argument presented in the proof of Lemma A.2, we have that inequality (46) must hold. Then, from $(F.2)$ we have that, given that $p_1 > p_2$, the last term in $(F.2)$'s RHS – $\left(\frac{r}{A_2\gamma_2 x_k}\right)^{\frac{\theta}{\theta-\gamma_2}}\left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right]x_k$ – is negative. We also know that since $A_2 > A_1$ and $\gamma_2 < \theta$, $\left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_2-\theta}} < 1$. Therefore, $(F.2)$ gives us:

$$\frac{M_2 x_2^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k\right]^{\frac{1}{\theta}}} > \left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}} k_1 \tag{78}$$

Given (48) we have that, once $\frac{p_2}{p_1} \in (0,1)$:

$$k_1\left(\frac{p_2}{p_1}\right)^{\left(1+\frac{\alpha\theta}{1-\theta}\right)} < \left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma}{\theta(1-\gamma)}} k_1$$

Consequently, (46) and (78) give us a contradiction. Following the same procedure we can easily show that $A_1 > A_2$ and $p_2 > p_1$ give us the same contradiction. Since lemma A.3 shows that price equality is only achieved through TFP equality, this concludes our proof. ■

Proof of Proposition 2

Proof. Without loss of generality, assume $A_1 > A_2$, Then, based on proposition 2, we have that $p_1 > p_2$. Then, from equation (58), we have:

$$\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta-1}} = \left[\frac{\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k}{\left(\frac{r}{A_2\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k}\right] \tag{79}$$

Then, since $\theta < 1$, we have $\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta-1}} < 1$. Consequently:

$$\left[\frac{\left(\frac{r}{A_1\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k}{\left(\frac{r}{A_2\gamma_2 x_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k}\right] < 1 \tag{80}$$

Rearranging it:

$$\left(\frac{k_1}{k_2}\right)^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} < \left(\frac{A_1}{A_2}\right)^{\frac{\theta}{\gamma_2-\theta}} \tag{81}$$

Since $\gamma_2 < \theta$, this implies that $\left(\frac{k_1}{k_2}\right)^{\frac{\theta(1-\gamma_2)}{\theta-\gamma_2}} > \left(\frac{A_1}{A_2}\right)^{\frac{\theta}{\theta-\gamma_2}}$. Since $A_1 > A_2$, we must have that $\frac{k_1}{k_2} > \frac{A_1}{A_2} \Rightarrow k_1 > k_2$. $\blacksquare$

Before we prove Theorem 1, let's prove some preliminary results that will be important for the theorems' proofs.

Lemma 1 *If $A_1 > A_2$ we must have that $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2} > 1$.*

Proof. From proposition 1 we have that $A_1 > A_2 \Rightarrow p_1 > p_2$. Now, let's focus on $(F.1)$'s RHS. This term is positive or negative depending on the following term:

$$\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\right]^{\frac{\gamma_i}{1-\gamma_i}}, \ \forall i \in \{1,3\} \tag{82}$$

Now, towards a contradiction, let's assume that $A_1 > A_2$ and $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2} < 1$. Consequently, the second term in expression (82) is less than one. Similarly, $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2} < 1 \Rightarrow \frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha} > 1$. Since $\alpha < 1$ and $\frac{p_1}{p_2} > 1$, this gives us that

$$\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\right]^{\frac{\gamma_i}{1-\gamma_i}} > 0, \ \forall i \in \{1,3\}$$

and the $(F.1)$'s RHS is positive. Then, $(F.1)$'s LHS must also be positive. Following the same argument presented in the proof of lemma A.2, we have that inequality (46) must hold.

Similarly, from $p_1 > p_2$, we have that the last term on $(F.2)$'s RHS is negative. Therefore, since $\gamma_2 < \theta$, we have:

$$\left\{ \frac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}} \right\} > \left(\frac{A_2}{A_1} \right)^{\frac{1}{1-\gamma_2}} \left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1-\theta} \times \frac{\gamma_2 - \theta}{(1-\gamma_2)}} k_1 \tag{83}$$

Then, we have that:

$$\frac{\text{RHS}(46)}{\text{RHS}(83)} = \left(\frac{p_2}{p_1} \right)^{1 + \frac{\alpha\theta}{1-\theta} \left[1 - \frac{\gamma_2 - \theta}{\theta(1-\gamma_2)} \right]} \left(\frac{A_1}{A_2} \right)^{\frac{1}{1-\gamma_2}} \tag{84}$$

Notice that $1 - \frac{\gamma_2 - \theta}{\theta(1-\gamma_2)} = \frac{\gamma_2(1-\theta)}{\theta(1-\gamma_2)}$. Consequently:

$$\frac{\text{RHS}(46)}{\text{RHS}(83)} = \left(\frac{p_2}{p_1} \right)^{1 + \frac{\gamma_2 \alpha}{(1-\gamma_2)}} \left(\frac{A_1}{A_2} \right)^{\frac{1}{1-\gamma_2}} = \left\{ \left(\frac{p_2}{p_1} \right)^{1-\gamma_2(1-\alpha)} \frac{A_1}{A_2} \right\}^{\frac{1}{1-\gamma_2}} \tag{85}$$

But then, notice that $1 - \gamma_2(1-\alpha) - \alpha = (1-\alpha)(1-\gamma_2) > 0$. Therefore, $1 - \gamma_2(1-\alpha) > \alpha$. Since $p_2 < p_1$, we have that:

$$\left(\frac{p_2}{p_1} \right)^{1-\gamma_2(1-\alpha)} \frac{A_1}{A_2} < \left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} < 1 \tag{86}$$

where the last inequality comes from our assumption for the contradiction. Then, since $\frac{1}{1-\gamma_2} > 0$, we have $\frac{\text{RHS}(46)}{\text{RHS}(83)} < 1$. But then inequalities (46) and (83) cannot both be satisfied and we have a contradiction. ■

Corollary 2 *If $A_1 > A_2$ we must have $m_{11} > m_{12}$ and $m_{31} > m_{32}$.*

Proof. From the expression for m_{11}, we have:

$$m_{11} = \frac{\left[\left(\frac{p_1}{p_2} \right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_1 - 1}} M_1}{\left\{ 1 + \left[\left(\frac{p_1}{p_2} \right)^{\alpha} \frac{A_2}{A_1} \right]^{\frac{1}{\gamma_1 - 1}} \right\}} = \frac{\left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma_1}} M_1}{\left\{ 1 + \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma_1}} \right\}} \tag{87}$$

Since from lemma 1 we have $\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} > 1$, we must have that $\frac{\left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma_1}} M_1}{\left\{ 1 + \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma_1}} \right\}} > \frac{M_1}{2}$. Consequently $m_{11} > m_{12}$. The identical argument shows that $m_{31} > m_{32}$. ■

Proof of Theorem 1

Proof. We already know that $m_{11} > m_{12}$ and $m_{31} > m_{32}$. So, the only way in which we may have $S_2 > S_1$ is that $m_{22} > m_{21}$. Therefore, towards a contradiction, assume that $m_{22} > m_{21}$. From (68):

$$\frac{M_2 x_2^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}} > k_1 \tag{88}$$

Then, back to $(F.2)$, we have:

$$\left\{ \frac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} = \tag{89}$$
$$= \left(\frac{A_2}{A_1} \right)^{\frac{\theta}{\gamma_2 - \theta}} \left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} + \left(\frac{r}{A_2 \gamma_2 x_k} \right)^{\frac{\theta}{\theta - \gamma_2}} \left[1 - \left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{1-\theta}} \right] x_k$$

Since $A_1 > A_2$ we know from previous results that $p_1 > p_2$. Consequently, the last term in $(F.2)$'s RHS is negative and we have:

$$\left\{ \frac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{\left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} > \left(\frac{A_2}{A_1} \right)^{\frac{\theta}{\gamma_2 - \theta}} \left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{1-\theta} \times \left[1 + \frac{1-\gamma_2}{\gamma_2 - \theta} \right]} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} \tag{90}$$

Now, from (88) we have that, since $\gamma_2 < \theta$:

$$\left\{ \frac{M_2 x_2^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1 \gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - x_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} < k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} \tag{91}$$

Now, substituting (91) into (90), we have:

$$k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} > \left\{ \frac{M_2 x_2^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1\gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1\gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} > \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{\theta}{\theta-\gamma_2}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} \quad (92)$$

From lemma 2 and the fact that $\theta > \gamma_2$, we have that $\left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{\theta}{\theta-\gamma_2}} > 1$. Consequently, we found a contradiction. Therefore, we must have $m_{21} > m_{22}$ and $S_1 > S_2$. ■

Before presenting the proof for theorem 2, let's consider a final intermediary result:

Claim 1 *Assume $\gamma_2 < \theta$. If $A_1 > A_2$ we must have $\frac{m_{21}}{m_{22}} < \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma_2}}$*

Proof. From lemma 1, we have that if $A_1 > A_2$ we must have $\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} > 1$. Then, from $(F.2)$, since $p_1 > p_2$, we must have:

$$\left\{ \frac{M_2 x_2^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}} k_1}{k_1 \left[\left(\frac{r}{A_1\gamma_2 x_k} \right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - x_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} < \left\{ \frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right\}^{\frac{\theta}{\gamma_2-\theta}}$$

From (68) and $\gamma_2 < \theta$, we have $\frac{m_{21}}{m_{22}} < \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma_2}}$, concluding the proof. ■

Proof of Theorem 2:

Proof. Assume that $\gamma_i \equiv \gamma, \forall i \in \{1,2,3\}$ and $\gamma < \theta$. Assume that $A_1 > A_2$ as well. From theorem 1 and claim 1 we have $S_1 < \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma_2}} S_2$. Then, notice that $pdf_{1i} = \frac{m_1 i}{S_i}$. Therefore $\frac{pdf_{11}}{pdf_{12}} = \frac{m_{11}}{m_{12}} \times \frac{S_2}{S_1}$. Since $\frac{m_{11}}{m_{12}} = \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma}}$ and $\frac{S_2}{S_1} > \frac{1}{\left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma}}}$, we have that:

$$\frac{pdf_{11}}{pdf_{12}} > \left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma}} \times \frac{1}{\left[\left(\frac{p_2}{p_1} \right)^{\alpha} \frac{A_1}{A_2} \right]^{\frac{1}{1-\gamma}}} \quad (93)$$

Consequently $pdf_{11} > pdf_{12}$. The same calculation gives us $pdf_{31} > pdf_{32}$. Since density functions must add to one, we must also have $pdf_{21} < pdf_{22}$ ■

Proof of Proposition 3

Proof. Towards a contradiction, assume that $A_1 > A_2$ and $p_2 > p_1$. Then, from (58), after some manipulations and using $\gamma_3 > \theta$, and $\frac{p_2}{p_1} > 1$ we have:

$$\left(\frac{r}{A_1\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} > \left(\frac{r}{A_2\gamma_3 x_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_2^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}$$

i.e.:

$$\frac{k_1}{k_2} > \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{94}$$

From equation (55), we have:

$$\frac{m_{31}}{m_{32}} > \frac{k_1}{k_2} \Rightarrow \frac{m_{31}}{m_{32}} > \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{95}$$

Then, from LHS(62), substituting (94) and (96), we have:

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta}x_3 + k_1^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}x_3 - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta}x_3 + k_2^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta}x_3 \end{array} \right\} > \left[\left(\frac{A_1}{A_2}\right)^{\frac{\gamma_3}{1-\gamma_3}} - \frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)\right]\left(m_{32}^{\theta}x_3 + k_2^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta}x_3 > 0 \tag{96}$$

While from RHS(62), we have that:

$$\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_i}{\gamma_i-1}}\right] < 0, \forall \gamma_i < 1$$

Consequently RHS(62)< 0, which gives us a contradiction. Since we showed in lemma A.2 that $p_1 = p_2$ only happens if $A_1 = A_2$, we must have that $A_1 > A_2 \Rightarrow p_1 > p_2$. Following the same procedure we can easily show that $A_2 > A_1 \Rightarrow p_2 > p_1$. ■

Proof of Proposition 4

Proof. Without loss of generality, assume that $A_1 > A_2$. From proposition 3 we have that $A_1 > A_2 \Rightarrow p_1 > p_2$. From (60) and (F.2), given that $p_1 > p_2$, we have – after some manipulations:

$$\frac{k_1}{k_2} > \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \left(\frac{p_2}{p_1}\right)^{\frac{\alpha(\gamma_3-\theta)}{(1-\gamma_3)(1-\theta)}}$$

While from (55), we have that:

$$\frac{m_{31}}{m_{32}} > \left(\frac{p_2}{p_1}\right)^{\frac{\alpha}{1-\theta}} \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \left(\frac{p_2}{p_1}\right)^{\frac{\alpha(\gamma_3-\theta)}{(1-\gamma_3)(1-\theta)}}$$

Simplifying it:

$$\frac{m_{31}}{m_{32}} > \left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right]^{\frac{1}{1-\gamma_3}} \tag{97}$$

Let's consider two cases:

Case 1: $\left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right] \geq 1$ – In this case, equation (97) already implies that $m_{31} \geq m_{32}$. From (55) and $\theta < 1$, we have that:

$$\frac{k_1}{k_2} \geq \left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} > 1 \tag{98}$$

Consequently, $k_1 > k_2$, concluding this part of the proof.

Case 2: $\left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right] < 1$ – In this case, from RHS(62), we have that:

$$\left\{\begin{array}{l} \left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} x_1\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right]^{\frac{\gamma_1}{1-\gamma_1}}\right] \\ +\left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} x_2\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right]^{\frac{\gamma_2}{1-\gamma_2}}\right] \end{array}\right\}$$

Given $\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha} < 1$, notice that:

$$\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha} < 1 \Rightarrow \frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha} > 1 \Rightarrow \frac{A_2}{A_1}\frac{p_1}{p_2} > 1$$

Consequently, $\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right]^{\frac{\gamma_i}{1-\gamma_i}}\right] > 0$ for $i \in \{1, 2\}$ and RHS(62)> 0.

But then, from (64), given that $\left(m_{31}^{\theta}x_3 + k_1^{\theta}x_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}x_3 > 0$, we would need to have:

$$1 - \left(\frac{p_1}{p_2}\right)^{1-\alpha}\frac{M_3 - m_{31}}{m_{31}} > 0$$

Rearranging it:

$$\frac{m_{31}}{m_{32}} > \left(\frac{p_1}{p_2}\right)^{1-\alpha} > 1 \tag{99}$$

From (99) and (55), we have:

$$\left(\frac{p_2}{p_1}\right)^{\frac{\alpha}{1-\theta}}\frac{k_1}{k_2} > \left(\frac{p_1}{p_2}\right)^{1-\alpha} \Rightarrow \frac{k_1}{k_2} > \left(\frac{p_1}{p_2}\right)^{1+\frac{\alpha\theta}{1-\theta}} \tag{100}$$

Consequently, (100) implies that $k_1 > k_2$, concluding our proof. ■

Proof of Corollary 1

Proof. Proof of proposition 4 already showed this result for all cases but $\left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right] = 1$. In this case, notice that:

$$\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha} = 1 \Rightarrow \frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha} = 1$$

Since $\alpha < 1$ and $p_1 > p_2$, we have that $\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right) > 1$. Again, we can show that the RHS(62)> 0. Following the same steps presented in the proof of proposition 4, we can conclude that $m_{3i} > m_{3j}$.
■

Proof of Theorem 3

Proof. Towards a contradiction, assume that $pdf_{31} \leq pdf_{32}$. In this case, we must have:

$$\frac{m_{31}}{m_{11} + m_{21} + m_{31}} \leq \frac{m_{32}}{m_{12} + m_{22} + m_{32}}$$

Rearranging and simplifying it, we obtain:

$$m_{31}m_{12} - m_{32}m_{11} + m_{31}m_{22} - m_{32}m_{21} \leq 0 \tag{101}$$

From equations (53) and (54) and labor market clearing conditions, we have:

$$m_{11} = \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}} m_{12} \quad \text{and} \quad m_{21} = \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}} m_{22} \tag{102}$$

As a result, we have:

$$m_{31}m_{12} - m_{32}m_{11} = m_{32}m_{12}\left\{\frac{m_{31}}{m_{32}} - \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}}\right\} > 0 \tag{103}$$

and

$$m_{31}m_{22} - m_{32}m_{21} = m_{32}m_{22}\left\{\frac{m_{31}}{m_{32}} - \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}}\right\} > 0 \tag{104}$$

where the inequalities come from $\frac{m_{31}}{m_{32}} > \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}}$ as shown in equation (97). Consequently, equations (101), (103), and (104) jointly show a contradiction. As a result, $pdf_{31} > pdf_{32}$.

Similarly, towards a contradiction, consider that $pdf_{21} \geq pdf_{22}$. In this case, we must have:

$$\frac{m_{21}}{m_{11} + m_{21} + m_{31}} \geq \frac{m_{22}}{m_{12} + m_{22} + m_{32}}$$

Rearranging and simplifying it, we obtain:

$$m_{12}m_{21} - m_{22}m_{11} + m_{32}m_{21} - m_{31}m_{22} \leq 0 \tag{105}$$

From (102), after some manipulations, we have:

$$m_{12}m_{21} - m_{22}m_{11} = 0 \tag{106}$$

and

$$m_{32}m_{21} - m_{31}m_{22} = m_{32}m_{22} \left\{ \left[\frac{A_1}{A_2} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{1-\gamma}} - \frac{m_{31}}{m_{32}} \right\} < 0 \tag{107}$$

where the inequalities come from $\frac{m_{31}}{m_{32}} > \left[\frac{A_1}{A_2} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{1-\gamma}}$ as shown in equation (97). Consequently, equations (105), (106), and (107) jointly show a contradiction. As a result, $pdf_{21} < pdf_{22}$.

Finally, towards a contradiction, assume that $pdf_{11} \geq pdf_{12}$. In this case, we must have:

$$\frac{m_{11}}{m_{11} + m_{21} + m_{31}} \geq \frac{m_{12}}{m_{12} + m_{22} + m_{32}}$$

Rearranging and simplifying it, we obtain:

$$m_{11}m_{22} - m_{12}m_{21} + m_{32}m_{11} - m_{31}m_{12} \leq 0 \tag{108}$$

In equation (106), we already showed that $m_{11}m_{22} - m_{12}m_{21} = 0$. Then, from (102) and (97), we have:

$$m_{32}m_{11} - m_{31}m_{12} = m_{32}m_{12} \left\{ \left[\frac{A_1}{A_2} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{1-\gamma}} - \frac{m_{31}}{m_{32}} \right\} < 0 \tag{109}$$

Consequently, equations (108), (106), and (109) jointly show a contradiction. As a result, $pdf_{11} < pdf_{12}$, concluding our proof that pdf_1 F.O.S.D. pdf_2. ■

C Empirical Evidence - City Size

C.1 Size

In the theory there is strong relationship between size and productivity of locations, which should also lead to a similar house price productivity relationship. In this section we consider, whether it holds that size has a similar relationship as house prices with automation technology adoption and the change in occupational shares.

Table 16: Entreprise Resource Planning Software

	(1) ERP 2015	(2) ERP 2015	(3) ERP 2015
log employment	0.0143***		0.0135***
	(0.00259)		(0.00331)
routine cognitive share 1980		0.508***	0.106
		(0.133)	(0.166)
Observations	253	253	253
R^2	0.261	0.082	0.264

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Each observation (an MSA) is weighted by its employment in 2015.

Table 17: Change in routine-cognitive share, 1980-2015

	(1) Δ rout-cog	(2) Δ rout-cog	(3) Δ rout-cog
log employment	-0.00941***		-0.00417
	(0.00238)		(0.00219)
routine cognitive share 1980		-0.831***	-0.707***
		(0.0765)	(0.0948)
Observations	253	253	253
R^2	0.252	0.490	0.529

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Each observation (an MSA) is weighted by its employment in 2015.

Qualitatively the relationship between rent and ERP adoption, as well as the change in the routine-cognitive share is the same as with size.

D Skill Biased Technological Change and City Size - Numerical Examples

Differently from the case of Automation, **SBTC does not imply that the high-TFP city is larger**. In this section, we present two examples that illustrate that results can go either way.

D.1 High-TFP city is smaller (the "Boulder" case)

Consider the following parameter values:

Example 1: Parameters

x_1	x_2	x_3	A_1	A_2	H	x_k	r
0.3189	1	1.4733	19118	19000	62559000	1.333	635.58
γ_1	γ_2	γ_3	θ	α	M_1	M_2	M_3
0.8	0.8	0.82	0.5	0.24	15836150	66973717	40745094

In this case, we obtain the following equilibrium prices and quantities:

Example 1: Equilibrium outcomes

	m_{1j}	m_{2j}	m_{3j}	S_j	p_j
City 1	7683472.87	32494687	21148855.3	61327015.1	496.69
City 2	8152677.13	34479030	19596238.7	62227945.9	460.71

	w_{1j}	w_{2j}	w_{3j}	k_j
City 1	204.68	481.03	5308.34	1207650344
City 2	201.02	472.42	5213.4	1020740138

As we can see, the high-TFP city, while paying higher wages, investing more in capital, having higher housing prices, and having more high-skill workers, it is still smaller than the low-TFP city. In particular, the high-TFP city has fewer low- and mid-skill workers than the low TFP city. Finally, as expected, the skill distribution in the High-TFP city skill dominates in first order the skill distribution in the Low-TFP city, as we see in figure 1.

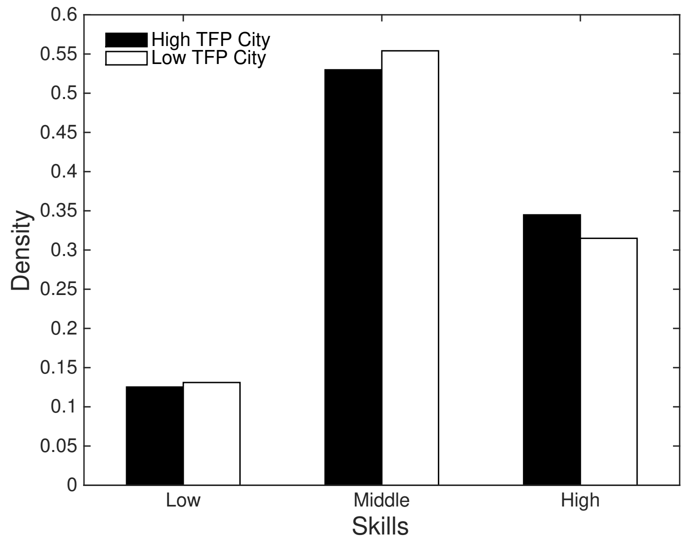


Figure 7: Skill Distribution: High vs. Low TFP cities - Example 1

D.1.1 High-TFP city is larger (the "NYC" case)

Consider the following parameter values:

Example 2: Parameters

x_1	x_2	x_3	A_1	A_2	H	x_k	r
0.3189	1	1.4733	21118	19000	62559000	1.333	635.58
γ_1	γ_2	γ_3	θ	α	M_1	M_2	M_3
0.8	0.8	0.82	0.5	0.24	15836150	66973717	40745094

In order to make a simple comparison, the parameters are the same of Example 1, apart from a higher A_1. In this case, we obtain the following equilibrium prices and quantities:

Example 2: Equilibrium outcomes

	m_{1j}	m_{2j}	m_{3j}	S_j	p_j
City 1	7841477.23	33162913.76	24434572.45	65438963.44	663.10
City 2	7994672.77	33810803.24	16310521.55	58115997.56	420.08

	w_{1j}	w_{2j}	w_{3j}	k_j
City 1	225.17	529.19	6283.30	1954872597.36
City 2	201.81	474.28	5631.30	995018365.67

Notice that the high-TFP city is larger. However, we still have fewer low- and mid-skill workers. As before, all other results follow through, including the F.O.S.D. of the skill distribution of the high-TFP city, as seen in figure 2.

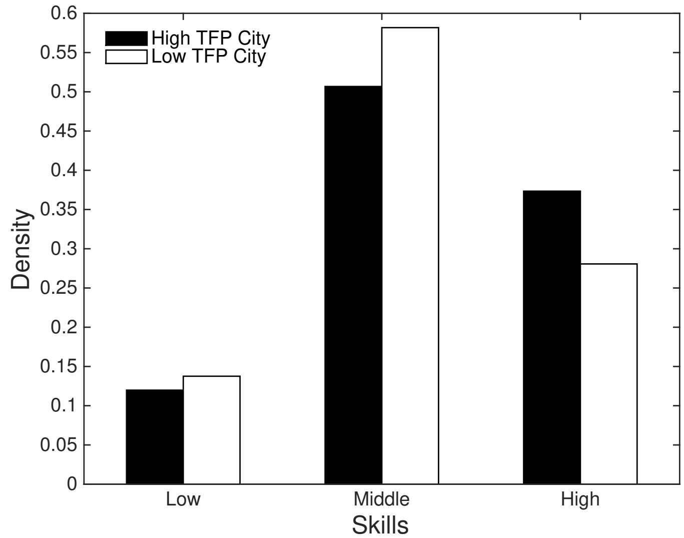


Figure 8: Skill Distribution: High vs. Low TFP cities - Example 2