

Spatial Polarization*

Fabio Cerina[†] Elisa Dienesch[‡] Alessio Moro[§] Michelle Rendall[¶]

July 2019

Abstract

We study the allocation of skills across space and time in the U.S. Employment polarization is stronger in larger cities and mainly driven by number of workers. Using a theory-based measure of skills obtained within a spatial general equilibrium model, we investigate how the skill distribution changes across space and time. Consistent with employment polarization by city size, during the 1980-2008 larger cities display a higher increase in the fraction of high- and low-skilled workers, while smaller cities display a higher increase in the fraction of medium-skilled. These patterns are largely determined by the emergence of skill-biased technological change.

JEL Classification: J21, O14, R12, R23.

Keywords: Employment Polarization, Spatial Sorting, City Sizes.

*We thank Nicola Fuchs-Schündeln, Nir Jaimovich, Joe Kaboski, Jordy Meeks, Eric Mengus, Kurt Schmidheiny and seminar participants at Universitat Autònoma de Barcelona, Universitat de Barcelona, Universitat Rovira y Virgili, University of Pau, the Workshop of the Australasian Macroeconomics Society (WAMS) in Queenstown, the VII Workshop on Structural Transformation and Macroeconomic Dynamics in Cagliari, the European Meeting of the Urban Economics Association in Amsterdam and the 18th Journées Louis-André Gerard-Varet in Aix-en-Provence for the useful comments and discussions. We also thank Sara Pau for excellent research assistance. The usual disclaimers apply.

[†]University of Cagliari and CRENoS. E-mail: fcerina@unica.it.

[‡]Sciences Po Aix. E-mail: elisa.dienesch@sciencespo-aix.fr.

[§]University of Cagliari. E-mail: amoro@unica.it.

[¶]Monash University. E-mail: michelle.rendall@monash.edu.

1 Introduction

In this paper we study the allocation of skills across space and time in the U.S. both at the occupational and the individual level. On the one hand, the broad literature on employment polarization documents how employment levels change over time along the skill distribution. On the other hand, the rising inequality of urban places (relative to rural ones) has been extensively documented by the literature on spatial sorting of skills. In this paper we provide a comprehensive analysis of the evolution of the skill distribution in the U.S. showing that employment polarization and rising inequality in large cities are strictly connected phenomena.

Our first step is to document employment polarization at the city level. We report two main results. First, employment polarization in the 1980-2008 period is more pronounced in larger cities with respect to smaller ones. This means that in larger cities, the increase of employment shares of the high- and low-skilled relative to the middle-skilled occurs at a faster pace than in smaller cities. Second, we find that this difference, as well as aggregate employment polarization, is largely accounted for by a change in the number of workers at each point of the distribution, rather than by a change in hours worked by each person. These findings suggest that over time larger cities attract both a larger fraction of high- and low-skilled workers relative to medium-skilled, and so that their skill distribution becomes more dispersed with respect to that of smaller cities.

To study the relationship between employment polarization and spatial sorting of skills, we build a spatial general equilibrium model. Similarly to the setup in [Eeckhout et al. \(2014\)](#), there is a multi-location environment in which agents with heterogeneous skills decide where to locate to maximize utility. In doing so agents consider both the wage they receive and the price of housing in the specific location. In addition, agents consume a tradable good that is produced in all locations, and by its nature follows the law of one price at the economy level. Utility equalization by skill type determines the allocation of workers across locations. We extend this setting by introducing a home/market labor time decision, and a multisector environment, in which each agent consumes, in addition to housing and the tradable good, services produced at home and services produced in the market, which are imperfect substitutes. Also, market services are assumed to be locally produced and non-tradable across locations.

We use the model for two purposes. First, we construct a model based measure of skills that can be used together with data on wages and prices to construct empirical skill distributions for different groups of cities in different years.¹ This measure only requires a

¹We follow the same approach as [Eeckhout et al. \(2014\)](#).

subset of model parameters to be computed, and allows us to construct the skill distribution without taking a stand on the type of technological change that is occurring in market sectors in the model. Second, we calibrate the model and use it to run counterfactual exercises to assess the role of technological change in generating differential patterns of employment polarization and overtime changes in spatial sorting at the city level.

We construct the model based measure of skills for the years 1980 and 2008 finding that, consistently with the empirical evidence on employment polarization, the skill distribution is similar between small and large cities in 1980, while in 2008 large cities display fatter tails with respect to smaller cities. We also compute the same measure for the 1960 and find that, as in 1980, the skill distribution is similar across city size. The results are robust to different definitions of “large” and “small” cities and suggest that only after 1980 the spatial sorting of workers starts changing, with larger cities attracting proportionally more high- and low-skilled workers with respect to small cities.

The above results suggest that stronger polarization and stronger emergence of fat tails in larger cities are intimately connected. We then use the model to investigate quantitatively the role of technological change in shaping these phenomena. We consider a version of the model with two locations and two equilibria, calibrated to the years 1980 and 2008. In the benchmark calibration, the data counterparts of the two locations correspond to the sets of cities with population below the first and above the third tercile of the distribution of city size. The data counterparts of the skill groups in the model are three occupational groups defined according to U.S. Census occupational classification and ranked according to their mean wage in 1980. We allow for three types of technological change: total factor productivity (TFP) growth in both the tradable and the non-tradable sector, and skill-biased technological change (SBTC) in the tradable sector.

There are two main channels in the model through which technological change can generate the emergence of fatter tails in larger cities (i.e. an increase in both the share of high- and low-skilled and a decline in the share of middle-skilled). The first one is that proposed in [Eeckhout et al. \(2014\)](#) which is driven by the assumption of production complementarities between high- and low-skilled workers in the tradable sector. Due to this mechanism, TFP differentials across cities in the tradable sector generate fatter tails in the city with larger TFP. The second channel is generated through the introduction of the non-tradable sector, which is assumed to employ only low-skilled workers, and is motivated by two main reasons. First, as observed by [Autor and Dorn \(2013\)](#), the increase of employment shares at the lower tail of the skill distribution is driven by rising hours in a single broad category of employment, that of service *occupations*. This category includes jobs like food service workers, security guards, janitors and gardeners, cleaners, home health aides, child care workers, hairdressers

and beauticians, and recreation occupations which display a mean wage and mean educational level significantly lower than average. In 1980, service occupations represent more than 50% of employment in the non-tradable sector compared to only 11% in the rest of the economy. In 2008 the figure is the same for the non-tradable sector while it is 15% for the rest of the economy.²

The second reason motivating the introduction of the non-tradable sector stems from the results in [Cerina et al. \(2017\)](#). These authors argue that a main driver of employment polarization in the U.S. is the reallocation of hours from home production to market work experienced by high-skilled women since the beginning of the 80s, and induced by the emergence of SBTC. This reallocation of labor time is thus responsible for the increase of the employment shares in both the upper tail, through a direct effect, and the lower tail of the skill distribution through an indirect effect. The indirect effect is due to high-skilled women who, after reducing home work because of higher wages in the market, increase the demand for market services which are substitutes to household production (i.e. non-tradable services) and, therefore, foster labor demand for service occupations. The authors show that through the direct and the indirect effects SBTC can account for a substantial part of employment polarization in the U.S. during the period 1980-2008.

The model presented here builds on a similar mechanism to explain the differential patterns of employment polarization across cities. The non-tradable sector generates the emergence of consumption spillovers when an individual experiences a rise in her market wage. After such an increase, home production becomes relatively more expensive, and the individual reacts by increasing the amount of non-tradable services purchased in the market and reducing home production. As market services are produced by means of only low-skilled workers and consumed locally, this effect generates an increase in the share of low-skilled workers in the city in which the rise in wage occurs. When SBTC occurs in a location, it attracts high-skilled workers by raising, *ceteris paribus*, their wages. For this type of workers the opportunity cost of working at home is high, so their demand for services produced in the market is also high. As low-skilled services are non-tradable and produced by low-skilled workers, the model generates a correlation between employment shares of high- and low-skilled workers within the same city, with a decline in the employment shares of middle-skilled workers. We emphasize that this mechanism is consistent with i) our finding according to which the skill distributions of small and large cities started to diverge in 1980,

²We follow the classification in [Moro et al. \(2017\)](#), and include in the non-tradable sector all those services having a counterpart that can be produced within the household and therefore being good substitutes of home production. Our definition of non-tradable sector in the data includes services like laundry, cleaning, taxi, food, child and elderly care, personal services, and represents a significant and increasing fraction of the U.S. employment (8.15% in 1980, 11.39% in 2008).

ii) the robust empirical evidence according to which the skill premium started increasing dramatically after 1980 (see [Acemoglu and Autor \(2011\)](#), for instance) and iii) the faster increase of the skill premium in larger cities (see for instance [Baum-Snow and Pavan \(2013a\)](#), [Baum-Snow et al. \(2018\)](#) and [Davis and Dingel \(2019\)](#))³.

The calibration for the 1980-2008 period suggests that both TFP growth and SBTC are faster in the large city. This result suggests that both types of technological change are relevant for the model to account for the larger emergence of fat tails in the larger city. We then run two counterfactuals to assess the role of each type of technological change. We first set SBTC to be the same across locations between 1980 and 2008 in the benchmark calibration and compute the difference in the change in the share of the three types of workers between the two cities. This difference is reduced by 10% for the low-skilled, 62% for the middle-skilled and 81% for the high-skilled in the counterfactual with respect to the benchmark. By setting the same growth of TFP in tradables in the two cities, instead, the corresponding numbers are 19% for each part of the skill distribution. We stress that the effect of SBTC on the lower tail has to be considered as a lower bound, as we use a conservative value of the elasticity of substitution between home and market services (2.3). While being commonly used in the literature, this value is an estimate of substitutability between home services and the whole basket of market consumption goods. As discussed for instance in [Ngai and Pissarides \(2011\)](#), a higher value of this elasticity is more appropriate in the context of a model distinguishing between services that are substitutable and non-substitutable to home production. With a higher value of this elasticity (4.6), by setting SBTC to be the same in both cities between 1980 and 2008, the difference in the change in the share of the three types of workers between the two cities is reduced by 23% for the low skilled, 66% for the middle skilled and 82% for the high skilled. The corresponding numbers for the TFP channels are 22%, 21% and 20%.

Finally, we propose a quantitative exercise aiming at disentangling the channel driven by the tradable sector and the one driven by the non-tradable one. Since the benchmark calibration supports the existence of production complementarity in the tradable sector, we run a counterfactual where we shut down the latter and we interpret the residual differences in the employment polarization across cities as generated by the channel driven by non-tradables only. By comparing the latter results to those of the benchmark calibration, we find that the difference in the change in the share of the three types of workers between the two cities is reduced by 33% for the low-skilled, 36% for the middle-skilled and 37% for the high-skilled. Hence, the model suggests that the contribution of the channel driven

³In Appendix B we also document that, consistent with the theory in [Cerina et al. \(2017\)](#), the stronger employment polarization in larger cities is mainly accounted for by women.

by production complementarity in the tradable sector, while being quantitatively relevant, is able to explain a smaller fraction of the differential employment polarization patterns between large and small cities with respect to that associated to the non-tradable sector.

The reminder of the paper is as follows. In Section 2 we discuss the background literature and identify the contribution of our paper. In Section 3 we present the evidence on employment polarization by city size; in Section 4 we present the model and in Section 5 we present the model based empirical distributions of skills across space and time; in Section 6 we present the calibration and the quantitative exercises. Finally, Section 7 concludes.

2 Related Work

This paper lies at the intersection between two strands of literature: that on employment polarization and that on spatial allocation of skills. Since the work on commuting zones by [Autor and Dorn \(2013\)](#), there is a growing interest on the spatial dimension of labor market polarization but there are still relatively few papers focusing on it. From an empirical perspective, the geography of employment polarization in the U.S. has been recently studied by [Autor \(2019\)](#), who finds stronger employment polarization in denser areas. Here we extend his results along two dimensions. First, we provide an analysis of employment polarization by city size based on a more disaggregated definition of occupations. This analysis confirms that the disappearance of middle-skill and the rise of high-skill occupations are found to be significantly more pronounced in areas with a larger population.⁴ Second, the classification of low-skilled occupations in our paper is guided by the model’s mechanism as explained above. Thus, our low-skilled jobs only includes service occupations while [Autor \(2019\)](#) also considers in this category transport, laborers and construction workers, which are included in middle-skilled occupations in our classification. With this alternative classification we also find that employment polarization is stronger in larger cities.⁵

⁴The running variable in [Autor \(2019\)](#) is urban density in 1970 while for us it is urban population in 1980. Moreover, his location units are 722 commuting zones in the U.S., while our analysis is based 218 metropolitan areas, which are significantly larger.

⁵[Autor \(2019\)](#) argues that the decline in middle-skilled occupations in urban areas is driven by the fact that in large cities non-college workers move from increasingly disappearing clerical/administrative/manufacturing occupations to rising services low-skilled occupations. Finding direct evidence for this hypothesis requires rich longitudinal data keeping track of the job history of workers’ cohorts and represents an intriguing future research agenda. Our model abstracts from occupational choice as we use occupation groups as invariant proxies for skills. For this reason, the faster increase in low-skilled occupations in large cities are more naturally interpreted as sorting of workers with innate low skills into large cities rather than a degrading of non-college workers into low-skilled occupations. Nevertheless, we emphasize that the role of consumption and production externalities generated through the non-tradable sector in explaining faster employment polarization in larger cities is not in contrast with the occupational downgrading hypothesis proposed by [Autor \(2019\)](#).

The spatial dimension of labor market polarization in U.S. is empirically investigated also in the state-level analysis of [Lindley and Machin \(2014\)](#). Interestingly, they find that between 1980 and 2010 employment polarization has been stronger in states where there was more education sorting and where both college premium and housing/amenities prices increased faster. Such high-polarization states also experienced bigger increases in the numbers of eating and drinking places, apparel stores and hair and beauty salons. This observation, coupled with the finding in [Moretti \(2013\)](#), who reports that house prices are higher and have risen faster in cities where wage inequality has risen by more, is in line with our idea that large cities are becoming increasingly polarized due to a rising concentration of more educated workers who demand more services which are supplied by low-skill labor.

As for the spatial allocation of skills at the individual level, there is a large number of papers either documenting and/or proposing explanations for the increasing inequality in large cities. Among the many of them, we emphasize the ones particularly related to our work.⁶ Using a model-based measure of skill based on real wages similar to ours, the above mentioned paper by [Eeckhout et al. \(2014\)](#) finds that U.S. large cities in 2009 display fat tails of the skill distribution and argue that extreme skill complementarities in production are the main driver of the spatial sorting of both high- and low-skilled workers in more populated areas.⁷ The main differences with respect to their work are the following. First, we add a time dimension by showing that city size did not affect the shape of the skill distributions until 1980; second we connect spatial changes in the skill distribution to spatial changes in occupational structure and employment polarization patterns across cities; finally, we study the role of the non-tradable service sector and quantitatively assess its contribution in generating fat tails and stronger employment polarization in large cities, comparing it to their channel based on production complementarities in the tradable sector.

Focusing more on the spatial rather than on the time dimension, [Roca and Puga \(2016\)](#) use a rich longitudinal administrative database of Spanish workers between 2004 and 2009 to argue that the greater dispersion of skills in large cities is better explained by a static earning premium upon arrival in a bigger city and by the accumulation of more valuable experience for high-skilled workers rather than by spatial sorting. Since these gains are stronger for workers with higher unobserved ability, their combination also explains higher dispersion of earnings. While spatial sorting based on extreme skill complementarities is quite natural *across* occupational groups, the static and dynamic advantages put forward by [Roca and Puga \(2016\)](#) appear as more natural *within* occupational groups.⁸ By documenting that

⁶Other notable examples are [Behrens and Robert-Nicoud \(2014\)](#); [Davis and Dingel \(2014, 2019\)](#); [Behrens et al. \(2014\)](#) and [Diamond \(2016\)](#).

⁷Production spillovers at the local level have been extensively documented by [Moretti \(2010\)](#).

⁸To use [Roca and Puga \(2016\)](#) words, it would be hard to argue that “the top surgeon benefiting par-

the occupational structure of large and small cities changed substantially since 1980, we indirectly support the idea that the *between* occupation component of the larger dispersion in urban earnings is important and, accordingly, that spatial sorting based on production complementarities (and consumption spillovers) plays an important role in explaining it.

Possibly, the paper which is closest to our spirit is that developed by [Davis et al. \(2019\)](#) in contemporaneous research. They build a model based on elements of [Autor and Dorn \(2013\)](#) and [Davis and Dingel \(2019\)](#) which predicts, for larger cities, a faster increase in employment shares for the high-skilled, a faster decrease in employment shares for the middle skilled, and a *slower* increase in employment shares for the low-skilled workers. They document that the empirical evidence for France supports these theoretical predictions. Thus, on the empirical side, the two papers suggest a different behavior of employment shares of low-skilled worker in the U.S. and in France.⁹ As for the theory side, while their mechanism is triggered by a decline in the price of capital/offshoring goods, ours focuses on the role of production and consumption complementarities. In addition, we bring the model to the data in order to provide some quantitative prediction on the comparative role of different kinds of technological changes.

Finally, [Baum-Snow et al. \(2018\)](#) find that the skill bias of agglomeration economies, by boosting the impact of skill-biased technical change in larger cities, can account for most of the increase in urban inequality since 1980. Their finding is consistent with the idea that SBTC has been faster in large cities and so provides a microfoundation for our calibrated model.

3 Employment Polarization and City Size

Employment polarization in the U.S., i.e. the relative disappearance of middle-skill occupations in favor of both high and low-skill ones since the beginning of the 80s is a well documented fact. Based on individual data from 1980 U.S. Census and the 2008 American Community Survey, we start our investigation by providing novel evidence showing that employment polarization is more pronounced in larger cities and so that there is a spatial

ticularly from working with a mediocre surgeon”, while it is quite natural to think that “a top surgeon or a top lawyer in New York City, given the value of her time, benefits greatly from the ease to hire in that city low-skilled services at her job (catering, administrative assistance) and home (child care, schooling and help in the household)” (p. 137). Their argument is consistent with that of [Baum-Snow and Pavan \(2013a\)](#) who find that most of the impact city size has had on the increase in inequality in the U.S. between 1979 and 2007 has come because wages have become more unequal within skill groups in larger cities than in smaller cities.

⁹[Davis et al. \(2019\)](#) use French data for the period 1994-2015, while we focus on U.S 1980-2008. Also, they use a classification of occupations based on [Goos et al. \(2014\)](#), which is different from ours.

Occupation Group	Avg hourly wage 1980	Emp. Share 1980	Change 1980-2008
Services	4.89	11.61%	+3.12%
Admin, Tech, etc.	6.80	62.72%	-11.66%
Prof. and Manag.	9.74	25.68%	+8.54%

Table 1: Employment polarization in the U.S. in the period 1980-2008.

dimension to this phenomenon.¹⁰ We adopt the same classification used in [Autor and Dorn \(2013\)](#) which harmonizes U.S. Census codes overtime and, guided by the model, we divide occupations into three broad skill groups. The group of low-skill occupations is that of Services (codes 405-472) which account for the increase of employment shares at the bottom of the skill distribution at the aggregate level in the U.S. between 1980 and 2008.¹¹ On the other spectrum of the skill distribution, we define as high-skill all Managerial and Professional Specialty Occupations (codes 004-199). All remaining occupations are in the middle-skill group (codes 203-889 except 405-472). Table 1 reports the well known pattern for the whole U.S. economy. The employment shares in terms of hours of occupations at the extremes of the distribution increase over time, while that of occupations in the middle shrink.

To perform the analysis by city size we consider 218 metropolitan statistical areas and we rank them according to their population in 1980. We then consider three different groupings: i) cities above the median city size and those below, ii) cities below the first tercile and above the second tercile of city size and iii) cities below the first quartile and above the third quartile of city size.¹² Figure 1 reports employment polarization for large and small cities for the three groupings. Consider first the median grouping (top-left panel of Figure 1). In this case the increase of employment shares in terms of hours of low-skill occupations and that of high-skill ones between 1980 and 2008 is bigger in large cities than in small ones (3.19 *percentage points* versus 3.03 and 9.15 versus 8.13 respectively). This implies that for middle-skill occupations the decline of employment shares is bigger in large cities (-12.34) than in small ones (-11.16). Importantly, the divergence in employment polarization increases with the difference in city size. To show this we compare the group of cities in the extreme terciles and quartiles of the distribution of city size. The results are reported in the top-right and bottom-left panels of Figure 1. In the case of terciles, there is an increase of employment shares of low-skill occupations of 3.05 percentage points in small cities compared to a 3.72 in large ones. For middle-skill occupations the figures are -10.90 for small cities and -13.39 for large cities and for high-skill occupations 7.85 of small cities versus 9.66 of large ones. In the case of quartiles, the increase of employment shares of low-skill occupations is 3.07 in small

¹⁰Data and sample description can be found in Appendix A.

¹¹See [Autor and Dorn \(2013\)](#).

¹²A detailed description of city size definition is provided in Appendix A.

cities and 4.24 in large ones. In middle-skill occupations small cities display a -10.26 versus a -13.37 of large ones and for high-skill occupations there is a 7.19 of small cities versus a 9.13 of large ones. The bottom-right panel of Figure 1 displays the difference in difference between the two cities (i.e. the difference between the two groups of cities in the change of employment shares of each group of occupations). This panel highlights that differences in employment polarization between small and large cities increase with differences in size.¹³

¹³Cerina et al. (2017) show that aggregate employment polarization in the U.S. is largely generated by the sharp rise in the education premium after 1980, which fostered *directly* women's participation at the top and, *indirectly*, at the bottom of the skill distribution, due to a larger demand for low-skilled services by skilled women. While the analysis in that work is performed at the aggregate level, as low-skilled services are produced and consumed locally, the mechanism should emerge at the level of metropolitan areas and it should be more evident in larger cities, where the education premium rose faster since 1980. In Appendix B we provide empirical evidence showing that this is the case. The spatial difference in employment polarization across city size is mainly driven by women, whose changes in the employment shares of the three occupational groups (positive for the high- and low-skill occupations and negative for the mid-skill ones) are substantially more pronounced in larger cities.

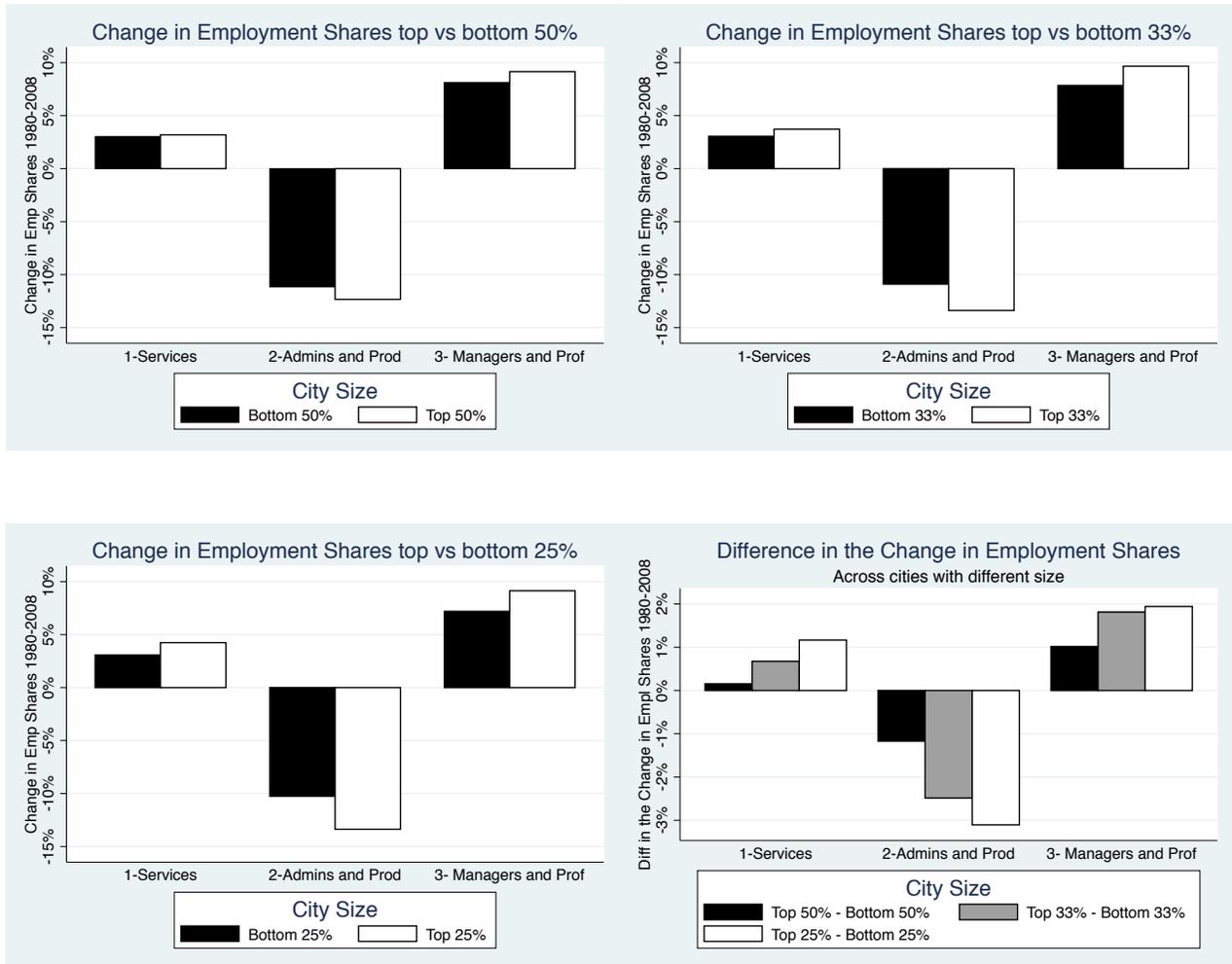


Figure 1: Employment polarization by city size. The upper left panel compares metropolitan areas with population above and below the median in 1980, the upper right panel metropolitan areas with population above the 2nd tercile and below the 1st tercile in 1980, the bottom left panel metropolitan areas with population above the 3rd quartile and below the 1st quartile in 1980 while the bottom right panel compares the difference in the change in employment shares across cities with different size.

The results for broad occupation categories confirm the well documented existence of employment polarization at the aggregate level, but suggest a spatial dimension of the phenomenon, which is more pronounced in large cities than in small ones. To provide further evidence on this distinction by city size, we use the same methodology as in [Acemoglu and Autor \(2011\)](#) to produce employment polarization graphs for each group of city (i.e. small and large). More precisely, we compute the average wage in 1980 of each occupation at the three digit level according to the 1990 occupational classification used by [Autor and Dorn \(2013\)](#). Then, we rank these occupations according to their average wage and construct

occupation percentiles. By keeping the same ranking in 2008 we construct, for each of the six groups of cities (largest and smallest according to median, terciles and quartiles), employment polarization graphs by measuring the change in employment share of each 1980 percentile and using a locally weighted smoothing regression. Results appear in Figure 2. As for broad occupation categories, employment polarization is more pronounced in larger cities than in smaller ones.

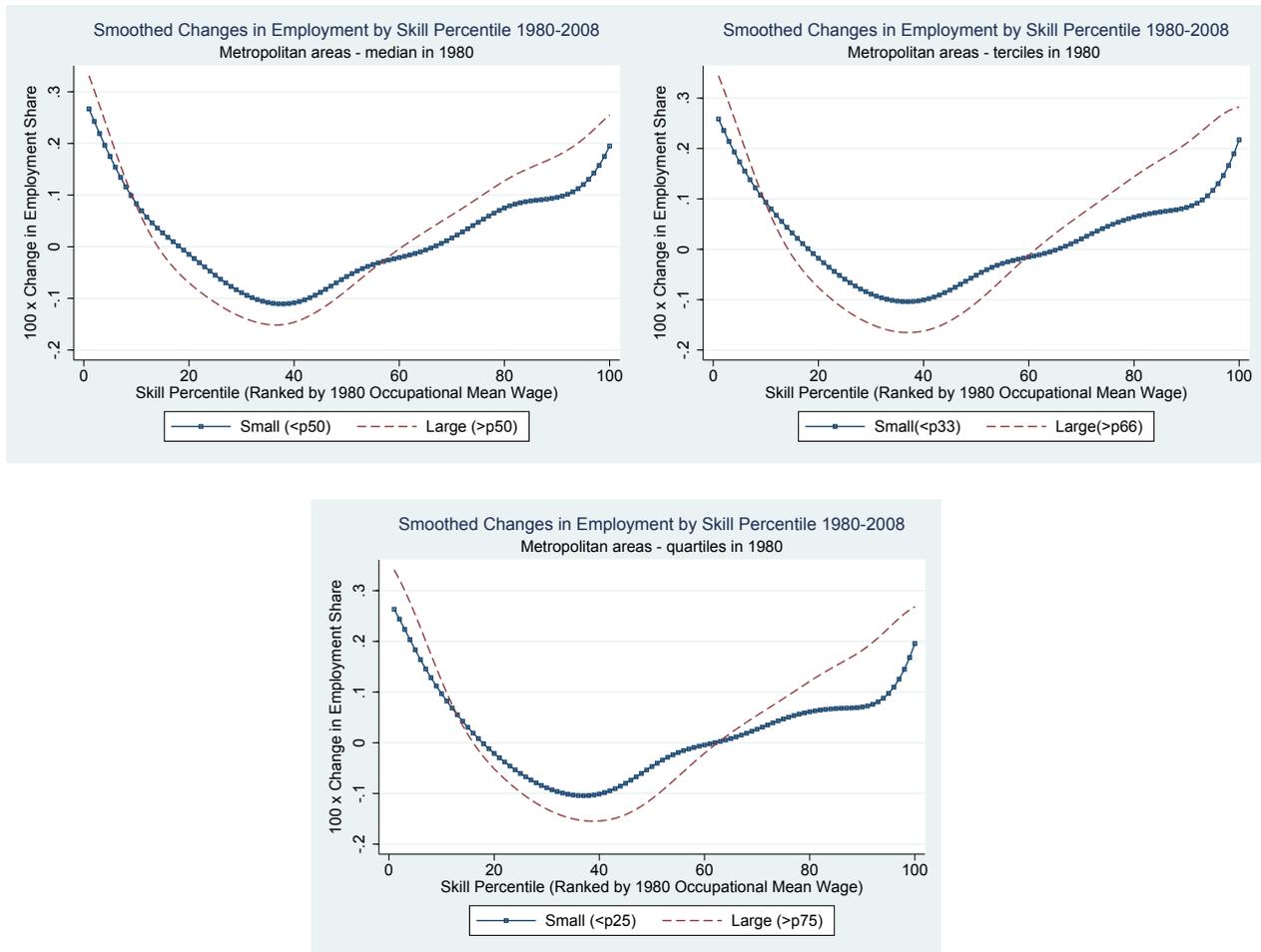


Figure 2: Employment polarization by city size. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel metropolitan areas with population above the 2nd tercile and below the 1st tercile in 1980 and the bottom panel metropolitan areas with population above the 3rd quartile and below the 1st quartile in 1980.

The measures presented above suggest that the employment share in terms of hours of high- and low-skill occupations increases more in large cities than in small ones. An important caveat here is that changes in employment shares include both the intensive and

the extensive margin of employment. To relate employment polarization to spatial sorting, which typically focuses on shares of individuals, we are interested in understanding to what extent the former phenomenon is driven by the extensive and the intensive margin. To do this, we modify the graphs in Figure 2 to consider only the change in the number of workers along the skill distribution, rather than the change in hours. Putting it differently, we reconstruct Figure 2 by assuming that there is no change in hours worked between 1980 and 2008 in any of the occupations used to construct Figure 2. Formally, we retain the same percentiles classification as in Figure 2 and measure, for each percentile, the change in the share of *workers* from 1980 to 2008.¹⁴ The results are reported in Figure 3 and show that the U-shape is driven by a change in the number of workers along the skill distribution. In addition, this measure confirms that large cities are more polarized than small ones, suggesting that the observed employment polarization is driven by a larger increase in the proportion of high- and low-skilled *individuals* in large cities than in small cities.

In addition to the distinction between individuals and hours, we also remark here that the employment polarization is a *dynamic* concept commonly based on the assumption that mean occupation wage in the initial year is a good proxy of the skills possessed by workers in that occupation. Spatial sorting, instead, is a *static* concept focusing on how the skill distribution is shaped across space, and the unit of analysis is usually a measure of the skill level of individual workers. With this distinction in mind, and as long as mean occupation wage in the initial year is a good proxy for the skills of workers performing that occupation, the results in this section suggest that between 1980 and 2008 we should also observe: i) more similar distributions of skills across city size in 1980 with respect to 2008; and ii) the tails of the distribution become fatter in all cities, but the larger the city, the more pronounced the phenomenon. To investigate whether these predictions hold in the data, a theory of spatial sorting is needed. In fact, it is not possible to directly infer the skill distribution by city from the wage distributions, as there are a number of factors that contribute to determine the utility (and, therefore, the mobility) of the individual in addition to the wage. Some of these factors are the housing market and the cost of local non-tradable goods. For this reason, in the next section we present a spatial general equilibrium model that explicitly considers these factors and allows us to investigate the empirical skill distributions across space and time. In addition, the model provides a laboratory to account for the role of different kinds of technological progress (both skill-biased and unbiased) in generating differences in overtime patterns of employment polarization across cities.

¹⁴Thus 1980 percentiles used to construct Figure 2 can be considered as bins of occupations that are kept constant over time. Using these bins we construct Figure 3.

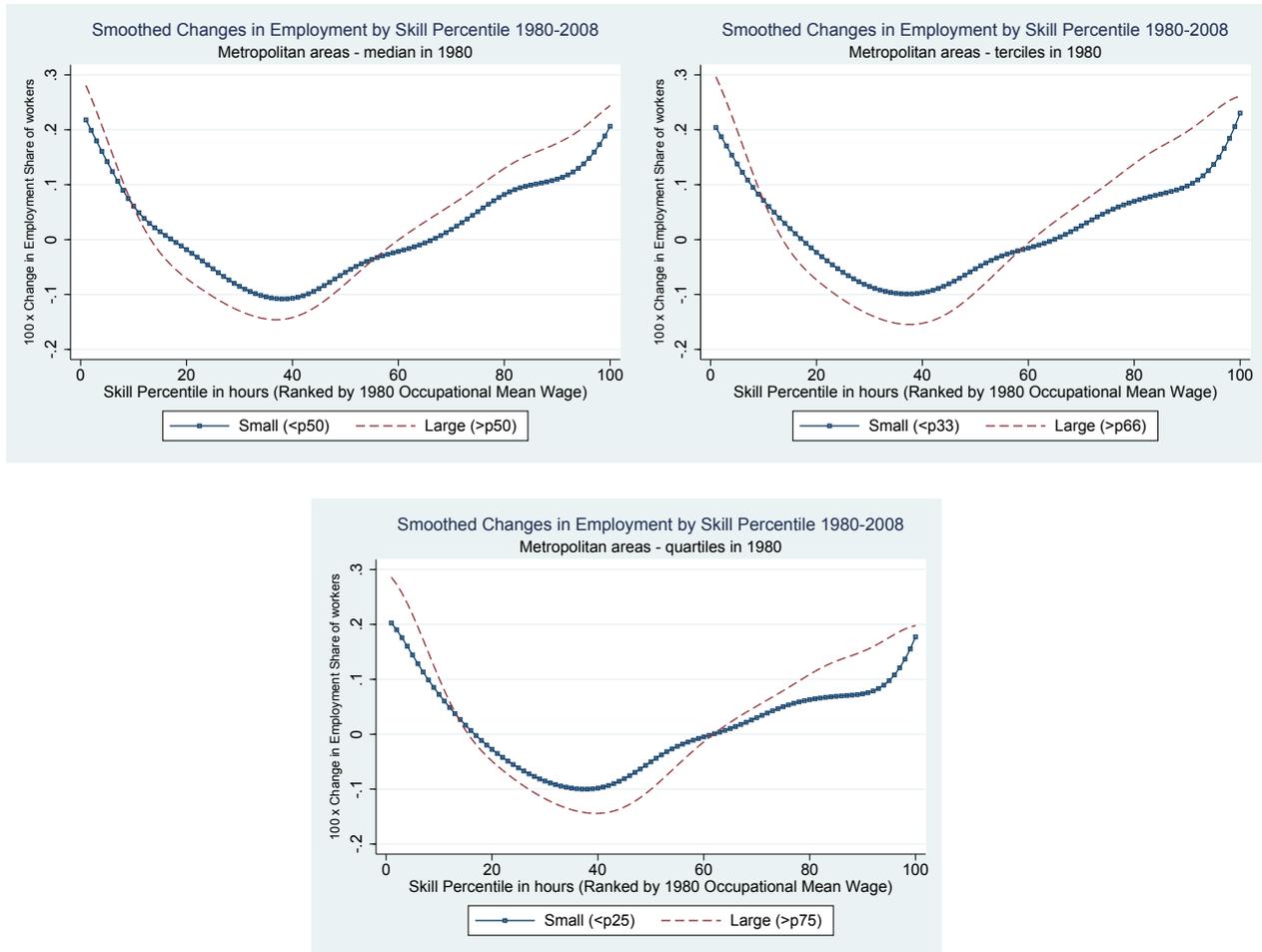


Figure 3: Employment polarization by city size in terms of workers. The ranking of occupations and the bins of occupations are the same as in Figure 2. The variable on the vertical axis is the change in the share of workers in each bin. The left panel compares metropolitan areas with population above and below the median in 1980 the right panel metropolitan areas with population above the 2nd tercile and below the 1st tercile in 1980 and the bottom panel metropolitan areas with population above the 3rd quartile and below the 1st quartile in 1980.

4 Theoretical Framework

In this section we develop a general equilibrium model to jointly study the time patterns of spatial sorting of workers with heterogeneous skills and employment polarization. Workers make a location decision based on their skill level, the wage rate paid to their skill type in each location and the cost of living, which differs across location because housing and non-tradable services have different prices. In equilibrium, the utility of two workers with

the same skill level but living in two different cities is equalized. The distributions of skills across locations and time are determined by the state of technology that we allow to vary in space and time due to total factor productivity growth and skill-biased technological change. The model builds on elements from the spatial setting in [Eeckhout et al. \(2014\)](#) and the multi-sector environment with a home production sector in [Cerina et al. \(2017\)](#).

4.1 The Environment

The economy consists of K locations (cities) indexed by $k \in (1, 2, \dots, K)$. In each location there is a fixed amount of housing H^k whose unit-price is location-specific and defined by p_H^k . As in [Eeckhout et al. \(2014\)](#) the expenditure on housing is the flow value that compensates for the depreciation and interest on capital. In a competitive rental market, the flow payment equals the rental price. To highlight the main mechanisms at work we restrict the number of cities to $K = 2$ but the model can be generalized to any number of cities.

Both cities are populated by workers with heterogeneous skills indexed by $i \in (1, 2, \dots, I)$ and associated with this skill order is a level of productivity a^{ik} . We focus on the case of three skills, $i = h, m, l$. At the economy wide level, there is a fixed amount of workers for each skill N^i for $i = h, m, l$.

There are two market sectors producing goods $j = g, s$. The first, g , broadly interpreted as manufacturing and modern services, is tradable across location while the second, s , interpreted as traditional services, is non-tradable and can only be consumed in the same location where it is produced. Also, there exists a non-marketable service h which is produced within the household and interpreted as home production.

By n_j^{ik} we define the number of workers of skill i working in sector $j = g, s$ in location k . Hence $S_k = \sum_i n^{ik} = \sum_i \sum_j n_j^{ik}$ is the population size of city k . Workers of each skill move towards the city where their utility is higher so that the size of city k is an endogenous equilibrium outcome pinned down by the equalization of utilities across cities for the same skill. Total population of the economy is then exogenously given by $S = \sum_k S^k = \sum_k \sum_i n^{ik}$.

4.2 Demand

Citizens of skill type i who live in city k have preferences over consumption of the tradable good c_g^{ik} , the amount of housing H^{ik} and consumption of services c_n^{ik} . We assume the latter is a CES bundle of home services c_h and market services c_s , which are assumed to be imperfect substitutes with elasticity of substitution equal to $\gamma > 1$.¹⁵ More precisely, a worker of skill

¹⁵See [Rogerson \(2008\)](#) and [Ngai and Pissarides \(2011\)](#)

i living in city k has the following preferences

$$\begin{aligned} U^{ik} &= (H^{ik})^\alpha (c_g^{ik})^\omega (c_n^{ik})^{1-\omega-\alpha} \\ c_n^{ik} &= \left(\psi (c_s^{ik})^{\frac{\gamma-1}{\gamma}} + (1-\psi) (c_h^{ik})^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}}, \end{aligned} \quad (1)$$

where c_j , with $j = g, n, s, h$, represents consumption of goods, services, market services and home services, respectively. We impose $\alpha + \omega < 1$ and $\psi \in (0, 1)$.

Home services are produced within the household according to the technology

$$c_h^{ik} = A_h l^{ik}, \quad (2)$$

where $l^{ik} \in (0, 1)$ is the fraction of time an agent of skill i in city k devotes to work at home, thus being $1 - l^{ik}$ the fraction of time dedicated to work in the firm. We assume that home productivity is invariant across skills and locations. The budget constraint for workers of ability i living in city k is

$$p_g c_g^{ik} + p_s^k c_s^{ik} + p_H^k H^{ik} = w^{ik}(1 - l^{ik}), \quad (3)$$

where p_s^k and p_H^k are, respectively, the price of market services and housing, which are both location-specific and, therefore, indexed by k . Instead, the price of the tradable good, p_g , is the same in the whole economy. In what follows, we choose good g as the numeraire and, therefore, we set $p_g = 1$. We also assume workers are perfectly mobile across sectors so that, in a given location and for a given skill i , the wage rate is equal across sectors and therefore $w_g^{ik} = w_s^{ik} = w^{ik}$ holds. Workers of skill i living in city k solve the following problem

$$\begin{aligned} \max_{c_g^{ik}, c_s^{ik}, c_h^{ik}, l^{ik}} U^{ik} &= (H^{ik})^\alpha (c^{ik})^\omega \left(\left(\psi (c_s^{ik})^{\frac{\gamma-1}{\gamma}} + (1-\psi) (c_h^{ik})^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}} \right)^{1-\omega-\alpha} \\ s.t. \quad &: c_g^{ik} + p_s^k c_s^{ik} + p_H^k H^{ik} = w^{ik}(1 - l^{ik}) \\ &: c_h^{ik} = A_h l^{ik}. \end{aligned}$$

From the demand functions it can be shown that labor supply at home is a negative function of $\frac{w^{ik}}{A_h p_s^k}$, which can be interpreted as the relative price between home services and market services. In cities in which wages relative to the price of market services are higher, workers devote less time to home production and increase the demand of market services. This is the channel of consumption spillovers, which contributes to the emergence of fat tails and employment polarization in the model.

4.3 Production

On the production side there are two sectors: the tradable sector, which produces in all cities *goods* that can be traded across locations; and the non-tradable sector which produces *market services* that can only be consumed in the same location where they are produced.

4.3.1 The Tradable Sector

There is a representative firm in each location which employs three kinds of labor, h , m and l . The production function of the representative firm in city k in the g sector is

$$Y_g^k = A_g^k F(e_g^{hk}, e_g^{mk}, e_g^{lk})$$

where e_g^i is the amount of hours worked by workers of skill i . In equilibrium, this amount of time is the product of an intensive margin - the individual labor supply $1 - l^{ik}$, and an extensive margin - the number of workers employed by the firm, n_g^{ik} . Since labor supply is chosen by the individual worker who maximizes utility, the equilibrium number of workers of each skill employed by the firm is pinned-down by the relationship $n_g^{ik} = e_g^{ik} / (1 - l^{ik})$. A_g^k is the location-specific TFP in the tradable sector. For comparison reasons in the quantitative section, we follow [Eeckhout et al. \(2014\)](#) in assuming that the production function of the representative firm has the following functional form:

$$Y_g^k = A_g^k \left[(a^{hk} e_g^{hk})^\eta + (a^l e_g^{lk})^\eta + (a^m e_g^{mk})^\eta \right].$$

We assume $\eta < 1$ so that there are decreasing returns to scale. We also assume that the firm is owned by absentee capitalists, such that the profits of the firm do not enter the budget constraint of the workers. The parameters a^m and a^l are economy wide productivities of middle- and low-skilled workers, respectively, and without loss of generality we normalize $a^l = 1$. In the quantitative exercises in Section 6, we allow both parameters A_g^k and a^{hk} to change over time, potentially at a different pace across cities. We interpret the time changes in a^{hk} as *skill-biased technological change*. Also, as in [Eeckhout et al. \(2014\)](#), we allow $\lambda > 0$ to be potentially different from one. With $\lambda > 1$ there is extreme-skill complementarity and when $\lambda < 1$ there is extreme-skill substitutability.

The representative firm solves the following problem

$$\max_{\{e_g^{hk}, e_g^{mk}, e_g^{lk}\}} \pi^k = Y_g^k - w^{hk} e_g^{hk} - w^{mk} e_g^{mk} - w^{lk} e_g^{lk}$$

where w^{ik} is wage per unit of time worked by a worker of skill i in location k . Note

that, despite workers' perfect spatial mobility, wages are not equalized across cities because workers decide their location according to their utility, which depends both on wages and on local prices of housing and services. Also, note that wages are not indexed by sector because workers are also mobile across sectors and, therefore, wages of the same type of workers are equalized.

4.3.2 The Non-Tradable Service Sector

The representative firm in the non-tradable service sector operates with the following production function

$$Y_s^k = A_s^k e_s^{lk}$$

where A_s^k is the location-specific TFP in the non-tradable sector.

Profit maximization implies equality between prices and marginal costs.

$$p_s^k = \frac{w^{lk}}{A_s^k} \tag{4}$$

The assumption that only low-skilled workers are employed in the services sector is motivated by the fact that in the data the hours share of this type of workers (i.e. individuals employed in service occupations, as defined in Section 3) in this sector (52.44% in 1980 and 51.25% in 2008) is substantially larger than in the overall economy (11.16% in 1980 and 14.73% in 2008)¹⁶. Also, conditional of being employed in a service occupation, the probability of working in the non-tradable sector is substantially larger (36.75% in 1980 and 39.58% in 2008) than the same probability computed for the overall economy (8.15% in 1980 and 11.38% in 2008).

5 Spatial Sorting

In this section we first use the model to construct a utility-based measure of skill for each individual that can be easily mapped to the data to construct empirical skill distributions. In this way, we are able to reproduce skill distributions by city size and over time to assess whether the results obtained are consistent with the evidence on employment polarization. Next, we resort to quantile regressions to formally assess whether there is a city size effect on the distribution of skills across cities.

¹⁶In the quantitative analysis below the list of sectors included in market services is the same as in [Moro et al. \(2017\)](#). See Appendix A for details.

5.1 Skill-Distributions

Workers of each skill choose the location which ensures the highest utility. Using the first order conditions of the household's problem we obtain the indirect utility for a worker of skill i in city k , which is given by

$$U^{ik} = \Omega (p_H^k)^{-\alpha} (w^{ik})^{\alpha+\omega} \left(1 + \left(\frac{\psi}{1-\psi} \right)^\gamma \left(\frac{w^{ik}}{A_h p_s^k} \right)^{\gamma-1} \right)^{\frac{1-\omega-\alpha}{\gamma-1}} \quad (5)$$

and where

$$\Omega = \alpha^\alpha \omega^\omega (1-\omega-\alpha)^{(1-\omega-\alpha)} (1-\psi)^{\frac{\gamma(1-\omega-\alpha)}{\gamma-1}} (A_h)^{(1-\omega-\alpha)}.$$

The assumption of workers mobility ensures that utility of two workers of the same type is the same across locations ($U^{i1} = U^{i2}$). Thus, there is one-to-one mapping between equilibrium utility and skill level for the worker of type i in any city k . As in [Eeckhout et al. \(2014\)](#), we can interpret (5) as the measure of skill implied by the model and use it to construct a model-based distribution of skills in a particular year by using data on p_H^k , p_s^k and w^{ik} .¹⁷ In doing this we depart from the assumption of three skills in the model, and allow for a generic number of them, identified in the empirical distributions by the actual combinations of observables in (5) in the data. The model-based measure of skills (5) only requires a subset of model parameters to be computed, and allows us to construct the skill distribution without taking a stand on the type of technological change that is occurring in market sectors in the model.

Our aim here is to investigate how the spatial sorting of workers with heterogeneous skills changes across time (between 1980 and 2008) and space (large and small cities). To do this, we first present the skill distributions for different city size and year. Second, we run quantile regressions to provide a formal assessment on the change in the shapes of both the skill and the wage distributions across time and space. Details of methodology, parametrization and the data used can be found in Appendix A.

¹⁷Note that if $\alpha + \omega = 1$ our setting coincides with that of [Eeckhout et al. \(2014\)](#), in which there is no home production and no market production of services.

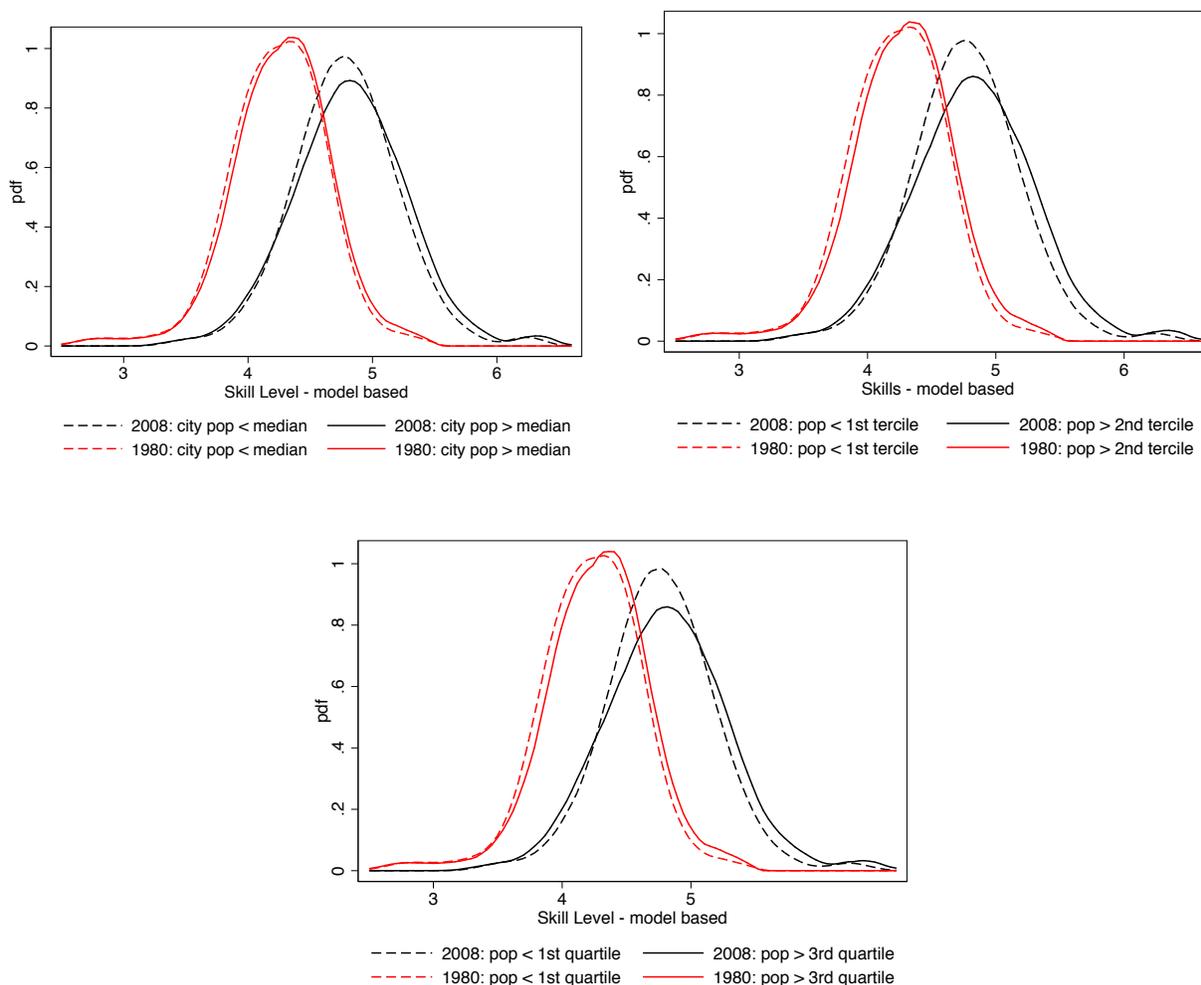


Figure 4: Skill distribution (logarithm of (5)) in 1980 (red) and 2008 (black) in small (dashed lines) and large (thick lines) cities. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel compares metropolitan areas with population above the 2nd and below the 1st tercile in 1980, and the bottom panel compares metropolitan areas with population above the 3rd quartile and below the 1st quartile in 1980.

Figure 4 reports the skill distribution across time and space. The first panel of Figure 4 shows that in 2008 cities with population above the median (black thick line) display fatter tails with respect to cities with population below the median (black dashed line). The middle panel shows that the divergence in the skill distribution between large and small cities is increasing in relative size: the difference in the tails' mass between cities with population above the 2nd tercile and cities with population below the 1st tercile is substantially larger than the same difference computed for the groups of cities with population above and below the median. By considering cities with population above the 3rd quartile and cities with

population below the 1st quartile the divergence in tails is even more pronounced.

Using a similar model-based measure of skills, [Eeckhout et al. \(2014\)](#) find that in 2009 the average and the median worker have the same level of skill in large and small cities but, crucially, the skill distribution in larger cities has fatter tails both at the top and at the bottom of the distribution. Thus, our observation for the year 2008 appear consistent with their results for 2009. However, the evidence in Figure 4 for 1980 is substantially different. In this year (red lines) the skill distributions of large and small cities are remarkably similar and almost overlap. Thus, there is no evidence of fat tails in larger cities, either by comparing cities with population above and below the median, above the second and below the first tercile, and above the third and below the first quartile. If anything, there is a slight first-order stochastic dominance of cities with population above the third quartile over those below the first quartile, and above the second tercile over those below the first tercile, while the skill distribution of cities with size above and below the median are virtually identical.

These results suggest that the emergence of fat tails in the skill distribution of large cities is a phenomenon which emerged in the last decades. This is confirmed by the analysis of the skill distribution in 1960. In Figure 5 we document for 1960 a similar picture as for 1980: the skill distribution is similar in small and large cities.¹⁸ The larger dispersion in 1980 relative to 1960 is a phenomenon which is common to all cities regardless of their size. Thus, the emergence of fat tails which increase with city size should be related to changes in the economic structure that occurred after 1980.¹⁹

¹⁸We report here only the results for the terciles grouping. However, results with the median and quartiles grouping are very similar and available upon request.

¹⁹We use city-level prices for non-tradables from [Carrillo et al. \(2014\)](#) as a measure of p_s^k in constructing the skill distributions of 1980 and 2008, but a similar procedure cannot be applied to the year 1960 due to a lack of data. To overcome this problem we use the first order condition of the model $p_s^k = w^{lk}/A_s^k$, which implies that the price of non-tradables in city k is proportional to the local wages in the non-tradable sector. We then compute the average of the wages of all workers in the non-tradable sector (weighted by hours worked) for each of the $k=218$ metropolitan areas in the sample for the years 1960, 1980 and 2008. As we do not have a measure for A_s^k across cities in 1960, we choose to set $A_{s,1960}^k = 1$ for all cities. While this is an arbitrary choice, we use the same assumption, that is $A_{s,1980}^k = A_{s,2008}^k = 1$ for each city k , to compute the skill distributions for 1980 and 2008 appearing in Figure 5. The figure shows that even in this case the evidence on fat tails across time and city size is similar to the one reported in Figure 4.

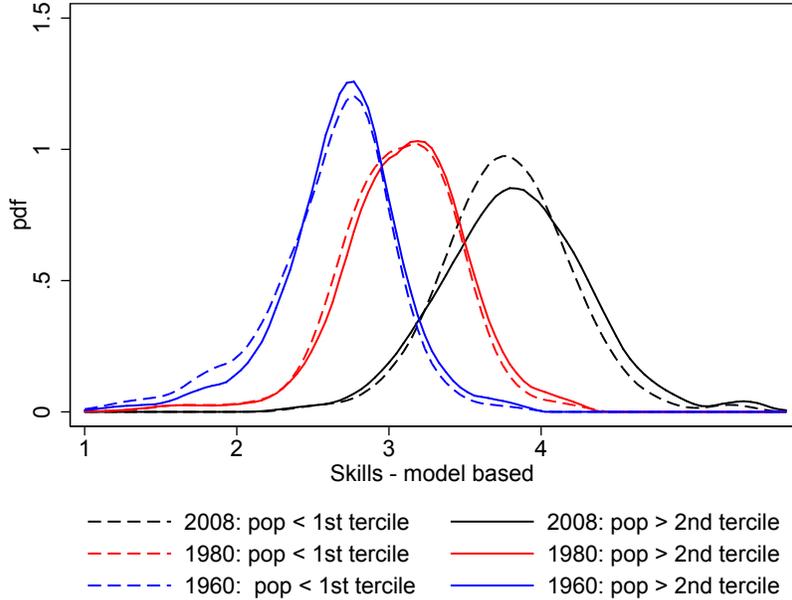


Figure 5: Skill distribution in 1960 (blue), 1980 (red) and 2008 (black) in small (dashed lines) and large (thick lines) cities. The figure compares metropolitan areas with population above the 2nd tercile and below the 1st tercile in 1980.

5.2 Quantile Regressions

To provide a formal quantitative assessment on the dynamics of the wage and the skill distribution in large and small cities we perform a set of quantile regressions. More precisely, we want to analyze how the effect of city size on both wages and skills changes at different points of the distribution. As discussed at the end of Section 3, wage distributions at the city level are a poor measure of skills. Here we show this fact by comparing them with the skill distributions constructed with the model based measure described in Subsection 5.1.

Formally, assuming a linear relation between the individual characteristic x^{ik} (representing either wage w^{ik} or skill U^{ik}), and population (S^k) in location k , we estimate the following specification for each quantile τ :

$$Q_\tau(x^{ik}|S^k) = \beta_0(\tau) + \beta_1(\tau)S^k,$$

where consistent estimators of $\beta_0(\tau)$ and $\beta_1(\tau)$ are obtained by minimizing an asymmetrically weighted sum of absolute errors. We perform this exercise for both the wage and skill distribution in 1980 and 2008.²⁰ Each of these four exercises is represented in a figure with

²⁰In Appendix B we also report the wage distributions across time and space.

two panels: on the left one we plot five quantiles of the distribution (the 10th, the 25th, the median, the 75th and the 90th) against city size, while in the right panel we plot the coefficient of each quantile against its rank. This procedure shows how the effect of city size on the shape of the wage and skill distributions changes from 1980 to 2008.

Wage distribution in 1980. Figure 6 shows that in 1980 the quantiles values increase with city size (i.e. city-size wage premium). Coefficients are all positive and homogeneous along most of the skill distribution, with the only exception of the extreme wage quantiles. This suggests that in 1980 the wage distributions shifts to the right with city size, without a change in its shape.

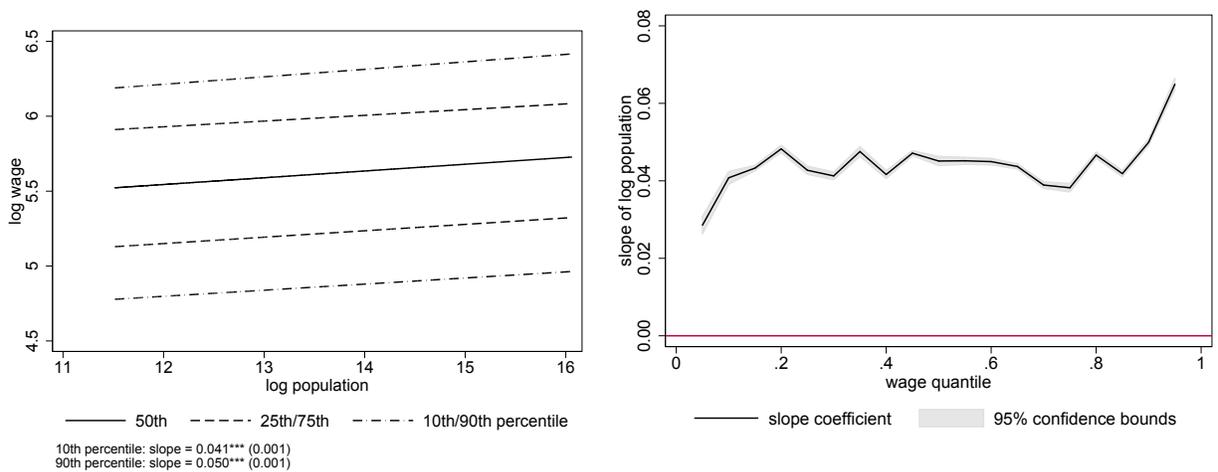


Figure 6: Quantile regression of wage on population in 1980: left, five selected quantiles; right, estimated slope for all quantiles.

Wage distribution in 2008. Figure 7 shows that each quantile of the wage distribution increases with city size (left panel) except the bottom one. The whole distribution shifts to the right (city-size wage premium) so that, like for 1980, coefficients of the relationships between quantiles and city size are positive (right panel). In this case, however, the distribution is also expanding, as coefficients are increasing in quantiles (right panel). This confirms results in [Eeckhout et al. \(2014\)](#), who report similar coefficients for the 2009.

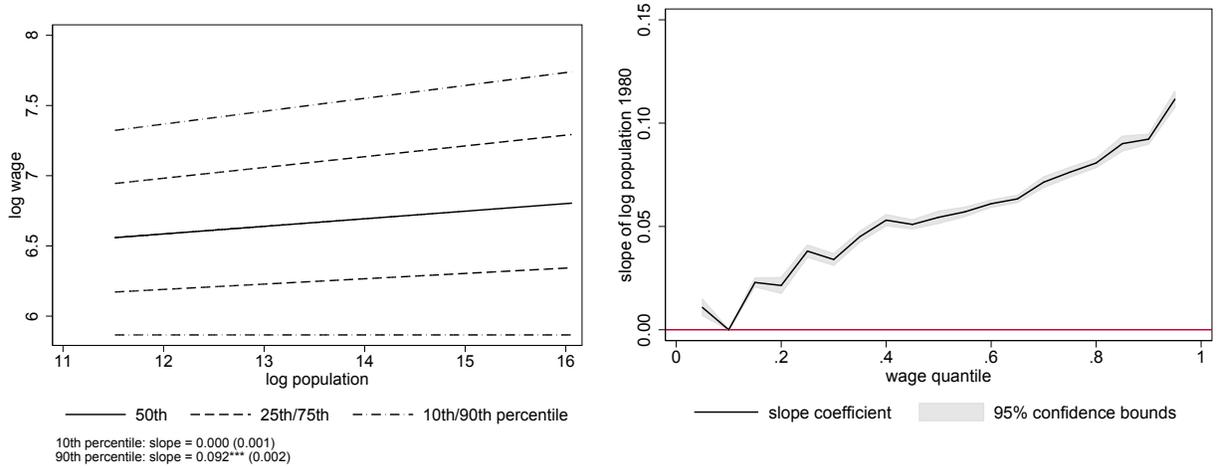


Figure 7: Quantile regression of wage 2008 on population in 1980: left, five selected quantiles; right, estimated slope for all quantiles.

Skill distribution in 1980. Figure 8 reports the result for the 1980 skill distribution, which appears similar to that for the wage distribution. There is no divergence across city size in 1980. Coefficients of the quantile regressions are slightly positive and similar for each quantile (except the very last quantiles). So the quantile regression confirm that in 1980 there is no evidence of fatter tails for larger cities.

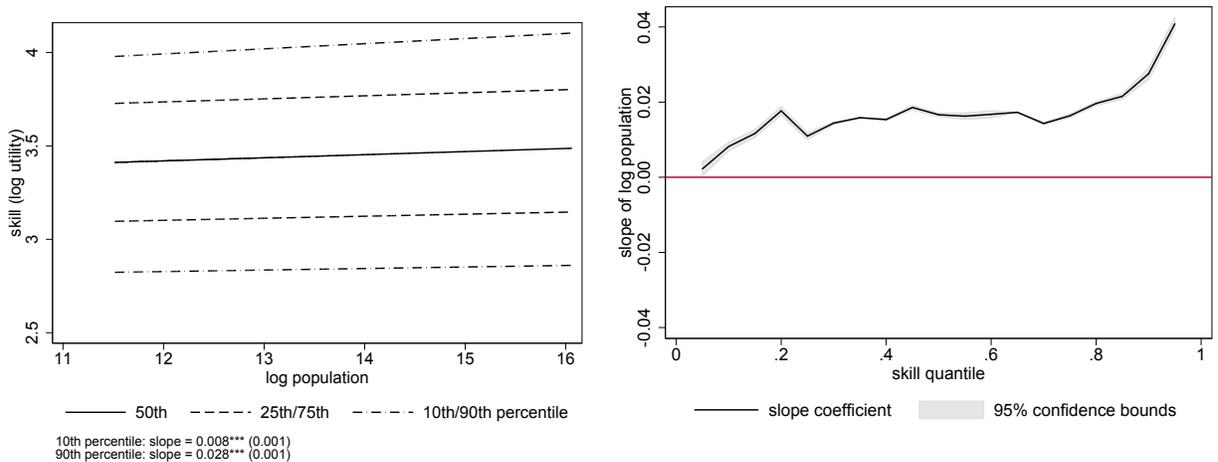


Figure 8: Quantile regression of utility on population in 1980 (i.e. model-based skill measure): left, five selected quantiles; right, estimated slope for all quantiles.

Skill distribution in 2008. Figure 9 reports the results of quantile regressions for the skill distribution in 2008. The right panel shows that slopes are increasing with the quantile rank, being negative up to the 30th percentile and positive otherwise. This confirms the visual result of Figures 4 and 5 for the year 2008: lower quantiles decrease with city size while the opposite happens for higher quantiles (left panel). This represents evidence of fatter tails in the skill distribution for larger cities relative to smaller ones.

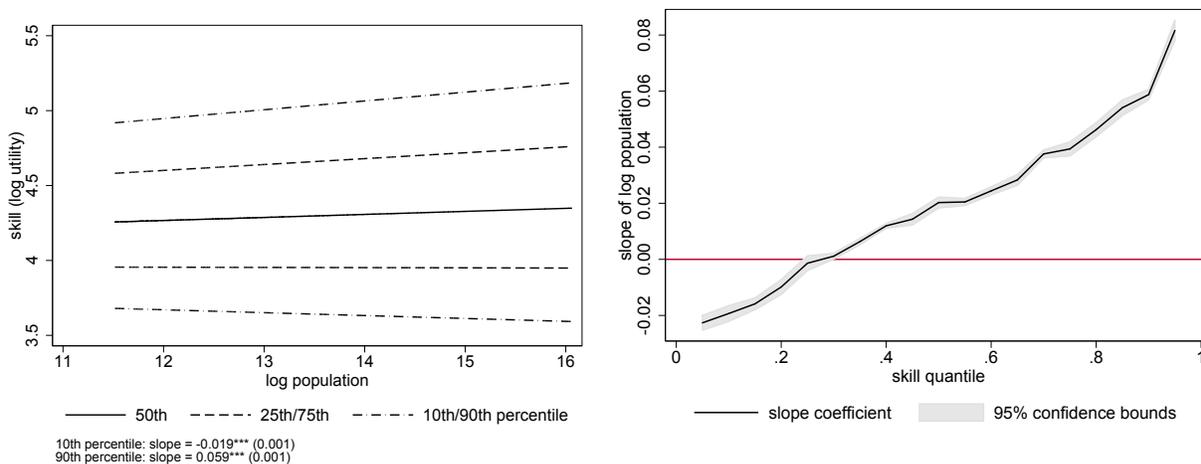


Figure 9: Quantile regression of utility in 2008 (i.e. model-based skill measure) on population in 1980: left, five selected quantiles; right, estimated slope for all quantiles.

6 Quantitative Analysis

The quantile regressions in Section 5.2 document that there is no difference in the shape of the skill distributions across city size in 1980, while in 2008 larger cities display fatter tails with respect to smaller ones. The aim of this section is to use a calibrated version of the model to investigate the role of technological change in shaping the differential evolution of the skill distribution across city size.

In particular, we allow for two types of technological change to potentially generate fatter tails in larger cities. First, we consider the role of skill-biased technological change (SBTC). [Cerina et al. \(2017\)](#) note that the increase in the skill premium coincides with the timing of employment polarization in the U.S. They show that SBTC, a typical driver of the increasing skill premium, can generate employment polarization in a general equilibrium setting through consumption spillovers. SBTC increases the productivity and so the wage of the high-skilled, who work little at home and purchase a substantial amount of market services. In [Cerina et al. \(2017\)](#), this process induces more people to educate, thus increasing the fraction

and the market working time of the high-skilled, who foster consumption spillovers in the economy. In the spatial equilibrium model presented in this paper, faster SBTC in a city relative to another implies that the first city attracts more high skilled workers who, through consumption spillover, attract more low-skilled workers to that location. The differential SBTC channel in the model is motivated by both the timing of the acceleration in the skill premium (the end of the 70's), and the evidence according to which the education premium grew faster in larger cities.²¹ In addition, faster SBTC in larger cities can be interpreted as stronger skill-biased agglomeration economies in those cities, something documented by the empirical literature on urban economics.²²

Second, we investigate the role of TFP growth in the tradable sector. [Eeckhout et al. \(2014\)](#) show that with top-skill complementarities, TFP differences across cities generate fatter tails in the city with larger TFP. Motivated by this result, we allow for a differential evolution of TFP in the two cities coupled with a value of λ different from one.

Note that, while allowing for a different evolution of technology in the two cities, we are not imposing any restriction of the growth of SBTC and TFP across cities, or in the value of λ . Thus, the calibration itself will provide an indication of the role of the two types of technological change in generating fatter tails in larger cities within the model. Next, by using the calibrated model we run counterfactual exercises that help assessing the role of technology in generating fat tails.

6.1 Calibration

The quantitative exercise is set up as a horse-race between different types of technical change in explaining the spatial differences in polarization, i.e., the emergence of fatter tails in bigger cities. We thus calibrate the model such that, given the types of technological change that we allow, it replicates two spatial equilibria at different points in time, intended to replicate the 1980 and the 2008 U.S. economies. In the two equilibria all preference and technology parameters are imposed to be the same except those defining SBTC and TFP in the two market sectors (a^{hk}, A_g^k, A_s^k). Also, the quantile regressions in Section 5.2 suggest that there is no difference in the skill distribution across city size in 1980. Thus, we require that the two cities in the model are symmetric in the 1980 equilibrium. This implies that all technological

²¹See for instance [Davis and Dingel \(2019\)](#); [Baum-Snow and Pavan \(2013b\)](#).

²²For instance [Baum-Snow et al. \(2018\)](#) find that a substantial part of the rise in urban inequality in U.S. cities between 1980 and 2007 is driven by skilled-biased agglomeration economies in a period of rapid skilled-biased technological change. These agglomeration economies create a stronger impact of economy level SBTC in larger cities with respect to smaller ones.

parameter are the same in the two cities in 1980.²³

The parameters, γ , α and ω are set from previous studies based on empirical evidence. Following the discussion in [Ngai and Pissarides \(2008\)](#) and [Moro et al. \(2017\)](#) we set the elasticity of substitution between home production and substitutable services to $\gamma = 2.3$. Note that the magnitude of this parameter is key for the emergence of consumption spillovers, because it determines the degree of substitutability between market services and home production. While we choose the upper bound of the range of estimates in previous literature, this is likely to be a conservative value. The reason is that this parameter is typically estimated by considering substitutability between *total market consumption* and home production. In our model, instead, γ governs substitutability between home production and market services representing *close substitutes* to home production. For this reason, while we use $\gamma = 2.3$ in our benchmark calibration, we also show results for a larger values of this parameter. Next, we obtain the values of α and ω by computing average consumption shares in housing and tradable goods between 1980 and 2008 using NIPA data and rescaling them to take into account that we also have home and market services in the utility function. This procedure gives a value of ω equal to 0.52, and of α equal to 0.13.

The relative supply of skills (i.e. the aggregate skill distribution) in 1980 and 2008 is taken from U.S. Census data. The definition of low-, middle- and high-skilled is the same as in Section 3. Low-skilled workers are those working in service occupations, high-skilled workers those in professional or managerial occupations and middle-skilled workers those in all remaining occupations.²⁴ Hence, following these definitions, we first normalize to one total population in 1980 ($\sum_i N_{1980}^i = 1$), then we compute population growth rates between 1980 and 2008 and derive $\sum_i N_{2008}^i = 1 + g_N$ and then finally we feed the model with the aggregate skills shares of low-, middle- and high skilled in 1980 $\{N_{1980}^i\}_{i=l,m,h}$ and in 2008 $\{N_{2008}^i\}_{i=l,m,h}$. In doing so, we are taking aggregate polarization as given. This is consistent with the aim of our quantitative exercise, which is that of accounting for the *differential patterns in employment polarization* across cities and not that of explaining aggregate employment polarization.²⁵ Lastly, we adopt the following normalizations/restrictions:

- Productivities of low-skilled workers is normalized to one, $a^l = 1$;

²³In particular, we have $a^{h1} = a^{h2}$, $A_g^1 = A_g^2$, $A_s^1 = A_s^2$. All other technology parameters are the same both across cities and over time.

²⁴Following [Autor and Dorn \(2013\)](#) we exclude agriculture and military occupations.

²⁵We stress that one could easily extend the current model by allowing aggregate shares of high-, middle- and low-skilled workers to be endogenized through an education and/or occupational decision, and account for the emergence of aggregate polarization through the same mechanisms at work in our model. For a model in which SBTC can generate employment polarization in a multisectoral environment with education and home/market work decision see [Cerina et al. \(2017\)](#).

Table 2: Calibrated Parameters

Preferences				Technology											
α	ω	γ	ψ	η	λ	a^h	a^m	g_{a^h1}	g_{a^h2}	A_g	$g_{A_g^1}$	$g_{A_g^2}$	A_s	$g_{A_s^1}$	$g_{A_s^2}$
0.13	0.52	2.3	0.16	0.69	1.09	4.62	4.43	39%	43%	1.53	76%	85%	2.19	0%	6%

- The amount of land in each location is normalized to one, $H = 1$;
- Following the evidence in [Bridgman \(2016\)](#) there is no home productivity change between 1980 and 2008, and we normalize it to 1 in both periods, $A_{h,1980} = A_{h,2008} = A_h = 1$;
- We don't allow market TFP to decline in any sector. This is because the calibration could imply negative TFP growth in low-skill services to better match the allocation of low-skilled workers across cities.

The remaining 13 parameters: (1) weight in preferences $\{\psi\}$, (2) productivity parameters $\{a^m, a^h, A_g, A_s\}$ (3) production parameters $\{\eta, \lambda\}$ and (4) technological change $\{g_{a^{hk}}, g_{A_g^k}, g_{A_s^k}\}_{k=1,2}$, where g indicates the growth rate between 1980 and 2008 of the variable at the subscript, are calibrated to match a number of moments: the difference in the change of (hours) employment shares between the two cities for the three types of workers (3 targets); the aggregate wage premiums middle/low and high/low for 1980 and 2008 (4 targets); the relative change in the price of housing between city 2 and city 1 (1 target); the aggregate growth of consumption of tradables and consumption of non-tradables (2 targets); the aggregate consumption share of non-tradables in 2008 (1 target); the aggregate employment share of low-skilled in tradables in 1980 and 2008 (2 targets).

All targets are computed using the 1980 Census and the 2008 American Community Survey unless noted. Table 2 reports the parameter values while table 3 reports the fit of the model.

6.2 Results

Despite its parsimonious structure, the model does a good job at replicating the data targets. In particular, the calibration matches perfectly the difference between the two cities in the change in the shares of the three types of workers between 1980 and 2008 (i.e. the emergence of fatter tails in city 2 relative to city 1). Thus, the values of the calibrated parameters in Table 2 provide an assessment of the role of technology in generating fat tails in the model. First, we note that both SBTC and TFP in tradables grow over time in both cities. This

Table 3: Model's fit

Moment		Data	Model
Diff. in change in emp. shares by cities	Low-skilled	0.68%	0.68%
	Middle-skilled	-2.49%	-2.49%
	High-skilled	1.82%	1.82%
Aggregate wage premiums	Medium/Low 1980	1.39	1.40
	Medium/Low 2008	1.44	1.42
	High/Low 1980	1.99	2.01
	High/Low 2008	2.51	2.35
Change in relative price of housing	$\frac{(p_{h,2008}^2/p_{h,1980}^2)}{(p_{h,2008}^1/p_{h,1980}^1)}$	1.16	1.16
Aggregate growth in consumption	Trad: $\frac{\sum_j \sum_k n_{2008}^{jk} c_{g,2008}^{jk}}{\sum_j \sum_k n_{1980}^{jk} c_{g,1980}^{jk}}$	2.71	2.71
	Non-trad: $\frac{\sum_j \sum_k n_{2008}^{jk} p_{s,2008}^k c_{s,2008}^{jk}}{\sum_j \sum_k n_{1980}^{jk} p_{s,1980}^k c_{s,1980}^{jk}}$	2.10	2.06
Aggr. consumption share non-trad 2008	$\frac{\sum_j \sum_k n_{2008}^{jk} p_{s,2008}^k c_{s,2008}^{jk}}{\sum_j \sum_k n_{2008}^{jk} (p_{s,2008}^k c_{s,2008}^{jk} + p_H^k H_{2008}^{jk} + c_{g,2008}^{jk})}$	9.7%	6.1%
Aggr. empl share low-skilled in trad	1980	7.5%	7.2%
	2008	9.2%	7.6%

suggests that both types of technological change are key for the model to match the data targets. Second, there is faster growth of both SBTC and TFP in tradables in larger cities over time. This suggests that both types of technological change are important to generate fatter tails in larger cities. The result is consistent with the fact that since the start of a rising skill-premium (around 1980), the rise has been faster in larger cities with respect to smaller ones as emphasized by [Baum-Snow and Pavan \(2013a\)](#), [Baum-Snow et al. \(2018\)](#) and [Davis and Dingel \(2019\)](#). Finally, to assess the performance of the model's calibration in terms of empirical validation, we note that the model behaves well in replicating some untargeted moments, as shown in Table 4. The calibration accounts well for the increase in the degree of concentration in large cities and both the direction and the magnitude of the change in the employment shares of the three skill groups at the city level.²⁶

6.3 Counterfactuals

We now describe two counterfactuals to disentangle the effect of SBTC and that of TFP growth in generating fatter tails. In the first one, we assume the same growth of SBTC between 1980 and 2008 in both cities, which is set to the average growth between the two cities in the benchmark calibration. In the second exercise we assume the same TFP growth

²⁶As explained above, we only target the *difference* in the change in the employment shares across cities but not the change within each city. Hence the latter is predicted by the model without any restriction.

Table 4: Untargeted moments

		Data	Model
Change employment shares city 1	Low-skilled	+3.05%	+2.52%
	Middle-skilled	-10.90%	-9.91%
	High-skilled	+7.85%	+7.4%
Change employment shares city 2	Low-skilled	+3.72%	+3.19%
	Middle-skilled	-13.39%	-12.41%
	High-skilled	+9.66%	+9.22%
Population	Relative size 2008	1.19	1.12

in tradables during the same period, again computed as the average across cities in the benchmark calibration. The SBTC counterfactual is displayed in the left panel of Figure 10. For all counterfactuals the orange bars represent the benchmark calibration and the grey one the counterfactual. The difference in the change in the share of the three types of workers between the two cities is reduced by 10% for the low-skilled, 62% for the middle-skilled and 81% for the high-skilled. This suggests that the existence of this type of technological change produces a large fraction of the asymmetry between the two cities. A key point here is that, while SBTC has a direct effect on the productivity of the high-skilled, it has a substantial impact also on the difference in the fraction of middle- and low-skilled across cities.

The middle panel of Figure 10 reports the effect of removing growth of TFP in tradables in the two cities. The numbers for the three types are now 19% for each part of the skill distribution. Thus, with respect to SBTC, removing TFP in tradables have both a smaller and a more homogeneous effect on the difference in the change in the share of the three types of workers between the two cities.

In addition to the above counterfactuals on SBTC and TFP in tradables, the right panel of Figure 10 also reports the effect of setting growth of TFP in non-tradables the same in both cities. In this case, the effect is mostly on low-skilled workers with a reduction of 73% in the difference in the change in the share between the two cities and partly on middle-skilled, with a reduction of 20%. Thus, this type of technological change alone cannot generate a divergence in fat tails across cities.

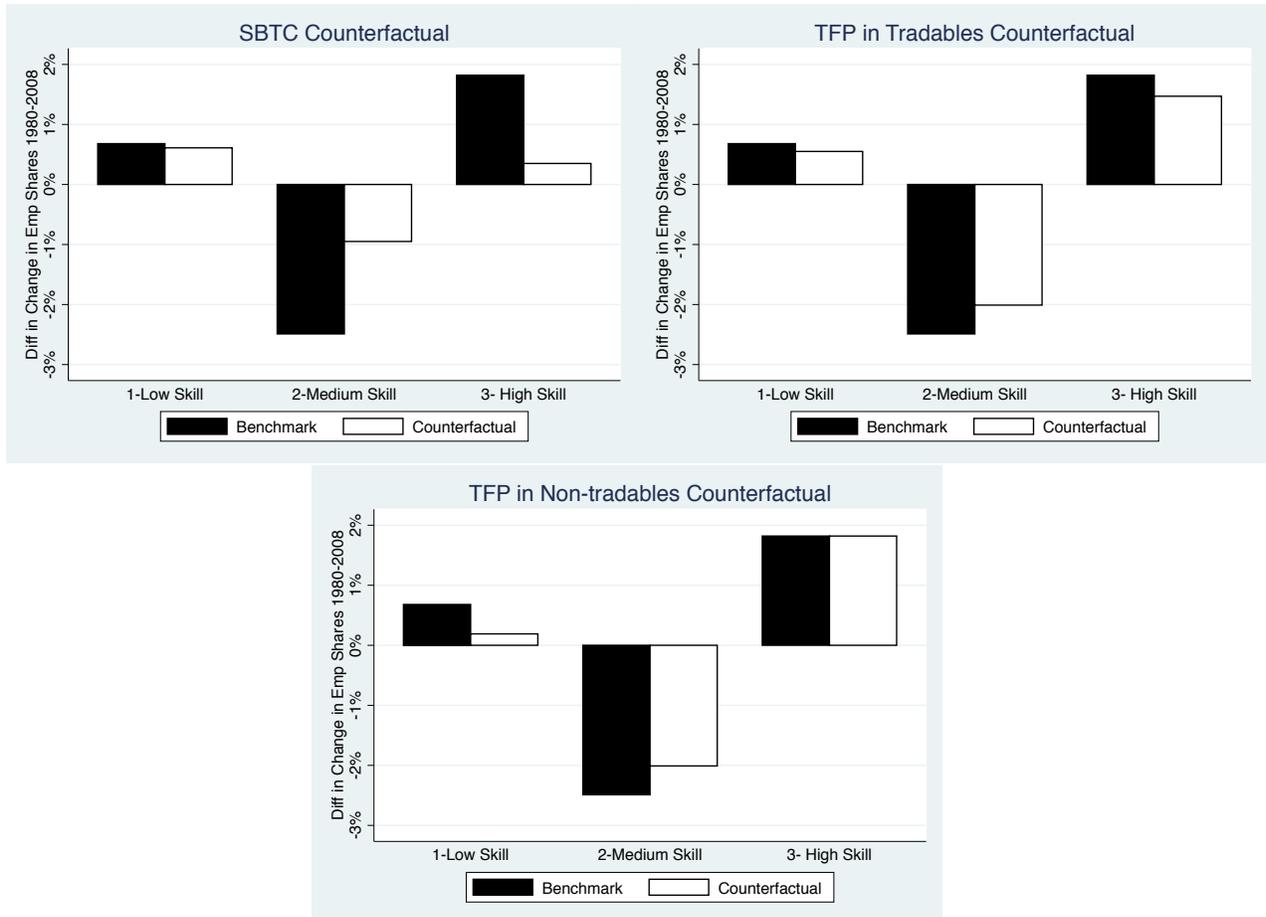


Figure 10: Counterfactual exercises. Black bars represent the benchmark calibration and white bars represent the counterfactual. Left panel: SBTC counterfactual. Middle panel: tradables TFP counterfactual. Right panel: non-tradables TFP counterfactual.

As discussed in the calibration Section 6.1, we set the elasticity of substitution between home production and substitutable services to $\gamma = 2.3$, which is likely to be a conservative value because this parameter is typically estimated by considering substitutability between *total market consumption* and home production. In our model instead, γ governs substitutability between home production and market services representing *close substitutes* to home production. For this reason, we report in Figure 11 the same counterfactual results as in Figure 10 but now imposing a value of $\gamma = 4.6$. With a higher value of γ , the same amount of SBTC induces the consumer to reduce home hours and increase consumption of market substitutes to a larger extent. As a result, the effect of SBTC in generating fatter tails increases. In this case the difference in the change in the share of the three types of workers between the two cities is reduced by 24% (10% in Figure 10) for the low-skilled, 64% (62%) for the middle-skilled and 79% (81%) for the high-skilled. Thus, when the value of γ

is increased to more plausible values, equalizing SBTC across cities has a more substantial effect in reducing the low tail.

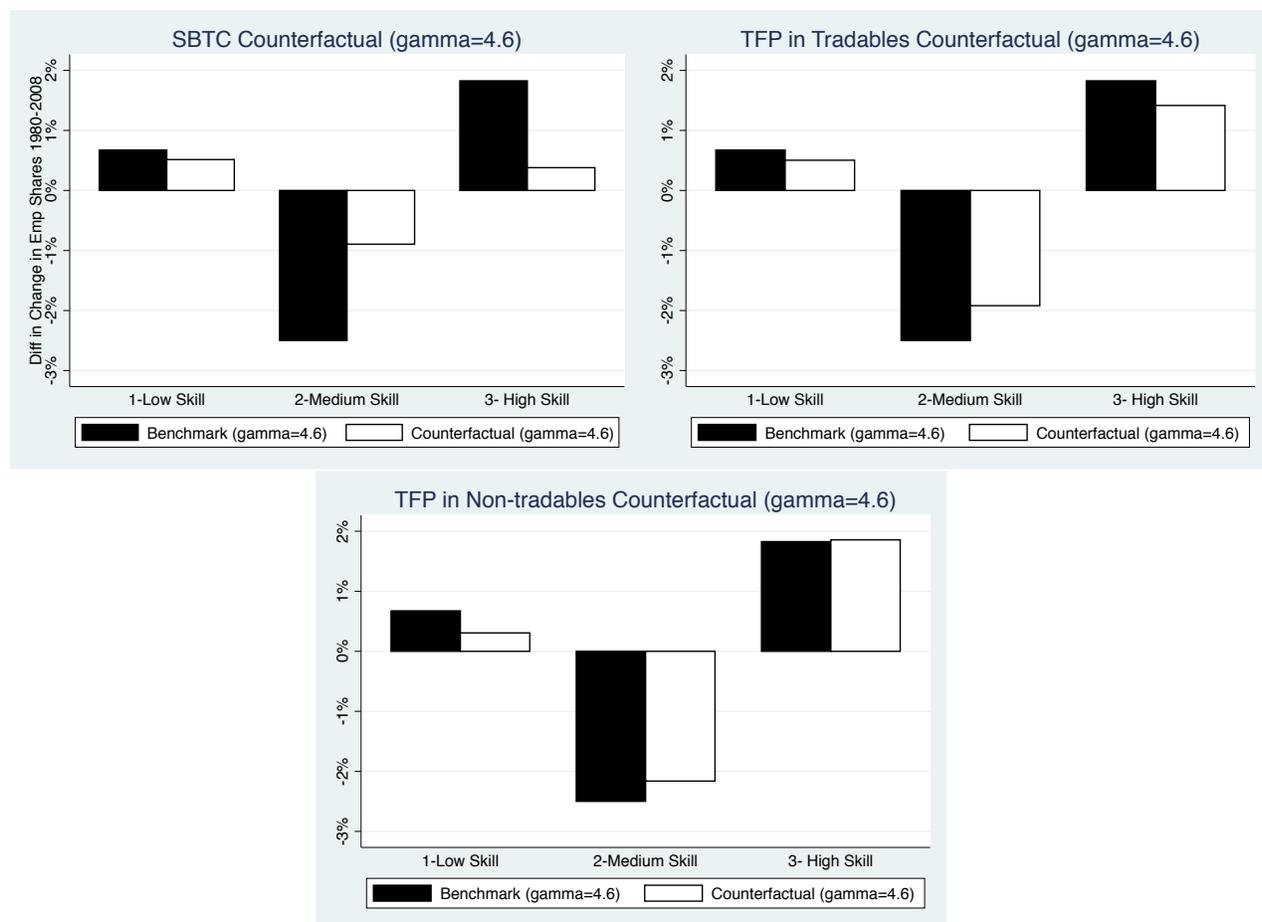


Figure 11: Counterfactual exercises with $\gamma = 4.6$. Black bars represent the benchmark calibration and white bars represent the counterfactual. Left panel: SBTC counterfactual. Middle panel: tradables TFP counterfactual. Right panel: non-tradables TFP counterfactual.

Finally, since the benchmark calibration supports the existence of production complementarity in the tradable sector (i.e. $\lambda > 1$), we run a counterfactual where we shut down the latter and we interpret the residual differences in the employment polarization across cities as generated by the introduction of non-tradables in the model. More precisely, other conditions equal, we set λ in 2008 equal to one.²⁷ The results are reported in Figure 12 for both the case of $\gamma = 2.3$ and $\gamma = 4.6$. In the former case, compared to the benchmark

²⁷An alternative exercise is to set λ equal to one in both 1980 and 2008. The results are virtually identical to those reported in the text. The reason is that the 1980 equilibrium is almost unaffected by the change in this parameter.

calibration, the difference in the change in the share of the three types of workers between the two cities is reduced by 33% for the low-skilled, 36% for the middle-skilled and 37% for the high-skilled. As the two panels in Figure 12 show, the output of this counterfactual is robust to changes in the elasticity of substitution: when $\gamma = 4.6$ the difference in the change in the share of the three types of workers between the two cities is reduced by 39% for the low-skilled, 37% for the middle-skilled and 36% for the high-skilled. This exercise then suggests that the contribution of the channel driven by production complementarity in the tradable sector, while being quantitatively relevant, accounts for a smaller fraction of the differential employment polarization patterns between large and small cities with respect to that associated to the non-tradable sector.

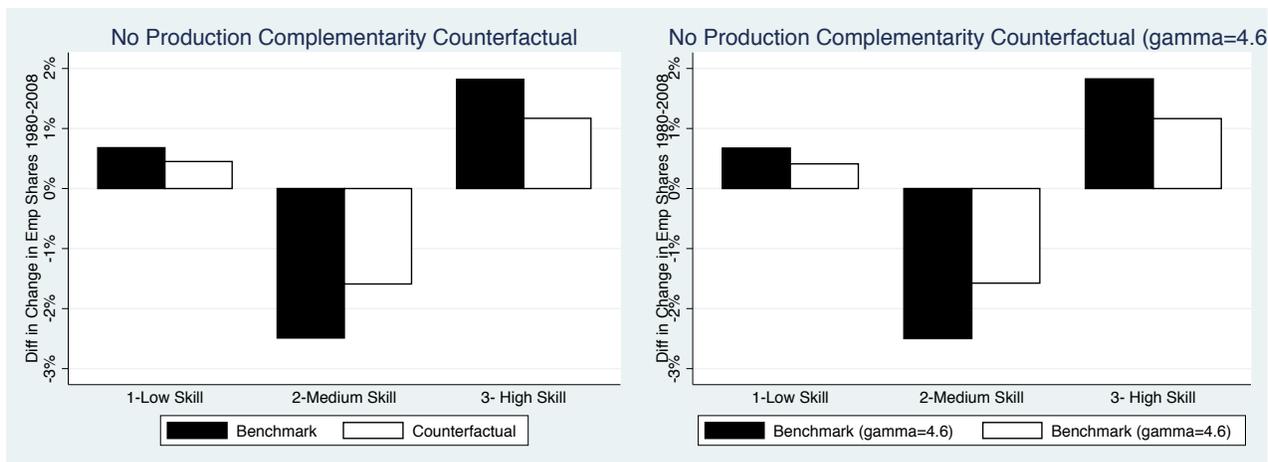


Figure 12: Counterfactual exercises. Black bars represent the benchmark calibration and white bars represent the counterfactual. Left panel: no production complementarity with $\gamma = 2.3$. Right panel: no production complementarity counterfactual with $\gamma = 4.6$.

7 Conclusion

In this paper we provide a comprehensive study on how the allocation of skills and the occupational structure of the U.S. labor market change across space and time during the 1980-2008 period. We first document that in this period employment polarization is stronger in cities whose size is larger in 1980, and that the intensity of this phenomenon increases with city size. Hence, initially larger cities display a faster increase in the employment shares of both high-skill and low-skill occupations and faster reduction in the employment shares of middle-skill occupations with respect to cities that were initially smaller. Importantly, we document that this pattern is driven by the extensive (heads) rather than the intensive (hours) margin, and it is confirmed when considering three broad occupation categories: non-

tradable service (low-skilled); administrative, clerical, sales and production (middle-skilled) and professional and managerial (high-skilled).

To account for the patterns observed in the data, we build a spatial general equilibrium model with location-specific skilled-biased technical change in the tradable sector, a low-skill intensive non-tradable sector and a home vs market labor decision. The model provides a theory-based measure of individual skill that can be used to construct empirical skill distributions. Using this tool, we show that the skill distribution is similar across city size in 1980 while in 2008 larger cities display fatter tails with respect to smaller ones. This finding supports the idea that the increasingly different occupational structure of more versus less urban areas has been driven by the sorting of both low- and high-skilled workers who have been largely attracted to large cities due to the relative increase in the labor demand for their skills.

To further investigate the technological channels that contributed to the occurrence of employment polarization and spatial sorting, we calibrate the model using two groups of cities and the three groups of skills. The benchmark calibration suggests that the role of both unbiased and biased technological change are quantitatively important and supports the existence of both consumption spillovers and production complementarities between high and low-skilled workers. We then perform a series of counterfactuals which show that faster skilled-biased technological change experienced in larger cities is responsible for most of the faster increase in the employment shares of high-skilled workers, for most of the faster reduction in the employment shares of middle-skilled and for a substantial part of the faster increase in the employment shares of low-skilled workers. By neutralizing the channel of production complementarities between high- and low-skilled workers extreme-skill complementarity we also find that the non-tradable sector can account for a substantial part of the different changes in employment shares between small and large cities. This finding suggests that consumption spillovers contribute significantly to the observed divergence in the occupational structure between small and large cities.

References

- Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. volume 4B, chapter 12, pages 1043–1171. Elsevier, 1 edition.
- Autor, D. H. (2019). Work of the past, work of the future. *AEA Papers and Proceedings*, 109:1–32.
- Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5):1553–97.
- Baum-Snow, N., Freedman, M., and Pavan, R. (2018). Why has urban inequality increased? *American Economic Journal: Applied Economics*, 10(4):1–42.
- Baum-Snow, N. and Pavan, R. (2013a). Inequality and city size. *Review of Economics and Statistics*, 95(5):1535–1548.
- Baum-Snow, N. and Pavan, R. (2013b). Inequality and city size. *The Review of Economics and Statistics*, 95(5):1535–1548.
- Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3):507–553.
- Behrens, K. and Robert-Nicoud, F. (2014). Survival of the fittest in cities: Urbanisation and inequality. *Economic Journal*, 124(581):1371–1400.
- Bridgman, B. (2016). Home productivity. *Journal of Economic Dynamics and Control*, 71:60–76.
- Carrillo, P. E., Early, D. W., and Olsen, E. O. (2014). A panel of interarea price indices for all areas in the United States 1982-2012. *Journal of Housing Economics*, 26:81–93.
- Cerina, F., Moro, A., and Rendall, M. (2017). The role of gender in employment polarization. *Working Paper of the Department of Economics - University of Zurich*, (250).
- Davis, D. R. and Dingel, J. I. (2014). The Comparative Advantage of Cities. NBER Working Papers 20602.
- Davis, D. R. and Dingel, J. I. (2019). A spatial knowledge economy. *American Economic Review*, 109(1):153–70.
- Davis, D. R., Mengus, E., and Michalski, T. (2019). Labor Market Polarization and The Great Divergence: Theory and Evidence. *mimeo*.

- Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.
- Eeckhout, J., Pinheiro, R., and Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, 122(3):554 – 620.
- Goos, M., Manning, A., and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review*, 104(8):2509–2526.
- Lindley, J. and Machin, S. (2014). Spatial changes in labour market inequality. *Journal of Urban Economics*, 79(C):121–138.
- Moretti, E. (2010). Local multipliers. *American Economic Review*, 100(2):373–77.
- Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.
- Moro, A., Moslehi, S., and Tanaka, S. (2017). Does home production drive structural transformation? *American Economic Journal: Macroeconomics*, 9(3):116–46.
- Ngai, L. R. and Pissarides, C. A. (2008). Trends in hours and economic growth. *Review of Economic Dynamics*, 11(2):239–256.
- Ngai, L. R. and Pissarides, C. A. (2011). Taxes, social subsidies, and the allocation of work time. *American Economic Journal: Macroeconomics*, 3(4):1–26.
- Roca, J. D. L. and Puga, D. (2016). Learning by Working in Big Cities. *The Review of Economic Studies*, 84(1):106–142.
- Rogerson, R. (2008). Structural transformation and the deterioration of european labor market outcomes. *Journal of Political Economy*, 116(2):235–259.

Appendix A: Data

This appendix discusses the data used in this paper and especially how we document the evolution of wage and skill distributions over time and across locations. One challenge is how to deal with comparability issues, as spatial boundaries of geographical statistical areas change over time.

Individual data

To construct information about workers of different skills and show empirical evidence of employment polarization, we use the national 5percent public-use micro data samples for the 1960 and 1980 Censuses of Population and the 1-percent American Community Survey for 2008. When constructing employment polarization figures, we use data for all individuals who report positive wages and salary income, considering both full and part-time workers, in order to obtain a complete image of changes in employment shares, especially at the bottom of the distribution. However, turning to wage and skill distribution analysis and in order to avoid any data mismeasurement on wages, consistently with the literature we restrict the sample to individuals that work at least 40 hours per week and 40 weeks per year.²⁸ Following [Eeckhout et al. \(2014\)](#), we drop the lowest 0.5 percent of wages to eliminate likely misreported wages close to zero. Instead of using the IPUMS version of the 1990 Census Bureau occupational classification scheme, we chose to work with a balanced set of occupations for 1980 and 2008 used in [Autor and Dorn \(2013\)](#). As a result the total number of workers considered is 1,674,247 in 1980 and 533,021 in 2008 while, when dealing with employment polarization, total observations rise to 3,117,644 in 1980 and 863,101 in 2008.

In addition to wages, we construct the skill distributions using a price-theoretic measure of skills formally represented by equation (5), which we report here for convenience

$$U^{ik} = \Omega (p_H^k)^{-\alpha} (w^{ik})^{\alpha+\omega} \left(1 + \left(\frac{\psi}{1-\psi} \right)^\gamma \left(\frac{w^{ik}}{A_h p_s^k} \right)^{\gamma-1} \right)^{\frac{1-\omega-\alpha}{\gamma-1}}, \quad (6)$$

and where

$$\Omega = \alpha^\alpha \omega^\omega (1-\omega-\alpha)^{(1-\omega-\alpha)} (1-\psi)^{\frac{\gamma(1-\omega-\alpha)}{\gamma-1}} (A_h)^{(1-\omega-\alpha)}.$$

To quantify this measure using individual wages w^{ik} , we need to provide values for the parameters α , ω , A_h , ψ , γ as well as for the prices p_H^k and p_s^k . The five parameters are set according to the benchmark calibration, described in Section 6.1 (table 2). We note here

²⁸Farmers activities and military have been excluded.

that an advantage of using (5) as a measure of skill is that, while this measure emerges from the general equilibrium of the model in Section 4, most of the parameters can be calibrated independently from the rest of model.²⁹

As for the price of housing, following the methodology in [Eeckhout et al. \(2014\)](#), we computed location-specific housing price indices using a hedonic regression model. While housing is a homogeneous good in the model, in the data housing differs in many characteristics that may affect prices. Thus, by relating the log of rent against a number of housing characteristics (number of rooms, age and size of the structure, etc.) and with *city-specific fixed effects*, we isolate the location-specific component of housing prices that can be used to index the difference in housing values across cities. Data on dwelling features comes from the American Community Survey (ACS) and are reported in the IPUMS database at the public use metropolitan area level (PUMA codes) after 2000 and at the metropolitan area level (METAREA) before 1990. Metro areas are “regions consisting of a large urban core together with surrounding communities that have a high degree of economic and social integration with the urban core”.

For the price of non-tradables p_s^k , we rely on the price indexes at the metropolitan area level for the period 1982-2012 provided by [Carrillo et al. \(2014\)](#). Since this paper provides only aggregate prices for goods and services, we use the value of the consumption share of non-tradable from the benchmark calibration ($1 - \alpha - \omega = 0.35$) to impute the variation of prices across location only to the non-tradable services assuming that for tradable goods the law of one price holds. We stress, however, that the measure of skill distribution obtained is very robust to different value of non-tradable prices.

Spatial boundaries

To analyse how the patterns of the distributions differ accross city size, we need to match census micro data to metropolitan areas. The main issue is that the variable “metro area” reports a combination of metropolitan area codes (MSA, primary MSA, central city or county) which has evolved considerably over time, and thus leads to difficulties in matching with PUMA codes or any other harmonized classification of cities. Thus, one issue is to define spatial boundaries of locations which are consistent over time, to identify a “constant” city size effect. The most common way to proceed is to use allocation factors between PUMA (or CBSA) codes in 2008 and metro areas in 1980. This step requires special attention and some manual correction when the county composition of each metro area has changed between 1980 and 2008. For this purpose, population data at the county level is useful in order to

²⁹The only parameter that cannot be independently calibrated in (5) is ψ .

check the consistency of geographical composition. Once this consolidation of spatial boundaries is done, it is possible to merge individual data with population data coming from the 1960, 1980 and 2008 National Censuses. We obtain a subset of 218 metro areas, representing 63% of the 1980 U.S. population and 71% of the 2008 U.S. population. To construct information about workers of different city size, we split these 218 areas into two groups “small” and “large” cities according to three different definitions: top vs. bottom 50% (median), 33% (terciles) and 25% (quartiles) of the population distribution in 1980.

Appendix B: Additional Evidence

In this appendix we provide some additional evidence of divergence between small and large cities overtime, based on some observable measures of skills. We also provide evidence of the time and spatial evolution of the wage distribution.

Changes in the spatial distribution of educational attainments

Table 5 shows how the distribution of educational attainments evolved differently in large and small cities between 1980 and 2008. Based on the sample of workers used to analyze employment polarization, we observe that while in 1980 the relative employment shares of the three different categories considered (less than high-school, less than college, college or more) were similar across city size, in 2008 larger cities display a relative increase in both low-skilled workers (less than high school) and high-skill workers (with a college degree or more) and a relative decrease in middle-skilled workers (less than college). We also observe how the relative increase in high-skilled workers and the relative decrease in medium-skilled workers increases with more extreme definitions of large and small cities (i.e. when we compare cities belonging to the 3rd and 1st quartile). We conclude that this evidence on observable skill measure complements the one presented in the main text and based on the wage level in 1980 as a proxy for skills.

Table 5: Overtime changes in education in large and small cities

		Group	1980	2008	Change	Ch. L-S
Median	Bottom 50%	Less than HS	18,50%	5,77%	-12,73%	
		Less than College	61,84%	58,79%	-3,05%	
		College or more	19,66%	35,44%	15,78%	
	Top 50%	Less than HS	18,60%	6,65%	-11,94%	0,79%
		Less than College	57,90%	50,01%	-7,89%	-4,84%
		College or more	23,51%	43,34%	19,83%	4,05%
Terciles	Bottom 33%	Less than HS	19,11%	6,68%	-12,43%	
		Less than College	62,13%	61,76%	-0,37%	
		College or more	18,77%	31,44%	12,67%	
	Top 33%	Less than HS	19,65%	9,61%	-10,04%	2,39%
		Less than College	56,70%	49,70%	-7,00%	-6,63%
		College or more	23,64%	40,69%	17,05%	4,38%
Quartiles	Bottom 25%	Less than HS	18,90%	6,20%	-12,70%	
		Less than College	62,45%	61,53%	-0,92%	
		College or more	18,65%	32,76%	13,61%	
	Top 25%	Less than HS	20,39%	8,09%	-12,30%	0,40%
		Less than College	57,18%	49,26%	-7,92%	-7,01%
		College or more	22,43%	42,65%	20,22%	6,61%

Employment shares non-tradables across cities and over time

To identify low-skilled employment in the data, we follow [Moro et al. \(2017\)](#). Accordingly, from the 1990 Census classification (3 digits) we select the following industries: Bakery products; Miscellaneous personal services; Beauty shops; Eating and drinking places; Laundry, cleaning, and garment services; Taxicab service; Food stores, n.e.c.; Private households; Child day care services; Retail bakeries; Nursing and personal care facilities; Miscellaneous repair services; Educational services, n.e.c.; Residential care facilities, without nursing; Bus service and urban transit; Personnel supply services; Liquor stores; Barber shops.

Our definition of non-tradable sectors employ a share of low-skilled workers (i.e. workers employed in low-skilled - service - occupations) which is about 5 times larger than the rest of the economy (52.44% vs 11.16% in 1980 and 51.25% vs 14.73% in 2008). A prediction of our theory is that employment shares of non-tradable sectors increase more in large rather than in small cities. These shares are reported in table 7. Consistent with the theory, the share of non-tradables increases over time both in small and large cities but such increase is stronger in the latter group. Moreover, once again the relative increase in large cities increases with more extreme definitions of large and small cities.

Table 6: Employment shares of the non-tradables across cities and overtime

		1980	2008	Change	Diff in Change L-S
Median	Bottom 50%	8.03%	11.07%	+3.04%	+0.41%
	Top 50%	8.25%	11.70%	+3.45%	
Terciles	Bottom 33%	8.18%	11.17%	+2.99%	+0.84%
	Top 33%	8.22%	12.05%	+3.83%	
Quartiles	Bottom 25%	8.23%	11.17%	+2.94%	+1.30%
	Top 25%	8.39%	12.63%	+4.24%	

Wage distributions

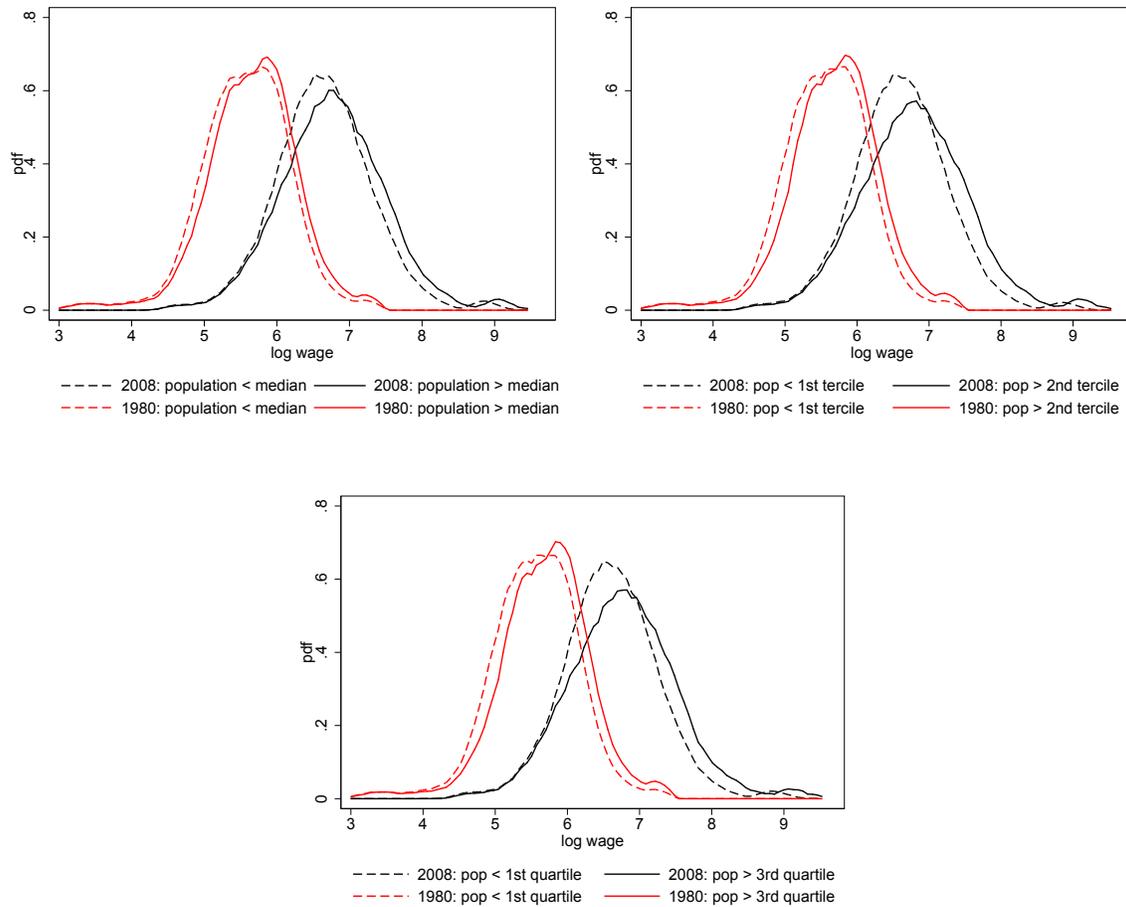


Figure 13: Wage distribution in 1980 (red) and 2008 (black) in small (dashed lines) and large (thick lines) cities. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel compares metropolitan areas with population above the 2nd and below the 1st tercile in 1980 and the bottom panel compares metropolitan areas with population above the 3rd and below the 1st quartile in 1980.

Figure 13 shows the wage distribution across time and space. As in the main text the three panels split cities into two groups. The first one groups cities into those above the median city size and those below. The second panel considers the group of cities below the first tercile and that above the second tercile while the third panel compares the group below the first quartile and above third quartile. Consistent with previous literature³⁰ there is a city-size wage premium both in 1980 and in 2008. Average wages are higher and there is a first-order stochastic dominance of the wage distribution in large cities relative to that of small ones. That is, for each wage level x , the fraction of people earning a wage lower than x is larger in small cities than in large cities. In addition, we observe a divergence in the shape of skill distributions overtime. In 1980 the wage distribution of large cities appears to have the same shape as that of small cities. In 2008 instead, the tails of the distribution are fatter in large cities than in small ones. This is formally confirmed by quantile regressions in Section 5.2 in the text. The result emerges in the three panels of Figure 13, but the difference is more pronounced when considering quartiles with respect to terciles, or terciles with respect to the median split, which suggests that the divergence between small and large cities is increasing with cities relative size.

The role of gender in spatial polarization

[Cerina et al. \(2017\)](#) document that a main driver of employment polarization in the U.S. is the reallocation of hours from home production to market work experienced by women since 1980s. They show how the sharp increase in the education premium in the 80s increased directly women's participation at the top and, indirectly, at the bottom of the skill distribution, due to a larger demand for low-skilled services by skilled women. By its nature, this mechanism should emerge at the level of metropolitan areas, because low-skilled services are produced and consumed locally. Also, it should be more evident in large cities, where education premium rose faster, so that the results in [Cerina et al. \(2017\)](#) suggest that a large fraction of the spatial differences in employment polarization should be driven by women. In this section we investigate to what extent this is the case.

³⁰[Eeckhout et al. \(2014\)](#) among the many.

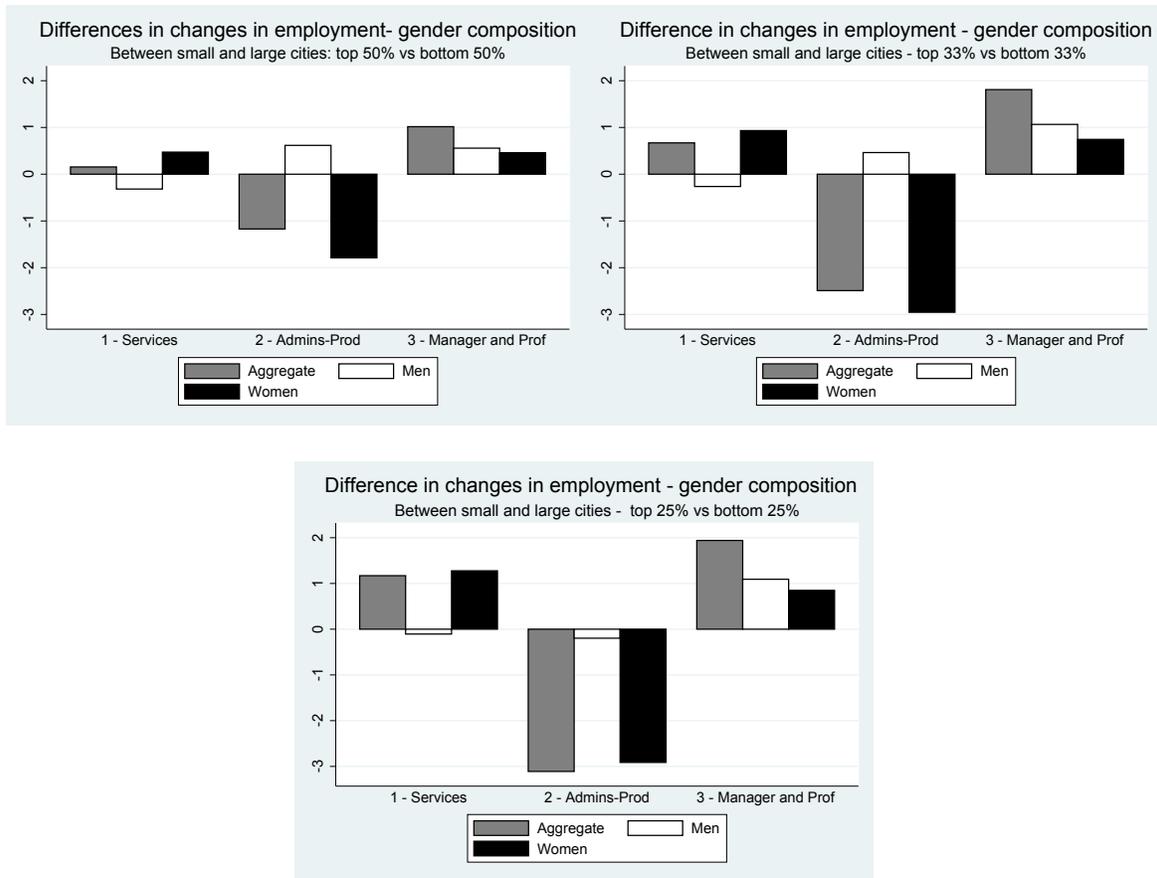


Figure 14: Difference in the change in employment shares between large and small cities in low-, middle- and high-skilled occupations across gender. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel compares metropolitan areas with population above the 2nd and below the 1st tercile in 1980 and the bottom panel compares metropolitan areas with population above the 3rd and below the 1st quartile in 1980.

Figure 14 decomposes the overall spatial difference in the change of employment shares in our three main occupational categories (bar in grey) between men (bar in white) and women (bar in black). The first panel presents the spatial differences in employment polarization between cities below and above the median of city size in 1980, while the second and the third present the same differential pattern for cities in the top and in the bottom 33% and 25%, respectively. The graphs reveal that the differential pattern in employment polarization between large and small cities is mainly driven by women, especially at the bottom and in the middle of the skill distribution. When comparing cities above and below the median (first panel), women display 304%, 153% and 45% of the difference in the change of employment shares for low-, middle- and high-skilled occupations, respectively. When comparing cities

above the second and below the first tercile, the corresponding figures are 140%, 119% and 41% (second panel), while for cities above the third and below the first quartile they are 109%, 94% and 44% (third panel). Thus, irrespective of the definition of small and large cities, women are responsible for the majority of the change at the bottom and in the middle of the skill distribution and for slightly less than half of the difference at the top. In low- and middle-skilled occupations men display a pattern which is either similar across city size or opposite with respect to the aggregate one. In particular, the increase in employment shares in low-skilled occupations for men is always higher in small than in large cities.

This section shows that the contribution of women to employment polarization is particularly strong in large cities, where i) there has been more employment polarization and ii) the increase in the education premium since the 80s has been sharper ([Baum-Snow and Pavan \(2013a\)](#), [Baum-Snow et al. \(2018\)](#) and [Davis and Dingel \(2019\)](#)). This finding provides support for the results in [Cerina et al. \(2017\)](#), showing the key role of women in generating aggregate employment polarization in the U.S. and produces additional empirical evidence for the mechanism proposed in this paper.