IT and Urban Polarization*

Jan Eeckhout[†] Christoph Hedtrich[‡] Roberto Pinheiro[§]

September 2021

Abstract

We show that differential IT investment across cities has been a key driver of job and wage polarization since the 1980s. Using a novel data set, we establish two stylized facts: IT investment is highest in firms in large and expensive cities, and the decline in routine cognitive occupations is most prevalent in large and expensive cities. To explain these facts, we propose a model mechanism where the substitution of routine workers by IT leads to higher IT adoption in large cities due to a higher cost of living and higher wages. We estimate the spatial equilibrium model to trace out the effects of IT on the labor market between 1990 and 2015. We find that the fall in IT prices explains 50 percent of the rising wage gap between routine and non-routine cognitive jobs. The decline in IT prices also accounts for 28 percent of the shift in employment away from routine cognitive towards non-routine cognitive jobs. Moreover, our estimates show that the impact of IT is uneven across space. Expensive locations have seen a stronger displacement of routine cognitive jobs and a larger widening of the wage gap between routine and non-routine cognitive jobs.

Keywords: IT investment. Job Polarization. Spatial Sorting. Urban Wage Premium.
JEL Codes: D21, J24, J31, R23.

*We would like to thank Stephen Redding, Stuart Rosenthal, Ed Glaeser, Morris Davis, Chris Taber, Henry Siu, Tim Dunne, Jim Davis, Bruce Fallick, James Feigenbaum, Domenico Ferraro, Pawel Krolikowski, Jeff Lin, Murat Tasci, as well as numerous seminar participants for comments. We are grateful to Albert Saiz and David Albouy for access to data and Ross Cohen-Kristiansen for research assistance. Jan Eeckhout acknowledges financial support from ERC Advanced Grant 339186 and from ECO2015-67655-P. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2018-05973. The views in this article are those of the authors and do not necessarily reflect the views of the Federal Reserve System, the Federal Reserve Bank of Cleveland, or the Board.
[†]UPF Barcelona (ICREA-GSE-CREi), jan.eeckhout@upf.edu
[‡]Uppsala University, christoph.hedtrich@nek.uu.se
[§]Federal Reserve Bank of Cleveland, Roberto.Pinheiro@clev.frb.org

1 Introduction

Polarization in the labor market is one of the main forces behind the rise in wage inequality (see Acemoglu and Autor (2011) and Cortes et al. (2017)). With the advent of the information age, new technologies make all workers more productive, but they affect different workers differently. In particular, the labor-saving investment has the highest return where information technology (IT) substitutes routine tasks. Those routine tasks are disproportionately performed by workers earning wages in the middle of the distribution, and as a result, the displacement of these jobs leads to polarization of the earnings distribution. Moreover, there is a marked geographical dimension to polarization (see Autor (2019), Autor and Dorn (2013)), with strong variation in polarization across metropolitan areas.

Yet, to date, little is known about the mechanism that links investment in IT and job displacement and how this mechanism explains the geographical variation. In this paper we make two contributions. First, we analyze a novel data set and we document two new stylized facts: 1. IT investment is highest in large, densely populated cities; 2. The decline in routine cognitive occupations is largest in large cities. We obtain these two facts from analyzing micro-data on IT usage at the establishment level and US Census data (Ruggles et al., 2020) on employment. We show that these empirical results are robust and hold under many different specifications, most notably after controlling for *firm* fixed effects and headquarter location.

Second, we propose an equilibrium mechanism that can rationalize these facts and that explains why polarization is a phenomenon with a strong urban component. The main insight is that the composition of the production factors that firms choose varies by geography: workers must be compensated for *local* housing prices, whereas IT is a highly tradable good that can be bought at similar prices *everywhere*. As a result, labor and IT demand varies significantly with cities' cost of living. Because across locations, housing prices comove with labor productivity and wages, it is beneficial for firms to use IT more intensively in expensive cities. Consequently, more productive areas are the ones prone to replace routine tasks with IT because those routine tasks disproportionately drive up the cost of labor.

The focus in our empirical analysis is on the distribution of employment and our stylized facts document the effect of IT investment on the displacement of routine workers across geographical locations. At the same time, the technological change that is at the origin of the change in the distribution has general equilibrium effects on wages. In our empirical analysis, we document in detail the pattern of wage inequality. First, we show the evolution of the city-size wage premium and find that the increase is most pronounced for non-routine cognitive occupations. Second, we analyze wage inequality within and between cities. Wage inequality *within* cities is higher in larger cities (see also Baum-Snow and Pavan (2013), Eeckhout et al. (2014), and Santamarıa (2018)), and the inequality increased between 1990 and 2015. At the same time, wage inequality *between* cities is almost constant over the same period. A variance decomposition of wages reveals that nearly all of the variance is due to within-city inequality, not between cities.

With these facts about the distribution of employment and wages in mind, we build an equilibrium model of production in cities with heterogeneous workers who optimally choose their location and occupation, given their exogenous abilities and their idiosyncratic taste for different locations. Moreover, representative firms in each city choose their optimal input combination given the region's total factor productivity (TFP), the differential cost of production inputs, and the degree of complementarity between each occupation and IT. Furthermore, labor and IT differ in their tradability. Labor must be provided locally; as a result, wages are determined by local labor market conditions. Instead, each city is seen as a small open economy in the market for IT. In other words, our model combines elements from Roy (1951), Rosen (1979), Roback (1982), and Krusell et al. (2000).

First, we derive analytical results for a simplified version of the model. This gives us crisp insights into the workings of the economic mechanism. We show that the impact of the cost of living on the distribution of occupations across space depends crucially on the elasticity of substitution between labor and IT. Labor in occupations that feature a high elasticity of substitution are reallocated towards cheaper cities when IT productivity rises. By contrast, occupations that are more complementary sort into expensive cities. Now in turn, the cost of living is an equilibrium outcome. We show that in equilibrium, more productive cities have higher housing prices, and we derive conditions under which there is more investment in IT in cities with higher TFP. This then allows us to establish that more productive cities are larger in population size and that there is spatial sorting by occupations consistent with the stylized facts.

Second, we estimate the model parameters for the full model – in particular the productivity parameters for the different routine and non-routine occupations as well as the parameters governing the distribution of amenities and housing supply – matching city- and occupation-level moments of the wage and employment distribution. We find that IT has a prominent role in explaining recent employment and wage trends across cities. A counterfactual exercise where we simulate a fall in IT prices by 65 percent – corresponding to a similar change in the data between 1990 and 2015 – explains both the fall in employment in routine cognitive jobs and the rise in non-routine cognitive jobs. Quantitatively, the exercise explains about 28 percent of the change in employment shares in cognitive occupations. Additionally, the simulation results imply that IT accounts for 50 percent of the widening wage gap between routine and non-routine cognitive jobs. The model simulation also highlights the strong urban component of polarization: the employment share of routine cognitive occupations falls substantially more in expensive locations, about 40 percent more relative to the average. Similarly, the wage gap between routine and non-routine cognitive jobs widens even more in expensive locations, about 20 percent more compared to the average. Overall the results indicate a strong role for IT in the displacement of routine cognitive employment and a rise in non-routine cognitive employment and the accompanying polarization of earnings across jobs and cities.

Related Literature. Our paper builds on a large literature on the polarization of the labor market and the disappearing routine jobs, for example, Autor and Dorn (2013), Goos et al. (2014), Cortes et al. (2017), and Acemoglu and Autor (2011). Much of the focus of this literature is on technological change as the main driver of polarization. We embrace this technological explanation but focus on the role of capital investment. The notion that capital investment affects different skilled workers is of course not new. Krusell et al. (2000) were the first to argue that the college premium has risen so much because technological investment affects the high skilled more than the low skilled. The drop in the cost of such new technologies then further widens the gap between skilled and unskilled workers. We rely on a similar mechanism to explain the polarization of the labor market.

In addition to the role of capital investment, our analysis focuses on differential technology adoption across cities. Beaudry et al. (2010) show that technology adoption – measured by PCs per worker – has occurred first in areas with a relatively high supply of skill (or with a low relative price of skill). They also show that these areas experienced the greatest increase in the return to education. Our analysis, while controlling for the relative supply of high skill workers in the MSA, highlights the importance of local prices in the sorting of workers and activities across space, which is mostly missing from the analysis by Beaudry et al. (2010). Moreover, by allowing more than two types of workers, our framework is better suited to address the issues of job polarization and the "disappearing middle" of the income distribution.

We also analyze the evolution of wage inequality across cities, both empirically and in the model. Baum-Snow and Pavan (2013) and Baum-Snow et al. (2018) document that wage inequality rose more in large cities in the US between 1980 and 2007, suggesting that the forces driving inequality have an urban bias. Our paper provides a mechanism for this finding: the endogenously more intensive adoption of IT in expensive locations. Our empirical findings go further by focusing on the evolution of inequality across different tasks. Baum-Snow et al. (2018) estimate production functions and find evidence for capital-skill complementarity, but also skill bias in agglomeration economies of technical change. Our paper instead focuses on the spatial implications of technological change in an equilibrium system of cities and highlights that the adoption of IT can explain, at least in part, the skill bias in agglomeration economies. Further, our paper documents results using novel data on IT usage across the whole economy and not just capital data from the manufacturing sector.

There is an extensive literature documenting geographical patterns of occupations that are related to our results. Rubinton (2020) finds that the adoption of IT is higher in larger cities. She uses data from the Annual Capital Expenditures Survey, thus complementing our findings. The focus of her paper is on the gap in wages between low- and high-skilled workers and business dynamism. In contrast, here the focus is on the role of technology in the polarization of employment and wages. Rossi-Hansberg et al. (2019) study cognitive hubs and find, as we do, that non-routine occupations are disproportionately represented in large cities. They use different data and propose an interesting mechanism that is based on a flexible technology specification that exhibits externalities, which leads to inefficient equilibrium allocations. They find crisp predictions regarding optimal policy, in an approach that is complementary to ours.

In a study of the role of technological change in regional convergence in the US, Giannone (2017) finds that skill-biased technological change can explain a substantial share of the decline in regional convergence across cities in the US. A key difference is our focus on the role of technology behind the evolution of wages and employment and the endogenous nature of adoption of said technology. Finally, for France, Davis et al. (2020) document similar patterns of the sorting of workers across cities, suggesting that our results may extend beyond the US economy. However, they do not use direct evidence on technology to determine its role for their findings.

The paper is organized as follows. Section 2 describes the data. Section 3 presents the empirical results and the two stylized facts. Section 4 contains the equilibrium mechanism that rationalizes the facts based on a general equilibrium model. The section contains the setup of the general model, a series of analytical results for a simplified version of the model, and the estimation of the full model. We use the estimated model to trace out the effects of IT on the labor market within and across cities under counterfactual scenarios. Finally, we make some concluding remarks in Section 5. All proofs are presented in the Appendix.

2 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 2014-2016 we combine the American Community Survey (ACS) yearly files. From these files, we construct labor force and price information at the metropolitan statistical area (MSA) level.[1] For each year we then construct information on the labor force, earnings, and the local price level in each MSA. We focus our attention to full-time, full-year workers aged 25-54.[2]

Our variable for the price at the MSA level is a simple price index including both consumption goods – which sell at 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

[1] The definition of a MSA we use is the Census Beaureau's 2000 combined metropolitan statistical areas (CMSA) for all MSAs that are part of an CMSA or 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 Beaureau's public use microdata area (PUMA) of each census sample to the 2000 Census Metropolitan Area definitions. The census data restricts us to consider only MSAs that are sufficiently large, as they are otherwise not identifiable due to the minimal size of a PUMA.

[2] In particular, we restrict our sample to workers who report working at least 40 weeks, 35 usual hours per week and who earn at least 75 percent of the federal minimum wage in each year. Our earnings measure is the log hourly wage calculated by subtracting log weeks times usual hours worked. Since the information on weeks worked in ACS 2013-2015 is presented in intervals, we use the same interval mid-points in order to calculate the usual hours worked for the census samples. Finally, to maintain comparability with the census data, we shift the wage distribution in each of the ACS sample years to have the same median as that for the 2015 sample. Similarly, we adjust all earnings data to reflect values in 2000 US dollars.

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). See the Appendix for further details.

Data on IT. The technology data come from the Ci Technology Database, produced by the Aberdeen Group (formerly known as Harte-Hanks). The data have detailed hardware and software information for over 200,000 sites in 2015,[3] including not only installed capacity but also expected future expenses in technology. Their data also include 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 consider several measures of investment in technology. Initially, we consider a broad measure of investment in technology: the total IT budget per worker. While this measure may overstate the investment in technology made to either boost the productivity or replace a given set of workers, it has several advantages. First, this measure is available for all the establishments in our sample. Second, the portion of our database that includes IT budget information covers a significant fraction of the employed labor force as well as establishments, when compared to other standard databases such as the National Establishment Time-Series (NETS) and the County Business Pattern (CBP). For example, compared to NETS our sample covers on average 53 percent of the MSA's employed labor force. An even larger share of the employed labor force is covered when compared to the CBP (73 percent). We find that there is nearly full geographical coverage, with only very few MSAs missing.[4] 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). Detailed descriptive statistics about the IT data are provided in the Appendix.

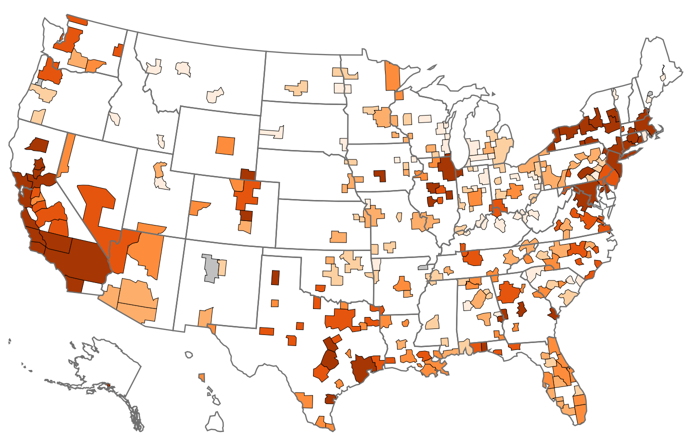


Figure 1: Avg. IT budget per worker

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

[4] See Figures A-3 and A-1 in the Appendix for the geographical dispersion of the IT budget per worker in 2015 relative to NETS and CBP, respectively.

We also focus on measures that target the degree of complementarity between a group of occupations and technology. In particular, we use the adoption of enterprise resource planning (ERP) software in order to measure the establishments' intent in automating routine cognitive tasks. As pointed out by Bloom et al. (2014), ERP software systems integrate several data sources and processes of an organization into a unified system, reducing the need for clerical and low-level white collar workers. We consider ERPs that help in managing the following areas: Accounting, Human Resources, Customer & Sales Force, Collaborative and Integration, Supply Chain Management, as well as bundle software like the ones produced by SAP, which are usually called enterprise applications.

The main benefit of using ERP is that it is a clear measure of an establishment's intent in automating. In this sense, the measure of ERP software is quite distinct from aggregate measures such as IT budget and other general purpose technologies, such as the adoption of personal computers. The key drawbacks are: first, there is a significant reduction in establishment coverage. Our information on ERP adoption covers on average only 16 percent of workers and 1 percent of establishments in the MSA, compared to NETS. (see Appendix Table A-8). Similarly, our ERP sample covers on average only 20 percent of workers and 1 percent of establishments in the MSA, compared to the CBP (see Appendix Table A-4). Second, we need to focus on coarser measures of technology adoption. Our leading measure of ERP adoption is the fraction of establishments in the MSAs that adopted ERP software. This measure does not capture the intensive margin of ERP adoption.[5] Due to the drawbacks of the ERP measure, we focus our analysis on the IT budget per worker in Section 3. However, we present the results for ERP measures in Appendix Section D. While results are understandably weaker for ERP – due to smaller sample size – they are qualitatively similar to the ones presented in section 3.

Data on Metropolitan Areas' Characteristics. In order to control for metropolitan area characteristics, we gather information on housing supply elasticity, natural amenities, and industry composition in the MSA. Our key measure for the housing supply elasticity is based on Saiz (2010). This measure takes into account both land use restrictions and geographical restrictions on building in different areas.[6]

We control for amenities using the climate and geographical measures presented in appendix B.4 of Albouy (2012). In particular, we focus on the measures that capture heating and cooling degree days (annual); average sunshine as a percentage of possible; average slope of the land in the

[5] For example, consider two establishments, A and B, that adopt ERP software to different degrees. Establishment A adopts a relatively simple accounting software that may replace the work of a few accounting assistants. Differently, establishment B adopts an integrated ERP software system that allows it to automate several processes within the firm – sales, HR, inventory, accounting, etc. Both establishments would be classified as "adopters" and contribute the same to our leading measure. Consequently, our leading measure will be biased towards finding no effect.

[6] In previous versions, we presented robustness considering two additional measures. The Wharton Residential Land Use Regulation Index (WRLURI), based on work by Gyourko et al. (2008), which takes into account building regulations. Ganong and Shoag (2017)'s Land regulation index, which is based on the number of state supreme and appellate court cases containing the phrase "land use" over time.

metropolitan area; and average distance to the closest coastline.[7]

We follow Beaudry et al. (2010) and include controls that reflect a city's employment mix across 12 industry groups in 1980 in order to control for the metropolitan areas' industry composition.[8]

3 Empirical Evidence

In this section we document the main findings regarding urban polarization. In the first subsection, we report the evidence on technology adoption and job polarization by city size. In the second subsection, we focus on the empirical implications for wage inequality.

3.1 Technology Adoption and Job Polarization by City Size

In describing the evidence on the adoption of automation technology and the occupational composition of cities, we report two main findings: (1) locations with higher housing costs adopt automation technology at higher rates; and (2) locations with higher housing costs see a decreasing share of their workforce employed in routine occupations, whose tasks are being replaced by automation technology.

Fact 1. Stronger IT Adoption in Expensive Cities. Figure 2 visualizes the positive correlation between local rental prices and the average IT budget per worker. Mere inspection shows that the magnitude of the change in IT spending as the rent index changes is sizable. Furthermore, Table 1 shows the results for MSA-level linear regression models of the log of the average IT budget per worker, adjusted for plant employment interacted with three-digit SIC industries, following Beaudry et al. (2010) and Doms and Lewis (2006). The regression results provide support for the hypothesis that IT expenditure per worker is increasing in the cost of housing. The elasticity is highly significant and its value barely changes under different regression specifications. The MSA's rental price index in 1980 helps to explain the variation in IT budget per worker across MSAs, even after controlling for the presence of natural amenities, housing supply elasticity, and industry composition.[9] In specification (1), a one standard deviation increase in the local price index (an increase of 21.4 percent in the 1980 local price index) is associated with an increase of $107.43 in the

[7]In previous versions, we considered natural amenities coming from the US Department of Agriculture (USDA). In particular, we focused on the following measures: mean temperature for January (1941-1970); mean temperature for July (1941-1970); mean hours of sunlight for January (1941-1970); ln(% of water area); mean relative humidity for July (1941-1970). Results were qualitatively similar.

[8]In particular, we control for the share of employment in industry categories that correspond roughly to one-digit SICs (public sector is the excluded category): Agriculture and Mining; Construction; Non-durable Manufacturing; Durable Manufacturing; Transportation and Utilities; Wholesale; Retail; Finance, Insurance, and Real Estate; Business and Repair Services; Other Low-Skill Services; Entertainment; Professional Services. To calculate this share, we gather information on employment across industry sectors within MSAs using the 1980 County Business Patterns (CBP).

[9]As pointed out by Beaudry et al. (2010), in this case the industry mix controls are on top of the detailed industry adjustment already preformed on the dependent variable (three-digit SIC $\times$ establishment size). The industry mix controls therefore capture any additional indirect or "spillover" effects of industry mix in the IT regressions.

MSA's average IT budget per worker. This magnitude corresponds to an increase of 3.67 percent in the average IT budget per worker. Specification (2) finds no statistically significant correlation between the MSA's share of routine cognitive jobs in 1980 and the average IT budget per worker in 2015. Specification (3) finds a statistically significant correlation between the area's ratio of college equivalents to non-college equivalents and the average IT budget per worker in 2015, constructed as suggested by Beaudry et al. (2010). However, as we include all controls presented in specifications (1)-(3) together in specification (4), the area's ratio of college equivalents to non-college equivalents loses statistical significance. Differently, the impact of local rent prices shows just minor change in statistical significance between specifications (1) and (4). Finally, specification (5) controls for the MSA's average degree of offshorability of local jobs in 1980 – using the task offshorability index presented by Autor and Dorn (2013). We find again that the impact of local housing prices is robust to the addition of the controls.

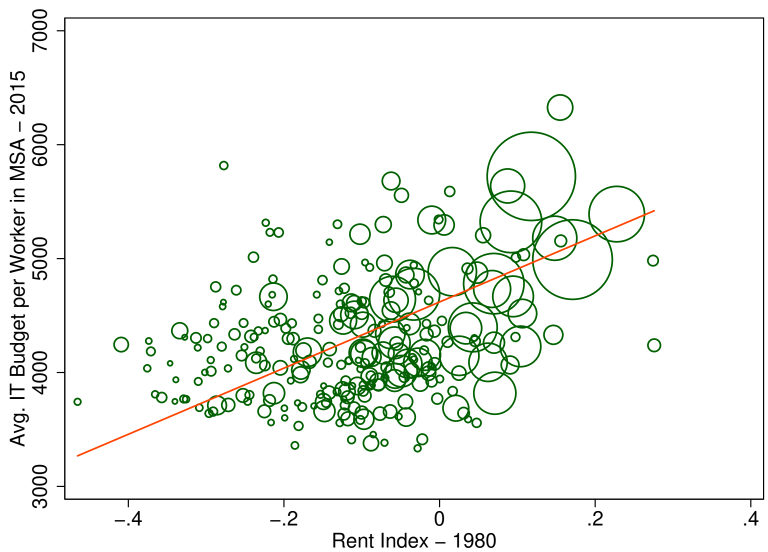


Figure 2: Avg. IT per worker vs local price level

Results from Table 2 highlight the importance of local prices for the establishment's IT budget per employee, even after controlling for firm and industry fixed effects. In fact, from specification (1), we observe that a one standard deviation increase in the local price index (an increase of 21.4 percent in the local price index) is associated with an increase in the establishment's average IT budget per worker of about \$60.97. This magnitude corresponds to an increase of 2.22 percent in the average IT budget per worker. While this effect seems small, we must keep in mind that we are already controlling for firm- and industry-fixed effects, as well as establishment's size and revenue and MSA's natural amenities, labor force composition, and industry mix. Moreover, notice that the coefficient of the local price index on IT budget per worker does not vary significantly across the different specifications presented in Table 2. Finally, the coefficients of the share of routine cognitive workers in 1980, MSA's average degree of offshorability of the local jobs in 1980, and MSA's ratio of college equivalent workers are all statistically insignificant.

Table 1: IT budget per worker – 2015

	(1)	(2)	(3)	(4)	(5)
	log(IT)	log(IT)	log(IT)	log(IT)	log(IT)
MSA log rent index 1980	0.182***			0.173**	0.183***
	(0.059)			(0.070)	(0.069)
MSA routine cognitive share 1980		0.225		0.189	0.377
		(0.405)		(0.395)	(0.406)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			0.0613*	0.013	0.018
			(0.0322)	(0.038)	(0.038)
MSA Offshorability 1980					-0.149
					(0.109)
Housing supply elasticity	0.000	-0.009	-0.0066	-0.001	0.001
	(0.007)	(0.007)	(0.0069)	(0.006)	(0.006)
Amenities	Yes	Yes	Yes	Yes	Yes
MSA's Industry Mix Controls	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	31.18	25.37	25.99	29.34	28.62
Adj. R^2	0.637	0.611	0.618	0.634	0.636
MSA	217	217	217	217	217

Standard errors in parentheses. The dependent variable in all columns is the logarithm of the average IT budget per employee in the metro area, adjusted for plant employment interacted with three-digit SIC industry. Each observation (an MSA) is weighted by its employment in 2015. MSA controls include the unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work States. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Fact 2. Routine Cognitive Occupations Decline Faster in Expensive Cities. We now turn to the second result: high cost locations feature a decline in the share of workers in routine occupations, whose tasks can presumably be automated after the introduction of new technology. We use 1980 as the pre-technology period in order to construct the control variables and compare it to the occupational composition in the period 1990–2015. Our focus on such a long span of time is motivated by the fact that in the model, we compare steady-state predictions and ignore short-term dynamics. Furthermore, the national trend shows a decline in the share of routine cognitive jobs starting in the late 1980s.

Table 3 presents the results of linear regressions of the change in the MSA's share of routine cognitive occupations between 1990 and 2015. Specification (1) indicates that a one standard deviation increase in the local price index (an increase of 21.4 percent in the local price index) is associated with a 0.8 percentage point larger drop in the routine cognitive share over 1990-2015. Thus, the most expensive places have about a 4.5 percentage point larger drop in the routine-cognitive share relative to the cheapest locations. This is one quarter lower than the average routine-cognitive share of 23 percent in 2015.

Table 2: IT Investment by Establishment - Firm and Industry FE

	log(IT)				
	(1)	(2)	(3)	(4)	(5)
MSA log rent index 1980	0.114***			0.131***	0.128***
	(0.040)			(0.048)	(0.048)
MSA routine cognitive share 1980		0.003		0.003	0.003
		(0.003)		(0.003)	(0.003)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			0.013	-0.025	-0.026
			(0.025)	(0.028)	(0.028)
MSA Offshorability 1980					0.048
					(0.090)
log(Site's Size)	-0.050***	-0.050***	-0.050***	-0.050***	-0.050***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
log(Site's Revenue)	2.214***	2.214***	2.214***	2.214***	2.214***
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
Headquarters dummy	0.043***	0.043***	0.043***	0.043***	0.043***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Housing elasticity	0.001	-0.005	-0.003	-0.000	-0.000
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	894.99	967.58	949.56	894.62	895.91
Adj. R^2	0.7487	0.7487	0.7487	0.7487	0.7487
No. of Sites	249,270	249,270	249,270	249,270	249,270
No. of Firms	126,180	126,180	126,180	126,180	126,180
No. of MSAs	218	218	218	218	218

Standard errors in parentheses. The dependent variable in all columns is the logarithm of the average IT budget per employee in the establishment. Each observation (an establishment) is weighted by the probability weight from a match between the Aberdeen data and the 2015 County Business Patterns. MSA controls include the unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work States. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Industry dummies are twp-digit SIC dummies. Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Specification (2) highlights the impact of the 1980 share of routine cognitive workers. Results show that the impact of the initial share of routine cognitive workers is both statistically and economically insignificant. A one standard deviation increase in the 1980 share of routine cognitive workers (an increase of 1.8 percentage points in the local share of routine cognitive jobs) is associated with a less than 0.1 percentage point larger drop subsequently and the effect is not statistically significant. Specification (3) shows that a one standard deviation increase in the share of "college-equivalent" workers relative to non-"college-equivalent" workers (representing a 26.6

Table 3: Change in routine-cognitive share, 1990-2015

	Δrout-cog				
	(1)	(2)	(3)	(4)	(5)
MSA log rent index 1980	-0.0413***			-0.0292**	-0.0283**
	(0.0114)			(0.0125)	(0.0127)
MSA routine cognitive share 1980		0.0172		0.0438	0.0604
		(0.0881)		(0.0867)	(0.0951)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			-0.0252***	-0.0175**	-0.0170**
			(0.0078)	(0.0087)	(0.0085)
MSA Offshorability 1980					-0.0129
					(0.0230)
Housing supply elasticity	-0.0032**	-0.0014	-0.0018	-0.0031**	-0.0030**
	(0.0014)	(0.0013)	(0.0012)	(0.0013)	(0.0013)
Amenities	Yes	Yes	Yes	Yes	Yes
MSA's Industry Mix Controls	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	6.87	6.48	6.88	6.78	6.59
Adj. R^2	0.308	0.252	0.305	0.322	0.319
$MSAs$	211	211	211	211	211

Standard errors in parentheses. The dependent variable in all columns is the change in the share of routine cognitive occupations in the MSA's employed labor force between 1990 and 2015. Each observation (an MSA) is weighted by its employment in 2015. MSA controls include the unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work states. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

percentage-point increase in this share) is associated with a 1.5 percentage point larger drop in the routine-cognitive share over 1990-2015. Specification (4) combines all three regressors plus controls in one regression. Both the local price index and the relative share of "college-equivalent" workers continue to be statistically significant, even after accounting for their covariation. However, the partial effect of each is smaller. The effect of a one standard deviation higher house price drops to 0.6 percentage point. Similarly, the effects of a one standard deviation higher ratio of "college equivalent" workers drops to 0.6 percentage point.

Finally, in specification (5) we control for the average degree of offshorability of the jobs in the MSA. Notice that our proxy for the offshorability of jobs in 1980 has a negligible and not statistically significant effect on the change in the routine cognitive share of MSAs. Nevertheless, our measure of offshorability only highlights the occupation's potential exposure to offshoring, and it is not unlikely that both offshoring and automation have happened concomitantly during the 1990-2015 period. Furthermore, results for the other variables of interest are in line with what we observed in specification (4). The effect of a higher local price index drops to about 70 percent of the observed effect in specification (1), though the difference in the coefficients is minor compared to specification (4). Similarly, the effects of the ratio of "college equivalent" workers drop by 33 percent. Overall, our results confirm the prediction that expensive locations have seen a substantially larger decline in their share of routine cognitive workers.

3.2 Patterns of Wage Inequality

So far, our focus has been on the composition of employment by occupation and its differential change across cities. The composition and the change are governed by equilibrium prices. We have already extensively analyzed housing prices, but equilibrium wages play an equally important role in balancing the impact of changes in the relative cost between technology and labor. Next, we first analyze the city-size wage premium and then we report in detail on the evolution of wage inequality, within and between cities.

The City-Size Wage Premium. When wages adjust in response to changes in the price of capital, the ratio of relative wages across regions will typically not be constant. Agents can optimally choose their occupation, and they have idiosyncratic tastes for cities and occupations as well as differences in innate abilities. In order to highlight the need for such potential extensions of the basic mechanism, we briefly present the impact of changes in housing costs and in investment in technology on the relative wages of routine cognitive and non-routine cognitive occupations across cities.

Table 4 shows how the relative MSA-level average wages for routine cognitive and non-routine cognitive occupations change over the period 1990–2015. As we can see, areas that were more expensive in 1980 have seen an increase in the wage premium observed by non-routine cognitive occupations. Specification (1) indicates that a one standard deviation increase in local price index (an increase of 21.4 percent in the local price index) is associated with a 4 percentage point increase in the wage premium of non-routine cognitive occupations relative to routine cognitive occupations over 1990-2015. Moreover, results are qualitatively and quantitatively robust to including the previously discussed controls, such as the relative share of "college-equivalent" workers, the share of offshorable jobs in the MSA, and the share of routine cognitive and non-routine cognitive jobs in the MSA in 1980. In particular, specification (5) shows that a one standard deviation increase in the local price index (an increase of 21.4 percent in the local price index) is associated with about a 3.5 percentage point larger increase in the wage premium of non-routine cognitive occupations relative to routine cognitive occupations over 1990-2015. Moreover, Table 5 shows that the driving force behind this result is the decline in real wages of routine cognitive occupations, which has been steeper in more expensive areas.[10] In particular, while on average mean wages for routine cognitive occupations drop by 2.6 percent in MSAs, this average dropped by 5.1 percent in MSAs with local price indexes one standard deviation above the average (an increase of 21.4 percent in the local price index). By contrast, we have seen no statistically significant difference in wage gains

[10] All our wage figures are in 1999 dollars.

for non-routine cognitive occupations across MSAs with different local price indexes.

Table 4: Wage ratios NRC-RC: 1990–2015

	$\Delta \ln \left(\frac{W_{NRC}}{W_{RC}}\right)$				
	(1)	(2)	(3)	(4)	(5)
MSA log rent index 1980	0.1997***			0.1829***	0.1688***
	(0.0306)			(0.0346)	(0.0335)
MSA RC share 1980		0.8103***		0.7150***	0.4564*
		(0.2580)		(0.2457)	(0.2690)
MSA NRC share 1980		0.2542*		-0.0125	-0.1343
		(0.1431)		(0.1781)	(0.1833)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			0.0745***	0.0201	0.0272
			(0.0235)	(0.0334)	(0.0340)
MSA Offshorability 1980					0.1794***
					(0.0688)
Housing supply elasticity	0.0032	-0.0074**	-0.0043	0.0003	-0.0015
	(0.0031)	(0.0030)	(0.0030)	(0.0031)	(0.0031)
Amenities	Yes	Yes	Yes	Yes	Yes
Industry Controls	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	15.61	10.62	10.49	17.40	19.43
Adj. R^2	0.616	0.563	0.563	0.632	0.643
$MSAs$	211	211	211	211	211

Standard errors in parentheses. The dependent variable in all columns is the change in the log ratio of nonroutine cognitive occupation and routine cognitive occupation real average wages between 1990 and 2015. Each observation (an MSA) is weighted by its employment in 2015. MSA controls include the unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work States. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Wage Inequality Within and Between Cities. In the previous section, we focused on average wages across occupations and the change in the wage premium across MSAs over time. Instead, in this section we look at the patterns of wage inequality within and between cities and how these patterns changed over time and across occupational groups. As a result, we are able to infer the role of within-occupational-group worker heterogeneity in explaining the variations observed in the data.

First, as pointed out in the literature (see Baum-Snow and Pavan (2013), Eeckhout et al. (2014), and Santamarıa (2018), among others), large cities are more unequal and inequality has gone up

Table 5: Wage changes 1990–2015: RC x NRC

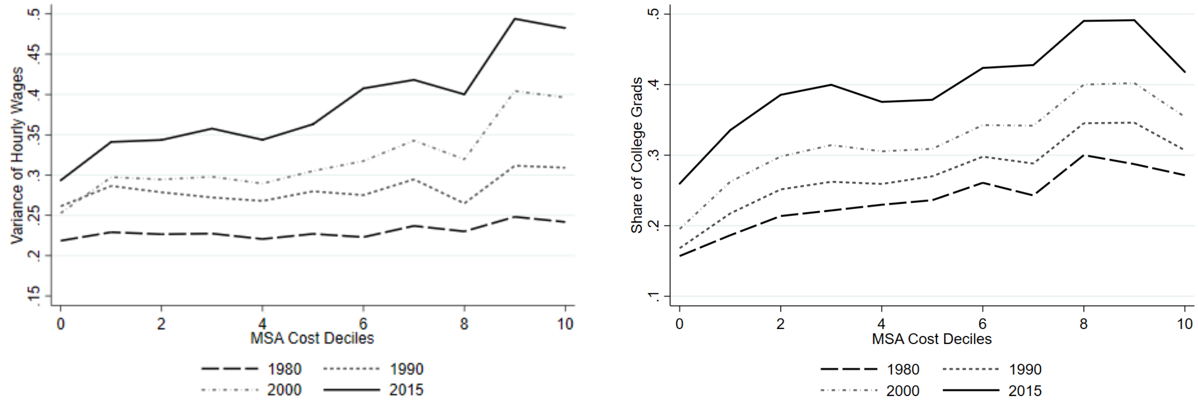
	$\ln\left(\frac{w_{RC,2015}}{w_{RC,1990}}\right)$	$\ln\left(\frac{w_{NRC,2015}}{w_{NRC,1990}}\right)$
MSA log rent index 1980	-0.1339***	0.0349
	(0.0482)	(0.0453)
MSA RC share 1980	0.1877	0.6441
	(0.3303)	(0.4291)
MSA NRC share 1980	0.2406	0.1063
	(0.2159)	(0.2433)
MSA's log $\left(\frac{S}{U}\right)$ in 1980	0.0336	0.0608
	(0.0428)	(0.0421)
MSA Offshorability 1980	-0.1460	0.0334
	(0.1025)	(0.1144)
Saiz (2010)'s housing supply elasticity	-0.0004	-0.0020
	(0.0053)	(0.0053)
Amenities	Yes	Yes
Industry Controls	Yes	Yes
CMSA Controls	Yes	Yes
F statistic	6.51	7.30
Adj. R^2	0.589	0.562
MSAs	211	211

Standard errors in parentheses. The dependent variable in columns 1 and 2 are the change in the log real average wages between 1990 and 2015 for routine cognitive and non-routine cognitive occupations, respectively. Each observation (an MSA) is weighted by its employment in 2015. MSA controls include the unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work States. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

over time. As we see in Figure 3a,[11] wage dispersion is larger in big cities.[12]

[11] Observe that in Figures 3 and 4, as well as the figures in appendix section F.2, we present deciles in terms of cities' cost of living, proxied by log(rent index) in 1980, and not in terms of the city size as in Baum-Snow and Pavan (2013) for example.

[12] In the Appendix we show that within-city inequality is driven by the wage gap between both the 90-50 percentiles and the 50-10 percentiles (see Figure A-11 in the online appendix). Most striking is the rise in the gap between 50 and 10 percentiles in large MSAs. However, as we see in Figure A-12 in the Appendix, we can attribute a significant share of the uptick in 50-10 gap for large cities to observable characteristics. Differently, the patterns for the 90-50 gap is qualitatively the same for wages and residual wages, as we can see by comparing Figures A-11b and A-12b in the Appendix.



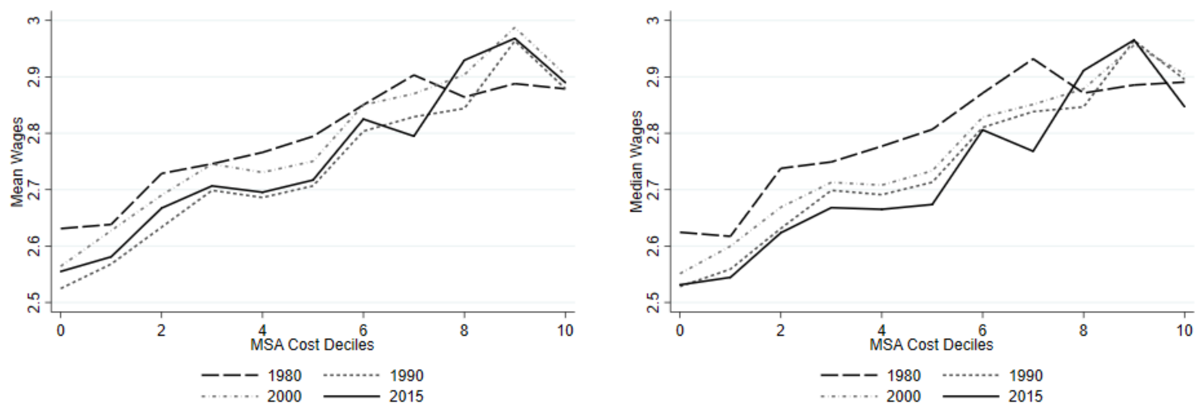
(a) Variance of Hourly Wages

(b) Share of College Graduates

Figure 3: Inequality within cities over time and by city's housing cost

Moreover, while we have seen that college attainment has been marginally higher in larger MSAs (Figure 3b), the results in Figure 3a still hold even after we control for several observable characteristics.

Instead, inequality *between* cities as measured by the city-size skill premium has not changed over time. Figure 4 shows that the increase in the mean and median wages with city size has not changed significantly over time.[13]



(a) Mean Hourly Wages

(b) Median Hourly Wages

Figure 4: Inequality between cities: the Urban Wage Premium over time

Finally, we decompose the overall variance in wages in terms of a within- and between-city contribution. Following the decomposition proposed by Lazear and Shaw (2009), the total variance in wages, σ^2, is given by

$$\sigma^2 = \sum_{j=1}^{J} s_j \sigma_j^2 + \sum_{j=1}^{J} s_j (\overline{w}_j - \overline{\overline{w}})^2. \tag{1}$$

[13]In fact, once we control for observable characteristics, as presented in Appendix section F.2, differences over time in mean and median residual wages are even smaller.

The first term on the RHS of equation (1) is the within-city component of the variance. s_j is the share of workers in the economy employed in city j, while σ_j^2 is the variance of wages in city j. The second term on the RHS of equation (1) represents the between-city component of the wage variance. In this expression, $\overline{w}_j$ is the mean wage in MSA j, and $\overline{\overline{w}}$ is the mean wage in the economy.

The results in Table 6 show that most of the wage dispersion is due to the within-city component (around 95 percent). Moreover, the decomposition in terms of within- and between-city components is persistent over time. Consequently, the contribution of each component to the overall increase in wage inequality has stayed proportional to each component's contribution to the overall dispersion. These results are preserved even when we focus on wage dispersion within occupational groups (See Table A-21 in Appendix Section F.2) as well as when we control for observables (Tables A-22 and A-23 in Appendix Section F.2).

Table 6: Variance decomposition log hourly wages

			Variance		
Year	Total	Within City	Between City	% Within	% Between
1980	0.238	0.227	0.011	95%	5%
1990	0.296	0.279	0.018	94%	6%
2000	0.337	0.320	0.017	95%	5%
2015	0.408	0.385	0.023	94%	6%

We need to keep in mind though that, while the bulk of wage dispersion is due to the within-MSA component, this does not mean that geographical components do not play a key role in explaining wage dispersion. As technology is adopted unevenly across space and workers and firms choose to search for workers and post jobs in different cities, these decisions affect both the within- and between-MSA components of wage inequality. Consequently, our decomposition exercise mostly says that, in terms of wage inequality, while cities vary in terms of wage inequality, the bulk of the wage inequality happens within the average city.

4 The Economic Mechanism

In order to highlight the economic mechanism and the general equilibrium effects of technological change on polarization, we present a model where firms located in heterogeneous cities hire heterogeneously skilled workers. Simultaneously, firms adopt IT technology. The objective is to analyze the equilibrium allocation of workers of different skills to cities and in different occupations (routine and non-routine) in the light of changing prices of IT technology. In Section 4.3, we estimate the model.

4.1 Model Setup

Cities and Population. Consider a static economy with heterogeneous cities and heterogeneously skilled workers. City $j \in \mathcal{J}$ is characterized by its production opportunities, housing supply, and amenities. Each city produces a single final output that is a combination of different occupations i. Each occupation produces output by combining labor in efficiency units m_{ij} with capital k_{ij}. Cities are heterogeneous in their city-wide productivity A_j as well as in their occupation-specific labor productivity $A_{l,ij}$. The productivity of capital is skill-specific but otherwise common economy-wide $A_{k,i}$.

Workers are heterogeneous in their skills $\mathbf{s}$, tastes for jobs $\mathbf{t}$, and tastes for locations $\mathbf{a}$. Each worker is endowed with a set of skills for each occupation i, summarized by the vector $\mathbf{s} = [s_1, \ldots, s_I]$. The skill vector represents how many efficiency units of labor a worker could supply in each occupation. The distribution of skills is given by $G(\mathbf{s})$. The income a worker earns in an occupation is the product of efficiency units and the wage per efficiency unit: $w_{i,j}(\mathbf{s}) = s_i \tilde{w}_{i,j}$.

Preferences. Workers consume the final good and housing, where housing must be consumed in the same city as the workplace. The preferences over consumption and housing are a Cobb-Douglas aggregate of the final good c and housing h given by

$$u(c, h) = c^{1-\alpha} h^{\alpha}, \tag{2}$$

where α governs the spending share on housing and the final good. Workers maximize utility subject to their budget constraint

$$c + p_j h \leq w_j^*, \tag{3}$$

where income $w_j^* \equiv w_j(\mathbf{s}, i^*)$ depends on local wages per efficiency unit of labor, the worker's skill, and the worker's occupation choice. The price of the final good is normalized to one and is the same everywhere. Housing has a location-specific price p_j.

A worker's utility from choosing a location and occupation depends not only on real income but also on the idiosyncratic taste for occupations and locations. The indirect utility V of a location-occupation pair for a worker with skills $\mathbf{s}$ and tastes $\mathbf{a}, \mathbf{t}$ is

$$V(i, j, \mathbf{s}, \mathbf{a}, \mathbf{t}) = a_j t_i v(\tilde{w}_{i,j} s_i, p_j). \tag{4}$$

The idiosyncratic taste for location a_j follows a Fréchet distribution with shape parameter τ and location parameter $\bar{a}_j$. The idiosyncratic taste for occupation t_i follows a Fréchet distribution with shape parameter η and location parameter $\bar{t}_i$. Idiosyncratic tastes are i.i.d. across individuals and locations. For simplicity we assume that the taste for a location is drawn first and after a worker has chosen a location, her taste for occupations is drawn followed by the occupation choice. This setup represents the idea that the location choice is relatively more permanent compared to the occupation choice. Given the specification of tastes we can derive the probability distribution of workers' occupation and location choices conditional on skills and prices in closed form.

Technology. In each city, there is a technology operated by a representative firm with access to the city-specific technology. The production function F has a nested CES structure

$$A_j F(\mathbf{m}_j, \mathbf{k}_j, \mathbf{A}_j) = A_j \left\{ \sum_i A_{l,ij}^{\frac{\gamma_i}{\lambda}} \left[m_{ij}^{\gamma_i} + A_{k,i} k_{ij}^{\gamma_i} \right]^{\frac{\lambda}{\gamma_i}} \right\}^{\frac{1}{\lambda}}. \tag{5}$$

A_j is the total factor productivity of city j. Production combines labor and capital (IT) within occupations with a finite elasticity of substitution, which is governed by γ_i.[14] The elasticity of substitution between capital and labor is occupation specific, allowing capital to complement labor in some occupations and substitute labor in others. Occupation-enhancing productivity $A_{l,ij}$ for each occupation i is allowed to vary across cities j, to capture preexisting specialization of cities. The capital productivity, relative to labor, $A_{k,i}$ is the same across cities, implying that two cities with the same relative price of capital and labor would employ capital and labor in the same ratio in an occupation. In other words, we assume that the capital technology used in a given occupation has no inherent bias towards specific cities.

The output of the different occupations is aggregated with a finite elasticity of substitution that is determined by λ. The final output is freely traded and its price is normalized to 1. Firms maximize profits and are price takers. Both output and factor markets are competitive; thus, both labor and capital are paid according to their marginal product. Each efficiency unit of labor costs $\tilde{w}_{i,j}$ and capital supply is fully elastic at the rental rate r, which is taken as given.

In section 4.2.1, we derive analytical results for a version of the model with a simplified technology, in which there are two cities, $j \in \{1, 2\}$ and three skill levels $i \in \{1, 2, 3\}$. We define the simplified technology as follows.

Definition 1 *The simplified technology satisfies:*

$$A_j F(m_{1j}, m_{2j}, m_{3j}, k) = A_j \left\{ m_{1j}^{\gamma_1} A_{l,1} + \left(m_{2j}^{\theta} A_{l,2} + k_j^{\theta} A_k \right)^{\frac{\gamma_2}{\theta}} + m_{3j}^{\gamma_3} A_{l,3} \right\} \quad , \textit{ where} \quad \gamma_2 < \theta. \tag{6}$$

Notice that, in this simplified technology case, capital and "middle-skill occupations" are substitutes (automation case). In Appendix Section G.2, we consider the case of top-skill and capital complementarity, i.e., skill-biased technological change.

Housing. The housing market is competitive. Housing supply follows the price-quantity schedule

$$p_j(H) = \phi_j H^{\epsilon_{p,j}}. \tag{7}$$

In equilibrium, housing supply H adjusts such that the housing amount demanded by workers equals the amount supplied. The inverse housing supply elasticity ϵ_p is finite and captures limitations to

[14] The within-occupation elasticity of substitution between capital and labor is $\frac{1}{1-\gamma_i}$

increasing the stock of housing in a given city. Furthermore, we allow the housing supply elasticity to vary across cities. The revenue from the housing market is consumed by absentee landowners. Housing demand in a city is given by

$$H_j^D = \frac{\alpha}{p_j} \int s_i \tilde{w}_{i,j} P(occ = i, city = j | skill = \mathbf{s}) dG(\mathbf{s}). \tag{8}$$

Spatial Equilibrium. Spatial Equilibrium is defined as a set of wages per efficiency unit of labor $\{\tilde{w}_{i,j}^*\}_{\substack{i\in\mathcal{I}\\ j\in\mathcal{J}}}$, housing prices p_j^*, a distribution of workers across space and occupations $P^*(occ = i, city = j|skill = \mathbf{s})$ and capital k_{ij}^* for all occupations $i \in \mathcal{I}$ and cities $j \in \mathcal{J}$, that fulfill the following conditions:

1. Labor markets clear, that is (16) holds with $m_{i,j} = \int s_i P^*(occ = i, city = j|skill = \mathbf{s}) dG(\mathbf{s})$. Labor supply satisfies (14) and (13).
2. Optimal capital demand (17) holds at given prices and labor demand.
3. Housing markets clear, that is (7) and (8) are satisfied at housing prices p_j^*.

4.2 Solution and Analytical Results

We now solve for the conditions that pin down the equilibrium allocation. Before we estimate the model, we then derive analytical results on a simplified version. This allows us to gain crisp insights into the economic mechanism underlying urban job polarization.

The Worker's Solution. Within a given city j and given a wage $w_{ij} = \tilde{w}_{ij} s_i$, a citizen chooses consumption bundles $\{c_{ij}, h_{ij}\}$ to maximize utility subject to the budget constraint:

$$\begin{aligned} \max_{\{c_{ij},h_{ij}\}} u(c_{ij}, h_{ij}) &= c_{ij}^{1-\alpha} h_{ij}^{\alpha} \\ \text{s.t. } c_{ij} + p_j h_{ij} &\leq w_{ij} \end{aligned} \tag{9}$$

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 $v(w_{ij}, p_j) = (1-\alpha)^{(1-\alpha)} \alpha^\alpha \frac{w_{ij}}{p_j^\alpha}$, which completes the derivation of the indirect utility of a location-occupation pair in equation (4).

Given the specification of tastes we can derive the probability distribution of workers' occupation and location choices conditional on skills and prices in closed form.

$$\bar{u}(i,j,\mathbf{s}) = \bar{a}_j \bar{t}_i v(s_i \tilde{w}_{ij}, p_j) \tag{10}$$

$$\mathbb{E}_{\mathbf{t}}[\max_i u(i,j,\mathbf{s})] = \sum_i \bar{u}(i,j,\mathbf{s}) \left(\frac{1}{\sum_{i'} \left(\frac{\bar{u}(i,j,\mathbf{s})}{\bar{u}(i',j,\mathbf{s})} \right)^{-\eta}} \right)^{1-\frac{1}{\eta}} \Gamma\left(1 - \frac{1}{\eta}\right) \tag{11}$$

$$P(\text{city} = j | \text{skill} = \mathbf{s}) = \frac{\mathbb{E}_{\mathbf{t}}[\max_i u(i,j,\mathbf{s})]^\tau}{\sum_{j'=1}^{J} \mathbb{E}_{\mathbf{t}}[\max_i u(i,j',\mathbf{s})]^\tau} \tag{12}$$

$$P(\text{occ} = i | \text{city} = j, \text{skill} = \mathbf{s}) = \frac{u(i,j,\mathbf{s})^\eta}{\sum_{i'=1}^{I} u(i',j,\mathbf{s})^\eta} \tag{13}$$

See Appendix H for the derivation. The joint distribution of skills and choices of occupation and location then follows as

$$P(\text{skill} = \mathbf{s}, \text{occ} = i, \text{city} = j) = G(\mathbf{s}) P(\text{city} = j | \text{skill} = \mathbf{s}) P(\text{occ} = i | \text{city} = j, \text{skill} = \mathbf{s}). \tag{14}$$

The Firm's Solution. All firms are price takers and do not affect wages or capital markets. 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:

$$\max_{m_{ij}, \forall i} A_j F(m_{1j}, ..., m_{Ij}, k_j) - \sum_{i=1}^{I} w_{ij} m_{ij} - r k_j, \tag{15}$$

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 $A_j F_{k_j}(m_{ij}, k_j) = r$.[15]

For the general model setup with the CES technology, optimal labor and capital demand obtained from profit maximization satisfies

$$\tilde{w}_{i,j} = A_j \left\{ \sum_i A_{l,ij}^{\frac{\gamma_i}{\lambda}} \left[m_{ij}^{\gamma_i} + A_{k,i} k_{ij}^{\gamma_i} \right]^{\frac{\lambda}{\gamma_i}} \right\}^{\frac{1}{\lambda}-1} A_{l,ij}^{\frac{\gamma_i}{\lambda}} \left[m_{ij}^{\gamma_i} + A_{k,i} k_{ij}^{\gamma_i} \right]^{\frac{\lambda}{\gamma_i}-1} m_{ij}^{\gamma_i - 1} \tag{16}$$

$$r = A_j \left\{ \sum_i A_{l,ij}^{\frac{\gamma_i}{\lambda}} \left[m_{ij}^{\gamma_i} + A_{k,i} k_{ij}^{\gamma_i} \right]^{\frac{\lambda}{\gamma_i}} \right\}^{\frac{1}{\lambda}-1} A_{l,ij}^{\frac{\gamma_i}{\lambda}} \left[m_{ij}^{\gamma_i} + A_{k,i} k_{ij}^{\gamma_i} \right]^{\frac{\lambda}{\gamma_i}-1} A_{k,i} k_{ij}^{\gamma_i - 1}. \tag{17}$$

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, as presented in Definition 1. In the Appendix, we present the firm's first order conditions for each j and all skill types i and capital.[16] Even without fully solving the system of equations for

[15] 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.

[16] We also solve the allocation under *separable* technology as a special case of the more general technologies

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 $A_{l,i}$ 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.[17]

4.2.1 Main Theoretical Results

We now derive analytical results for the simplified technology defined above. We also impose additional constraints on the worker's problem in order to gain tractability. In particular, we consider the following special case

1. Workers have fixed occupations: a type i worker has $s_i = 1$ and $s_{i'} = 0$ for $i' \neq i$ and $i \in \{1, 2, 3\}$.

2. Workers have no idiosyncratic preferences for locations.

3. Housing supply is fixed.

Given these simplifications, we first establish the relationship between TFP and house prices. As mentioned before, we focus on the automation case. When cities have the same amount of land, we can establish the following result.

Proposition 1 (TFP and Housing Prices) *Assume the simplified technology presented in Definition 1. Then the more productive city has higher housing prices:* $A_i > A_j \Rightarrow p_i > p_j$.

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 such a city even if the productivity is lower. Because in our empirical application we consider large metropolitan areas (New York City MSA for example includes large parts of New Jersey and Connecticut), we believe that this assumption does not lead to much loss of generality.[18]

presented in the paper.

[17]For a given order i, wages may not be monotonic as they depend on the relative supply of skills as well as on $A_{l,i}$. 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-skill workers can perform lower-skill 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.

[18]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.

We now focus on the relation between 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 stronger, leading to an increase in the use of capital in order to replace skilled workers. Hence, high-TFP cities demand more capital.

Proposition 2 (TFP and Capital Demand) *Assume the simplified technology presented in Definition 1. Then the more productive city has higher investment in IT:* $A_i > A_j \Rightarrow k_i > k_j$.

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 (IT and City Size) *Assume the simplified technology, and* $\gamma_2 < \theta$. *Then the more productive city has a larger population:* $A_1 > A_2 \Rightarrow S_1 > S_2$.

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

Theorem 2 (IT and Spatial Sorting) *Assume the simplified technology,* $\gamma_i \equiv \gamma,\ \forall i \in \{1, 2, 3\}$, *and* $\gamma < \theta$. *Then the larger city has a more unequal skill distribution:* $A_1 > A_2$ *implies that city 1 has a thicker tailed skill distribution.*

4.3 Estimation

In this section we quantify the economic mechanism using our model. Our model is set up to capture key features of the data. A key feature for the analysis of the impact of IT on the labor market is to allow for potentially heterogeneous effects based on the type of job or workers' skills (Autor and Dorn, 2013; Krusell et al., 2000). We do this by making the elasticity of substitution between labor and IT occupation specific. To model workers' responses in terms of labor supply across jobs, we model the occupation choice in the spirit of Roy (1951) and allow for idiosyncratic tastes for occupations. To capture a more realistic supply elasticity of workers across cities, we model idiosyncratic tastes for cities, where workers trade off local wages, housing costs, and their valuation of local amenities when they choose where to live. Finally, following the evidence in Saiz (2010), the housing market is modeled as having a finite supply elasticity that varies across cities to capture differences in housing supply restrictions.

4.3.1 Estimation Approach

We estimate the main model parameters by indirect inference (Gourieroux et al., 1993), after we calibrate some parameters based on external evidence. In particular, we estimate a vector of model parameters θ by minimizing the weighted square distance between a vector of moments estimated in the data $\widehat{m}$ and the corresponding model-implied moments $m(\theta)$. The model moments are directly calculated from the equilibrium distribution of workers and prices. To capture heterogeneity across cities in the data, we bin them into three distinct groups based on their rent index calculated from the 1980 census. Each bin approximately represents one third of workers. Occupations are grouped as described in Section 2. See Appendix H for details on the estimation procedure and the calculation of moments.

Externally Calibrated Parameters. Scale parameters of the Fréchet distributions of location and occupation tastes are externally calibrated. Following Kennan and Walker (2011) and Monte et al. (2018), we set the scale parameter of the location tastes within their range of estimates to $\tau = 4$. The scale parameter of the taste for occupations is set to $\eta = 5$ following evidence by Berger et al. (2019).[19] Further, we set the elasticity of substitution of output across occupation nests at $\frac{2}{3}$, implying a value of $\lambda = -0.5$. We pick this value to fall within the range of estimates by Goos et al. (2014), who estimate an elasticity of substitution of 0.9 between tasks of differing routine intensity, and Lee and Shin (2017), who estimate an elasticity of substitution of 0.7 between different tasks and an elasticity of substitution of 0.34 between managers and other workers. Further, we calibrate the housing supply price-quantity elasticity ϵ_j directly to the values estimated by Saiz (2010).

Estimated Parameters and Targeted Moments The moments used in the estimation are summarized in Table 7 and the corresponding parameter estimates are shown in Table 8. The skill distribution is parameterized as a multivariate log-normal with zero mean and a diagonal covariance matrix. We target:

1. the average wage in each city to estimate the average productivity by city A_j (3 moments and 3 parameters),

2. the rent index in each city to estimate the intercept of the housing supply function ϕ_j (3 moments and 3 parameters),

3. the relative size of cities to estimate the parameter governing the average taste for a city $\bar{a}_j$, with the normalization $\bar{a}_1 = 1$ (2 moments and 2 parameters),

4. the share of workers by occupation group and the difference across cities to inform the relative productivity of each occupation group across cities $A_{l,ij}$ with the normalization that $A_{l,RCj} = 1 \,\forall j$ (9 moments and 9 parameters),

[19] Their estimate is 5.38 for "within market" substitutability of firms.

5. the average log wage (per week in full-time jobs) by occupation to inform the average taste for occupations $\bar{t}_i$, with the normalization $\bar{t}_1 = 1$ (3 moments and 3 parameters),

6. the standard deviation of log wages by occupation to inform the standard deviations σ_i of the log-normal skill distribution (4 moments and 4 parameters),

7. the relative importance of PCs across occupations calculated from O*NET as a measure of relative IT usage per worker across occupations, combined with the aggregate IT share out of labor and IT spending as calculated in Eden and Gaggl (2018) to estimate the productivity of capital relative to labor $A_{k,i}$ by occupation group (4 moments and 4 parameters),

8. the elasticity of employment shares with respect to IT prices as implied by the calibrated model in vom Lehn (2020) to target the elasticity of substitution between capital and labor by occupation group. Following the classification in vom Lehn (2020) we set γ to be equal for routine cognitive and routine manual jobs, but keep them as distinct categories in our model. (3 moments and 3 parameters).

This makes a total of 31 moments targeted and 31 parameters estimated. The overall fit is very good, albeit at the estimated parameters the model slightly over-predicts the share of non-routine manual workers and under-predicts the share of routine manual workers.

Panel A in Table 7 shows that expensive cities are not only expensive in terms of rent, but wages are also substantially higher and they are several times larger in terms of population. Furthermore, there is non-trivial sorting of occupations across space, with non-routine jobs being more prevalent in expensive cities. This is striking for non-routine cognitive jobs, whose share of employment is 6.8 percentage points higher in expensive cities compared to cheap cities. In contrast, the share of routine jobs, both manual and cognitive, is lower in more expensive cities. This pattern highlights that job polarization in terms of employment not only is an aggregate phenomenon over time, but also presents itself in the cross-section of cities.

In Panel B, we present moments calculated at the occupation level. Average wages vary substantially across occupation groups. Wages in non-routine cognitive jobs are more than twice as large as in non-routine manual jobs. Wage inequality is, however, not only substantial between occupations but also within. Within-occupation group log wage standard deviations range from 0.55 in non-routine manual jobs to 0.7 in non-routine cognitive jobs. The importance of PC usage, as measured in O*NET, is larger in cognitive jobs compared to manual jobs. As the scale of the measure is not in units of the final good, we use it only to compare across occupations, with the normalization that it sums to one. To measure the overall importance of IT in the economy, we calculate the share of aggregate costs of IT out of labor and IT (Eden and Gaggl, 2018).

The remaining targets are the elasticities of occupation employment shares with respect to IT prices based on the model and calibration of vom Lehn (2020). We target an elasticity that is negative for non-routine cognitive jobs, as expected if non-routine cognitive jobs are complements to IT, and positive for routine jobs. This is in line with IT substituting labor in routine jobs, while complementing labor in non-routine cognitive jobs. Finally, the implied elasticity of the non-routine manual share is also positive, but slightly smaller than for routine jobs.

Table 7: Moments 2015 and model fit

Panel A: City-level Moments

	low rent index		mid rent index		high rent index	
Moment	**Data**	**Model**	**Data**	**Model**	**Data**	**Model**
Average log wage	6.69	6.69	6.85	6.85	6.93	6.93
	(0.0141)		(0.031)		(0.0501)	
Log rent (index)	-0.227	-0.227	-0.0147	-0.0147	0.328	0.328
	(0.0207)		(0.0574)		(0.0701)	
Log size difference to expensive cities	-2.06	-2.06	-0.796	-0.796		
	(0.314)		(0.325)			
Occupation Share (relative to low rent cities in pp)						
non-routine manual			-0.34	-0.33	1.5	1.5
			(0.66)		(0.6)	
routine manual			-4.0	-4.1	-6.2	-6.4
			(0.9)		(1.4)	
routine cognitive			-1.6	-1.4	-2.0	-1.9
			(0.51)		(0.67)	
non-routine cognitive			5.9	5.9	6.8	6.8
			(1.3)		(2.0)	

Panel B: Occupation-level Moments

	non-routine manual		routine manual		routine cognitive		non-routine cognitive	
Moment	**Data**	**Model**	**Data**	**Model**	**Data**	**Model**	**Data**	**Model**
Share in %	11.0	13.0	22.0	21.0	24.0	24.0	43.0	43.0
	(0.24)		(0.52)		(0.23)		(0.64)	
$log(w)$	6.2	6.2	6.6	6.6	6.7	6.7	7.2	7.2
	(0.011)		(0.011)		(0.015)		(0.019)	
$\sigma(log(w))$	0.55	0.54	0.59	0.59	0.68	0.68	0.7	0.7
	(0.0034)		(0.0048)		(0.0073)		(0.0052)	
PC importance	0.16	0.16	0.17	0.17	0.32	0.32	0.34	0.34

Panel C: Additional External Moments

Moment	**Data**	**Model**
ICT Share	0.1	0.1
ICT price index, Base 2015 = 1	1.0	1.0
Elasticity NRM share - IT price	0.1	0.1
Elasticity R share - IT price	0.12	0.12
Elasticity NRC share - IT price	-0.16	-0.15

Note: Data moments calculated from the American Community Survey 2014-2016 (Ruggles et al., 2020) and O*NET. Standard errors in parentheses calculated by bootstrap resampling of MSAs. City classifications are assigned based on the 1980 rent index calculated for each MSA. The number of cities in each group is the same in the data and the model. Thresholds are pinned down such that each class represents 33% of the population. IT share calculated as in Eden and Gaggl (2018) and elasticity of employment with respect to IT price calculated from vom Lehn (2020). See Appendix H for details.

Table 8: Estimated Parameters 2015

Panel A: City level Parameters

Parameter	**low rent index**	**mid rent index**	**high rent index**
TFP A_j	16348	23817	25944
	(1047)	(1898)	(4049)
Amenity a_j	1.0	1.4	2.0
		(0.092)	(0.2)
House price shifter ϕ_j	0.04	0.027	0.03
	(0.0032)	(0.0028)	(0.0026)
House price elasticity ϵ_j	0.49	0.65	1.2
Occupation Productivity $A_{l,ij}$			
non-routine manual	4.4	4.1	2.8
	(0.76)	(0.53)	(0.18)
routine manual	1.9	2.7	3.5
	(0.14)	(0.25)	(0.54)
routine cognitive	1.0	1.0	1.0
non-routine cognitive	2.2	3.2	3.3
	(0.13)	(0.22)	(0.45)

Panel B: Occupation level Parameters

Parameter	**non-routine manual**	**routine manual**	**routine cognitive**	**non-routine cognitive**
Occupation Amenity $\bar{t}_i$	1.0	0.46	0.61	0.29
		(0.11)	(0.15)	(0.069)
Std dev. Skills $\sigma_i(log(s))$	2.0	1.1	1.4	0.94
	(0.48)	(0.027)	(0.044)	(0.0088)
Capital Productivity $A_{k,i}$	0.048	0.0099	0.015	2.1
Capital-Labor substitution parameter γ_i	0.65	0.69	0.69	-0.69

Panel C: Additional Parameters

Parameter	**Value**	**Source/Explanation**
τ	4.0	Dispersion of tastes for locations (Kennan and Walker, 2011; Monte et al., 2018).
η	5.0	Dispersion of tastes for jobs (Berger et al., 2019)
λ	-0.5	Occupation output elasticity $\frac{2}{3}$ (Goos et al., 2014; Lee and Shin, 2017)

Note: Standard errors in parentheses. Housing supply elasticity from Saiz (2010). See Appendix H for details.

Using these targeted moments we estimate the parameters of the model, reported in Table 8. For the elasticity of substitution between capital and labor by occupation we find: 2.85 in non-routine manual jobs, 3.2 in routine jobs and 0.59 in non-routine cognitive jobs (see the value of γ_i in Table 8). Thus, IT is estimated to be complementary to labor in non-routine cognitive occupations, while it substitutes labor in routine cognitive and manual jobs. This suggests that IT impacts the labor market through both substituting away labor in routine cognitive and manual jobs and by complementing labor in non-routine cognitive jobs. We discuss the quantitative implications of our estimates for the allocation of labor to jobs and cities in the next section. There are, to the best of our knowledge, no directly comparable estimates of the elasticity of substitution between IT and labor by occupation in the literature. There are, however, recent estimates of substitution elasticities between other types of capital and labor by occupation. Caunedo et al. (2021) estimate the elasticity of substitution between labor and "capital embodied technological change" to be between 0.5 and 2.0 across occupations, suggesting that other forms of technology may feature a bias across occupations similarly large as IT. Adachi (2021) estimates an elasticity of substitution between robots and labor that range between 0.8 and 4 depending on the occupation group, thus also suggesting a large variation in the substitutability between robots and labor by type of job.

Expensive cities are more productive, as the estimated TFP is larger. But the estimates indicate that they are also more desirable places to live, meaning that in the absence of the higher estimated "average amenity" for expensive locations, their populations would decrease. The calibrated housing supply elasticity, combined with the estimated housing supply shifter, determines the housing supply of a city. The estimates indicate that the differences in the housing supply elasticity are a key determinant of the different levels of housing prices between medium priced and expensive cities, as the supply shifter is similar across those cities. These differences highlight that for expensive cities the expansion of housing supply is relatively costly, while for medium and cheap cities a rise in the population could be accommodated with a smaller rise in house prices.

The productivity parameters $A_{l,ij}$ reflect the allocation of employment across cities and occupations. The productivity patterns are closely related to the employment allocation across space, indicating that it is not only IT that leads to sorting in space, but other factors are also important in determining the differences in occupational composition of labor demand across cities. The remaining parameters in Panel B are the workers' average taste for occupations and the standard deviation of skills in each type of job. The parameter governing the location of the Fréchet distribution of tastes for jobs varies inversely with an occupation's average wage, since the model features a competitive labor market; wages would otherwise (almost) equalize.[20] The estimated standard deviation of skills is estimated to fit the standard deviation of wages *within* occupations, given the normalization that the mean of the log skill distribution is zero in each dimension. This leads to large estimated standard deviations of skills for occupations with a smaller employment share, e.g. non-routine manual jobs. The within-occupation variability in skills on the one hand replicates the within-occupation variability of wages, but it also captures to what extent workers'

[20] There is a finite supply elasticity of labor, so wages would not equalize exactly.

skills determine their occupation choice.

4.3.2 The Rise of IT in the Estimated Model

We consider an experiment where the quality-adjusted price of IT capital falls from its 1990 level to its value in 2015, which is by approximately 65 percent. We take the model as estimated for 2015 to be the baseline, and compare it to the model-implied allocation when only IT prices change back to their 1990 level. With this exercise, we evaluate to what extent the fall in prices of IT can explain employment and wage trends in the US between 1990 and 2015. We show results not only for the economy as a whole but also to what extent they are heterogeneous across low- and high-rent cities.

Table 9: The rise of IT and urban job polarization

	Δ Data 1990-2015	**Δ Model IT price ↓ 65%**
Employment Share by Occupation in pp		
non-routine manual	3.3	-0.57
routine manual	-7.6	-0.47
routine cognitive	-5.8	-1.6
non-routine cognitive	10.0	2.7
Avg log wage by Occupation		
non-routine manual	-0.032	0.064
routine manual	-0.074	0.09
routine cognitive	0.042	0.066
non-routine cognitive	0.14	0.12
High vs. Low Rent City Gaps		
Log Size	-0.34	-0.0014
Log House Price	0.021	0.0046
Average Log Wage	-0.0011	0.01
High vs. Low Rent City Employment Shares in pp		
non-routine manual	0.65	-0.47
routine manual	1.8	0.042
routine cognitive	-2.5	-0.71
non-routine cognitive	0.14	1.1
High vs. Low Rent City Wages		
non-routine manual	-0.073	-0.0034
routine manual	-0.071	0.0022
routine cognitive	-0.026	-0.0039
non-routine cognitive	0.021	0.0068

Note: The column Δ Model is calculated as the difference between the model allocation at the estimated parameters in Table 8 with IT prices normalized to $r = 1$, and the model allocation at $\hat{r} = \frac{r_{1990}}{r_{2015}}$ with the parameters otherwise held constant. Here r_y denotes the price of IT relative to the GDP deflator in year y, calculated as in Eden and Gaggl (2018).

Table 9 summarizes the results. Overall employment has reallocated substantially from routine to non-routine jobs; the share of routine employment fell by 13 percentage points between 1990 and 2015. The fall in IT prices can explain a substantial share of this reallocation, particularly for cognitive jobs. We find that the fall in IT prices explains 28 percent of the fall in the routine-cognitive employment share and 27 percent of the rise in non-routine cognitive employment share. The overall predicted changes for manual jobs are much smaller. These results confirm the hypothesis that IT has been particularly important for jobs that involve cognitive-intensive tasks tasks.

Turning to wages across occupations, the fall in IT prices predicts average wage changes in both routine cognitive and non-routine cognitive jobs that are almost the same as in the data, albeit slightly too large for routine cognitive jobs. The model also predicts a rise in average wages for manual jobs, which is not true in the data. This highlights that to explain the evolution of employment and wages across *all* occupation groups, the rise of IT by itself is not sufficient. Still, IT can explain a substantial share of the changes in cognitive occupations, where it is used more widely, indicating a potentially prominent role for future changes in employment and wages.

When we compare the effects of IT on cities we find relatively small effects on the gap between low- and high-rent cities in terms of average wages, house prices and city size. This is in line with the data, except that small cities have gained substantially in population relative to large cities in the US. This may partially be driven by overall population growth combined with the higher elasticity of housing supply in cheaper locations. This change is not captured in the counterfactual exercise presented in this section.

For the impact of IT on the sorting of jobs across cities we find that the model predicts a substantial drop of 0.71 percentage point in the routine cognitive employment share in high-rent cities compared to low-rent cities. Thus, IT can explain about 28% of the increased sorting of routine cognitive jobs to cheap cities. Further, the model predicts increased sorting for non-routine cognitive employment to expensive cities by 1.1 percentage points. In the data, the increase in the concentration of non-routine cognitive employment in expensive cities has been smaller. The model predicts that wage premium of expensive cities falls for routine cognitive jobs by about 0.4% compared to a fall of 2.6% in the data. Non-routine cognitive jobs are predicted to gain approximately 0.7% in terms of average wages relative to cheap cities. In the data, this gain was even larger at 2.1%.

Overall, the calibrated model implies that the rise of IT can explain an important part of the variation in employment and wages for cognitive occupations both in space and over time. This reallocation is predicted to be stronger in high-rent cities, which are using IT more intensively due to higher labor costs. Our results highlight that the impact of technology on workers depends strongly on their specialization into jobs and that this impact is heterogeneous across cities, because of the more intensive usage of IT in expensive locations.

Discussion. Given that IT can hardly explain *all* the changes in wages and employment over time and across cities, we highlight some other important factors here that potentially also interact with IT usage. Other technologies, in particular automation technology like robots (Graetz and

Michaels, 2018; Acemoglu and Restrepo, 2020), and trade (Autor et al., 2016) have been shown to be important factors in explaining employment trends. Given the strong geographic concentration of the manufacturing sector, and the potential spatial impact of changes in trade technology (Ducruet et al., 2019), these are complementary mechanisms that affect not only the aggregate changes in employment, but also the distribution of jobs in space. Furthermore, the long-run effects of technology may also work through the interaction with labor supply by changing skill accumulation and education decisions as in Dvorkin and Monge-Naranjo (2019). While the current setup is in itself not intended to make predictions about the *future* impact of new technologies, it highlights that the impact of further automation of cognitive tasks may not only lead to a substantial reallocation of workers from the jobs we labelled non-routine cognitive (Frank et al., 2019), but also turn over the sorting pattern in space that have been dominant in the last decades. However, in order to make predictions in this direction one needs credible estimates of the substitution elasticities between labor and *current* technologies.

5 Conclusion

Inequality through polarization has an important urban component, and this urban dimension is key in the investment decision of firms to adopt new technologies (IT). In this paper, we have used a novel data set about IT expenditure at the firm level to establish two robust facts about urban polarization. We find, first, that IT investment is increasing in local housing costs, second, we find that there is a relatively larger decline in routine cognitive occupations in expensive cities. In addition we document the evolution of wage inequality by occupation across cities with different housing costs.

We then used these facts to build and estimate an equilibrium model that elucidates the underlying mechanism of urban polarization. Workers locate in cities where the bundle of wages, housing prices, and amenities gives them the highest utility. This continuous arbitrage pins down equilibrium wages and prices for given productivity differences across cities. At these equilibrium wage and price bundles, the incentives for firms to invest in IT vary substantially across high productivity cities with high wages and high housing prices and low productivity cities with low wages and low housing prices. We find that IT investment that substitutes routine tasks depends crucially on the properties of the production technology. The estimated model can explain more than half of the rising wage gap between routine and non-routine cognitive jobs, and 28 percent of the shift of employment away from routine towards non-routine cognitive jobs.

There are substantial differences across cities in the impact of IT investment both on wage inequality and on the reallocation of routine tasks. This confirms that job polarization is a predominantly urban phenomenon that determines both the employment distribution between and within cities and inequality.

Appendix

Table of Contents

A Descriptive Statistics: Additional Description and Tables **34**

A.1 Construction: Occupation Categories . . . 34

A.2 Adjusted IT Measures . . . 36

B Employment and Establishment Coverage: Comparison to CBP's Data **37**

B.1 CBP Data . . . 37

B.2 Comparison to Ci Aberdeen Data: IT Budget Sample . . . 37

C Employment, Establishment, and Sales Coverage: Comparison to NETS Data **42**

C.1 NETS Data . . . 42

C.2 Comparison to Ci Aberdeen Data: IT Budget Sample . . . 42

C.3 Comparison to Ci Aberdeen Data: ERP Sample . . . 45

D Empirical Evidence - Alternative technology measures **46**

D.1 Enterprise Resource Planning (ERP) software . . . 46

D.2 Descriptive Statistics . . . 46

D.3 Empirical Results . . . 46

E Measures of Skill Concentration **50**

E.1 Location Quotient . . . 50

E.2 Ellison-Glaeser (1997) Index . . . 51

E.3 Oyer-Schaefer (2016) Index . . . 53

F Additional Reduced-Form Empirical Results **55**

F.1 Changes in Occupation Shares . . . 55

F.2 Residual Wage Distributions . . . 56

G Theoretical Results **62**

G.1 Automation . . . 62

G.2 Skill-Biased Technological Change 73

H Estimation: Additional Derivations and Supporting Information 85

H.1 Additional Derivations 85

H.2 Standard Errors 85

H.3 Preparation of Data on IT Usage 86

H.4 The Elasticity of Substitution between IT and Labor 86

A Descriptive Statistics: Additional Description and Tables

A.1 Construction: Occupation Categories

Table A-1 shows the classification into groups by task components and the corresponding titles of occupation groups in the Census 2010 Occupation Classification system[21].

Table A-1: 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

Table A-2 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: occupation shares, employment levels, and our MSA rent index.

Table A-3 presents the summary statistics for IT budget per worker across MSAs. First, notice that there is a difference in the definition of the unit of count between the first row and rows 2-4 in Table A-3. In the first row, we calculate the MSA's IT budget per worker by dividing the sum of the total IT budget of all establishments in the MSA by the sum of these establishments' labor force. In this sense, we obtain an average IT budget per worker that puts more weight on larger establishments. Differently, for the summary statistics presented in rows 2-4, we first calculate the IT budget per worker for each establishment and then look at the average, median, and standard deviation of IT budget per worker across establishments within a given MSA. Consequently, rows 2-4 have an establishment as the unit of measure, reducing the weight of larger establishments in the overall count. In this sense, rows 2-4 allow us to evaluate within- and between-MSA IT budget per worker dispersion across establishments. While our analysis focuses on the definition of MSA's IT budget per worker presented in Table A-3's row 1, rows 2-4 show that there is significant

[21] 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 A-2: Descriptive statistics

	1980 mean (st. dev.)	2015 mean (st. dev.)
MSA's Occupation Shares		
Non-Routine Cognitive	34.6%	45.3%
	(3.95)	(5.46)
Non-Routine Manual	9.9%	14.8%
	(2.43)	(2.38)
Routine Cognitive	29.8%	22.9%
	(2.12)	(1.96)
Routine Manual	25.3%	16.7%
	(4.71)	(3.08)
MSA's Rent and Size		
log rent index	0.01	0.01
	(0.13)	(0.23)
Employment in 000s	861.61	1535.77
	(1049.25)	(1678.15)
No. of MSAs	261	261

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

within-MSA variation of IT budget per worker across establishments. Moreover, our empirical results are robust to the different ways to calculate the IT budget per worker presented in Table A-3. As we can see in row 1 of Table A-3, there is significant variation in IT budget per worker across MSAs.

Table A-3: Descriptive statistics of technology adoption across MSAs – 2015

	Mean	**Median**	**S.D.**	**Min**	**Max**	**N**
IT Budget						
MSA's IT Budget/Emp.	4,919	4,381	2,436	2,710	33,905	272
Avg. IT Budget/Emp. by site	4,238	4,159	515	3,293	5,817	272
Median IT Budget/Emp. by site	2,888	2,860	342	2,062	3,750	272
St. Dev. IT Budget/Emp. by site	8,865	4,917	11,453	3,123	97,557	272

A.2 Adjusted IT Measures

Let $\gamma_{i,c,t}$ be the technology for establishment i in city c and time t. We estimate, using OLS, the following model:

$$\gamma_{i,c,t} = \sum_t \left[\beta_{I,t} Ind_{i,t} \times Size_{i,t} + \beta_{C,t} City_{i,t} + \beta_{Y,t} Year_{i,t}\right] + \varepsilon_{i,t} \tag{A.1}$$

where *Ind*, *Size*, and *City* are vectors of dummy variables of industry (3-digit SIC) of the establishment, size of the establishment (8 employment size classes, following CBP[22]). In this case, $\beta_{C,t}$ is the key measure, capturing the differences in technology use across cities, after controlling for over 950 industry/size interactions.

[22]Doms and Lewis (2006) are not clear about which categories they are. However, since they weight their regression based on the CBP and limit their sample to establishments with 5 employees or more, the class sizes are likely: 5 to 9 employees,10 to 19 employees, 20 to 49 employees, 50 to 99 employees, 100 to 249 employees, 250 to 499 employees, 500 to 999 employees, and more than 1000 employees.

B Employment and Establishment Coverage: Comparison to CBP's Data

B.1 CBP Data

The County Business Patterns is an annual series that provides sub national economic data by industry. This series includes the number of establishments, employment during the week of March 12, first quarter payroll, and annual payroll. The CBP excludes from its data self-employed individuals (a lot of them are in the 1-3 establishment size categories), as well as contract or temporary employees, counting only "hired employees." While we don't have a detailed description of Aberdeen's employment data, we believe that it follows a similar pattern of the National Establishment Time Series (NETS) which uses Dun & Bradstreet data (D&B). As we see below, our comparison of the Aberdeen data with NETS corroborates this result. In the NETS, data contract and temporary workers are included, as well as self-employed workers.

Moreover, establishments in the following NAICS industries are not included in the CBP: Crop and Animal Production (NAICS 111, 112), Rail Transportation (NAICS 482), Postal Service (NAICS 491), Pension, Health, Wealfare, and Vacation Funds (NAICS 52110, 525120, 525190), Trust, Estates, and Agency Accounts (NAICS 525920), Private Households (NAICS 814), and Public Administration (NAICS 92).

Furthermore, the CBP defines establishments as *"a single physical location at which business is conducted or services or industrial operations are performed (...) with paid employees."*. Differently, in the NETS, an establishment is defined as a "unique line of business (SIC8) at a unique location." So, it is possible to have more than one establishment at a location. As Aberdeen uses DUNS numbers to identify establishments, it is likely that it follows NETS' definition.

Finally, in order to properly compare the industry×MSA composition of the employment and establishment data between CBP and Ci Aberdeen, we use the imputed CBP files for 2015 by Eckert et al. (2021). However, results are qualitatively the same if we use CBP's raw files.

B.2 Comparison to Ci Aberdeen Data: IT Budget Sample

In order to properly compare the two samples, we restrict the IT budget sample to private establishments in industries covered by the CBP. Furthermore, we aggregate the establishment and employment counts at the MSA level. Notice that while our sample covers on average only 34 percent of the MSA's establishments (table A-4), table A-5 shows that this is mostly due to a low coverage of establishments with 1 to 4 employees. In fact, the coverage is on average above 60 percent for establishments with 10 employees or more. In contrast, our sample seems to have too many large establishments (500+). This is likely due to contract and temporary workers, which are not counted by the CBP. These patterns are in line with what Barnatchez et al. (2017) find when comparing NETS to the CBP, corroborating the idea that the Ci Aberdeen data have features similar to the NETS.

Table A-4: Coverage Ci Aberdeen relative to CBP

	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
Fraction Emp. in Ci	73%	16%	62%	66%	73%	79%	85%	277
Fraction Est. in Ci	34%	5%	29%	32%	35%	37%	39%	277
ERP Sample								
Fraction Emp. in Ci	25%	9%	16%	20%	25%	29%	33%	277
Fraction Est. in Ci	2%	1%	2%	2%	2%	3%	3%	277

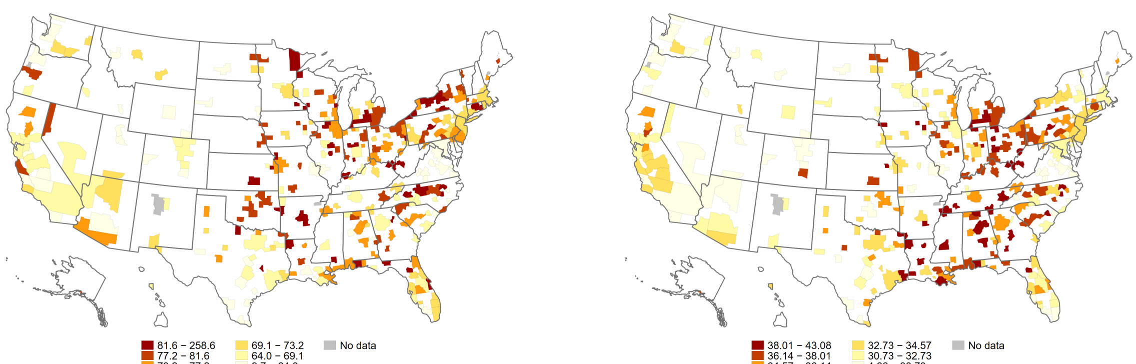
Table A-5: Coverage Ci Aberdeen relative to CBP by establishment size

	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
1 to 4 Employees	11%	2%	9%	10%	11%	13%	14%	277
5 to 9 Employees	45%	7%	38%	41%	45%	48%	53%	277
10 to 19 Employees	68%	9%	59%	63%	68%	73%	78%	277
20 to 49 Employees	63%	9%	55%	59%	64%	69%	73%	277
50 to 99 Employees	64%	11%	52%	59%	64%	71%	77%	277
100 to 249 Employees	74%	16%	58%	66%	73%	82%	93%	277
250 to 499 Employees	89%	29%	62%	71%	86%	100%	119%	277
500 to 999 Employees	115%	69%	60%	75%	100%	133%	200%	273
1,000 or more Employees	142%	94%	71%	100%	119%	167%	200	274
ERP Sample								
1 to 4 Employees	0%	0%	0%	0%	0%	0%	0%	277
5 to 9 Employees	0%	0%	0%	0%	0%	1%	1%	277
10 to 19 Employees	2%	1%	1%	1%	2%	2%	3%	277
20 to 49 Employees	6%	2%	4%	5%	6%	8%	9%	277
50 to 99 Employees	15%	5%	10%	12%	15%	18%	22%	277
100 to 249 Employees	32%	10%	22%	26%	31%	38%	45%	277
250 to 499 Employees	43%	19%	25%	32%	41%	50%	66%	277
500 to 999 Employees	66%	49%	25%	39%	54%	75%	120%	273
1,000 or more Employees	88%	66%	40%	53%	73%	100%	150%	274

In terms of industry coverage, we see that our sample has low coverage in leisure and hospitality, trade, transportation, and utility, as well as other services in both establishment and employment coverage (see Tables A-6 and A-7). On the other hand, our sample seems to overstate the employment in several sectors, in particular mining, manufacturing, and information. The excessive coverage in manufacturing and mining has also been documented in NETS by Barnatchez et al. (2017). Apart from the already mentioned differences in the types of employment covered by NETS (and probably Ci Aberdeen) and the CBP, another issue highlighted by Barnatchez et al. (2017) is the difficulty in industry assignment. Consequently, a significant share of the differences may be due not to measurement error, but to differences in industry assignment methods.

Table A-6: Ci Coverage relative to CBP: Establishments by industry

	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
Manufacturing	96%	18%	76%	85%	96%	107%	116%	277
Construction	26%	7%	17%	21%	25%	30%	35%	277
Information	84%	23%	63%	70%	80%	94%	110%	277
Finance	53%	11%	41%	45%	53%	60%	67%	277
Professional & Bus Services	34%	6%	27%	31%	35%	38%	42%	277
Education and Health	60%	11%	48%	54%	60%	66%	73%	277
Leisure and Hospitality	12%	3%	9%	10%	12%	14%	16%	277
Trade, Transp., and Util.	19%	3%	15%	18%	19%	21%	23%	277
Mining	66%	47%	21%	35%	56%	88%	117%	271
Other Services	18%	4%	14%	15%	18%	20%	22%	277
ERP Sample								
Manufacturing	12%	5%	6%	9%	13%	15%	18%	277
Construction	1%	1%	1%	1%	1%	2%	2%	277
Information	9%	4%	4%	7%	8%	11%	14%	277
Finance	1%	1%	1%	1%	1%	2%	2%	277
Professional & Bus Services	2%	1%	1%	2%	2%	3%	4%	277
Education and Health	4%	1%	2%	3%	4%	5%	5%	277
Leisure and Hospitality	1%	1%	1%	1%	1%	2%	2%	277
Trade, Transp., and Util.	1%	1%	1%	1%	1%	2%	2%	277
Mining	6%	10%	0%	0%	2%	8%	14%	271
Other Services	1%	1%	1%	1%	1%	2%	2%	277



(a) Employment Coverage (b) Establishment Coverage

Figure A-1: Geographical distribution of Ci Coverage relative to CBP: IT Budget Sample

Finally, in terms of geographic coverage, the IT budget sample shows a higher coverage in the Midwest and East Coast regions, while coverage rates are somewhat lower in the West Coast and Western regions. Patterns are quite similar for both employment (figure A-1a) and establishments

Table A-7: Ci Coverage relative to CBP: Employment by industry

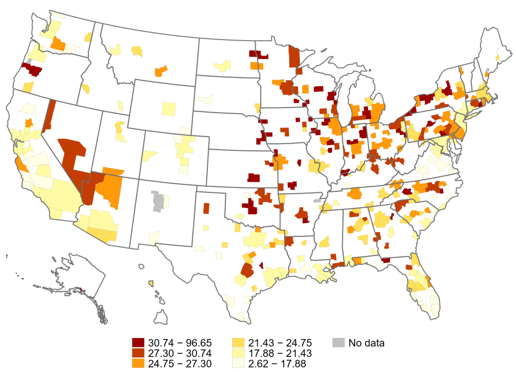
	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
Manufacturing	129%	48%	86%	102%	123%	145%	168%	277
Construction	87%	28%	61%	70%	84%	98%	115%	277
Information	126%	69%	68%	90%	115%	141%	179%	277
Finance	98%	27%	70%	83%	97%	111%	129%	277
Professional & Bus Services	69%	27%	47%	55%	63%	77%	96%	277
Education and Health	104%	62%	71%	85%	97%	114%	133%	277
Leisure and Hospitality	25%	12%	14%	17%	22%	28%	36%	277
Trade, Transp., and Util.	45%	10%	34%	39%	45%	50%	55%	277
Mining	178%	392%	14%	55%	104%	192%	326%	277
Other Services	64%	37%	41%	47%	57%	68%	90%	277
ERP Sample								
Manufacturing	56%	24%	29%	42%	54%	68%	80%	277
Construction	12%	11%	3%	7%	10%	15%	22%	277
Information	43%	35%	16%	25%	36%	53%	69%	277
Finance	18%	13%	4%	8%	16%	25%	37%	277
Professional & Bus Services	15%	13%	5%	9%	13%	19%	26%	277
Education and Health	45%	31%	24%	32%	43%	51%	63%	277
Leisure and Hospitality	8%	8%	2%	3%	6%	10%	15%	277
Trade, Transp., and Util.	10%	6%	4%	6%	9%	13%	17%	277
Mining	34%	120%	0%	0%	1%	31%	72%	277
Other Services	16%	30%	3%	7%	11%	16%	26%	277

(figure A-1b). That said, coverage rates even in areas with low coverage are still meaningful (above 64 percent for employment and above 30 percent for establishments).

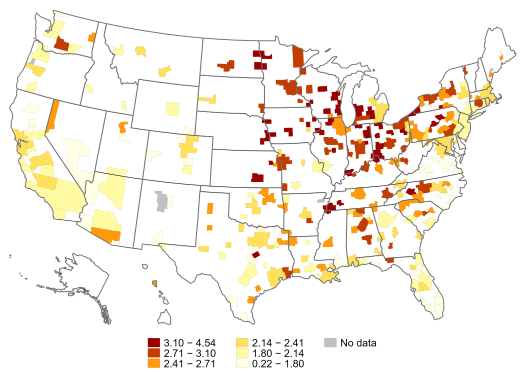
B.2.1 Comparison to Ci Aberdeen Data: ERP Sample

As discussed in Section 2, our ERP sample is limited. Our information on ERP adoption covers on average only 25 percent of workers and 2 percent of establishments in the MSA, compared to the CBP (see Table A-4). Moreover, as presented in Table A-5, even after controlling for establishment size, MSA average coverage is above 30 percent only for establishments that have 100 employees or more. Finally, Table A-6 shows that the ERP sample covers less than 15 percent of establishments in all industry sectors. However, since the coverage is tilted towards larger establishments, employment coverage varies from 8 (Leisure and Hospitality) to 56 percent (Manufacturing).

Finally, in terms of geographic coverage, the ERP sample shows a higher coverage in the Midwest and East Coast regions, while coverage rates are somewhat lower in the West Coast and Western regions. Patterns are quite similar for both employment (Figure A-2a) and establishments (Figure A-2b). That said, coverage rates even in areas with low coverage are still meaningful (above 15 percent for employment and above 1.8 percent for establishments).



(a) Employment Coverage



(b) Establishment Coverage

Figure A-2: Geographical distribution of Ci coverage relative to CBP: ERP sample

C Employment, Establishment, and Sales Coverage: Comparison to NETS Data

C.1 NETS Data

The National Establishment Time Series (NETS) is an annual series consisting of establishment-level longitudinal microdata covering, in principle, the universe of US Business. The starting point for the NETS database was 25 annual snapshots (taken every January) of the full Duns Marketing Information (DMI) file that followed over 58.8 million establishments between January 1990 and January 2013. These snapshots actually used the DMI file to determine which establishments were active in January of each year in question. The database includes information on: business name, address and contact information, headquarters ID, number of establishments per firm, industry classification, type of proprietorship, employment by location and estimated annual establishment sales. Finally, NETS includes unique firm and establishment identifiers through D&B hqduns and duns numbers.

As highlighted in Section B.1, there are some key distinctions between NETS and the databases provided by official sources such as the CBP. First, NETS information is not collected at a particular time of year, but throughout the year. Second, in NETS an establishment is defined as a "unique line of business (SIC8) at a unique location." So, it is possible to have more than one establishment at a location. Third, NETS data include not only firm owners among establishment employees, but also self-employed, contract, and temporary workers. Finally, there are some drawbacks to the data, highlighted by Barnatchez et al. (2017) and Crane and Decker (2020), in particular due to data staleness as well as issues with data imputation. While the data are reported to be regularly collected, some employment level information seems to be updated less frequently than official sources counterparts. Similarly, imputed data points differ quite significantly from their administrative data counterparts.

C.2 Comparison to Ci Aberdeen Data: IT Budget Sample

Ci Aberdeen data have lots of similarities to the NETS data. First, both are indexed by duns numbers and have imputed values for establishment sales. Second, establishment and employment data tend to follow similar definitions in both samples, including non-employment establishments. However, as we compare the two samples, we do observe some key distinctions. First, headquarter IDs are quite distinct between the two databases. Second, employment levels are quite distinct among large establishments. We present more details below.

Similar to the comparison to the CBP presented in Section B, while our IT budget sample covers more than 50 percent of employment in NETS, it only covers about 11% of establishment (see Table A-8. However, the low establishment coverage is due to low coverage of small establishments. In fact, our sample covers above 50 percent of NETS establishments for establishments with 10

Table A-8: Coverage Ci Aberdeen relative to NETS

	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
Fraction Emp. in Ci	53%	12%	43%	48%	54%	59%	62%	279
Fraction Est. in Ci	11%	2%	8%	10%	11%	13%	13%	277
Fraction Sales in Ci	53%	10%	44%	50%	54%	58%	62%	279
ERP Sample								
Fraction Emp. in Ci	20%	7%	13%	16%	20%	23%	27%	279
Fraction Est. in Ci	1%	0%	1%	1%	1%	1%	1%	277
Fraction Sales in Ci	17%	6%	9%	13%	17%	20%	24%	279

employees or more (see Table A-9).

In terms of industry coverage, we see that our sample has a low coverage in leisure and hospitality, trade, transportation, and utility, as well as other services in both establishment and employment coverage (see Tables A-10 and A-11).

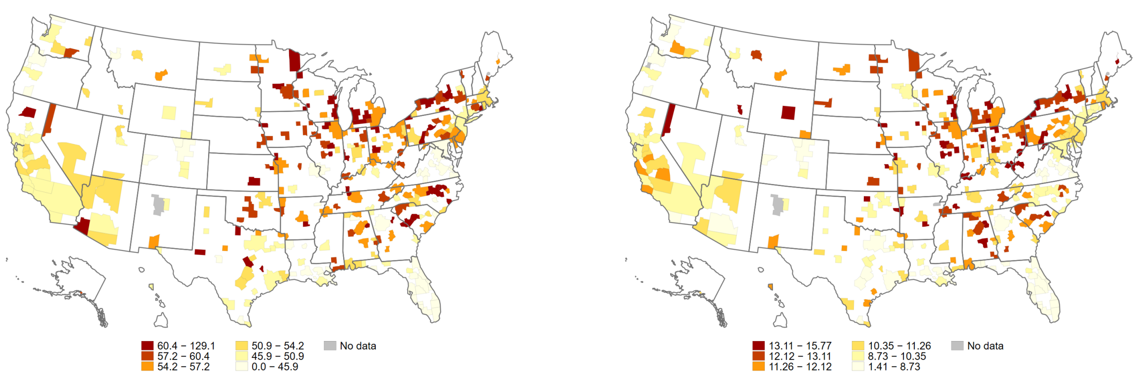
Table A-9: Coverage Ci Aberdeen relative to NETS by establishment size

	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
1 to 4 Employees	2%	1%	2%	2%	2%	3%	3%	279
5 to 9 Employees	24%	4	21%	22%	24%	26%	28%	279
10 to 19 Employees	53%	8%	48%	52%	54%	57%	59%	279
20 to 49 Employees	54%	8%	49%	52%	56%	58%	61%	279
50 to 99 Employees	56%	9%	50%	54%	57%	60%	64%	279
100 to 249 Employees	60%	12%	49%	55%	61%	67%	73%	279
250 to 499 Employees	73%	21%	50%	63%	72%	84%	95%	279
500 to 999 Employees	92%	39%	56%	72%	85%	108%	133%	279
1,000 or more Employees	111%	50%	67%	85%	100%	125%	160%	277
ERP Sample								
1 to 4 Employees	0%	0%	0%	0%	0%	0%	0%	279
5 to 9 Employees	0%	0%	0%	0%	0%	0%	0%	279
10 to 19 Employees	2%	1%	1%	1%	2%	2%	2%	279
20 to 49 Employees	6%	1%	4%	5%	6%	7%	7%	279
50 to 99 Employees	13%	4%	9%	11%	13%	15%	18%	279
100 to 249 Employees	28%	7%	21%	24%	28%	32%	38%	279
250 to 499 Employees	38%	16%	21%	29%	36%	45%	55%	279
500 to 999 Employees	56%	30%	29%	38%	50%	67%	89%	279
1,000 or more Employees	75%	47%	36%	50%	64%	89%	117%	277

Finally, in terms of geographic coverage, the IT budget sample shows a higher coverage in the Midwest and East Coast regions, while coverage rates are somewhat lower in the West Coast and Western regions. Patterns are quite similar for both employment (figure A-3a) and establishments (figure A-3b). That said, coverage rates even in areas with low coverage are still meaningful (above

Table A-10: Ci coverage relative to NETS: Employment by Industry

	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
Manufacturing	76%	22%	52%	65%	76%	85%	101%	279
Construction	48%	10%	38%	42%	47%	55%	60%	279
Information	64%	21%	44%	53%	63%	74%	83%	279
Finance	55%	19%	38%	45%	54%	63%	74%	279
Professional & Bus Services	39%	16%	23%	30%	38%	46%	56%	279
Education and Health	79%	18%	62%	72%	79%	87%	95%	279
Leisure and Hospitality	22%	11%	13%	16%	21%	27%	33%	279
Public Adm	73%	57%	47%	57%	66%	78%	92%	279
Trade, Transp., and Util.	37%	12%	25%	31%	36%	41%	47%	279
Mining	60%	41%	12%	37%	59%	76%	97%	279
Other Services	36%	19%	23%	28%	33%	40%	48%	279
ERP Sample								
Manufacturing	36%	18%	14%	24%	34%	44%	56%	279
Construction	7%	6%	2%	4%	6%	10%	14%	279
Information	25%	17%	8%	15%	22%	31%	44%	279
Finance	14%	11%	3%	6%	12%	19%	28%	279
Professional & Bus Services	11%	12%	3%	5%	9%	14%	19%	279
Education and Health	35%	13%	22%	27%	35%	41%	50%	279
Leisure and Hospitality	8%	9%	2%	3%	6%	9%	14%	279
Public Adm	30%	41%	10%	17%	24%	32%	45%	279
Trade, Transp., and Util.	11%	9%	3%	6%	9%	13%	17%	279
Mining	12%	23%	0%	0%	0%	14%	36%	279
Other Services	10%	15%	2%	4%	7%	12%	18%	279



(a) Employment Coverage

(b) Establishment Coverage

Figure A-3: Geographical distribution of Ci coverage relative to NETS: IT budget sample

45 percent for employment and above 8 percent for establishments).

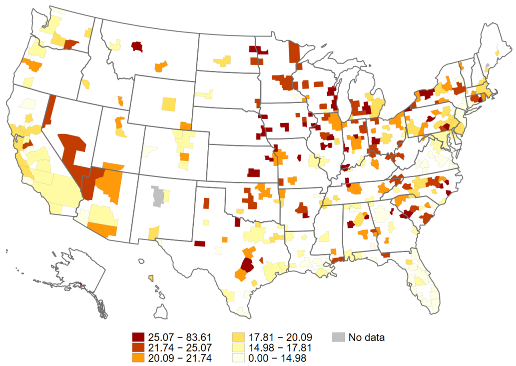
Table A-11: Ci coverage relative to NETS: Establishments by industry

	Mean	S.D.	p10	p25	p50	p75	p90	N
IT Budget Sample								
Manufacturing	36%	18%	14%	24%	34%	44%	56%	279
Construction	7%	6%	2%	4%	6%	10%	14%	279
Information	25%	17%	8%	15%	22%	31%	44%	279
Finance	14%	11%	3%	6%	12%	19%	28%	279
Professional & Bus Services	11%	12%	3%	5%	9%	14%	19%	279
Education and Health	35%	13%	22%	27%	35%	41%	50%	279
Leisure and Hospitality	8%	9%	2%	3%	6%	9%	14%	279
Public Adm	30%	41%	10%	17%	24%	32%	45%	279
Trade, Transp., and Util.	11%	9%	3%	6%	9%	13%	17%	279
Mining	12%	23%	0%	0%	0%	14%	36%	279
Other Services	10%	15%	2%	4%	7%	12%	18%	279
ERP Sample								
Manufacturing	31%	8%	22%	26%	31%	36%	41%	279
Construction	8%	2%	5%	7%	8%	9%	11%	279
Information	23%	7%	14%	19%	23%	28%	32%	279
Finance	18%	5%	13%	16%	18%	21%	24%	279
Professional & Bus Services	5%	1%	3%	4%	5%	6%	6%	279
Education and Health	29%	6%	23%	26%	29%	32%	35%	279
Leisure and Hospitality	7%	2%	5%	6%	7%	8%	10%	279
Public Adm	62%	9%	54%	59%	63%	66%	70%	279
Trade, Transp., and Util.	8%	2%	6%	7%	8%	9%	10%	279
Mining	24%	13%	9%	16%	24%	33%	40%	279
Other Services	5%	1%	4%	4%	5%	6%	7%	279

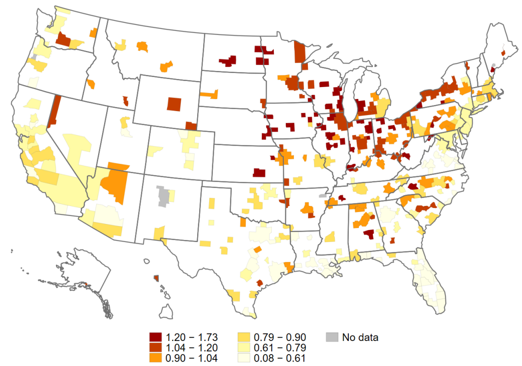
C.3 Comparison to Ci Aberdeen Data: ERP Sample

As discussed in Section 2, our ERP sample is limited. Our information on ERP adoption covers on average only 20 percent of workers and 1 percent of establishments in the MSA, compared to NETS (see table A-8). Moreover, as presented in Table A-5, even after controlling for establishment size, MSA average coverage is above 28 percent only for establishments that have 100 employees or more. Finally, Table A-11 shows that the ERP sample covers less than 35 percent of establishments in all industry sectors but public administration. However, since the coverage is tilted toward larger establishments, employment coverage varies from 10 (Leisure and Hospitality) to 36 percent (Manufacturing) of the NETS industry employment (Table A-10).

Finally, in terms of geographic coverage, the ERP sample shows a higher coverage in the Midwest and East Coast regions, while coverage rates are somewhat lower in the West Coast and Western regions. Patterns are quite similar for both employment (Figure A-4a) and establishments (Figure A-4b). That said, coverage rates even in areas with low coverage are still meaningful (above 15 percent for employment and above 0.6 percent for establishments).



(a) Employment Coverage



(b) Establishment Coverage

Figure A-4: Geographical distribution of Ci coverage relative to NETS: ERP sample

D Empirical Evidence - Alternative technology measures

D.1 Enterprise Resource Planning (ERP) software

In this section, we discuss the empirical evidence on the relationship between ERP adoption and local rental price index as well as 1980's share of routine cognitive jobs in the local labor force.

D.2 Descriptive Statistics

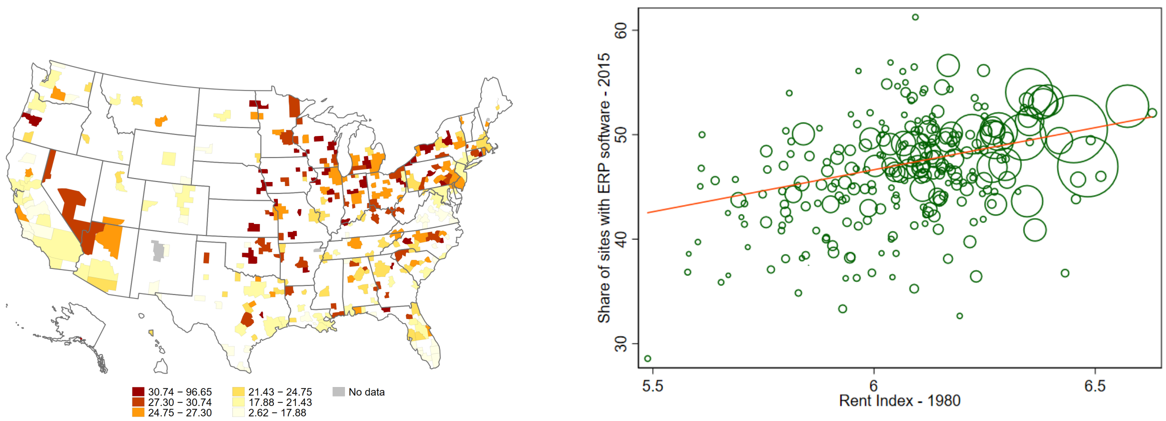
Table A-12 shows that there is a lot of dispersion in the ERP shares across MSAs even in 2015, when we should expect a more widespread use of technology. As we can see, we have at least some information on 277 MSAs across the country. Moreover, we can see that, while on average about 46 percent of the establishments have at least some form of ERP, there is substantial variation across the country. Some MSAs have a fraction as low as 29 percent, while others have more than 61 percent of establishments with some form of ERP. Even more, as we show in Figure A-5b, the degree of adoption seems closely tied to the size as well as cost of living in the MSA, proxied by the rental index. Finally, figure A-5a shows the geographical dispersion of ERP concentration across the country in 2015. First, geographical coverage is quite good, with only very few MSAs completely missing. In fact, the missing MSAs are due to the procedure for matching the census PUMA to the 2000 census metropolitan area definitions as described by Baum-Snow and Pavan (2013).

D.3 Empirical Results

Table A-13 presents the same specifications as presented in Table 1, replacing IT budget per worker with the fraction of establishments in the MSA with at least one ERP software. As we can observe, results for local price indexes are similar to the ones observed in Table 1, i.e., establishments in more expensive areas are more likely to have at least one installed ERP software. In particular, in specification (5), a one standard deviation increase in the local price index (an increase of 21.4

Table A-12: Descriptive statistics of technology adoption across MSAs – 2015

	Mean	Median	S.D.	Min	Max	N
ERP Share						
Share of Workers in Est. w/ ERP	51.97%	52.74%	12.87%	13.87%	86.64%	277
Share of Establishments w/ ERP	46.39%	46.67%	5.08%	28.57%	61.25%	277
No. of ERPs						
Avg. No. of ERPs per Est.	0.77	0.78	0.11	0.41	1.17	277
Median No. of ERPs per Est.	0.24	0	0.42	0	1	277
St. Dev. of No. ERP per Est.	1.05	1.06	0.11	0.73	1.36	277



(a) Share of workers exposed to ERP

(b) ERP share vs. local price level

Figure A-5: Geographical distribution of ERP across MSAs – 2015

percent in the 1980 local price index) is associated with an increase of about 1 percent in the share of establishments with ERP. In fact, moving from the cheapest to the most expensive MSA is associated with a 5 percent increase in the share of establishments in the MSA with at least one ERP software installed.

Table A-14 presents the results for a logit model on the presence of an installed ERP software in the establishment, after controlling for firm and industry fixed effects. Controls are the same as presented in Table 2. As expected, due to a significant decrease in sample size, results are weaker and lose statistical significance in some cases. However, the overall pattern is still the same as the one presented in Table 2, i.e., establishments in more expensive MSAs are more likely to adopt ERP software.

Table A-13: Share of Establishments with ERP – 2015

	(1)	(2)	(3)	(4)	(5)
	ERP Share	ERP Share	ERP Share	ERP Share	ERP Share
MSA log rent index 1980	0.074***			0.052**	0.047*
	(0.020)			(0.024)	(0.024)
MSA routine cognitive share 1980		-0.040		0.257	0.166
		(0.168)		(0.169)	(0.175)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			0.046***	0.030	0.027
			(0.015)	(0.018)	(0.017)
MSA Offshorability 1980					0.072
					(0.059)
Housing supply elasticity	0.006***	0.005**	0.004*	0.005**	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Amenities	Yes	Yes	Yes	Yes	Yes
MSA's Industry Mix Controls	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	21.99	14.83	23.31	27.53	27.69
Adj. R^2	0.599	0.503	0.599	0.610	0.613
MSA	217	217	217	217	217

Standard errors in parentheses. The dependent variable in all columns is the share of establishments with at least one ERP software in the metro area. Each observation (a MSA) is weighted by its employment in 2015. MSA controls include its unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work states. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A-14: ERP presence by establishment - Firm and industry FE

	ERP Dummy				
	(1)	(2)	(3)	(4)	(5)
lrentindexaltmsa	0.318			0.408*	0.439*
	(0.206)			(0.227)	(0.229)
MSA routine cognitive share 1980		0.030**		0.032**	0.039**
		(0.015)		(0.015)	(0.016)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			0.003	-0.136	-0.124
			(0.129)	(0.143)	(0.144)
MSA Offshorability 1980					-0.565
					(0.448)
log(Site's Size)	0.249***	0.249***	0.249***	0.249***	0.249***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
log(Site's Revenue)	0.327***	0.328***	0.327***	0.328***	0.328***
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
Headquarters dummy	1.156***	1.155***	1.156***	1.155***	1.154***
	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)
housing_elasticity	0.035	0.012	0.022	0.024	0.028
	(0.022)	(0.021)	(0.021)	(0.023)	(0.023)
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
No. of Sites	31,427	31,427	31,427	31,427	31,427
No. of Firms	4,318	4,318	4,318	4,318	4,318
No. of MSAs	218	218	218	218	218

Standard errors in parentheses. The dependent variable in all columns is the a dummy variable that indicates the presence of at least one ERP software in the establishment. MSA controls include its unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work States. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Industry dummies are 2-digit SIC dummies. Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

E Measures of Skill Concentration

We now calculate measures of the 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 25 years. Moreover, these measures abstract from issues of long-run trends in the composition of the labor force. Consequently, we are able to focus on the correlation between the spatial dispersion of skills and an MSA's 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").

E.1 Location Quotient

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

$$\lambda_{ij} = \frac{\frac{m_{ij}}{S_j}}{\frac{M_i}{\sum_{l=1}^{N} M_l}} \tag{A.2}$$

Intuitively, if am 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 or decrease in concentration across regions as well as how it correlates to these regions' characteristics.

Following what has been show in other sections, we consider two time periods: 1990 and 2015. Moreover, following Cortes et al. (2017), we divide the occupations in four groups: non-routine manual, routine manual, routine cognitive, and non-routine cognitive. We divide the regions into two groups around the median. We use the log rent index in 1980, i.e. cheap vs. expensive, as the measure to separate the MSAs. Results are presented in Table A-15.

As we can see from Table A-15, in 1990, cheaper cities had on average a higher concentration in routine manual jobs, a lower concentration in cognitive jobs (both routine and non-routine), and close to at par in non-routine manual jobs when compared to expensive cities. Differently, in 2015 we see cheap cities being on average more concentrated in routine cognitive jobs, while we see minor changes in the other occupation categories. These results are in line with what our theoretical results would predict.

Finally, Figures A-6 and A-7 present the density distributions of the location quotients for

Table A-15: Simple measure of concentration across skill and city size groups

Panel A: 1990								
	Non-Routine Manual		*Routine Manual*		*Routine Cognitive*		*Non-Routine Cognitive*	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Expensive City	1.05	1.00	1.04	1.01	0.99	1.00	0.97	0.96
Cheap City	1.06	1.06	1.22***	1.22†††	0.96***	0.95†††	0.90***	0.89†††
Panel B: 2015								
	Non-Routine Manual		*Routine Manual*		*Routine Cognitive*		*Non-Routine Cognitive*	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Expensive City	1.02	0.986	1.08	1.01	1.01	1.01	0.96	0.97
Cheap City	1.02	1.00	1.17***	1.16†††	1.04**	1.04†††	0.91***	0.90†††

***,**,* represent significance at 1, 5, and 10% respectively in a t-test of means with unequal variances.
†††,††, † represent significance at 1, 5, and 10% respectively in a Wilcoxon rank-sum test of medians.

small and large cities across occupation groups and time. While we observe that there is significant variance in this index across MSAs, the overall message is the same as the one presented in Table A-15.

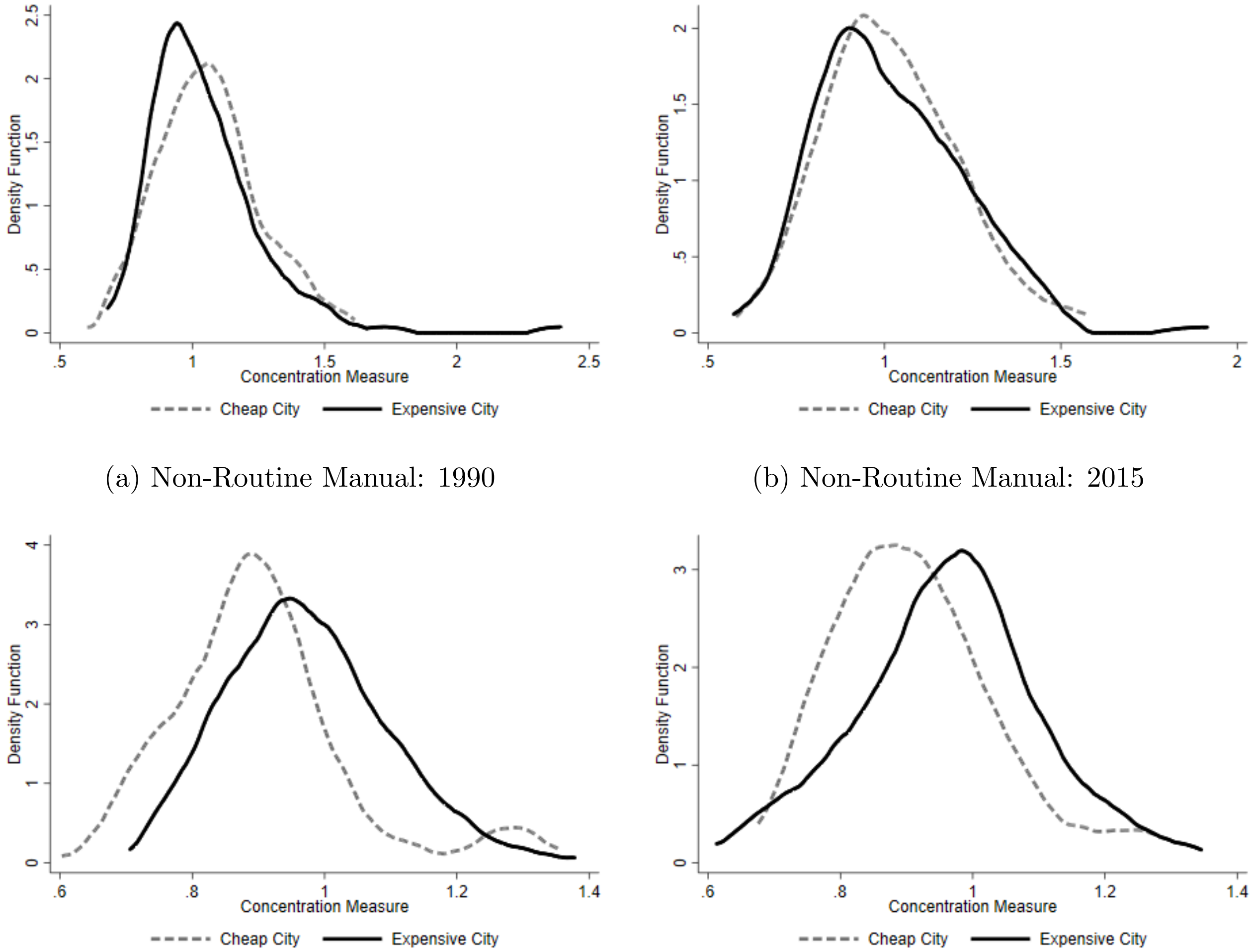
E.2 Ellison-Glaeser (1997) Index

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:

$$\gamma_i = \frac{\sum_j (s_{ij} - x_j)^2}{1 - \sum_j x_j^2} \tag{A.3}$$

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 are 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 to compare expensive and cheap cities. However, we are able to see if skill groups became more or less concentrated across cities over time.

Results are presented in Table A-16. As we can see, routine manual occupations have seen a minor decline in concentration, and all other occupational groups have seen an increase in concentration. These results complement the findings regarding the location quotient, by indicating how the concentration of each occupation group has changed across cities. While these results



(a) Non-Routine Manual: 1990

(b) Non-Routine Manual: 2015

(c) Non-Routine Cognitive: 1990

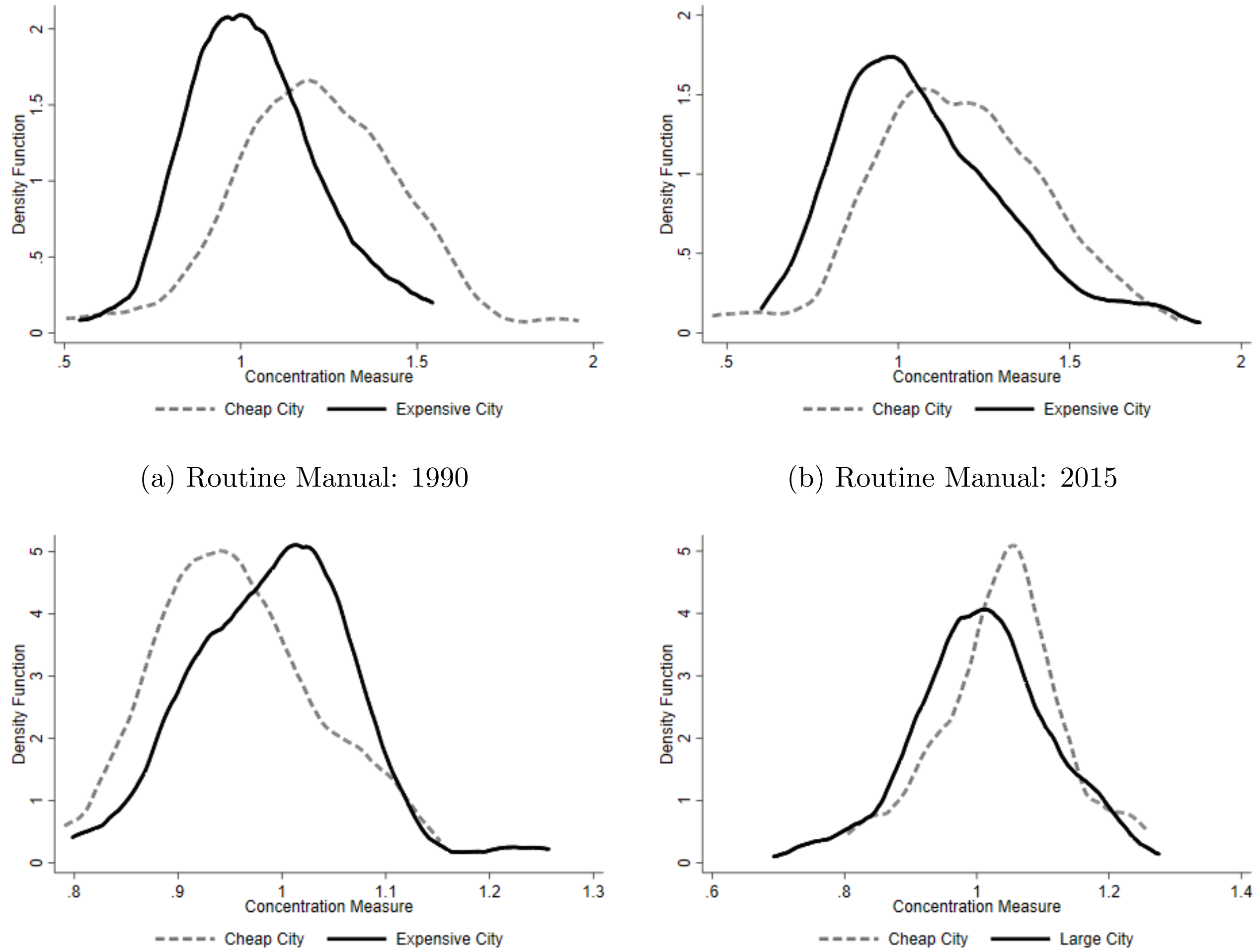
(d) Non-Routine Cognitive: 2015

Figure A-6: Non-routine occupations LQ distributions

Table A-16: Ellison-Glaeser Index

	1990	**2015**	**% Change**
Non-Routine Manual	0.00032	0.00045	42.00
Routine Manual	0.00055	0.00055	-1.26
Routine Cognitive	0.00005	0.00017	219.61
Non-Routine Cognitive	0.00018	0.00024	32.56

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.



(a) Routine Manual: 1990

(b) Routine Manual: 2015

(c) Routine Cognitive: 1990

(d) Routine Cognitive: 2015

Figure A-7: Routine occupations LQ distributions

E.3 Oyer-Schaefer (2016) Index

We now consider an adapted version of the EG concentration index based on Oyer and Schaefer (2016), which 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 $\tilde{x}_i$ as 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 defined as:

$$\zeta_j = \frac{S_j}{S_j - 1} \frac{\sum_i \left(\tilde{s}_{ij} - \tilde{x}_i\right)^2}{1 - \sum_i \tilde{x}_i^2} - \frac{1}{S_j - 1} \tag{A.4}$$

Differently from the EG index, in the OS index we are able to compare the degree of concentration 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, although the EG and OS indexes complement each other, both have its weaknesses and do not give a complete picture of the changes in concentration.

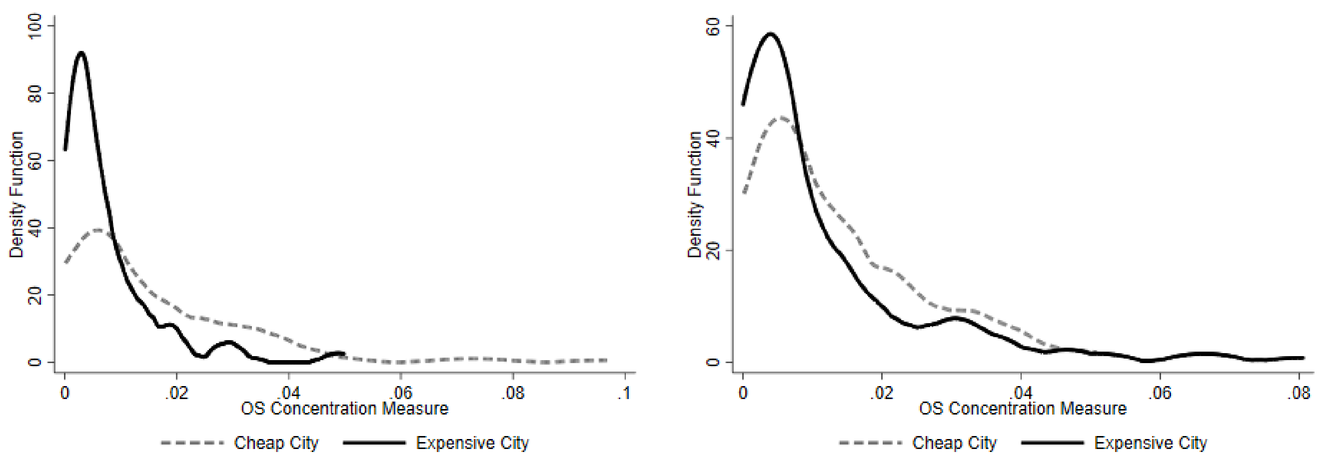
Table A-17 presents the results for 1990 and 2015. As we can see, in both periods, cheap cities are consistently more concentrated than expensive cities, although the statistical significance of the difference has decreased over time. Furthermore, while cheap cities have seen a reduction in concentration, expensive cities have become more concentrated over time.

Table A-17: OS index across city cost and time

Panel A: 1990					
	Mean	Median	St. Dev.	Min	Max
Expensive City	0.00737	0.00467	0.00863	0.00008	0.04983
Cheap City	0.01494***	0.01031†††	0.01597	0.00009	0.09682
Panel B: 2015					
	Mean	Median	St. Dev.	Min	Max
Expensive City	0.01138	0.00539	0.01465	0.00000	0.08061
Cheap City	0.01297	0.00950††	0.01204	0.00011	0.05087

***,**,* represent significance at 1, 5, and 10% respectively in a t-test of means. †††,††, † represent significant at 1, 5, and 10% respectively in a Wilcoxon rank-sum test of medians.

Finally, we present the changes in the density distribution of the OS index in Figure A-8. Notice that Figures A-8(a) and A-8(b) corroborate the results from Table A-17, showing an increase in concentration among expensive cities and a decrease in concentration among cheap cities.



(a) 1990

(b) 2015

Figure A-8: Distribution of OS index across city sizes and time

F Additional Reduced-Form Empirical Results

F.1 Changes in Occupation Shares

In this section, we present the changes over time of different occupation classes, based on the classification presented by Cortes et al. (2014). Our specification includes the same controls as the ones presented in Table 3.

Based on the results for Tables A-18, A-19, and A-20, we see that cost of living is not correlated to changes in these occupation categories in the period 1990-2015. Results from these tables corroborate our findings in Appendix Section E.1, in which location quotients do not show a significant change in the concentration across cities of different costs for all but routine cognitive occupations.

Table A-18: Change in non-routine cognitive share, 1990-2015

	Δnonrout-cog				
	(1)	(2)	(3)	(4)	(5)
MSA log rent index 1980	-0.0103			-0.0187	-0.0222
	(0.0181)			(0.0184)	(0.0181)
MSA routine cognitive share 1980		0.3394***		0.3336***	0.2623**
		(0.1069)		(0.1042)	(0.1182)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			0.0070	0.0096	0.0073
			(0.0110)	(0.0109)	(0.0106)
MSA Offshorability 1980					0.0554*
					(0.0323)
Housing supply elasticity	-0.0009	-0.0016	-0.0003	-0.0023	-0.0027*
	(0.0017)	(0.0014)	(0.0015)	(0.0016)	(0.0017)
Amenities	Yes	Yes	Yes	Yes	Yes
MSA's Industry Mix Controls	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	10.10	11.45	9.62	11.38	11.47
Adj. R^2	0.491	0.514	0.492	0.512	0.517
MSAs	211	211	211	211	211

Standard errors in parentheses. The dependent variable in all columns is the change in the share of non-routine cognitive occupations in the MSA's employed labor force between 1990 and 2015. Each observation (an MSA) is weighted by its employment in 2015. MSA controls include its unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work states. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A-19: Change in routine manual share, 1990-2015

	Δrout-man				
	(1)	(2)	(3)	(4)	(5)
MSA log rent index 1980	0.0089			0.0023	0.0014
	(0.0105)			(0.0123)	(0.0125)
MSA routine cognitive share 1980		0.0703		0.0600	0.0408
		(0.0831)		(0.0822)	(0.0879)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			0.0100	0.0089	0.0082
			(0.0066)	(0.0078)	(0.0078)
MSA Offshorability 1980					0.0149
					(0.0218)
Housing supply elasticity	0.0037***	0.0031**	0.0035***	0.0034**	0.0033**
	(0.0014)	(0.0013)	(0.0013)	(0.0014)	(0.0014)
Amenities	Yes	Yes	Yes	Yes	Yes
MSA's Industry Mix Controls	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	5.28	5.40	5.41	5.13	5.08
Adj. R^2	0.230	0.230	0.235	0.229	0.227
MSAs	211	211	211	211	211

Standard errors in parentheses. The dependent variable in all columns is the change in the share of routine manual occupations in the MSA's employed labor force between 1990 and 2015. Each observation (an MSA) is weighted by its employment in 2015. MSA controls include its unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work states. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

F.2 Residual Wage Distributions

We calculate residual wages as the residual of a Mincer regression. In particular, we estimate a separate Mincer regression for each year:

$$\log(w_{it}) = \alpha_t + \beta_t X_{i,t} + \varepsilon_{i,t} \tag{A.5}$$

We include the typical controls in a Mincer regression (age, age squared, a gender dummy, and a full set of race fixed effects). We also control for educational groups (less than high school, high school graduate, some college, college and more), a dummy for foreign born, and industry groups. Results are qualitatively the same if we do not include industry or educational groups. Results are presented in Figure A-9. As we can see, results are qualitatively the same as the ones presented in Figure 3a.[23] Similarly, we can calculate the mean and median residual wages, as well as the inter-quantiles residual wage differences.

[23]Notice that, while our results are qualitatively the same, some of the controls absorb part of the contribution of city's cost of living to wage inequality. This result is similar to differences in the industrial composition of cities of different sizes explaining up to one-third of the city size effect, as pointed out by Baum-Snow and Pavan (2013).

Table A-20: Change in non-routine manual share, 1990-2015

	Δnonrout-man				
	(1)	(2)	(3)	(4)	(5)
MSA log rent index 1980	-0.0126			-0.0018	-0.0005
	(0.0096)			(0.0115)	(0.0117)
MSA routine cognitive share 1980		0.0086		0.0260	0.0510
		(0.0586)		(0.0562)	(0.0602)
MSA's log $\left(\frac{S}{U}\right)$ in 1980			-0.0157***	-0.0155**	-0.0146**
			(0.0054)	(0.0064)	(0.0061)
MSA Offshorability 1980					-0.0195
					(0.0194)
Housing supply elasticity	-0.0005	0.0001	-0.0002	-0.0004	-0.0002
	(0.0012)	(0.0010)	(0.0010)	(0.0011)	(0.0011)
Amenities	Yes	Yes	Yes	Yes	Yes
MSA's Industry Mix Controls	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
F statistic	10.37	10.82	12.12	11.07	12.09
Adj. R^2	0.414	0.406	0.435	0.430	0.430
MSAs	211	211	211	211	211

Standard errors in parentheses. The dependent variable in all columns is the change in the share of non-routine manual occupations in the MSA's employed labor force between 1990 and 2015. Each observation (an MSA) is weighted by its employment in 2015. MSA controls include its unemployment rate in 1980, the share of the working age population that is female, African American, and Mexican born in 1980, and a dummy for right-to-work states. Industry mix controls include the share of area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public-sector share is excluded). Stars represent: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

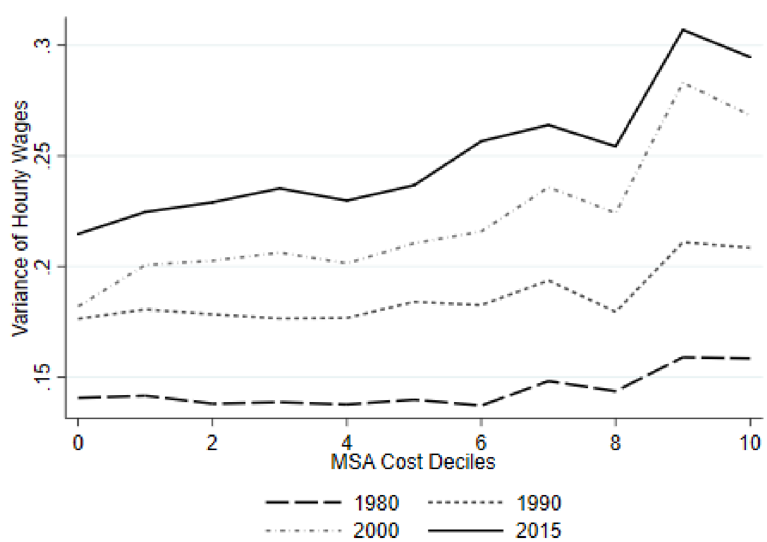
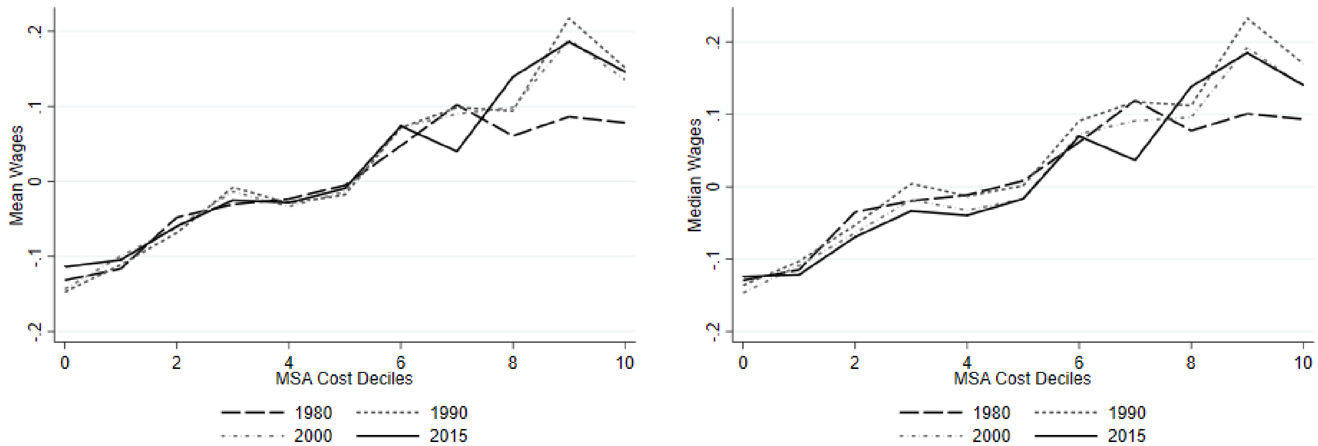
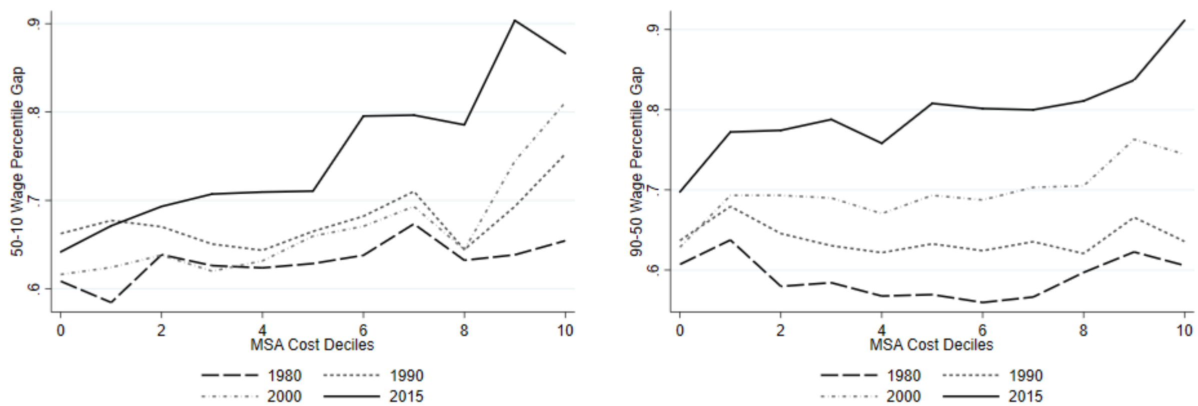


Figure A-9: Variance Residual Wages



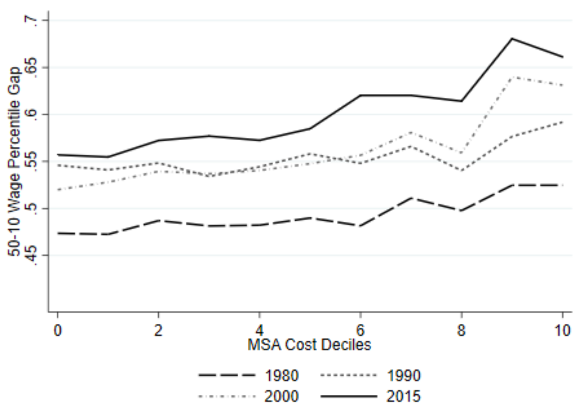
(a) Mean Residual Wages (b) Median Residual Wages

Figure A-10: Mean and median residual wages across city costs and time

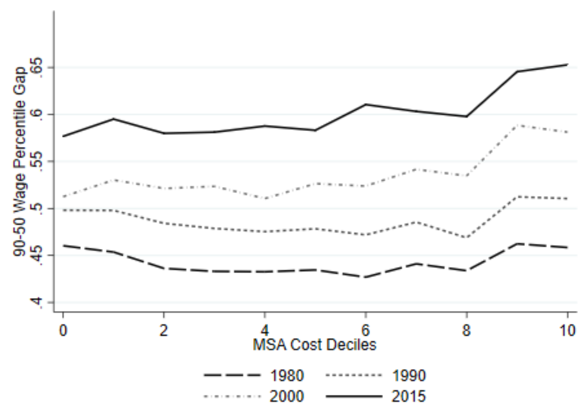


(a) 50-10 Percentile Gap (b) 90-50 Percentile Gap

Figure A-11: Wage gaps across city costs and time



(a) 50–10 Percentile Gap



(b) 90–50 Percentile Gap

Figure A-12: Residual wage gaps across city costs and time

Table A-21: Variance Decomposition: Log hourly wages – Occupational groups

Routine Cognitive

Year	Total	Within City	Variance Between City	% Within	% Between
1980	0.194	0.186	0.008	96%	4%
1990	0.244	0.229	0.016	93%	7%
2000	0.268	0.253	0.015	94%	6%
2015	0.329	0.311	0.018	94%	6%

Non-Routine Cognitive

Year	Total	Within City	Variance Between City	% Within	% Between
1980	0.228	0.217	0.010	96%	5%
1990	0.275	0.259	0.016	94%	6%
2000	0.324	0.308	0.016	95%	5%
2015	0.373	0.348	0.024	93%	7%

Routine Manual

Year	Total	Within City	Variance Between City	% Within	% Between
1980	0.195	0.176	0.019	90%	10%
1990	0.239	0.222	0.017	93%	7%
2000	0.237	0.225	0.012	95%	5%
2015	0.267	0.255	0.012	96%	4%

Non-Routine Manual

Year	Total	Within City	Variance Between City	% Within	% Between
1980	0.212	0.197	0.015	93%	7%
1990	0.275	0.252	0.024	92%	9%
2000	0.279	0.264	0.015	95%	5%
2015	0.289	0.277	0.012	96%	4%

Table A-22: Variance Decomposition: Log hourly residual wages

Year	Total	Within City	Variance Between City	% Within	% Between
1980	0.194	0.187	0.007	96%	4%
1990	0.228	0.216	0.012	95%	5%
2000	0.262	0.251	0.010	96%	4%
2015	0.286	0.276	0.010	96%	4%

Table A-23: Variance Decomposition: Log residual wages – Occupational Groups

Routine Cognitive

			Variance		
Year	Total	Within City	Between City	% Within	% Between
1980	0.172	0.166	0.006	97%	3%
1990	0.211	0.199	0.012	94%	6%
2000	0.237	0.227	0.010	96%	4%
2015	0.264	0.255	0.010	96%	4%

Non-Routine Cognitive

			Variance		
Year	Total	Within City	Between City	% Within	% Between
1980	0.190	0.183	0.007	96%	4%
1990	0.228	0.216	0.012	95%	5%
2000	0.274	0.263	0.011	96%	4%
2015	0.294	0.281	0.013	96%	4%

Routine Manual

			Variance		
Year	Total	Within City	Between City	% Within	% Between
1980	0.189	0.176	0.013	93%	7%
1990	0.208	0.195	0.014	93%	7%
2000	0.220	0.210	0.009	96%	4%
2015	0.241	0.234	0.008	97%	3%

Non-Routine Manual

			Variance		
Year	Total	Within City	Between City	% Within	% Between
1980	0.183107	0.175373	0.0080105	96%	4%
1990	0.205594	0.189578	0.0162198	92%	8%
2000	0.218852	0.207826	0.011194	95%	5%
2015	0.20666	0.198255	0.0086302	96%	4%

G Theoretical Results

G.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:

$$F_j(m_{1j}, m_{2j}, m_{3j}, k_j) = A_j \left[m_{1j}^{\gamma_1} A_{l,1} + \left(m_{2j}^{\theta} A_{l,2} + k_j^{\theta} A_k \right)^{\frac{\gamma_2}{\theta}} + m_{3j}^{\gamma_3} A_{l,3} \right] \tag{A.6}$$

FOCs:

$$\begin{aligned}
&(m_{1j}) : A_j \gamma_1 m_{1j}^{\gamma_1 - 1} A_{l,1} = w_{1j} \\
&(m_{2j}) : A_j \gamma_2 \left(m_{2j}^{\theta} A_{l,2} + k_j^{\theta} A_k \right)^{\frac{\gamma_2}{\theta} - 1} m_{2j}^{\theta - 1} A_{l,2} = w_{2j} \\
&(m_{3j}) : A_j \gamma_3 m_{3j}^{\gamma_3 - 1} A_{l,3} = w_{3j} \\
&(k_j) : A_j \gamma_2 \left(m_{2j}^{\theta} A_{l,2} + k_j^{\theta} A_k \right)^{\frac{\gamma_2}{\theta} - 1} k_j^{\theta - 1} A_k = r
\end{aligned}$$

Since from utility equalization, we have:

$$\frac{w_{ij}}{w_{ij'}} = \left(\frac{p_j}{p_{j'}} \right)^{\alpha}, \quad \forall i \in \{1, 2, 3\} \text{ and } \forall j \in \{1, 2\} \tag{A.7}$$

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

$$m_{11} = \frac{\left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_1 - 1}} M_1}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_1 - 1}}} \tag{A.8}$$

Similarly, for skill 3:

$$m_{31} = \frac{\left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_3 - 1}} M_3}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_3 - 1}}} \tag{A.9}$$

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

$$\left(\frac{m_{21}}{m_{22}} \right) = \left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{\theta - 1}} \frac{k_1}{k_2} \tag{A.10}$$

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

$$m_{21} = \left\{ \frac{1}{A_{l,2}} \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1 - \gamma_2)}{\gamma_2 - \theta}} - A_k \right] \right\}^{\frac{1}{\theta}} k_1 \tag{A.11}$$

Similarly, for (k_2), we have:

$$m_{22} = \left\{ \frac{1}{A_{l,2}} \left[\left(\frac{r}{A_2 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right] \right\}^{\frac{1}{\theta}} k_2 \tag{A.12}$$

Dividing (A.11) by (A.12)and substituting (A.10), we have:

$$\left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{\theta - 1}} = \left\{ \frac{\left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]}{\left[\left(\frac{r}{A_2 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]} \right\} \tag{A.13}$$

Manipulating and simplifying it, we have:

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

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

$$M_2 A_{l,2}^{\frac{1}{\theta}} = \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}} k_1 + \left[\left(\frac{r}{A_2 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}} k_2 \tag{A.14}$$

Substituting (A.13) and manipulating, we have:

$$k_2 = \frac{M_2 A_{l,2}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \tag{A.15}$$

Substituting (A.15) into (A.14) and manipulating, we have:

$$\begin{aligned} &\left\{ \frac{M_2 A_{l,2}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} = \\ &= \left(\frac{A_2}{A_1} \right)^{\frac{\theta}{\gamma_2 - \theta}} \left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} + \left(\frac{r}{A_2 \gamma_2 A_k} \right)^{\frac{\theta}{\theta - \gamma_2}} \left[1 - \left(\frac{p_1}{p_2} \right)^{\frac{\alpha\theta}{1-\theta}} \right] A_k \end{aligned} \tag{A.16}$$

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 for the wage and labor demand and rearranging it, we have:

$$\left\{\begin{array}{c}\left(m_{21}^{\theta}A_{l,2}+k_1^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}}m_{21}^{\theta}A_{l,2}-\\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{22}^{\theta}A_{l,2}+k_2^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}}m_{22}^{\theta}A_{l,2}\end{array}\right\}= \\ \left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}A_{l,1}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right]\\ +\left(\frac{M_3}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3}A_{l,3}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_3}{\gamma_3-1}}\right]\end{array}\right\} \tag{A.17}$$

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

$$\left(m_{22}^{\theta}A_{l,2}+k_2^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}}=\left(\frac{p_2}{p_1}\right)^{\alpha}\left(m_{21}^{\theta}A_{l,2}+k_1^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}}\times\left(\frac{m_{21}}{m_{22}}\right)^{\theta-1}\times\left(\frac{A_1}{A_2}\right) \tag{A.18}$$

Substituting (A.18) into (A.17) and rearranging, we have:

$$\left\{\left[1-\left(\frac{p_1}{p_2}\right)^{1-\alpha}\frac{M_2-m_{21}}{m_{21}}\right]\left(m_{21}^{\theta}A_{l,2}+k_1^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}}m_{21}^{\theta}A_{l,2}\right\}= \\ \left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}A_{l,1}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right]\\ +\left(\frac{M_3}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3}A_{l,3}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_3}{\gamma_3-1}}\right]\end{array}\right\} \tag{A.19}$$

But then, from equation (A.11), we have that:

$$m_{21}^{\theta}A_{l,2}=\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\theta)}{\gamma_2-\theta}}-k_1^{\theta}A_k \tag{A.20}$$

Similarly, from (k_1), we have:

$$\left(m_{21}^{\theta}A_{l,2}+k_1^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}}=\left(\frac{r}{A_1\gamma_2A_k}\right)k_1^{1-\theta} \tag{A.21}$$

Then, from (A.20) and (A.21), we have:

$$\left(m_{21}^{\theta}A_{l,2}+k_1^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}}m_{21}^{\theta}A_{l,2}=\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\gamma_2}{\gamma_2-\theta}}k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2-\theta}}-\frac{r}{A_1\gamma_2}k_1 \tag{A.22}$$

Substituting equation (A.15) into (A.10) and manipulating, we have:

$$\frac{M_2 - m_{21}}{m_{21}} = \frac{M_2 A_{l,2}^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}}}{k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \tag{A.23}$$

Consequently:

$$\left[1 - \left(\frac{p_1}{p_2} \right)^{1-\alpha} \frac{M_2 - m_{21}}{m_{21}} \right] = \frac{\left\{ \begin{array}{c} \left(1 + \left(\frac{p_1}{p_2} \right)^{1-\alpha} \right) k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2} \right)^{1-\alpha} M_2 A_{l,2}^{\frac{1}{\theta}} \end{array} \right\}}{k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \tag{A.24}$$

Then, from equations (A.22) and (A.24), we have that:

$$\begin{gathered} \left[1 - \left(\frac{p_1}{p_2} \right)^{1-\alpha} \frac{M_2 - m_{21}}{m_{21}} \right] \left(m_{21}^{\theta} A_{l,2} + k_1^{\theta} A_k \right)^{\frac{\gamma_2 - \theta}{\theta}} m_{21}^{\theta} A_{l,2} = \\ \frac{\left\{ \begin{array}{c} \left(1 + \left(\frac{p_1}{p_2} \right)^{1-\alpha} \right) k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2} \right)^{1-\alpha} M_2 A_{l,2}^{\frac{1}{\theta}} \end{array} \right\}}{k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \times \left\{ \left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\gamma_2}{\gamma_2 - \theta}} k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2 - \theta}} - \frac{r}{A_1 \gamma_2} k_1 \right\} \end{gathered} \tag{A.25}$$

Notice that the LHS of equation (A.25) is the same as the LHS of equation (A.19). Substituting it back, we have:

$$\begin{gathered} \frac{\left\{ \begin{array}{c} \left(1 + \left(\frac{p_1}{p_2} \right)^{1-\alpha} \right) k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2} \right)^{1-\alpha} M_2 A_{l,2}^{\frac{1}{\theta}} \end{array} \right\}}{k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\theta}{\gamma_2 - \theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \times \left\{ \left(\frac{r}{A_1 \gamma_2 A_k} \right)^{\frac{\gamma_2}{\gamma_2 - \theta}} k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2 - \theta}} - \frac{r}{A_1 \gamma_2} k_1 \right\} = \\ \left\{ \begin{array}{c} \left(\frac{M_1}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_1 - 1}}} \right)^{\gamma_1} A_{l,1} \left[\frac{A_2}{A_1} \frac{p_1}{p_2} - \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{\gamma_1}{\gamma_1 - 1}} \right] \\ + \left(\frac{M_3}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_3 - 1}}} \right)^{\gamma_3} A_{l,3} \left[\frac{A_2}{A_1} \frac{p_1}{p_2} - \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{\gamma_3}{\gamma_3 - 1}} \right] \end{array} \right\} \end{gathered} \tag{A.26}$$

Finally, notice that equations (A.26) and (A.16) generate a system with two equations and two unknowns (k_1 and $\frac{p_1}{p_2}$):

$$\begin{cases} \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_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-A_k\right]^{\frac{1}{\theta}}\\ -\left(\frac{p_1}{p_2}\right)^{1-\alpha}M_2A_{l,2}^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-A_k\right]^{\frac{1}{\theta}}}\times\left\{\left(\frac{r}{A_1\gamma_2A_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}A_{l,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}A_{l,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\} & (F.1) \\ \left\{\dfrac{M_2A_{l,2}^{\frac{1}{\theta}}-\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-A_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_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}}k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}-A_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}}= & (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_2A_k}\right)^{\frac{\theta}{\theta-\gamma_2}}\left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right]A_k \end{cases}$$

G.1.1 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.:

$$\frac{m_{i1}}{m_{11}+m_{21}+m_{31}} = \frac{m_{i2}}{m_{12}+m_{22}+m_{32}} \tag{A.27}$$

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:

$$\left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}A_{l,1}\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right]\\ +\left(\frac{M_3}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3}A_{l,3}\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_3\alpha}{\gamma_3-1}}\right]\end{array}\right\}>0$$

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 (A.22), we have:

$$\left(m_{21}^{\theta}A_{l,2} + k_1^{\theta}A_k\right)^{\frac{\gamma_2-\theta}{\theta}} m_{21}^{\theta}A_{l,2} = \left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\gamma_2}{\gamma_2-\theta}} k_1^{\frac{\gamma_2(1-\theta)}{\gamma_2-\theta}} - \frac{r}{A_1\gamma_2}k_1$$

So the second term on the LHS of $(F.1)$ must be positive. Moreover, from (A.21), we have that:

$$k_1\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}} = m_{21}A_{l,2}^{\frac{1}{\theta}} > 0$$

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

$$\frac{M_2A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}} < k_1\left(\frac{p_1}{p_2}\right)^{\alpha-1}$$

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

$$\frac{M_2A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}} < k_1\left(\frac{p_2}{p_1}\right)^{\left(1+\frac{\alpha\theta}{1-\theta}\right)} \tag{A.28}$$

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

$$\frac{M_2A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}} > \left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}} k_1 \tag{A.29}$$

Then, notice that:

$$1 + \frac{\alpha\theta}{1-\theta} - \frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)} = 1 + \frac{\alpha\theta}{1-\theta}\left[1 - \frac{\theta-\gamma_2}{\theta(1-\gamma_2)}\right] = 1 + \frac{\alpha\theta}{1-\theta}\left[\frac{\gamma_2(1-\theta)}{\theta(1-\gamma_2)}\right] > 0 \tag{A.30}$$

Therefore the exponent at $\frac{p_2}{p_1}$ is higher on the RHS of (A.28). 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 (A.28) and (A.29) 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:

$$\left\{\begin{array}{l} \left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} A_{l,1}\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right] \\ +\left(\frac{M_3}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_3-1}}}\right)^{\gamma_3} A_{l,3}\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_3\alpha}{\gamma_3-1}}\right] \end{array}\right\} < 0$$

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 $\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}$:

$$\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}} > k_1\left(\frac{p_2}{p_1}\right)^{\left(1+\frac{\alpha\theta}{1-\theta}\right)} \tag{A.31}$$

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

$$\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}} < \left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}} k_1 \tag{A.32}$$

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 (A.31) and (A.32) give us a contradiction. Therefore, we have that $p_1 = p_2 \Leftrightarrow A_1 = A_2$. ■

G.1.2 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 (A.28) must hold. Then, from $(F.2)$ we have that, given that $p_1 > p_2$, the last term in $(F.2)$'s RHS – $\left(\frac{r}{A_2\gamma_2 A_k}\right)^{\frac{\theta}{\theta-\gamma_2}}\left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right]A_k$ – 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:

$$\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}} > \left(\frac{p_2}{p_1}\right)^{\frac{\alpha\theta}{1-\theta}\times\frac{\theta-\gamma_2}{\theta(1-\gamma_2)}} k_1 \tag{A.33}$$

Given (A.30) 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, (A.28) and (A.33) 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.1 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 1, we have that $p_1 > p_2$. Then, from equation (A.13), we have:

$$\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta-1}} = \left[\frac{\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k}{\left(\frac{r}{A_2\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k}\right] \tag{A.34}$$

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

$$\left[\frac{\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k}{\left(\frac{r}{A_2\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_2^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k}\right] < 1 \tag{A.35}$$

Rearranging it:

$$\left(\frac{k_1}{k_2}\right)^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} < \left(\frac{A_1}{A_2}\right)^{\frac{\theta}{\gamma_2-\theta}} \tag{A.36}$$

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 $\frac{k_1}{k_2} > \frac{A_1}{A_2} \Rightarrow k_1 > k_2$. ■

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:

$$\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\right]^{\frac{\gamma_i}{1-\gamma_i}}, \ \forall i \in \{1,3\} \tag{A.37}$$

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 (A.37) 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 $(F.1)$'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 (A.28) 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:

$$\left\{\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}}\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}\right\} > \left(\frac{A_2}{A_1}\right)^{\frac{1}{1-\gamma_2}}\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}\times\frac{\gamma_2-\theta}{(1-\gamma_2)}} k_1 \tag{A.38}$$

Then, we have that:

$$\frac{\text{RHS(A.28)}}{\text{RHS(A.38)}} = \left(\frac{p_2}{p_1}\right)^{1+\frac{\alpha\theta}{1-\theta}\left[1-\frac{\gamma_2-\theta}{\theta(1-\gamma_2)}\right]}\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_2}} \tag{A.39}$$

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

$$\frac{\text{RHS(A.28)}}{\text{RHS(A.38)}} = \left(\frac{p_2}{p_1}\right)^{1+\frac{\gamma_2\alpha}{(1-\gamma_2)}}\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_2}} = \left\{\left(\frac{p_2}{p_1}\right)^{1-\gamma_2(1-\alpha)}\frac{A_1}{A_2}\right\}^{\frac{1}{1-\gamma_2}} \tag{A.40}$$

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:

$$\left(\frac{p_2}{p_1}\right)^{1-\gamma_2(1-\alpha)} \frac{A_1}{A_2} < \left(\frac{p_2}{p_1}\right)^{\alpha} \frac{A_1}{A_2} < 1 \tag{A.41}$$

where the last inequality comes from our assumption for the contradiction. Then, since $\frac{1}{1-\gamma_2} > 0$, we have $\frac{\text{RHS(A.28)}}{\text{RHS(A.38)}} < 1$. But then inequalities (A.28) and (A.38) cannot both be satisfied and we have a contradiction. ■

Corollary 1 *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:

$$m_{11} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma_1-1}} M_1}{\left\{1 + \left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma_1-1}}\right\}} = \frac{\left[\left(\frac{p_2}{p_1}\right)^{\alpha} \frac{A_1}{A_2}\right]^{\frac{1}{1-\gamma_1}} M_1}{\left\{1 + \left[\left(\frac{p_2}{p_1}\right)^{\alpha} \frac{A_1}{A_2}\right]^{\frac{1}{1-\gamma_1}}\right\}} \tag{A.42}$$

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 (A.23):

$$\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1 \gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1 \gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}} > k_1 \tag{A.43}$$

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

$$\begin{aligned} &\left\{\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1 \gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1 \gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} = \\ &= \left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_2-\theta}} \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} + \left(\frac{r}{A_2 \gamma_2 A_k}\right)^{\frac{\theta}{\theta-\gamma_2}} \left[1 - \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right] A_k \end{aligned} \tag{A.44}$$

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:

$$\left\{\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1}{\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} > \left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_2-\theta}} \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}\times\left[1+\frac{1-\gamma_2}{\gamma_2-\theta}\right]} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} \tag{A.45}$$

Now, from (A.43) we have that, since $\gamma_2 < \theta$:

$$\left\{\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} < k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} \tag{A.46}$$

Now, substituting (A.46) into (A.45), we have:

$$k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} > \left\{\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - k_1\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}{\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} > \left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\right]^{\frac{\theta}{\theta-\gamma_2}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} \tag{A.47}$$

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:

$$\left\{\frac{M_2 A_{l,2}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1}{k_1\left[\left(\frac{r}{A_1\gamma_2 A_k}\right)^{\frac{\theta}{\gamma_2-\theta}} k_1^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} - A_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_2)}{\gamma_2-\theta}} < \left\{\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right\}^{\frac{\theta}{\gamma_2-\theta}}$$

From (A.23) 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:

$$\frac{pdf_{11}}{pdf_{12}} > \left[\left(\frac{p_2}{p_1}\right)^{\alpha} \frac{A_1}{A_2}\right]^{\frac{1}{1-\gamma}} \times \frac{1}{\left[\left(\frac{p_2}{p_1}\right)^{\alpha} \frac{A_1}{A_2}\right]^{\frac{1}{1-\gamma}}} \tag{A.48}$$

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 $pdf_{21} < pdf_{22}$ ■

G.2 Skill-Biased Technological Change

We now consider the case of skill-biased technological change (henceforth SBTC) in which capital and high-skill workers are complements. Consequently, the technology is given by the following production function:

$$A_j F(m_{1j}, m_{2j}, m_{3j}, k) = A_j \left\{ m_{1j}^{\gamma_1} A_{l,1} + \left(m_{3j}^{\theta} A_{l,3} + k_j^{\theta} A_k\right)^{\frac{\gamma_3}{\theta}} + m_{2j}^{\gamma_2} A_{l,2} \right\} \quad \text{where} \quad \gamma_3 > \theta \tag{A.49}$$

As a result, the FOCs for each city j, skill type i, and capital, respectively are:

$$\begin{aligned}
&(m_{1j}) : A_j \gamma_1 m_{1j}^{\gamma_1 - 1} A_{l,1} = w_{1j} \\
&(m_{2j}) : A_j \gamma_2 m_{2j}^{\gamma_2 - 1} A_{l,2} = w_{2j} \\
&(m_{3j}) : A_j \gamma_3 \left(m_{3j}^{\theta} A_{l,3} + k_j^{\theta} A_k\right)^{\frac{\gamma_3}{\theta} - 1} m_{3j}^{\theta - 1} A_{l,3} = w_{3j} \\
&(k_j) : A_j \gamma_3 \left(m_{3j}^{\theta} A_{l,3} + k_j^{\theta} A_k\right)^{\frac{\gamma_3}{\theta} - 1} k_j^{\theta - 1} A_k = r
\end{aligned} \tag{A.50}$$

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:

$$m_{11} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma_1 - 1}} M_1}{\left\{1 + \left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma_1 - 1}}\right\}} \quad \text{and} \quad m_{21} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma_2 - 1}} M_2}{\left\{1 + \left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma_2 - 1}}\right\}} \tag{A.51}$$

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

$$m_{31} = \frac{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{\theta-1}} \frac{k_1}{k_2}}{\left[1+\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{\theta-1}} \frac{k_1}{k_2}\right]} M_3 \quad \text{and} \quad k_2 = \frac{M_3 A_{l,3}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}}} \tag{A.52}$$

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 such a city even if the productivity is lower. Because in our empirical application we consider large metropolitan areas (New York 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 2 shows that the high TFP city also attracts more high-skill workers.

Corollary 2 (SBTC and Demand for High Skill Workers) *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 $\gamma > \theta$, a 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.*

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:

$$F_j(m_{1j}, m_{2j}, m_{3j}, k_j) = A_j \left[m_{1j}^{\gamma_1} A_{l,1} + \left(m_{3j}^{\theta} A_{l,3} + k_j^{\theta} A_k \right)^{\frac{\gamma_3}{\theta}} + m_{2j}^{\gamma_2} A_{l,2} \right] \tag{A.53}$$

FOCs:

$$\begin{aligned}
&(m_{1j}) : A_j \gamma_1 m_{1j}^{\gamma_1 - 1} A_{l,1} = w_{1j} \\
&(m_{2j}) : A_j \gamma_2 m_{2j}^{\gamma_2 - 1} A_{l,2} = w_{2j} \\
&(m_{3j}) : A_j \gamma_3 \left(m_{3j}^{\theta} A_{l,3} + k_j^{\theta} A_k \right)^{\frac{\gamma_3}{\theta} - 1} m_{3j}^{\theta - 1} A_{l,3} = w_{3j} \\
&(k_j) : A_j \gamma_3 \left(m_{3j}^{\theta} A_{l,3} + k_j^{\theta} A_k \right)^{\frac{\gamma_3}{\theta} - 1} k_j^{\theta - 1} A_k = r
\end{aligned}$$

Since from utility equalization, we have:

$$\frac{w_{ij}}{w_{ij'}} = \left(\frac{p_j}{p_{j'}} \right)^{\alpha}, \quad \forall i \in \{1, 2, 3\} \text{ and } \forall j \in \{1, 2\} \tag{A.54}$$

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

$$m_{11} = \frac{\left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_1 - 1}} M_1}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_1 - 1}}} \tag{A.55}$$

Similarly, for skill 2:

$$m_{21} = \frac{\left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_2 - 1}} M_2}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_2 - 1}}} \tag{A.56}$$

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

$$\left(\frac{m_{31}}{m_{32}} \right) = \left(\frac{p_1}{p_2} \right)^{\frac{\alpha}{\theta - 1}} \frac{k_1}{k_2} \tag{A.57}$$

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

$$m_{31} = \left\{ \frac{1}{A_{l,3}} \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1 - \gamma_3)}{\gamma_3 - \theta}} - A_k \right] \right\}^{\frac{1}{\theta}} k_1 \tag{A.58}$$

Similarly, for (k_2), we have:

$$m_{32} = \left\{ \frac{1}{A_{l,3}} \left[\left(\frac{r}{A_2 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_2^{\frac{\theta(1 - \gamma_3)}{\gamma_3 - \theta}} - A_k \right] \right\}^{\frac{1}{\theta}} k_2 \tag{A.59}$$

Dividing (A.11) by (A.59)and substituting (A.57), we have:

$$\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{\theta-1}} = \left\{\frac{\left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]}{\left[\left(\frac{r}{A_2\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_2^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]}\right\} \tag{A.60}$$

Manipulating and simplifying it, we have:

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

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

$$M_3 A_{l,3}^{\frac{1}{\theta}} = \left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1 + \left[\left(\frac{r}{A_2\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_2^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}} k_2 \tag{A.61}$$

Substituting (A.60) and manipulating, we have:

$$k_2 = \frac{M_3 A_{l,3}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}}} \tag{A.62}$$

Substituting (A.62) into (A.61) and manipulating, we have:

$$\begin{aligned}&\left\{\frac{M_3 A_{l,3}^{\frac{1}{\theta}} - \left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}} k_1}{\left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} \left[\left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} - A_k\right]^{\frac{1}{\theta}}}\right\}^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} = \\ &= \left(\frac{A_2}{A_1}\right)^{\frac{\theta}{\gamma_3-\theta}} \left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}} + \left(\frac{r}{A_2\gamma_3 A_k}\right)^{\frac{\theta}{\theta-\gamma_3}} \left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right] A_k\end{aligned} \tag{A.63}$$

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 for the wage and labor demand and rearranging it, we have:

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta}A_{l,3}+k_1^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}A_{l,3}- \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta}A_{l,3}+k_2^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta}A_{l,3} \end{array} \right\} =$$
$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} A_{l,1}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right] \\ +\left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} A_{l,2}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_2}{\gamma_2-1}}\right] \end{array} \right\} \tag{A.64}$$

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

$$\left(m_{32}^{\theta}A_{l,3}+k_2^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} = \left(\frac{p_2}{p_1}\right)^{\alpha}\left(m_{31}^{\theta}A_{l,3}+k_1^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} \times \left(\frac{m_{31}}{m_{32}}\right)^{\theta-1} \times \left(\frac{A_1}{A_2}\right) \tag{A.65}$$

Substituting (A.65) into (A.64) and rearranging, we have:

$$\left\{ \left[1-\left(\frac{p_1}{p_2}\right)^{1-\alpha}\frac{M_3-m_{31}}{m_{31}}\right]\left(m_{31}^{\theta}A_{l,3}+k_1^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}A_{l,3} \right\} =$$
$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} A_{l,1}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_1}{\gamma_1-1}}\right] \\ +\left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} A_{l,2}\left[\frac{A_2}{A_1}\frac{p_1}{p_2}-\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{\gamma_2}{\gamma_2-1}}\right] \end{array} \right\} \tag{A.66}$$

But then, from equation (A.58), we have that:

$$m_{31}^{\theta}A_{l,3} = \left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\theta}{\gamma_3-\theta}} k_1^{\frac{\theta(1-\theta)}{\gamma_3-\theta}} - k_1^{\theta}A_k \tag{A.67}$$

Similarly, from (k_1), we have:

$$\left(m_{31}^{\theta}A_{l,3}+k_1^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} = \left(\frac{r}{A_1\gamma_3 A_k}\right)k_1^{1-\theta} \tag{A.68}$$

Then, from (A.67) and (A.68), we have:

$$\left(m_{31}^{\theta}A_{l,3}+k_1^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}A_{l,3} = \left(\frac{r}{A_1\gamma_3 A_k}\right)^{\frac{\gamma_3}{\gamma_3-\theta}} k_1^{\frac{\gamma_3(1-\theta)}{\gamma_3-\theta}} - \frac{r}{A_1\gamma_3}k_1 \tag{A.69}$$

Substituting equation (A.62) into (A.57) and manipulating, we have:

$$\frac{M_3 - m_{31}}{m_{31}} = \frac{M_3 A_{l,3}^{\frac{1}{\theta}} - k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}}}{k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \tag{A.70}$$

Consequently:

$$\left[1 - \left(\frac{p_1}{p_2} \right)^{1-\alpha} \frac{M_3 - m_{31}}{m_{31}} \right] = \frac{\left\{ \begin{array}{c} \left(1 + \left(\frac{p_1}{p_2} \right)^{1-\alpha} \right) k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2} \right)^{1-\alpha} M_3 A_{l,3}^{\frac{1}{\theta}} \end{array} \right\}}{k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \tag{A.71}$$

Then, from equations (A.69) and (A.71), we have that:

$$\left[1 - \left(\frac{p_1}{p_2} \right)^{1-\alpha} \frac{M_3 - m_{31}}{m_{31}} \right] \left(m_{31}^{\theta} A_{l,3} + k_1^{\theta} A_k \right)^{\frac{\gamma_3 - \theta}{\theta}} m_{31}^{\theta} A_{l,3} =$$
$$\frac{\left\{ \begin{array}{c} \left(1 + \left(\frac{p_1}{p_2} \right)^{1-\alpha} \right) k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2} \right)^{1-\alpha} M_3 A_{l,3}^{\frac{1}{\theta}} \end{array} \right\}}{k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \times \left\{ \left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\gamma_3}{\gamma_3 - \theta}} k_1^{\frac{\gamma_3(1-\theta)}{\gamma_3 - \theta}} - \frac{r}{A_1 \gamma_3} k_1 \right\} \tag{A.72}$$

Notice that the LHS of equation (A.72) is the same as the LHS of equation (A.66). Substituting it back, we have:

$$\frac{\left\{ \begin{array}{c} \left(1 + \left(\frac{p_1}{p_2} \right)^{1-\alpha} \right) k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}} \\ - \left(\frac{p_1}{p_2} \right)^{1-\alpha} M_3 A_{l,3}^{\frac{1}{\theta}} \end{array} \right\}}{k_1 \left[\left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\theta}{\gamma_3 - \theta}} k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3 - \theta}} - A_k \right]^{\frac{1}{\theta}}} \times \left\{ \left(\frac{r}{A_1 \gamma_3 A_k} \right)^{\frac{\gamma_3}{\gamma_3 - \theta}} k_1^{\frac{\gamma_3(1-\theta)}{\gamma_3 - \theta}} - \frac{r}{A_1 \gamma_3} k_1 \right\} =$$
$$\left\{ \begin{array}{c} \left(\frac{M_1}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_1 - 1}}} \right)^{\gamma_1} A_{l,1} \left[\frac{A_2}{A_1} \frac{p_1}{p_2} - \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{\gamma_1}{\gamma_1 - 1}} \right] \\ + \left(\frac{M_2}{1 + \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{\gamma_2 - 1}}} \right)^{\gamma_2} A_{l,2} \left[\frac{A_2}{A_1} \frac{p_1}{p_2} - \left[\frac{A_2}{A_1} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{\gamma_2}{\gamma_2 - 1}} \right] \end{array} \right\} \tag{A.73}$$

Finally, notice that equations (A.73) and (A.63) 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_3A_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-A_k\right]^{\frac{1}{\theta}}\\-\left(\frac{p_1}{p_2}\right)^{1-\alpha}M_3A_{l,3}^{\frac{1}{\theta}}\end{array}\right\}}{k_1\left[\left(\frac{r}{A_1\gamma_3A_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-A_k\right]^{\frac{1}{\theta}}}\times\left\{\left(\frac{r}{A_1\gamma_3A_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}A_{l,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}A_{l,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\}\\ \left\{\dfrac{M_3A_{l,3}^{\frac{1}{\theta}}-\left[\left(\frac{r}{A_1\gamma_3A_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-A_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_3A_k}\right)^{\frac{\theta}{\gamma_3-\theta}}k_1^{\frac{\theta(1-\gamma_3)}{\gamma_3-\theta}}-A_k\right]^{\frac{1}{\theta}}}\right\}^{\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_3A_k}\right)^{\frac{\theta}{\theta-\gamma_3}}\left[1-\left(\frac{p_1}{p_2}\right)^{\frac{\alpha\theta}{1-\theta}}\right]A_k\end{array}\right.\begin{array}{r}(F.1)\\ \\ (F.2)\end{array}$$

G.2.1 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.:

$$\frac{m_{i1}}{m_{11}+m_{21}+m_{31}} = \frac{m_{i2}}{m_{12}+m_{22}+m_{32}} \tag{A.74}$$

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 (A.60), we have $k_1 < k_2$. But then, from equation (A.57), we obtain $m_{31} < m_{32}$. Finally, from the RHS of (A.17), we have:

$$\left\{\begin{array}{c}\left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1}A_{l,1}\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right]\\+\left(\frac{M_2}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2}A_{l,2}\left[\frac{p_1}{p_2}-\left(\frac{p_1}{p_2}\right)^{\frac{\gamma_2\alpha}{\gamma_2-1}}\right]\end{array}\right\}>0$$

Since $p_1 > p_2$, $\gamma_1 < 1$, and $\gamma_2 < 1$. However, given the results we obtained from (A.60) and (A.57), the LHS of (A.64) gives us:

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta} A_{l,3} + k_1^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta} A_{l,3} - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta} A_{l,3} + k_2^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta} A_{l,3} \end{array} \right\} < 0$$

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 (A.60), we have $k_1 > k_2$. Similarly, from (A.57), we obtain $m_{31} > m_{32}$. But then, from (A.64), we have that:

$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} A_{l,1}\left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_1\alpha}{\gamma_1-1}}\right] \\ +\left(\frac{M_2}{1+\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} A_{l,2}\left[\frac{p_1}{p_2} - \left(\frac{p_1}{p_2}\right)^{\frac{\gamma_2\alpha}{\gamma_2-1}}\right] \end{array} \right\} < 0$$

given $p_1 < p_2$. Then RHS(A.64) < 0. While

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta} A_{l,3} + k_1^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta} A_{l,3} - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta} A_{l,3} + k_2^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta} A_{l,3} \end{array} \right\} > 0.$$

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 (A.57), we have:

$$\frac{m_{31}}{m_{32}} = \frac{k_1}{k_2} \tag{A.75}$$

From (A.60) , we have

$$\frac{k_1}{k_2} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{A.76}$$

Combining (A.75) and (A.76), we have:

$$\frac{m_{31}}{m_{32}} = \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{A.77}$$

But then, from LHS(A.64), substituting (A.75) and (A.77) given $p_1 = p_2$, we have:

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta} A_{l,3} + k_1^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta} A_{l,3} - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta} A_{l,3} + k_2^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta} A_{l,3} \end{array} \right\} = \left[\left(\frac{A_1}{A_2}\right)^{\frac{\gamma_3}{1-\gamma_3}} - \frac{A_2}{A_1}\right]\left(m_{32}^{\theta} A_{l,3} + k_2^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta} A_{l,3} \tag{A.78}$$

while the RHS(A.64) gives us:

$$\left\{ \begin{array}{c} \left(\frac{M_1}{1+\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_1}}} \right)^{\gamma_1} A_{l,1} \left[\frac{A_2}{A_1} - \left(\frac{A_1}{A_2}\right)^{\frac{\gamma_1}{1-\gamma_1}} \right] \\ + \left(\frac{M_2}{1+\left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_2}}} \right)^{\gamma_2} A_{l,2} \left[\frac{A_2}{A_1} - \left(\frac{A_1}{A_2}\right)^{\frac{\gamma_2}{1-\gamma_2}} \right] \end{array} \right\} \tag{A.79}$$

Then, consider the case in which $A_1 > A_2$. From (A.78), we have that LHS(A.64)> 0, while (A.79) gives us RHS(A.64)< 0. Similarly, if $A_1 < A_2$, (A.78) gives us LHS(A.64)< 0 while (A.79) gives us RHS(A.64)> 0. Consequently, (A.64) is only satisfied if $A_1 = A_2$, concluding our proof.■

G.2.2 Proofs

Proof of Proposition 3

Proof. Towards a contradiction, assume that $A_1 > A_2$ and $p_2 > p_1$. Then, from (A.60), after some manipulations and using $\gamma_3 > \theta$, and $\frac{p_2}{p_1} > 1$ we have:

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

i.e.:

$$\frac{k_1}{k_2} > \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{A.80}$$

From equation (A.57), we have:

$$\frac{m_{31}}{m_{32}} > \frac{k_1}{k_2} \Rightarrow \frac{m_{31}}{m_{32}} > \left(\frac{A_1}{A_2}\right)^{\frac{1}{1-\gamma_3}} \tag{A.81}$$

Then, from LHS(A.64), substituting (A.80) and (A.81), we have:

$$\left\{ \begin{array}{c} \left(m_{31}^{\theta} A_{l,3} + k_1^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta} A_{l,3} - \\ -\frac{A_2}{A_1}\frac{p_1}{p_2}\left(m_{32}^{\theta} A_{l,3} + k_2^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta} A_{l,3} \end{array} \right\} > \left[\left(\frac{A_1}{A_2}\right)^{\frac{\gamma_3}{1-\gamma_3}} - \frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)\right] \left(m_{32}^{\theta} A_{l,3} + k_2^{\theta} A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{32}^{\theta} A_{l,3} > 0 \tag{A.82}$$

While from RHS(A.64), 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(A.17)< 0, which gives us a contradiction. Since we showed in lemma A.4 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 (A.62) 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 (A.57), 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:

$$\frac{m_{31}}{m_{32}} > \left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right]^{\frac{1}{1-\gamma_3}} \tag{A.83}$$

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 (A.83) already implies that $m_{31} \geq m_{32}$. From (A.57) and $\theta < 1$, we have that:

$$\frac{k_1}{k_2} \geq \left(\frac{p_1}{p_2}\right)^{\frac{\alpha}{1-\theta}} > 1 \tag{A.84}$$

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(A.64), we have that:

$$\left\{ \begin{array}{l} \left(\frac{M_1}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_1-1}}}\right)^{\gamma_1} A_{l,1}\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right]^{\frac{\gamma_1}{1-\gamma_1}}\right] \\ + \left(\frac{M_2}{1+\left[\frac{A_2}{A_1}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{\gamma_2-1}}}\right)^{\gamma_2} A_{l,2}\left[\frac{A_2}{A_1}\frac{p_1}{p_2} - \left[\frac{A_1}{A_2}\left(\frac{p_2}{p_1}\right)^{\alpha}\right]^{\frac{\gamma_2}{1-\gamma_2}}\right] \end{array} \right\}$$

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(A.64)> 0.

But then, from (A.66), given that $\left(m_{31}^{\theta}A_{l,3} + k_1^{\theta}A_k\right)^{\frac{\gamma_3-\theta}{\theta}} m_{31}^{\theta}A_{l,3} > 0$, 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:

$$\frac{m_{31}}{m_{32}} > \left(\frac{p_1}{p_2}\right)^{1-\alpha} > 1 \tag{A.85}$$

From (A.85) and (A.57), we have:

$$\left(\frac{p_2}{p_1}\right)^{\frac{\alpha}{1-\theta}} \frac{k_1}{k_2} > \left(\frac{p_1}{p_2}\right)^{1-\alpha} \Rightarrow \frac{k_1}{k_2} > \left(\frac{p_1}{p_2}\right)^{1+\frac{\alpha\theta}{1-\theta}} \tag{A.86}$$

Consequently, (A.86) implies that $k_1 > k_2$, concluding our proof. ■

Proof of Corollary 2

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(A.17)> 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:

$$m_{31}m_{12} - m_{32}m_{11} + m_{31}m_{22} - m_{32}m_{21} \leq 0 \tag{A.87}$$

From equations (A.55) and (A.56) and labor market clearing conditions, we have:

$$m_{11} = \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}} m_{12} \quad \text{and} \quad m_{21} = \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}} m_{22} \tag{A.88}$$

As a result, we have:

$$m_{31}m_{12} - m_{32}m_{11} = m_{32}m_{12}\left\{\frac{m_{31}}{m_{32}} - \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}}\right\} > 0 \tag{A.89}$$

and

$$m_{31}m_{22} - m_{32}m_{21} = m_{32}m_{22}\left\{\frac{m_{31}}{m_{32}} - \left[\frac{A_1}{A_2}\left(\frac{p_1}{p_2}\right)^{\alpha}\right]^{\frac{1}{1-\gamma}}\right\} > 0 \tag{A.90}$$

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 (A.83). Consequently, equations (A.87), (A.89), and (A.90) 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:

$$m_{12}m_{21} - m_{22}m_{11} + m_{32}m_{21} - m_{31}m_{22} \leq 0 \tag{A.91}$$

From (A.88), after some manipulations, we have:

$$m_{12}m_{21} - m_{22}m_{11} = 0 \tag{A.92}$$

and

$$m_{32}m_{21} - m_{31}m_{22} = m_{32}m_{22} \left\{ \left[\frac{A_1}{A_2} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{1-\gamma}} - \frac{m_{31}}{m_{32}} \right\} < 0 \tag{A.93}$$

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 (A.83). Consequently, equations (A.91), (A.92), and (A.93) 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:

$$m_{11}m_{22} - m_{12}m_{21} + m_{32}m_{11} - m_{31}m_{12} \leq 0 \tag{A.94}$$

In equation (A.92), we already showed that $m_{11}m_{22} - m_{12}m_{21} = 0$. Then, from (A.88) and (A.83), we have:

$$m_{32}m_{11} - m_{31}m_{12} = m_{32}m_{12} \left\{ \left[\frac{A_1}{A_2} \left(\frac{p_1}{p_2} \right)^{\alpha} \right]^{\frac{1}{1-\gamma}} - \frac{m_{31}}{m_{32}} \right\} < 0 \tag{A.95}$$

Consequently, equations (A.94), (A.92), and (A.95) jointly show a contradiction. As a result, $pdf_{11} < pdf_{12}$, concluding our proof that pdf_1 F.O.S.D. pdf_2. ■

H Estimation: Additional Derivations and Supporting Information

H.1 Additional Derivations

The following derivations use the fact that integrals involving the Fréchet distribution often have a closed-form solution, so we can calculate expectations analytically. In the model workers choose first location and then occupation. Therefore, we need to calculate the expected value of a location for a worker of a given type, without knowing yet the realization of the occupation-specific value. Here $f(t_i)$ is the pdf and $F(t_i)$ the cdf of the Fréchet distribution with shape parameter η and scale parameter 1.

$$\mathbb{E}_{\mathbf{t}}[\max_i \bar{u}(i,j,\mathbf{s})t_i] = \int \cdots \int \max\{\bar{u}(1,j,\mathbf{s})t_1, \ldots, \bar{u}(I,j,\mathbf{s})t_I\} \prod_i f(t_i) dt_1 \ldots dt_I \tag{A.96}$$

$$= \sum_i \int \bar{u}(i,j,\mathbf{s})t_i f(t_i) \prod_{i' \neq i} F\left(\frac{\bar{u}(i,j,\mathbf{s})t_i}{\bar{u}(i',j,\mathbf{s})}\right) dt_i \tag{A.97}$$

$$= \sum_i \bar{u}(i,j,\mathbf{s}) \left(\frac{1}{\sum_{i'}\left(\frac{\bar{u}(i,j,\mathbf{s})}{\bar{u}(i',j,\mathbf{s})}\right)^{-\eta}}\right)^{1-\frac{1}{\eta}} \Gamma\left(1-\frac{1}{\eta}\right) \tag{A.98}$$

The second line follows from the decision rule that a worker chooses occupation i if $\bar{u}(i,j,\mathbf{s})t_i > \bar{u}(i',j,\mathbf{s})t_{i'}\ \forall i'$ and that the draws of $t_i \forall i$ are i.i.d. from a Fréchet distribution. Thus, the probability of choosing i given t_i is $\prod_{i' \neq i} F\left(\frac{\bar{u}(i,j,\mathbf{s})t_i}{\bar{u}(i',j,\mathbf{s})}\right)$. The third line uses the definition of the Fréchet distribution, which allows us to calculate the analytical solution of the integral.

The choice probabilities in equations (12) and (13) follow from standard results and use the fact that draws of a_j and t_i are i.i.d. from given Fréchet distributions', see, e.g., Eaton and Kortum (2002).

H.2 Standard Errors

The estimator $\hat{\theta}$ solves

$$min_\theta \left(\bar{m} - m(\theta)\right)' \Omega \left(\bar{m} - m(\theta)\right).$$

where Ω is the weight matrix. In standard estimation problems the efficient choice for Ω would be to set it to the inverse of the covariance matrix of the data moment vector $\bar{m}$. However, we do not have an estimate for the full covariance matrix. Instead, we fix Ω to be diagonal with the weights

- 100 for city wage, city size, house prices, wage by occupation
- 1000 for occupation shares, IT share, standard deviation of wages by occupation, elasticity of employment share with respect to IT prices.

The estimator of the variance covariance matrix of the estimated parameters $\hat{\theta}$ is

$$\hat{V} = (\hat{M}'\Omega\hat{M})^{-1}\hat{M}'\Omega\hat{\Sigma}\Omega\hat{M}(\hat{M}'\Omega\hat{M})^{-1} \tag{A.99}$$

where $\hat{\Sigma}$ is the variance covariance matrix of the moments m_i. $\hat{M}$ is the jacobian of the moments with respect to the parameters. The jacobian is calculated numerically by finite differences.

The variances of several moments (and their covariances with the remaining moments) are not defined as we take those estimates from other sources. Further, we calibrate some parameters independently. The calculation of standard errors of the estimated parameters is done within the subset of parameters that are identified from moments for which we have estimates of the full covariance matrix. In other words, the standard errors are conditional on the remaining parameters being calibrated at their current value.

The subset of parameters for which we do not calculate the standard errors is: the elasticity of substitution between IT and labor and productivity of IT.

H.3 Preparation of Data on IT Usage

Using O*NET version 22 we calculate the employment weighted average PC Importance by occupation group as a measure of IT usage. The importance scale in O*NET starts at 1, so we subtract one from the raw measure before calculating the employment weighted average. Then we normalize the measure by its sum over the four occupation groups.

The data on the overall IT usage in the economy are constructed following Eden and Gaggl (2018).

H.4 The Elasticity of Substitution between IT and Labor

We target the elasticity of substitution between IT and labor in vom Lehn (2020). This measure was calibrated using time-series variation. Here we show the necessary derivations for calculating this elasticity of substitution in vom Lehn (2020) and our model.

H.4.1 The elasticity of substitution between IT and labor in vom Lehn (2020)

The representative firm production function is:

$$Y_t = \left[\mu_m N_{mt}^{\frac{\gamma_m-1}{\gamma_m}} + (1-\mu_m)\left[\mu_a N_{at}^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a)\left[(1-\mu_r)K_t^{\frac{\gamma_r-1}{\gamma_r}} + \mu_r N_{rt}^{\frac{\gamma_r-1}{\gamma_r}}\right]^{\frac{\gamma_r(\gamma_a-1)}{(\gamma_r-1)\gamma_a}}\right]^{\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}}\right]^{\frac{\gamma_m}{\gamma_m-1}} \tag{A.100}$$

Then, the firm's problem is given by:

$$\max_{K_t,N_{mt},N_{rt},N_{at}} \pi_t = Y_t - r_t K_t - w_{mt}N_{mt} - w_{rt}N_{rt} - w_{at}N_{at} \tag{A.101}$$

Then, the first order conditions are:

$$(N_{mt}): \quad \left(\frac{\gamma_m}{\gamma_m-1}\right)[\cdot]^{\frac{\gamma_m}{\gamma_m-1}-1} \times \mu_m N_{mt}^{\frac{\gamma_m-1}{\gamma_m}-1} \times \left(\frac{\gamma_m-1}{\gamma_m}\right) - w_{mt} = 0$$

Notice that: $\frac{\gamma_m}{\gamma_m-1} - 1 = \frac{\gamma_m-\gamma_m+1}{\gamma_m-1} = \frac{1}{\gamma_m-1}$ and $\frac{\gamma_m}{\gamma_m-1} \times \frac{1}{\gamma_m} = \frac{1}{\gamma_m-1}$. Therefore:

$$[\cdot]^{\frac{\gamma_m}{\gamma_m-1}-1} = Y_t^{\frac{1}{\gamma_m}}$$

while $\frac{\gamma_m-1}{\gamma_m} - 1 = \frac{\gamma_m-1-\gamma_m}{\gamma_m} = -\frac{1}{\gamma_m}$. Therefore, we have:

$$(N_{mt}): \quad w_{mt} = \mu_m \left(\frac{Y_t}{N_{mt}}\right)^{\frac{1}{\gamma_m}} \tag{FOC$_m$}$$

Before continuing, to simplify notation, define:

$$R_t = \left[(1-\mu_r)K_t^{\frac{\gamma_r-1}{\gamma_r}} + \mu_r N_{rt}^{\frac{\gamma_r-1}{\gamma_r}}\right]^{\frac{\gamma_r}{\gamma_r-1}}$$

and

$$\Omega_t = \left[\mu_a N_{at}^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a)R_t^{\frac{\gamma_a-1}{\gamma_a}}\right]^{\frac{\gamma_a}{\gamma_a-1}}$$

Then the F.O.C. w.r.t. N_{rt} becomes:

$$(N_{rt}): \quad \left\{ \begin{array}{l} \left(\frac{\gamma_m}{\gamma_m-1}\right)[\cdot]^{\frac{\gamma_m}{\gamma_m-1}-1} \times (1-\mu_m)\left(\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}\right)\left[\mu_a N_{at}^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a)R_t^{\frac{\gamma_a-1}{\gamma_a}}\right]^{\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}-1} \\ \times(1-\mu_a)\left(\frac{\gamma_r(\gamma_a-1)}{(\gamma_r-1)\gamma_a}\right)\left[(1-\mu_r)K_t^{\frac{\gamma_r-1}{\gamma_r}} + \mu_r N_{rt}^{\frac{\gamma_r-1}{\gamma_r}}\right]^{\frac{\gamma_r(\gamma_a-1)}{(\gamma_r-1)\gamma_a}-1} \mu_r\left(\frac{\gamma_r-1}{\gamma_r}\right)N_{rt}^{\frac{\gamma_r-1}{\gamma_r}-1} \end{array} \right\} - w_{rt} = 0$$

Simplifying it, we have:

$$(N_{rt}): \quad \left\{ \begin{array}{l} Y_t^{\frac{1}{\gamma_m}} \times (1-\mu_m)\left[\mu_a N_{at}^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a)R_t^{\frac{\gamma_a-1}{\gamma_a}}\right]^{\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}-1} \\ \times(1-\mu_a)\left[(1-\mu_r)K_t^{\frac{\gamma_r-1}{\gamma_r}} + \mu_r N_{rt}^{\frac{\gamma_r-1}{\gamma_r}}\right]^{\frac{\gamma_r(\gamma_a-1)}{(\gamma_r-1)\gamma_a}-1} \mu_r N_{rt}^{\frac{\gamma_r-1}{\gamma_r}-1} \end{array} \right\} - w_{rt} = 0$$

Then, notice that $\frac{\gamma_a(\gamma_m-1)}{\gamma_m(\gamma_a-1)} - 1 = \frac{\gamma_m-\gamma_a}{(\gamma_a-1)\gamma_m}$. In order to simplify the notation, we would like to find an exponent z such that:

$$\frac{\gamma_a}{\gamma_a-1}z = \frac{\gamma_m-\gamma_a}{(\gamma_a-1)\gamma_m} \Rightarrow z = \frac{\gamma_m-\gamma_a}{\gamma_a\gamma_m} = \frac{1}{\gamma_a} - \frac{1}{\gamma_m}$$

Similarly $\frac{\gamma_r(\gamma_a-1)}{(\gamma_r-1)\gamma_a}-1=\frac{\gamma_r\gamma_a-\gamma_r-\gamma_r\gamma_a+\gamma_a}{(\gamma_r-1)\gamma_a}=\frac{\gamma_a-\gamma_r}{(\gamma_r-1)\gamma_a}$. Again, in order to simplify the notation, we would like to find z_1 such that:

$$\frac{\gamma_r}{\gamma_r-1}z_1=\frac{\gamma_a-\gamma_r}{(\gamma_r-1)\gamma_a}\Rightarrow z_1=\frac{1}{\gamma_r}-\frac{1}{\gamma_a}$$

Then, the F.O.C. (N_{rt}) becomes:

$$(N_{rt}):\quad \left\{\begin{array}{c} Y_t^{\frac{1}{\gamma_m}}\times(1-\mu_m)\Omega_t^{\frac{1}{\gamma_a}-\frac{1}{\gamma_m}} \\ \times(1-\mu_a)R_t^{\frac{1}{\gamma_r}-\frac{1}{\gamma_a}}\mu_r N_{rt}^{-\frac{1}{\gamma_r}} \end{array}\right\}-w_{rt}=0$$

Rearranging it:

$$(N_{rt}):\quad w_{rt}=(1-\mu_m)(1-\mu_a)\mu_r\left(\tfrac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\tfrac{\Omega_t}{R_t}\right)^{\frac{1}{\gamma_a}}\left(\tfrac{R_t}{N_{rt}}\right)^{\frac{1}{\gamma_r}} \tag{FOC$_r$}$$

Now, the F.O.C. w.r.t. N_{at} is:

$$(N_{at}):\left\{\begin{array}{c}\left(\frac{\gamma_m}{\gamma_m-1}\right)[\cdot]^{\frac{\gamma_m}{\gamma_m-1}-1}\times(1-\mu_m)\left(\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}\right)\left[\mu_a N_{at}^{\frac{\gamma_a-1}{\gamma_a}}+(1-\mu_a)R_t^{\frac{\gamma_a-1}{\gamma_a}}\right]^{\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}-1} \\ \times\mu_a\left(\frac{\gamma_a-1}{\gamma_a}\right)N_{at}^{\frac{\gamma_a-1}{\gamma_a}-1}\end{array}\right\}-w_{at}=0$$

Simplifying it:

$$(N_{at}):\left\{\ Y_t^{\frac{1}{\gamma_m}}\times(1-\mu_m)\Omega_t^{\frac{1}{\gamma_a}-\frac{1}{\gamma_m}}\times\mu_a N_{at}^{-\frac{1}{\gamma_a}}\ \right\}-w_{at}=0$$

i.e.:

$$(N_{at}):w_{at}=(1-\mu_m)\mu_a\left(\tfrac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\tfrac{\Omega_t}{N_{at}}\right)^{\frac{1}{\gamma_a}} \tag{FOC$_a$}$$

Finally, the F.O.C. w.r.t. K_t is quite similar to the one for N_{rt}. Therefore, the F.O.C. is given by:

$$(K_t):\quad r_t=(1-\mu_m)(1-\mu_a)(1-\mu_r)\left(\tfrac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\tfrac{\Omega_t}{R_t}\right)^{\frac{1}{\gamma_a}}\left(\tfrac{R_t}{K_t}\right)^{\frac{1}{\gamma_r}} \tag{FOC$_K$}$$

Then, putting together all the F.O.C.s, we have:

$$\left\{\begin{array}{lll} w_{mt}= & \mu_m\left(\frac{Y_t}{N_{mt}}\right)^{\frac{1}{\gamma_m}} & (\text{FOC}_m) \\ w_{rt}= & (1-\mu_m)(1-\mu_a)\mu_r\left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\frac{\Omega_t}{R_t}\right)^{\frac{1}{\gamma_a}}\left(\frac{R_t}{N_{rt}}\right)^{\frac{1}{\gamma_r}} & (\text{FOC}_r) \\ w_{at}= & (1-\mu_m)\mu_a\left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\frac{\Omega_t}{N_{at}}\right)^{\frac{1}{\gamma_a}} & (\text{FOC}_a) \\ r_t= & (1-\mu_m)(1-\mu_a)(1-\mu_r)\left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\frac{\Omega_t}{R_t}\right)^{\frac{1}{\gamma_a}}\left(\frac{R_t}{K_t}\right)^{\frac{1}{\gamma_r}} & (\text{FOC}_K) \end{array}\right.$$

where:

$$R_t = \left[(1-\mu_r)K_t^{\frac{\gamma_r-1}{\gamma_r}} + \mu_r N_{rt}^{\frac{\gamma_r-1}{\gamma_r}}\right]^{\frac{\gamma_r}{\gamma_r-1}}$$

and

$$\Omega_t = \left[\mu_a N_{at}^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a)R_t^{\frac{\gamma_a-1}{\gamma_a}}\right]^{\frac{\gamma_a}{\gamma_a-1}}$$

Before we continue, let's define the share of labor demand for occupation j as:

$$s_{jt} \equiv \frac{N_{jt}}{N_{at}+N_{rt}+N_{mt}}$$

Then, dividing the top and the bottom by N_{jt}, s_{jt} can be rewritten as:

$$s_{jt} \equiv \frac{1}{\frac{N_{at}}{N_{jt}}+\frac{N_{rt}}{N_{jt}}+\frac{N_{mt}}{N_{jt}}}$$

Proposition 5 *The elasticities of the share of workers demanded in each occupation with respect to the rental rate of capital, $\frac{\partial s_{jt}}{\partial r_t}\frac{r_t}{s_{jt}}$ are given by:*

$$\begin{aligned}
\frac{\frac{\partial s_{at}}{\partial r_t}}{r_t} &= \xi_{kt}[s_{rt}(\gamma_a-\gamma_r)+\xi_{rt}s_{mt}(\gamma_a-\gamma_m)] \\
\frac{\frac{\partial s_{rt}}{\partial r_t}}{r_t} &= \xi_{kt}[(1-s_{rt})(\gamma_r-\gamma_a)+\xi_{rt}s_{mt}(\gamma_a-\gamma_m)] \\
\frac{\frac{\partial s_{mt}}{\partial r_t}}{r_t} &= \xi_{kt}[s_{rt}(\gamma_a-\gamma_r)+\xi_{rt}(1-s_{mt})(\gamma_m-\gamma_a)]
\end{aligned}$$

where ξ_{kt} is the share of all routine income paid to capital, i.e.:

$$\xi_{kt} = \frac{r_tK_t}{r_tK_t+w_{rt}N_{rt}}$$

and ξ_{rt} is the share of income paid to routine tasks in the CES nest combining abstract and routine tasks, i.e.:

$$\xi_{kt} = \frac{r_tK_t+w_{rt}N_{rt}}{r_tK_t+w_{rt}N_{rt}+w_{at}N_{at}}$$

Given these conditions, a decline in the rental rate of capital will generate an increase in the share of labor demand from abstract jobs if $\gamma_r-\gamma_a > \frac{\xi_{rt}s_{mt}}{s_{rt}}(\gamma_a-\gamma_m)$ and a decrease in the share of labor demand from routine jobs if $\gamma_r-\gamma_a > \frac{\xi_{rt}s_{mt}}{1-s_{rt}}(\gamma_m-\gamma_a)$.

Proof. First of all, from the FOCs (FOC_a) and (FOC_r), we have:

$$\frac{w_{at}}{w_{rt}} = \frac{(1-\mu_m)\mu_a\left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\frac{\Omega_t}{N_{at}}\right)^{\frac{1}{\gamma_a}}}{(1-\mu_m)(1-\mu_a)\mu_r\left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\frac{\Omega_t}{R_t}\right)^{\frac{1}{\gamma_a}}\left(\frac{R_t}{N_{rt}}\right)^{\frac{1}{\gamma_r}}}$$

Simplifying it, we have:

$$\frac{w_{at}}{w_{rt}} = \left(\frac{\mu_a}{1-\mu_a}\right)\mu_r^{-1}\left(\frac{R_t}{N_{at}}\right)^{\frac{1}{\gamma_a}}\left(\frac{N_{rt}}{R_t}\right)^{\frac{1}{\gamma_r}} \tag{A.102}$$

while FOCs (FOC_r) and (FOC_K) give us:

$$\frac{w_{rt}}{r_t} = \frac{(1-\mu_m)(1-\mu_a)\mu_r\left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\frac{\Omega_t}{R_t}\right)^{\frac{1}{\gamma_a}}\left(\frac{R_t}{N_{rt}}\right)^{\frac{1}{\gamma_r}}}{(1-\mu_m)(1-\mu_a)(1-\mu_r)\left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}}\left(\frac{\Omega_t}{R_t}\right)^{\frac{1}{\gamma_a}}\left(\frac{R_t}{K_t}\right)^{\frac{1}{\gamma_r}}}$$

Simplifying

$$\frac{w_{rt}}{r_t} = \left(\frac{\mu_r}{1-\mu_r}\right)\left(\frac{K_t}{N_{rt}}\right)^{\frac{1}{\gamma_r}}$$

Rearranging:

$$K_t = \left(\frac{w_{rt}}{r_t}\right)^{\gamma_r}\left(\frac{1-\mu_r}{\mu_r}\right)^{\gamma_r} N_{rt}$$

Substituting it back into the expression for R_t, we have:

$$R_t = \left\{\left[(1-\mu_r)^{\gamma_r}(r_t)^{1-\gamma_r} + \mu_r^{\gamma_r}(w_{rt})^{1-\gamma_r}\right]\right\}^{\frac{\gamma_r}{\gamma_r-1}}\left(\frac{w_{rt}}{\mu_r}\right)^{\gamma_r} N_{rt}$$

Rearranging:

$$\frac{N_{rt}}{R_t} = \left[(1-\mu_r)^{\gamma_r}(r_t)^{1-\gamma_r} + \mu_r^{\gamma_r}(w_{rt})^{1-\gamma_r}\right]^{\frac{\gamma_r}{1-\gamma_r}}\left(\frac{\mu_r}{w_{rt}}\right)^{\gamma_r}$$

Then, define:

$$C_{R,t} = \left[(1-\mu_r)^{\gamma_r}(r_t)^{1-\gamma_r} + \mu_r^{\gamma_r}(w_{rt})^{1-\gamma_r}\right]^{\frac{1}{1-\gamma_r}}$$

Substituting it back, we have:

$$\left(\frac{N_{rt}}{R_t}\right)^{\frac{1}{\gamma_r}} = \left(\frac{\mu_r C_{R,t}}{w_{rt}}\right) \tag{A.103}$$

Similarly, we have that:

$$\left(\frac{R_t}{N_{rt}}\right)^{\frac{1}{\gamma_a}} = \left(\frac{w_{rt}}{\mu_r C_{R,t}}\right)^{\frac{\gamma_r}{\gamma_a}}$$

rearranging:

$$\left(\frac{R_t}{N_{at}}\right)^{\frac{1}{\gamma_a}} = \left(\frac{w_{rt}}{\mu_r C_{R,t}}\right)^{\frac{\gamma_r}{\gamma_a}} \times \left(\frac{N_{rt}}{N_{at}}\right)^{\frac{1}{\gamma_a}}$$

Substituting it back into equation (A.102), we have:

$$\frac{w_{at}}{w_{rt}} = \left(\frac{\mu_a}{1-\mu_a}\right)\mu_r^{-1}\left(\frac{w_{rt}}{\mu_r C_{R,t}}\right)^{\frac{\gamma_r}{\gamma_a}}\left(\frac{\mu_r C_{R,t}}{w_{rt}}\right)\left(\frac{N_{rt}}{N_{at}}\right)^{\frac{1}{\gamma_a}} \tag{A.104}$$

simplifying it:

$$\left(\frac{N_{at}}{N_{rt}}\right)^{\frac{1}{\gamma_a}} = \left(\frac{\mu_a}{1-\mu_a}\right)^{\gamma_a} \mu_r^{-\gamma_r} \left(\frac{w_{rt}}{C_{R,t}}\right)^{\gamma_r} \left(\frac{C_{R,t}}{w_{at}}\right)^{\gamma_a} \tag{A.105}$$

Similarly, from (FOC$_a$)/(FOC$_m$), we have:

$$\frac{w_{at}}{w_{mt}} = \frac{(1-\mu_m)\mu_a \left(\frac{Y_t}{\Omega_t}\right)^{\frac{1}{\gamma_m}} \left(\frac{\Omega_t}{N_{at}}\right)^{\frac{1}{\gamma_a}}}{\mu_m \left(\frac{Y_t}{N_{mt}}\right)^{\frac{1}{\gamma_m}}}$$

Rearranging it:

$$\frac{w_{at}}{w_{mt}} = \left(\frac{1-\mu_m}{\mu_m}\right) \mu_a \frac{\left(\frac{\Omega_t}{N_{at}}\right)^{\frac{1}{\gamma_a}}}{\left(\frac{\Omega_t}{N_{mt}}\right)^{\frac{1}{\gamma_m}}}$$

$$\Omega_t = \left[\mu_a + (1-\mu_a)\left(\frac{R_t}{N_{at}}\right)^{\frac{\gamma_a - 1}{\gamma_a}}\right]^{\frac{\gamma_a}{\gamma_a - 1}} N_{at}$$

Since from equations (A.102) and (A.103), we have:

$$\begin{aligned} \left(\frac{R_t}{N_{at}}\right)^{\frac{\gamma_a-1}{\gamma_a}} &= \left(\frac{1-\mu_a}{\mu_a}\right)^{\gamma_a - 1} \left(\frac{w_{at}}{C_{Rt}}\right)^{\gamma_a - 1} \\ &= \left(\frac{1-\mu_a}{\mu_a}\right)^{\gamma_a - 1} \left(\frac{C_{Rt}}{w_{at}}\right)^{1-\gamma_a} \end{aligned}$$

Substituting it back, we have:

$$\Omega_t = \left[\mu_a + (1-\mu_a)\left(\frac{1-\mu_a}{\mu_a}\right)^{\gamma_a - 1} \left(\frac{C_{Rt}}{w_{at}}\right)^{1-\gamma_a}\right]^{\frac{\gamma_a}{\gamma_a - 1}} N_{at}$$

Rearranging it:

$$\Omega_t = \left[\frac{\mu_a^{\gamma_a} w_{at}^{1-\gamma_a} + (1-\mu_a)^{\gamma_a} C_{Rt}^{1-\gamma_a}}{\mu_a^{\gamma_a - 1} w_{at}^{1-\gamma_a}}\right]^{\frac{\gamma_a}{\gamma_a - 1}} N_{at}$$

i.e.:

$$\frac{\Omega_t}{N_{at}} = \left[\mu_a^{\gamma_a} w_{at}^{1-\gamma_a} + (1-\mu_a)^{\gamma_a} C_{Rt}^{1-\gamma_a}\right]^{\frac{\gamma_a}{\gamma_a - 1}} \left(\frac{w_{at}}{\mu_a}\right)^{\gamma_a}$$

Define:

$$C_{\Omega,t} \equiv \left[\mu_a^{\gamma_a} w_{at}^{1-\gamma_a} + (1-\mu_a)^{\gamma_a} C_{Rt}^{1-\gamma_a}\right]^{\frac{\gamma_a}{\gamma_a - 1}}$$

Then:

$$\frac{\Omega_t}{N_{at}} = C_{\Omega,t}\left(\frac{w_{at}}{\mu_a}\right)^{\gamma_a}$$

and

$$\frac{\Omega_t}{N_{mt}} = C_{\Omega,t}\left(\frac{w_{at}}{\mu_a}\right)^{\gamma_a} \frac{N_{at}}{N_{mt}}$$

Substituting it back into the expression for $\frac{w_{at}}{w_{mt}}$, we have:

$$\frac{w_{at}}{w_{mt}} = \left(\frac{1-\mu_m}{\mu_m}\right)\mu_a \frac{\left(C_{\Omega,t}\left(\frac{w_{at}}{\mu_a}\right)^{\gamma_a}\right)^{\frac{1}{\gamma_a}}}{\left(C_{\Omega,t}\left(\frac{w_{at}}{\mu_a}\right)^{\gamma_a}\frac{N_{at}}{N_{mt}}\right)^{\frac{1}{\gamma_m}}}$$

■

H.4.2 Elasticities in our model

The labor demand and capital demand in our model are given by (16) and (17). For simplicity, call the solution to (16) and (17): $m_{ij}(\mathbf{w},\mathbf{r})$ and $k_{ij}(\mathbf{w},\mathbf{r})$. Given labor and capital demand we calculate the elasticity of the share of labor employed in an occupation numerically. To target the elasticities given in vom Lehn (2020) we sum together routine manual and routine cognitive jobs when calculating the elasticity.

References

Acemoglu, Daron and David H. Autor (2011). "Skills, tasks and technologies: Implications for employment and earnings." *Handbook of labor economics*, 4, pp. 1043–1171. doi:10.1016/S0169-7218(11)02410-5.

Acemoglu, Daron and Pascual Restrepo (2020). "Robots and jobs: Evidence from US labor markets." *Journal of Political Economy*, 128(6), pp. 2188–2244. doi:10.1086/705716.

Adachi, Daisuke (2021). "Robots and wage polarization: The effects of robot capital across occupations." Mimeo.

Albouy, David (2012). "Are big cities bad places to live? Estimating quality of life across metropolitan areas." Working Paper 14472, National Bureau of Economic Research. doi:10.3386/w14472.

Autor, David H. (2019). "Work of the past, work of the future." *AEA Papers and Proceedings*, 109, pp. 1–32. doi:10.1257/pandp.20191110.

Autor, David H. and David Dorn (2013). "The growth of low-skill service jobs and the polarization of the US labor market." *The American Economic Review*, 103(5), pp. 1553–1597. doi:10.1257/aer.103.5.1553.

Autor, David H., David Dorn, and Gordon H. Hanson (2016). "The China shock: Learning from labor-market adjustment to large changes in trade." *Annual Review of Economics*, 8, pp. 205–240. doi:10.1146/annurev-economics-080315-015041.

Barnatchez, Keith, Leland Dod Crane, and Ryan Decker (2017). "An assessment of the National Establishment Time Series (NETS) database." Finance and Economics Discussion Series 2017-110, Board of Governors of the Federal Reserve System. doi:10.17016/FEDS.2017.110.

Baum-Snow, Nathaniel, Matthew Freedman, and Ronni Pavan (2018). "Why has urban inequality increased?" *American Economic Journal: Applied Economics*, 10(4), pp. 1–42. doi:10.1257/app.20160510.

Baum-Snow, Nathaniel and Ronni Pavan (2013). "Inequality and city size." *Review of Economics and Statistics*, 95(5), pp. 1535–1548. doi:10.1162/REST_a_00328.

Beaudry, Paul, Mark Doms, and Ethan Lewis (2010). "Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas." *Journal of Political Economy*, 118(5), pp. 988–1036. doi:10.1086/658371.

Berger, David W., Kyle F. Herkenhoff, and Simon Mongey (2019). "Labor market power." Working Paper 25719, National Bureau of Economic Research. doi:10.3386/w25719.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen (2014). “The distinct effects of information technology and communication technology on firm organization.” *Management Science*, 60(12), pp. 2859–2885. doi:10.1287/mnsc.2014.2013.

Caunedo, Julieta, David Jaume, and Elisa Keller (2021). “Occupational exposure to capital-embodied technical change.” Discussion Paper 15759, CEPR. URL `https://ideas.repec.org/p/cpr/ceprdp/15759.html`.

Cortes, Guido Matias, Nir Jaimovich, Christopher J. Nekarda, and Henry E. Siu (2014). “The micro and macro of disappearing routine jobs: A flows approach.” Working Paper 20307, National Bureau of Economic Research. doi:10.3386/w20307.

Cortes, Guido Matias, Nir Jaimovich, and Henry E. Siu (2017). “Disappearing routine jobs: Who, how, and why?” *Journal of Monetary Economics*, 91, pp. 69–87. doi:10.1016/j.jmoneco.2017.09.006.

Crane, Leland D. and Ryan A. Decker (2020). “Research with private sector business microdata: The case of NETS/D&B.” Technical report, Working paper.

Davis, Donald R., Eric Mengus, and Tomasz K. Michalski (2020). “Labor market polarization and the great divergence: Theory and evidence.” Working Paper 26955, National Bureau of Economic Research. doi:10.3386/w26955.

Doms, Mark and Ethan Lewis (2006). “Labor supply and personal computer adoption.” Working Paper 06-10, Federal Reserve Bank of Philadelphia. URL `https://ideas.repec.org/p/fip/fedpwp/06-10.html`.

Ducruet, César, Réka Juhász, Dávid Krisztián Nagy, and Claudia Steinwender (2019). “All aboard: The aggregate effects of port development.” Working Paper 1718, Universitat Pompeu Fabra. Departament d’Economia i Empresa. URL `http://hdl.handle.net/10230/44690`.

Dvorkin, Maximiliano A. and Alexander Monge-Naranjo (2019). “Occupation mobility, human capital and the aggregate consequences of task-biased innovations.” Working Paper 2019-13, Federal Reserve Bank of St Louis. doi:10.20955/wp.2019.013.

Eaton, Jonathan and Samuel Kortum (2002). “Technology, geography, and trade.” *Econometrica*, 70(5), pp. 1741–1779. doi:10.1111/1468-0262.00352.

Eckert, Fabian, Teresa C. Fort, Peter K. Schott, and Natalie J. Yang (2021). “Imputing missing values in the US Census Bureau’s County Business Patterns.” Working Paper 26632, National Bureau of Economic Research. doi:10.3386/w26632.

Eden, Maya and Paul Gaggl (2018). “On the welfare implications of automation.” *Review of Economic Dynamics*, 29, pp. 15–43. doi:10.1016/j.red.2017.12.003.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny (2014). "Spatial sorting." *Journal of Political Economy*, 122(3), pp. 554–620. doi:10.1086/676141.

Ellison, Glenn and Edward L. Glaeser (1997). "Geographic concentration in US manufacturing industries: a dartboard approach." *Journal of Political Economy*, 105(5), pp. 889–927. doi:10.1086/262098.

Frank, Morgan R., David H. Autor, James E. Bessen, Erik Brynjolfsson, Manuel Cebrian, David J. Deming, Maryann Feldman, Matthew Groh, José Lobo, and Esteban Moro (2019). "Toward understanding the impact of artificial intelligence on labor." *Proceedings of the National Academy of Sciences*, 116(14), pp. 6531–6539. doi:10.1073/pnas.1900949116.

Ganong, Peter and Daniel Shoag (2017). "Why has regional income convergence in the US declined?" *Journal of Urban Economics*, 102, pp. 76–90. doi:10.1016/j.jue.2017.07.002.

Giannone, Elisa (2017). "Skill-biased technical change and regional convergence." In *2017 Meeting Papers*, 190. Society for Economic Dynamics. URL `https://ideas.repec.org/p/red/sed017/190.html`.

Goos, Maarten, Alan Manning, and Anna Salomons (2014). "Explaining job polarization: Routine-biased technological change and offshoring." *American Economic Review*, 104(8), pp. 2509–26. doi:10.1257/aer.104.8.2509.

Gourieroux, Christian, Alain Monfort, and Eric Renault (1993). "Indirect inference." *Journal of Applied Econometrics*, 8(S1), pp. S85–S118. doi:10.1002/jae.3950080507.

Graetz, Georg and Guy Michaels (2018). "Robots at work." *Review of Economics and Statistics*, 100(5), pp. 753–768. doi:10.1162/rest_a_00754.

Gyourko, Joseph, Albert Saiz, and Anita Summers (2008). "A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies*, 45(3), pp. 693–729. doi:10.1177/0042098007087341.

Kennan, John and James R. Walker (2011). "The effect of expected income on individual migration decisions." *Econometrica*, 79(1), pp. 211–251. doi:10.3982/ECTA4657.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante (2000). "Capital-skill complementarity and inequality: A macroeconomic analysis." *Econometrica*, 68(5), pp. 1029–1053. doi:10.1111/1468-0262.00150.

Lazear, Edward P. and Kathryn L. Shaw (2009). *The Structure of Wages: An International Comparison*, chapter Wage Structure, Raises and Mobility: International Comparison of the Structure of Wages Within and between Firms, pp. 1–57. The University of Chicago Press. URL `http://www.nber.org/chapters/c2365`.

Lee, Sang Yoon Tim and Yongseok Shin (2017). "Horizontal and vertical polarization: Task-specific technological change in a multi-sector economy." Technical report, National Bureau of Economic Research. doi:10.3386/w23283.

Lucas, Robert E. and Esteban Rossi-Hansberg (2002). "On the internal structure of cities." *Econometrica*, 70(4), pp. 1445–1476. doi:10.1111/1468-0262.00338.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg (2018). "Commuting, migration, and local employment elasticities." *American Economic Review*, 108(12), pp. 3855–90. doi:10.1257/aer.20151507.

Oyer, Paul and Scott Schaefer (2016). "Firm/employee matching: An industry study of US lawyers." *ILR Review*, 69(2), pp. 378–404. doi:10.1177/0019793915605506.

Roback, Jennifer (1982). "Wages, rents, and the quality of life." *Journal of Political Economy*, 90(6), pp. 1257–1278. doi:10.1086/261120.

Rosen, Harvey S. (1979). "Housing decisions and the US income tax: An econometric analysis." *Journal of Public Economics*, 11(1), pp. 1–23. doi:10.1016/0047-2727(79)90042-2.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman (2019). "Cognitive hubs and spatial redistribution." Working Paper 26267, National Bureau of Economic Research. doi:10.3386/w26267.

Roy, Andrew Donald (1951). "Some thoughts on the distribution of earnings." *Oxford Economic Papers*, 3(2), pp. 135–146. doi:10.1093/oxfordjournals.oep.a041827.

Rubinton, Hannah (2020). "The geography of business dynamism and skill biased technical change." *FRB St Louis Working Paper*, (2020-020). doi:10.20955/wp.2020.020.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, and Erin Meyer (2020). "IPUMS USA: Version 10.0 [data set]." doi:https://doi.org/10.18128/D010.V10.0.

Saiz, Albert (2010). "The geographic determinants of housing supply." *The Quarterly Journal of Economics*, 125(3), pp. 1253–1296. doi:10.1162/qjec.2010.125.3.1253.

Santamarıa, Clara (2018). "Small teams in big cities: Inequality, city size, and the organization of production." URL `https://www.cemfi.es/ftp/pdf/papers/Seminar/Santamaria_JMP.pdf`, mimeo.

vom Lehn, Christian (2020). "Labor market polarization, the decline of routine work, and technological change: A quantitative analysis." *Journal of Monetary Economics*, 110, pp. 62–80. doi:10.1016/j.jmoneco.2019.01.004.