CONTRIBUTI DI RICERCA CRENOS

SPATIAL POLARIZATION

Fabio Cerina
Elisa Dienesch
Alessio Moro
Michelle Rendall

WORKING PAPERS

2019/09

arkadia

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENOS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.unica.it
crenos@unica.it

CRENOS – CAGLIARI
VIA SAN GIORGIO 12, I-09124 CAGLIARI, ITALIA
TEL. +39-070-6756397; FAX +39-070- 6756402

CRENOS - SASSARI
VIA MURONI 25, I-07100 SASSARI, ITALIA
TEL. +39-079-213511

Title: SPATIAL POLARIZATION

ISBN: 978 88 68512 347

Prima Edizione: Luglio 2019
Ultima Edizione: Marzo 2020

Arkadia Editore © 2019
Viale Bonaria 98 - 09125 Cagliari
Tel. 070/6848663 - info@arkadiaeditore.it
www.arkadiaeditore.it

Spatial Polarization*

Fabio Cerina+
University of Cagliari & CRENoS
Elisa Dienesch°
Sciences Po Aix
Alessio Moro•
University of Cagliari
Michelle Rendall±
Monash University

Abstract

We document the emergence of spatial polarization in the U.S. during the 1980-2008 period. This phenomenon is characterized by two facts: i) employment polarization is stronger in larger relative to smaller cities and it is mainly driven by heads rather than hours; and ii) while the skill distribution of cities is remarkably similar across city-size until 1980, after that date larger cities experience a faster increase in the share of both high- and low-skilled workers and a faster decline in the share of middle-skilled ones, i.e. the skill distribution of larger cities becomes "fatter" with respect to smaller cities. We quantitatively evaluate the role of technology in generating these patterns by using a spatial general equilibrium model, and find that faster skill-biased technological change in larger cities can account for a substantial fraction of spatial polarization in the U.S. Counterfactual excercises suggest that the differential increase in the share of low-skilled workers across city size is due, in similar proportions, to both the large demand by high-skilled workers for low-skilled services, and the higher complementarity between low- and high-skilled workers in production, relative to middle-skilled.

Keywords: Employment Polarization, Spatial Sorting, City Sizes.
Jel Classification: J21, O14, R12, R23.

---

* We thank Nicola Fuchs-Schündeln, Nezih Guner, Nir Jaimovich, Joe Kaboski, Jordy Meeks, Eric Mengus, Andrew Rendall, Frederic Robert-Nicoud, Kurt Schmidheiny, Benoit Schmutz and seminar participants at Universitat Autonoma de Barcelona, Universitat de Barcelona, Universitat Rovira y Virgili, University of Pau, the Workshop of the Australasian Macroeconomics Society (WAMS) in Queenstown, the Workshop on Macroeconomics and Development in Dalian, the VII Workshop on Structural Transformation and Macroeconomic Dynamics in Cagliari, the European and American Meetings of the Urban Economics Association in Amsterdam and Philadelphia, the 18th Journees Louis-Andre Gerard-Varet in Aix-en-Provence and the 60th meeting of the Society of Italian Economists in Palermo for the useful comments and discussions. We also thank Sara Pau for excellent research assistance. The usual disclaimers apply.

+ E-mail: fcerina@unica.it.
° E-mail: elisa.dienesch@sciencespo-aix.fr.
• E-mail: amoro@unica.it
± E-mail: michelle.rendall@monash.edu.

1 Introduction

Since 1980, the education premium in the U.S. increased sharply (Acemoglu and Autor, 2011). This increase has been spatially heterogeneous, being stronger in larger cities than smaller ones (Baum-Snow and Pavan, 2013a). Therefore, the former locations became more attractive to high skilled workers relative to smaller ones (Glaeser and Resseger, 2010). In turn, additional high-skilled jobs generate substantial local multiplier effects on unskilled jobs in the non-tradable sector (Moretti, 2012). Taken together, these findings suggest that we should observe: (i) faster employment polarization (i.e. the relative increase in the employment shares of high and low-skilled jobs, combined to shrinking middle-skill jobs) in larger cities in the post 1980 period and (ii) that the skill distribution in large and small cities displays a similar shape before 1980, while larger cities experience a faster increase in the tails of the distribution after that date. We dub these two joint phenomena as *spatial polarization.* In this paper we document the emergence of this phenomenon in the U.S. and provide a quantitative theory that can account for it.

From a theoretical standpoint, we consider two mechanisms through which high-skilled jobs can potentially foster the emergence of local employment opportunities for low-skilled individuals. The first implies that the productivity of high-skilled workers is enhanced by low-skilled workers and vice-versa. In this view, for instance, the opening of a new investment bank, a law office or a hi-tech ICT company would generate new demand for cleaning services, cafeterias, security and reception services, etc. Thus, an increase in the number of high-skilled workers within the firm would foster the demand for low-skilled workers through an optimal choice of profit maximizing firms. We refer to this mechanism as *production complementarity.* The second mechanism is due to the high opportunity cost of time of high-skilled workers, which makes *home production* of services (e.g. babysitting, food, cleaning and transportation) particularly expensive for these individuals, inducing them to substitute this production with purchases in the market of those services. Thus, a local increase in the number of high-skilled workers also fosters employment demand of low-skilled workers through an optimal market/home time decision of high skilled individuals. We refer to this mechanism as *consumption spillovers.* Both production complementarity and consumption spillovers are generated through the interaction of demand and supply forces induced by technological change. However, the key margins determining them are different. Production complementarities are mainly determined by the degree of substitutability of different skills in production, while consumption spillovers depend on the substitutability of home production with market produced services displayed by agents. Thus, while both mechanisms have the potential to generate spatial polarization as technological change occurs, their quantitative

role can be substantially different. To quantify these effects in the U.S. we build a quantitative spatial general equilibrium model designed to capture both mechanisms.

Our first contribution is to document that during the period 1980-2008, while there has been an overall increase of employment shares of high-skilled and low-skilled workers relative to middle-skilled ones (i.e. aggregate employment polarization), this phenomenon has been stronger in larger cities. In addition, we provide evidence that this difference, as well as aggregate employment polarization, is largely accounted for by a change in the extensive margin (i.e. number of workers) rather than in the intensive margin (i.e. hours worked by each worker). This suggests that over time larger cities attract both a larger fraction of high- and low-skilled workers relative to medium-skilled ones.

The second contribution of this paper is to show that that the increase in the tails of the skill distribution after 1980 has been faster in larger cities. To do this, we use the model to derive an equilibrium measure of skill for each agent, which is uniquely identified by the wage of the agent and the prices she faces for tradable and non-tradable goods and services. By using this theoretical measure of skill, which accounts for both observed and unobserved characteristics, together with data on location, wage, and prices, we construct model based empirical skill distributions. By using the same approach in a model without home production and substitutable services, Eeckhout et al. (2014) find that in 2009 large cities display fatter tails of the skill distribution with respect to smaller cities. As our reference period is the employment polarization era (1980-2008), we first confirm a similar result in the context of our model for the end of the sample, i.e. that the skill distribution displays more dispersion in large cities than in small ones in 2008. Crucially, we also find that, in contrast to 2008, in 1960 and in 1980 small and large cities display a remarkably similar skill distributions. This finding suggests that only after 1980 the process of spatial polarization starts emerging, with larger cities attracting proportionally more high- and low-skilled workers with respect to smaller cities.

The third contribution of this paper is to use the model to investigate and quantify the role of technological change in shaping spatial polarization. The model displays 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. Importantly, we allow for the high- and low-skilled workers to be either complements or substitutes in production. We introduce a home/market labor time decision and a multi-sector 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, non-tradable across locations and highly intensive in low-skilled labor.

The model posits differential skill-biased technological change (SBTC) growth across cities, which implies that a larger fraction of high-skilled workers is attracted to cities with faster SBTC over time. Since for this type of workers the opportunity cost of working at home is higher, their demand for market services, which are good substitutes of home production, is also higher. As these 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.[1]

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. In addition to SBTC in the tradable sector we also allow for two other types of technological change: total factor productivity (TFP) growth in both the tradable and the non-tradable sector. As in Eeckhout et al. (2014), TFP growth differentials across cities in the tradable sector generate fatter tails in the city with larger TFP due to production complementarity between high- and low-skilled workers in the tradable sector. By contrast, differential SBTC across cities and the presence of a non-tradable sector which is intensive in low-skilled labor creates an additional channel which, by itself, is able to generate spatial polarization.

We calibrate the model to match the observed differences in employment shares of the three types of workers between large and small cities during the 1980-2008 period. The calibration accounts for the targets well and shows that to match them, both TFP growth and SBTC must be faster in the larger city. This result suggests that both types of technological change are relevant for the model to account for spatial polarization. We then run a series of counterfactuals to assess the role of the different types of technological change. To do this, we impose the two cities to differ only with respect to one type of technological change at the time. We first set TFP in the two sectors to be the same across locations between 1980 and 2008 in the benchmark calibration, allowing only SBTC to remain city-specific, and compute the residual difference in the change in the share of the three types of workers between the two cities with respect to the benchmark calibration. This difference is 25% for

[1]The mechanism of *consumption spillovers* was first highlighted by Manning (2004) and Mazzolari and Ragusa (2013). Here we embed it in a general equilibrium model, explore its spatial implications and provide a quantitative estimate of its contribution in shaping the spatial difference in the bottom tail of the skill distribution.

the low-skilled, 65% for the middle-skilled and 80% for the high-skilled in the counterfactual with respect to the benchmark calibration. By only allowing for differences in the TFP in tradables in the two cities, instead, the corresponding numbers are respectively 25%, 22% and 21%. The result suggests that the SBTC channel is quantitatively more relevant than the TFP one to explain spatial polarization.

Finally, the calibration pins down a value for the elasticity of substitution between high- and low-skilled workers in the production function of the non-tradables that implies that they are complements. This result opens the door for a quantitative exercise aimed at quantifying the relative importance of the two mechanisms that in our model connect the upper and the lower part of the skill distribution, namely production complementarity and consumption spillovers. We note here that both mechanism emerge in the general equilibrium of the model in response to any type of technological change. Thus, for instance, consumption spillovers can emerge both due to differential TFP and differential skill-biased technological change across city size. To disentangle the role of the two mechanisms in association with each kind of technological change, we perform the same counterfactuals as above by also removing complementarity in the production function of the tradable sector between high and low-skilled workers. As expected, allowing only for spatial differences in TFP growth in tradables does not generate any difference in the skill distributions, confirming that the production complementarity is needed for the TFP channel to generate spatial polarization. By contrast, allowing only for city-specific SBTC, the residual spatial difference in the change in the share of the three types of workers amounts to 11% for the low-skilled, 82% for the middle skilled and 105% for the high-skilled relative to the new benchmark with no production complementarity. These results suggest that even without production complementarity the SBTC channel drives a large fraction of the spatial difference in the middle and upper part of the distribution (indeed the whole fraction in the latter case) while it keeps a substantial role in generating spatial differences in the bottom of the skill distribution.

The reminder of the paper is as follows. In Section 2 we discuss the background literature and 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 and 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 the the employment share of middle-skilled jobs shrink faster 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.[2] Second, the classification of low-skilled occupations in this paper is guided by the theory. 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.[3]

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 most related to our work.[4]

[2]The 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 US., while our analysis is based 218 metropolitan areas, which are on average significantly larger than the typical Commuting Zone.

[3]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).

[4]Other notable examples are Behrens and Robert-Nicoud (2014); Davis and Dingel (2014, 2019); Behrens

Using a model-based measure of skill based on real wages similar to the one we use in this paper, 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. The main differences with respect to their work are the following. First, we add a *time dimension* to the analysis, and show 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; third, we consider the role of SBTC, which is a key type of technological change over the period considered (1980-2008) and study its role in generating consumption spillovers when coupled with a home sector that allows substitutability with the non-tradable sector in the market; finally, and most importantly, we calibrate the model using U.S. data to perform a quantitative analysis to assess i) the role of technology and ii) the contribution of each of two channels, production complementarity and consumption spillovers, in generating the emergence of fatter tail in larger cities over time. This quantitative exercise indicates that, while faster TFP growth has a non negligible role in generating the emergence of fatter tail in larger cities, the impact of skill-biased technological change is quantitatively more relevant.[5]

The role of consumption spillovers in a context of a rising skill premium are also investigated in a general equilibrium context in Cerina et al. (2017). They 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

et al. (2014) and Diamond (2016).

[5]Appendix C in Eeckhout et al. (2014) also considers the role of home produced services in a model with extreme skill complementarity in production. In contrast to the results in this paper, their main conclusion is that the expenditure share on non-tradables must be unlikely high in order for consumption spillovers to generate fatter tails in larger cities. We note here that there are key differences between the two approaches, that allow us to reach an opposite conclusion, i.e. that consumption spillovers can produce fatter tails in larger cities with an empirically relevant expenditure share of non-tradables. First, we explicitly model home-production and low-skilled market services as two distinct sectors, while they assume that services are produced only at home and can be traded in the market. Thus, our approach allows to control the value of the elasticity of susbtitution between home production and non-tradables, whose value turns out to be key to assess quantitatively the role of consumption spillovers. Second, we discipline the quantitative role of consumption spillovers in a calibration exercise which targets the observed differential emergence of fat tails between large and small cities in the U.S. *between 1980 and 2008*, together with number of other moments as the change in wage premia and hours worked by different types of workers. Thus, our calibration identifies the role of consumption spillovers by using both spatial and time patterns. In contrast, Eeckhout et al. (2014) search for parameter values under which the observed difference in thick tails between large and small cities *in 2009* can be obtained through consumption spillovers, and discuss the empirical plausibility of those values.

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 mainly emerge at the level of metropolitan areas and it should be more evident in larger cities, where the education premium rose faster since 1980, as we emphasize here.[6] This idea is also supported by the results in Leonardi (2015), who estimates an "education elasticity of demand" finding evidence of consumption polarization, i.e. the fact that more educated consumers favor both skill-intensive goods and services like education, health, and professional services and very low-skill-intensive services like food preparation, cleaning.

Possibly, the most related paper 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 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.[7] 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 complementarities and consumption spillovers. In addition, we bring the model to the data in order to provide quantitative predictions on the comparative role of different kinds of technological changes.

Finally, this paper relates to two recent and related papers investigating the consequences of skilled-biased technological change.[8] The first is Baum-Snow et al. (2018) who 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. We complement their analysis by finding that SBTC in large cities does not only drive the rise in urban inequality but also contributes substantially to the emergence of fat tails and stronger employment polarization in large cities. The second is Giannone (2017). She first documents that the convergence process in wages between metropolitan areas in the U.S.,

[6] In Appendix B we also provide empirical evidence showing that, consistent with the theory in Cerina et al. (2017), the spatial differences in employment polarization across city size are 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 middle-skilled ones) are substantially more pronounced in larger cities.

[7] Davis et al. (2019) use French data for the period 1994-2015, while we focus on the U.S in 1980-2008. Also, they use a classification of occupations based on Goos et al. (2014), which is different from ours. Finally, while we focus on rather large notions of "cities" (our median city in 1980, Phoenix, is 1 million and a half inhabitants while the 3rd tercile threshold is Miami, with 2 millions and 700 thousands inhabitants), this is not the case for Davis et al. (2019) where a large city is defined as a location with more than 500'000 inhabitants

[8] In 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.

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.

which took place between 1940 and 1980, was reversed at the beginning of the 80s and then emphasizes how this change in the convergence trend was mostly driven by the fact that wages of college workers started to increase relatively faster in cities where they were initially higher. She rationalizes this evidence with a model where the positive effect of agglomeration externalities on the wages of educated workers is enhanced by the emergence of SBTC. Her counterfactual analysis concludes that had SBTC not taken place in 1980, the convergence process would have not reverted but only slightly mitigated. In a similar vein, our results suggest that the differential evolution of SBTC since 1980 in large and small cities largely contributed to the emergence of fatter tails in the former relative to the latter.

3 Employment Polarization and City Size

Employment polarization in the U.S., i.e. the relative disappearence 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 dimension to this phenomenon.[9] We adopt the same classification used in Autor and Dorn (2013) which harmonizes U.S. Census codes overtime and, guided by the theory, 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, as documented in Autor and Dorn (2013). 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).[10] 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.

[9]Data and sample description can be found in Appendix A.

[10]Following Autor and Dorn (2013) we exclude agriculture and military occupations.

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.[11] 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 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.

[11] A detailed description of city size definition is provided in Appendix A.

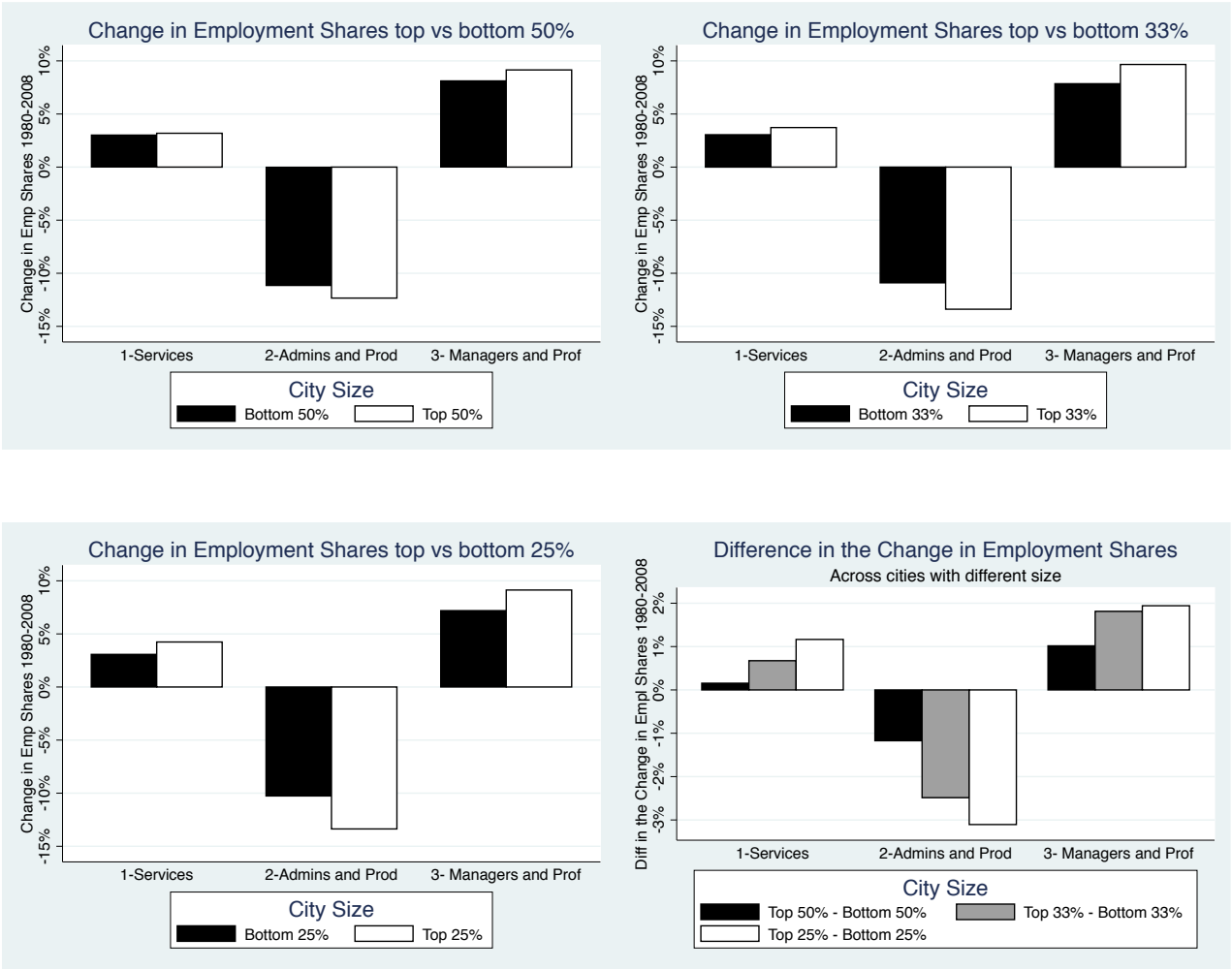

Figure 1: Employment polarization by city size. The upper left panel compares metropolitan areas with population above and below the median in 1980, the upper right panel metropolitan areas with population above the 2nd tercile and below the 1st tercile in 1980, the bottom left panel metropolitan areas 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.[12] To provide further evidence on this distinction by city size, we use the same methodology as in Acemoglu and

[12] Appendix B reports evidence of *educational* polarization: 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). Also, as for occupations, this difference is more pronounced when comparing more extreme definitions of small and large cities.

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.

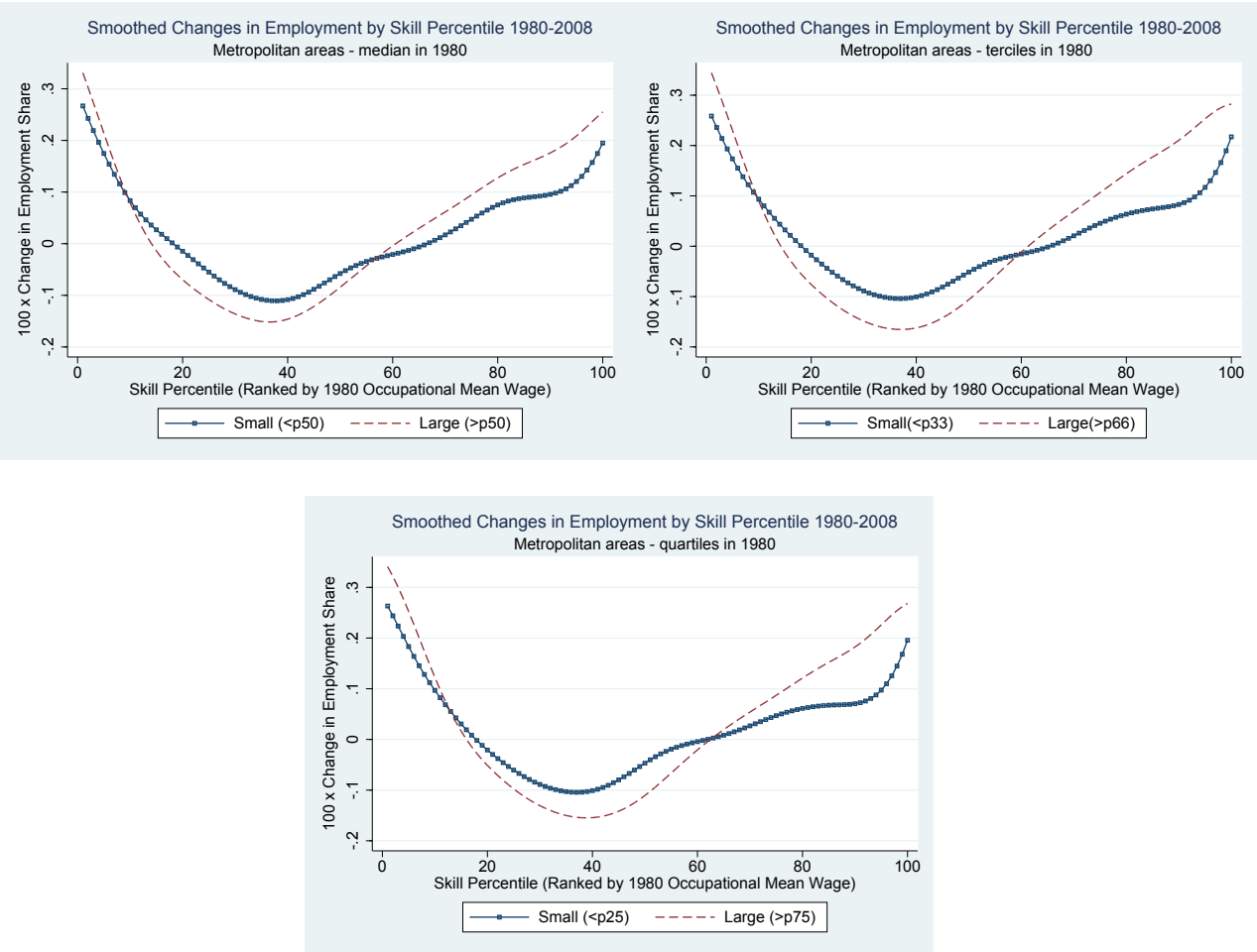

Figure 2: Employment polarization by city size. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel metropolitan areas with population above the 2nd tercile and below the 1st tercile in 1980 and the bottom panel metropolitan areas with population 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. To put 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.[13] 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. Also, 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.[14]

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) that 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, we use a theory of spatial sorting which allows to construct a theoretical measure of skill accounting for both observed and unobserved characteristics. This measure, together with data on location, wage, and prices, allows us to construct model based empirical skill distributions.. 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.

[13]Thus 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.

[14]This result holds also at the aggregate level, i.e. overall employment polarization is driven by the extensive margin rather than the intensive one. Results are available upon request.

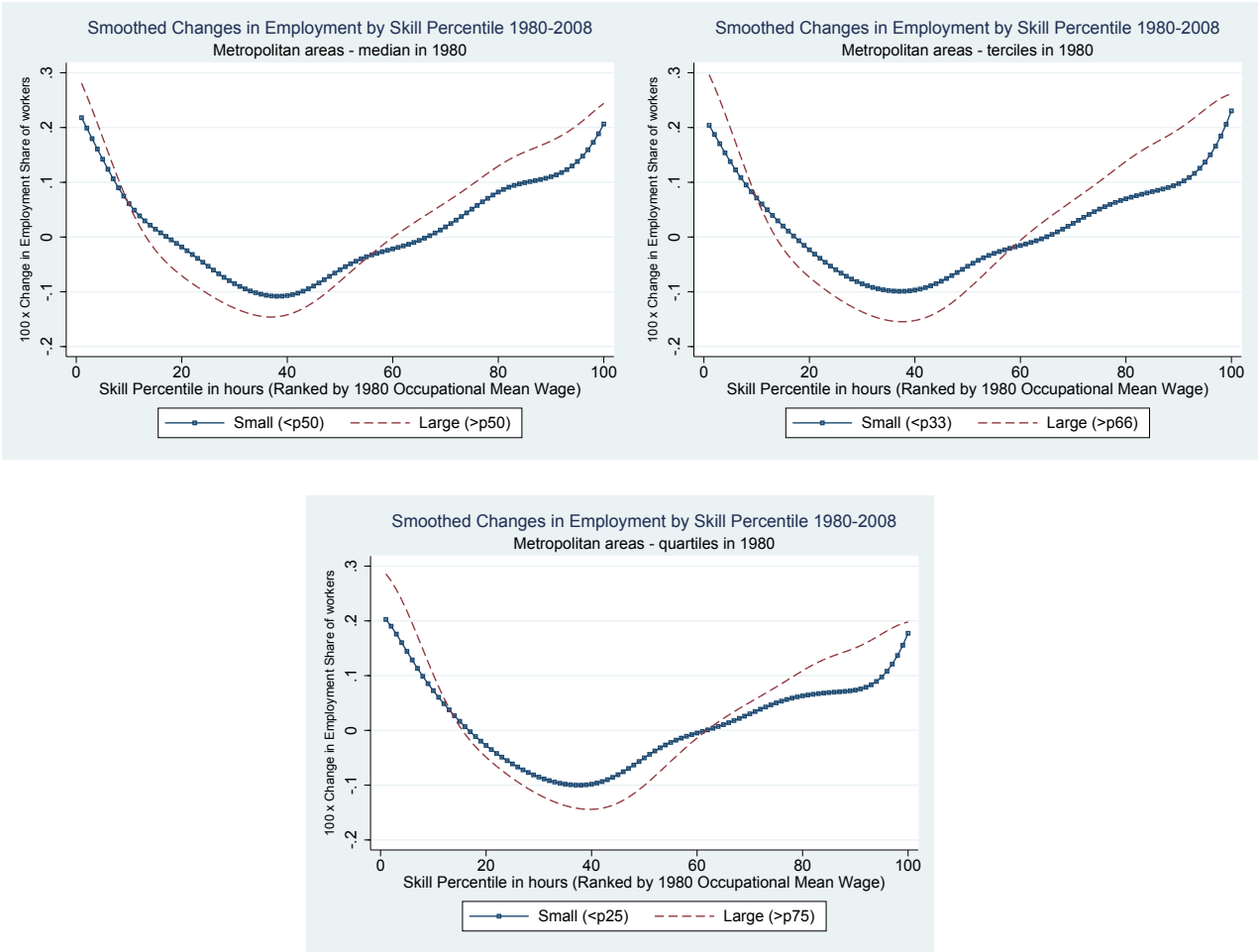

Figure 3: Employment polarization by city size in terms of workers. The ranking of occupations and the bins of occupations are the same as in Figure 2. The variable on the vertical axis is the change in the share of workers in each bin. The left panel compares metropolitan areas with population above 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 $k \in (1, 2, ...K)$. 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 $i \in (1, 2....I)$ 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 $S_k = \sum_i^I n^{ik} = \sum_i^I \sum_j^J n_j^{ik}$ 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 $S = \sum_k^K S^k = \sum_k^K \sum_i^I n^{ik}$.

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

[15]See Rogerson (2007) and Ngai and Pissarides (2011).

i living in city k has the following preferences

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

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

$$c_h^{ik} = A_h l^{ik}, \tag{2}$$

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 market. We assume that home productivity is invariant across skills and locations. The budget constraint for workers of ability i living in city k is

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

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

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

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 spatial 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

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

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:

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

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* (SBTC)[16] 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.

[16] Baum-Snow and Pavan (2013b) 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 SBTC. These agglomeration economies create a stronger impact of economy level SBTC in larger cities with respect to smaller ones. Differential SBTC at the spatial level in our model can be interpreted as a reduced-form version of the mechanism proposed and estimated by Baum-Snow et al. (2018).

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).[17] 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

The aim of this section 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

[17] In 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.

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.

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

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

and where

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

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 a one-to-one mapping between equilibrium utility and skill level for the worker of type i in any city k. 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}.[18] 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.

[18]Note that if $\alpha + \omega = 1$ our setting coincides (except for location-specific productivity of high-skilled workers) with that of Eeckhout et al. (2014), in which there is no home production and no market production of services.

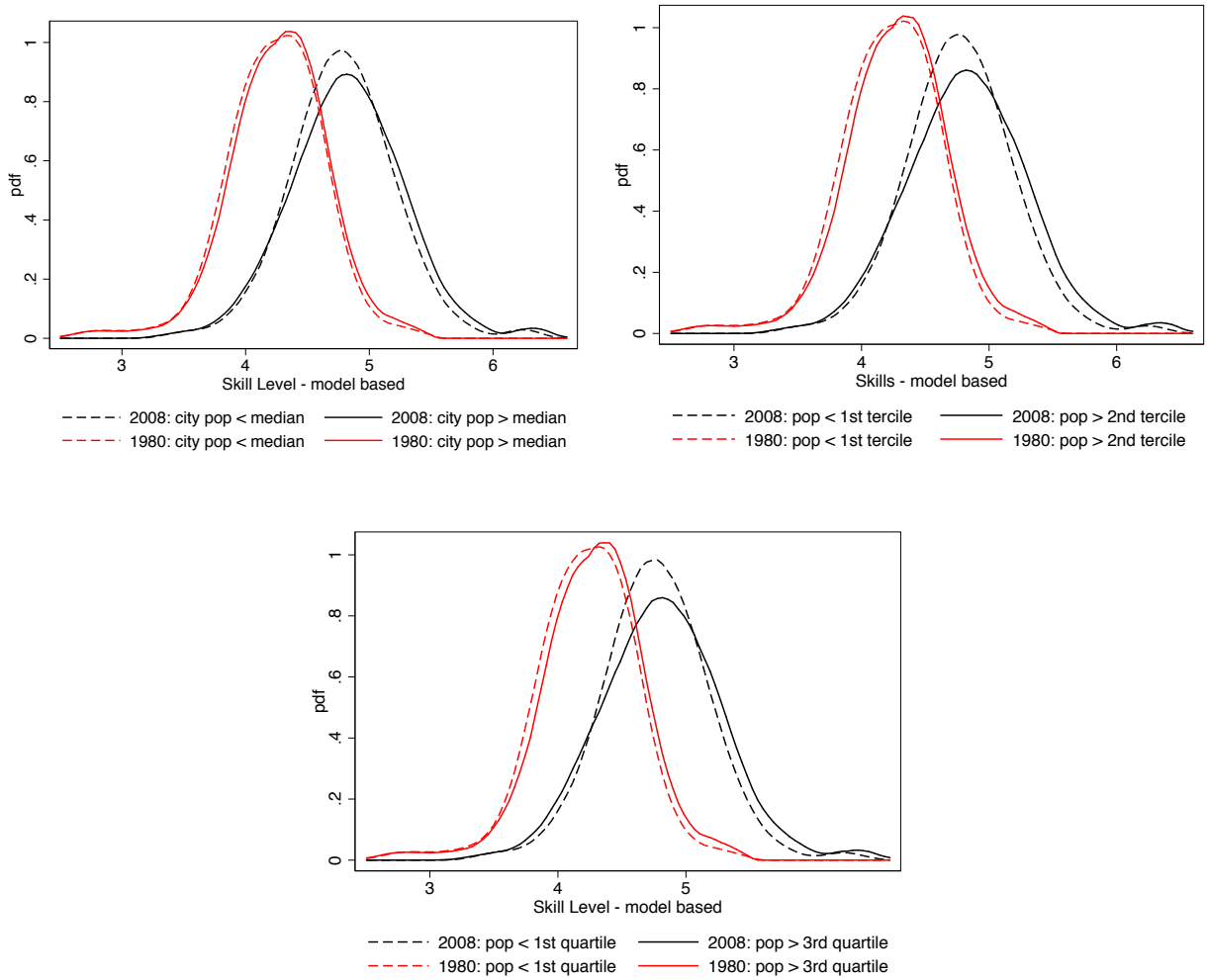

Figure 4: Skill distribution (logarithm of (5) in 1980 (red) and 2008 (black) in small (dashed lines) and large (thick lines) cities. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel compares metropolitan areas with population above the 2nd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

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 is consistent with their results for 2009 in a model without home production and substitutable services. 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 fatter tails in the skill distribution of larger cities is a phenomenon of 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.[19] 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.[20]

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

[20] We 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.

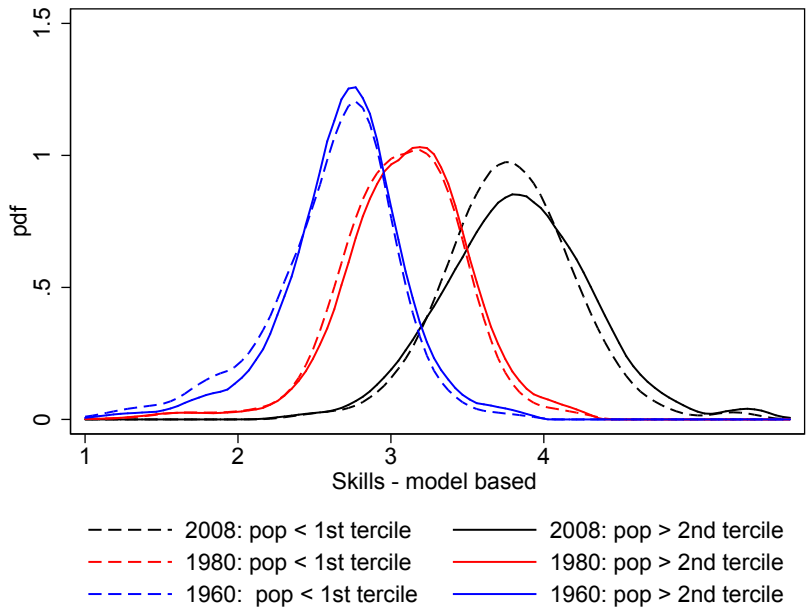

Figure 5: Skill distribution in 1960 (blue), 1980 (red) and 2008 (black) in small (dashed lines) and large (thick lines) cities. The figure compares metropolitan areas with population above the 2nd tercile and below the 1st tercile in 1980.

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.[21] Each of these four exercises is represented in a figure with

[21] In 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.

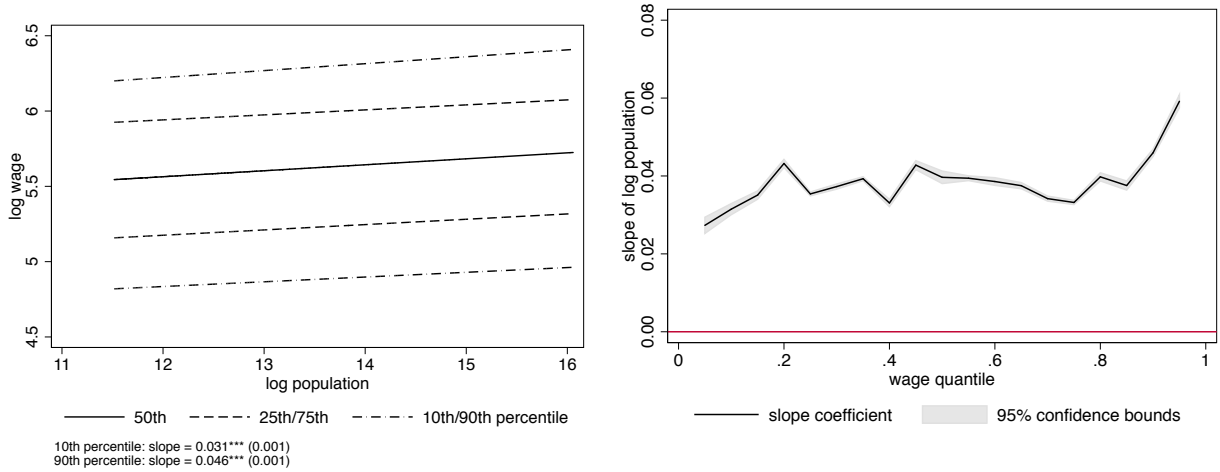

Figure 6: Quantile regression of wage on population in 1980: left, five selected quantiles; right, estimated slope for all quantiles.

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).[22]

[22] Eeckhout et al. (2014) report similar coefficients for the 2009.

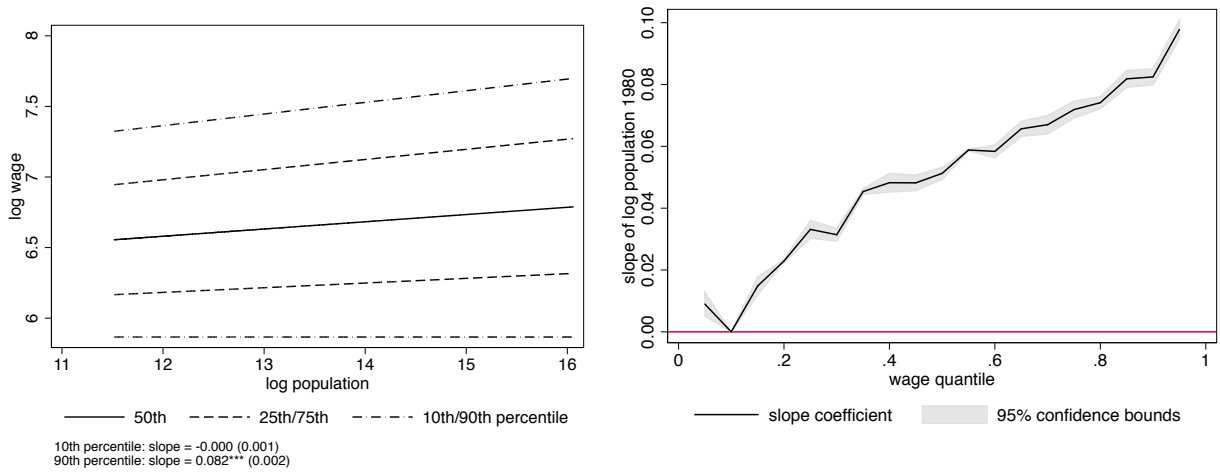

Figure 7: Quantile regression of wage 2008 on population in 1980: left, five selected quantiles; right, estimated slope for all quantiles.

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). Thus, the quantile regression confirms that in 1980 there is no evidence of fatter tails for larger cities.

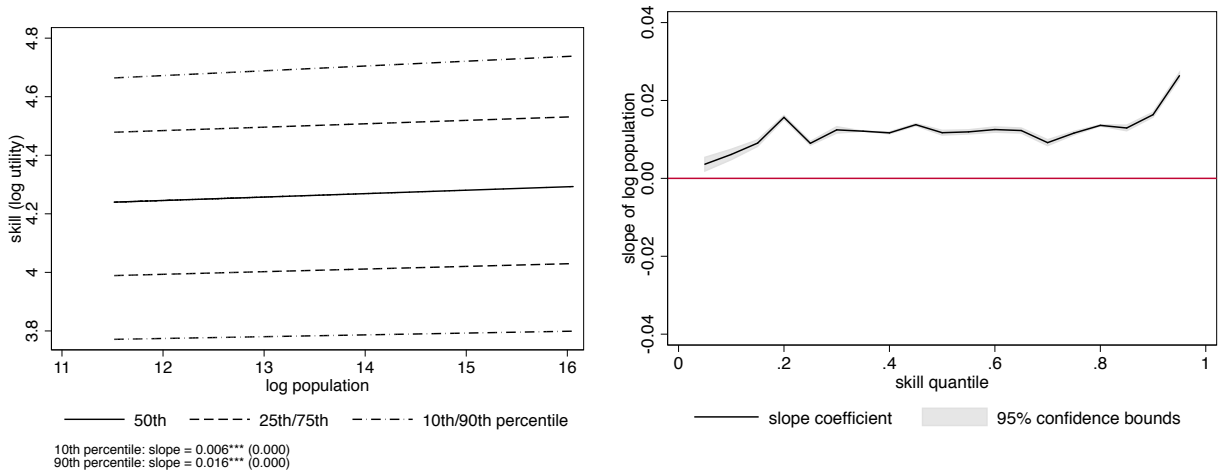

Figure 8: Quantile regression of utility on population in 1980 (i.e. model-based skill measure): left, five selected quantiles; right, estimated slope for all quantiles.

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 25th 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.

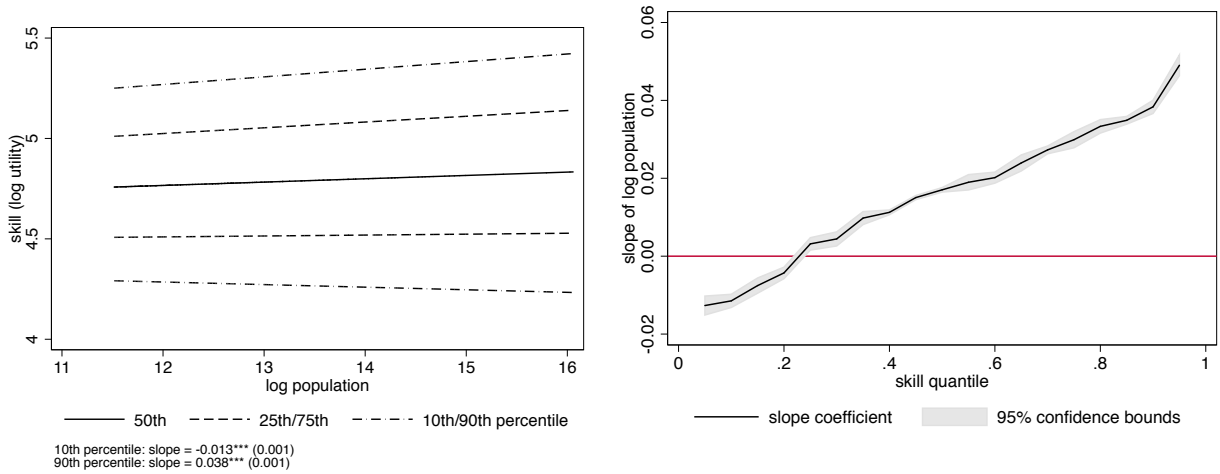

Figure 9: Quantile regression of utility in 2008 (i.e. model-based skill measure) on population in 1980: left, five selected quantiles; right, estimated slope for all quantiles.

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. On the other hand, in section 3, we document that the aggregate phenomena of employment polarization has been stronger in larger than in smaller cities. Hence, our analysis supports the emergence of *spatial polarization* after 1980. The aim of this section is to use a calibrated version of the model to investigate the role of technological change in generating the latter.

In particular, we allow for three types of technological change to potentially evolve differently across 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 the spatial equilibrium model presented in this paper, faster SBTC in a city relative to another would imply that the first city attracts more high skilled workers who, through consumption spillovers, also attract more low-skilled workers to that location. Second, we investigate the role of TFP growth in the tradable sector. Eeckhout et al. (2014) show that with extreme-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. Lastly, for completeness, we also allow for spatial changes in the TFP growth of non-tradables in the quantitative analysis.

Note that, while allowing for a *potentially 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 provides a measure of the differential growth of technological change needed to generate fatter tails in larger cities in the model. Next, by using the calibrated model we run counterfactual exercises to quantify the role of each type of technological change 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, namely the 1980 and the 2008 U.S. economies. In the two equilibria all preference and technology parameters are imposed to be the same except for the levels of SBTC and TFP in the two market sectors, which are allowed to grow differentially across cities according to the growth rates $g_{a^{hk}}$, $g_{A_g^k}$, $g_{A_s^k}$, where g indicates the total growth rate between 1980 and 2008 in city k of the variable at the subscript. Also, motivated by the quantile regressions in Section 5.2, which suggest that there is no difference in the skill distribution across city size in 1980, we require the two cities in the model to be symmetric in the 1980 equilibrium. This implies that all technological 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. The elasticity of substitution between home production and non-tradable market services γ is key for the emergence of consumption spillovers. Following the discussion in Rogerson

[23] In 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.

(2007), we set its value to 5.[24]

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, by introducing home produced services in the utility function, we have to consider the concept of *extended* total consumption expenditure in the data, i.e. the value of market consumption plus the market value of home production.[25] The nominal value of home production is taken from estimates in Bridgman (2016). 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.[26] Hence, following these definitions, we first normalize to one total population in 1980 ($\sum_i N^i_{1980} = 1$), then we compute population growth rates between 1980 and 2008, derive $\sum_i N^i_{2008} = 1 + g_N$ and then finally we feed the model with the aggregate skills shares of low-, middle- and high skilled in 1980 $\{N^i_{1980}\}_{i=l,m,h}$ and in 2008 $\{N^i_{2008}\}_{i=l,m,h}$ obtained using the two previous restrictions. 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.[27] Lastly, we adopt the following normalizations/restrictions:

- Productivity of low-skilled workers is normalized to one, $a^l = 1$;
- The amount of land in each location is normalized to one, $H = 1$;

[24] While there are several studies providing estimates of the elasticity of substitution between home services and *total market consumption* (these estimates range from 1.8 as in Aguiar and Hurst, 2007, up to 2.5 as in Rupert et al., 1995, and McGrattan et al., 1997), we are aware of only two other papers that calibrate the elasticity of substitution between home services and market substitutes only. The first is Olivetti (2006) who finds a value of 4, the second is Ragan (2013) who uses a value of 6.66. We performed alternative calibrations with these values of the elasticity of substitution which, as expected, deliver a lower and higher role of consumption spillovers in generating spatial polarization. However, even in the most conservative case of $\gamma = 4$ consumption spillovers remain quantitatively relevant. Results are available upon requests. In general, several works discuss how the elasticity of substitution between home production and market services should be substantially higher than the one between home production and total market consumption. See for instance the discussion in Rogerson (2007), Ngai and Pissarides (2011) and Moro et al. (2017).

[25] See Moro et al. (2017) for a discussion of the concept of extended total consumption expenditure.

[26] As already noted we exclude agriculture and military occupations.

[27] We stress that one could 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	5	0.26	0.7	1.09	4.70	4.24	22%	26%	1.25	82%	91%	1.25	0%	1.5%

- 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 do not allow market TFP to decline in any sector. This is because the calibration could imply negative TFP growth in low-skilled 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 $\{g_{a^{hk}},\ g_{A_g^k},\ g_{A_s^k}\}_{k=1,2}$, 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 spatial polarization in the model. First, we note that both SBTC and TFP in tradables grow over time in both cities. This 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

Table 3: Model's fit

Moment		Data	Model
Diff. in change in emp. shares by cities	Low-skilled	0.68%	0.67%
	Middle-skilled	-2.49%	-2.49%
	High-skilled	1.82%	1.82%
Aggregate wage premiums	Medium/Low 1980	1.39	1.41
	Medium/Low 2008	1.44	1.41
	High/Low 1980	1.99	2.08
	High/Low 2008	2.51	2.20
Change in relative price of housing	$\frac{\left(p^2_{h,2008}/p^2_{h,1980}\right)}{\left(p^1_{h,2008}/p^1_{h,1980}\right)}$	1.16	1.16
Aggregate growth in consumption	Trad: $\frac{\sum_j \sum_k n^{jk}_{2008} c^{jk}_{g,2008}}{\sum_j \sum_k n^{jk}_{1980} c^{jk}_{g,1980}}$	2.71	2.71
	Non-trad: $\frac{\sum_j \sum_k n^{jk}_{2008} p^k_{s,2008} c^{jk}_{s,2008}}{\sum_j \sum_k n^{jk}_{1980} p^k_{s,1980} c^{jk}_{s,1980}}$	2.10	2.14
Aggr. consumption share non-trad 2008	$\frac{\sum_j \sum_k n^{jk}_{2008} p^k_{s,2008} c^{jk}_{s,2008}}{\sum_j \sum_k n^{jk}_{2008} \left(p^k_{s,2008} c^{jk}_{s,2008} + p^k_H H^{jk}_{2008} + c^{jk}_{g,2008}\right)}$	9.7%	7.5.%
Aggr. empl share low-skilled in trad	1980	7.5%	7.6%
	2008	9.2%	7.4%

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 (2013b), Baum-Snow et al. (2018) and Davis and Dingel (2019). Third, the calibration delivers a value for λ which is larger than 1, which confirms that extreme-skill complementarity contribute to explain the spatial polarization observed in the data. 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. As 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.

6.3 Counterfactuals

6.3.1 Skill-biased technological change vs TFP growth

We now describe three counterfactuals to disentangle the effect of SBTC and that of TFP growth in generating fatter tails. In order to do this, we allow spatial differences of one type

Table 4: Untargeted moments

		Data	Model
Change employment shares city 1	Low-skilled	+3.05%	+2.38%
	Middle-skilled	-10.90%	-10.10%
	High-skilled	+7.85%	+7.77%
Change employment shares city 2	Low-skilled	+3.72%	+3.06%
	Middle-skilled	-13.39%	-12.58%
	High-skilled	+9.66%	+9.52%
Population	Relative size 2008	1.19	1.12

of technological change at a time. In the first counterfactual, we impose the same growth of TFP in tradables and non-tradables between 1980 and 2008 in both cities, which is set to the average growth between the two cities in the benchmark calibration, so that the only source of spatial polarization is the city-specific SBTC. More precisely, unlike the benchmark calibration reported in table 2, we set $g_{A_g^1} = g_{A_g^2} = 86.5\%$ and $g_{A_s^1} = g_{A_s^2} = 0.75\%$ while we keep the value of $g_{a_h^1} = 26\%$ and $g_{a_h^2} = 22\%$. In the second and in the third exercise we apply the same approach for TFP growth in tradables (setting $g_{A_s^1} = g_{A_s^2} = 0.75\%$ and $g_{a_h^1} = g_{a_h^2} = 24\%$ but $g_{A_g^1} = 82\%$ and $g_{A_g^2} = 91\%$) and non-tradables ($g_{a_h^1} = g_{a_h^2} = 24\%$ and $g_{A_g^1} = g_{A_g^2} = 86.5\%$ but $g_{A_s^1} = 0\%$ and $g_{A_s^2} = 1.5\%$) respectively. We focus on the three moments we are interested in: the *difference in the change in employment shares in high-, middle- and low-skilled workers between large and small cities*. For all counterfactuals the black bars represent these moments in the benchmark calibration while the white ones represent those in the counterfactual.

The SBTC counterfactual is displayed in the left panel of Figure 10. When SBTC is the only difference between large and small cities, the difference in the change in the share of the three types of workers between the two cities is 25% for the low-skilled, 65% for the middle-skilled and 80% for the high-skilled with respect to the benchmark calibration.[28] This suggests that the existence of this type of technological change alone 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 right panel of Figure 10 reports the effect of allowing only differences in the TFP growth in tradables across cities. The numbers for the three types relative to the benchmark are now 25% for the low-skilled, 22% for the middle-skilled and 21% for the high-skilled. Thus, with respect to SBTC, removing TFP in tradables have both a substantially smaller

[28]For each counterfactual, we report the percentage of each bar accounted for by the counterfactual with respect to the benchmark.

(at least for the high- and middle-skilled workers) 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 bottom panel of Figure 10 also reports the effect of allowing only for differential growth of TFP in non-tradables in the two cities. In this case, the effect is mostly on low-skilled workers, with a difference in employment share which is 50% of the benchmark, partly on middle-skilled, for whom the difference in employment shares generated is 20%, and a slightly negative effect on the upper tail. Thus, this type of technological change alone cannot generate spatial polarization.

Taken together, these countefactual exercises suggest that faster SBTC in large cities is the main driver of spatial polarization.

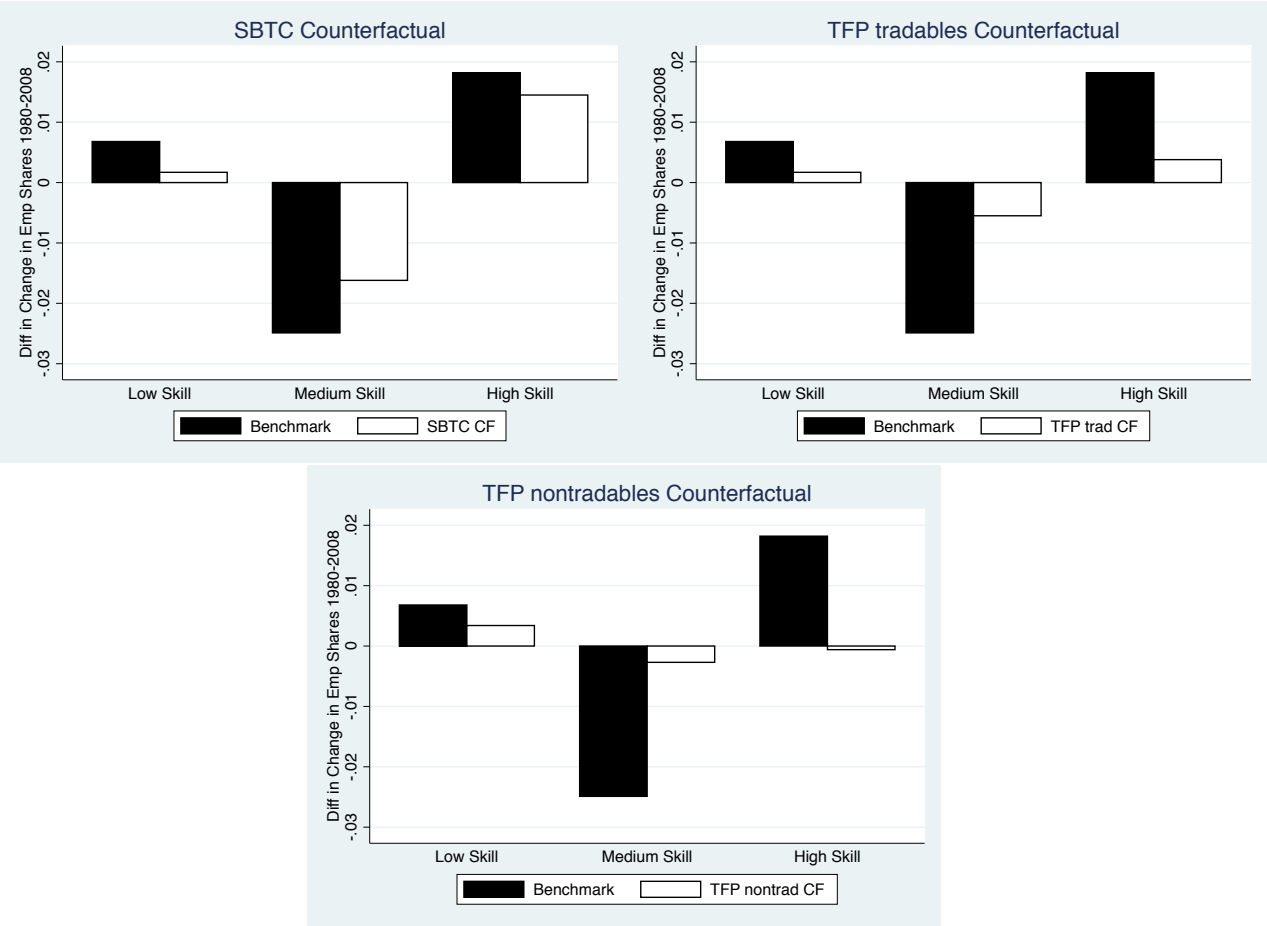

Figure 10: Counterfactual exercises. Black bars represent the benchmark calibration and white bars represent the counterfactual. Left panel: SBTC counterfactual. Right panel: tradables TFP counterfactual. Bottom panel: non-tradables TFP counterfactual.

6.3.2 Production complementarities vs consumption spillovers

The benchmark calibration supports the existence of production complementarity in the tradable sector (i.e. $\lambda > 1$), as proposed in Eeckhout et al. (2014). In the latter, this mechanism - triggered by spatial differences in TFP - is the only driver for the emergence of thicker tails in large cities. But in our model there is another mechanism connecting the upper and bottom tail of the skill distribution, that of consumption spillovers generated by the substitution of household services with low-skilled non-tradable services. In this section, we ask what is the relative contribution of these two mechanisms in generating spatial differences in changes in the employment shares of low-skilled workers. To answer this question, our first step is to shut down the production complemetarity channel by setting λ equal to one in 2008. This enables us to interpret the residual spatial polarization as generated by the joint effect of technological change and the existence of the non-tradable sector.[29]

The results are reported in figure 11, in light grey bars. Relative to the benchmark calibration (black bars), the difference in the change in the share of the three types of workers between the two cities is 44% for the low-skilled, 40% for the middle-skilled and 38% for the high-skilled. A first conclusion that can be drawn from this exercise is that the mechanism of production complementarity in the tradable sector, while being quantitatively relevant, is not a necessary condition for spatial polarization to be accounted for. This result is in contrast with that of Eeckhout et al. (2014) who suggest that a channel based on low-skilled services in combination with home production is not quantitatively relevant in accounting for fatter tails in larger cities.[30]

However, not all the residual spatial disparities in each part of the skill distribution (56%, 60% and 62% respectively) is linked to consumption spillovers. As described above, spatial differences in the TFP growth of the non-tradables do not generate any spatial difference at the top of the skill distribution, their role being confined to the bottom tail. Hence, to quantify the role of consumption spillovers, we perform the same counterfactuals described in section 6.3.1 on the SBTC and the TFP growth in tradables imposing $\lambda = 1$ and thereby removing production complementarity. This exercise allows us to interpret the residual spatial disparity in the bottom tail (with respect to the new benchmark) as generated by the consumption spillovers mechanism only. Results are reported in Figure 11 in white bars for the SBTC counterfactuals (left panel) and the TFP counterfactual (right panel). To allow for an easy comparison, we also report in the same figure the results for the counterfactuals

[29] An 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.

[30] We refer to footnote 7 for the list of differences between their approach and ours that lead to opposite conclusions regarding the role of consumption spillovers.

with $\lambda > 1$ (dark grey bars).

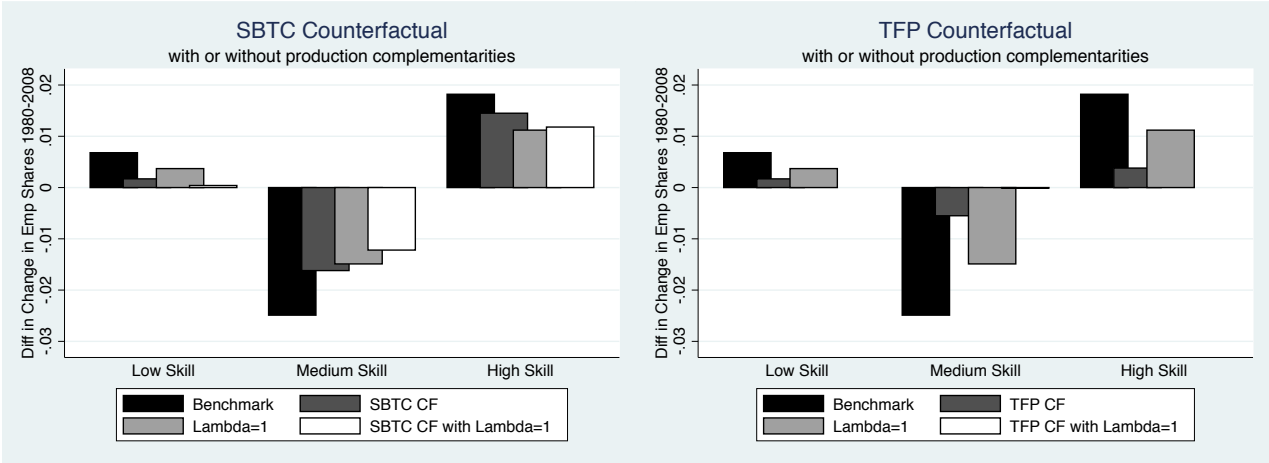

Figure 11: Counterfactual exercises. Black bars represent the benchmark calibration, dark grey bars represent the counterfactual allowing for spatial differences in SBTC (left panel) or TFP growth (right panel) only, light grey bars represent the counterfactual imposing $\lambda = 1$, and white bars represent the counterfactual allowing for spatial differences in SBTC (left panel) or TFP growth (right panel) only *and* imposing $\lambda = 1$.

As expected, the role of spatial disparities in TFP growth in tradables is canceled out without production complementarity. By looking at the white bars in the right panel, we can see that when $\lambda = 1$ and only differences in TFP growth in tradables across cities are allowed, there is no spatial polarization. This is not the case for SBTC: imposing $\lambda = 1$, allowing only for differences in a_h across cities, as suggested by the benchmark calibration (26% in large cities, 22% in small cities), and imposing the same TFP growth either in tradables (86%) and non-tradables (0.75%), the remaining amount of spatial polarization is quantitatively relevant. More precisely, the residual spatial difference in the change of employment shares of high-, medium- and low-skilled workers is respectively 81%, 75% and 24% of the corresponding values in the SBTC counterfactual with the value of λ predicted by the benchmark calibration (1.09).[31] Another way to look at the results is by considering that when consumption spillovers is the only mechanism linking the top and the bottom part of the skill distribution, spatial differences in SBTC alone still explain around 11% of the total difference at the bottom (the remaining difference being explained by the direct effect of TFP growth in non-tradables).[32] Since with production complementarity spatial differences in SBTC alone explain around 25% of the difference in the bottom tail (see section 6.3.1),

[31]Note that the percentages here refer to the fraction of the bar in the counterfactual "SBTC CF with Lambda=1" relative to the counterfactual "SBTC CF".

[32]Compare "SBTC CF with Lambda=1" and "Lambda=1" in Figure 11.

we conclude that consumption spillovers have a quantitatively relevant role in explaining the faster growth in the employment shares in low-skilled workers in large cities generated by faster SBTC therein. This results points to the importance of the demand channel associated to the substitution between home and market services in generating spatial polarization.

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, labeling this phenomenon *spatial polarization*. 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, 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 smaller cities. 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 investigate 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 1960 and 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 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.

Aguiar, M. and Hurst, E. (2007). Life-cycle prices and production. *American Economic Review*, 97(5):1533–1559.

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. *The Review of Economics and Statistics*, 95(5):1535–1548.

Baum-Snow, N. and Pavan, R. (2013b). Inequality and city size. *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.

Giannone, E. (2017). Skill-Biased Technical Change and Regional Convergence. *mimeo*.

Glaeser, E. L. and Resseger, M. G. (2010). The complementarity between cities and skills*. *Journal of Regional Science*, 50(1):221–244.

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.

Leonardi, M. (2015). The effect of product demand on inequality: Evidence from the united states and the united kingdom. *American Economic Journal: Applied Economics*, 7(3):221–47.

Lindley, J. and Machin, S. (2014). Spatial changes in labour market inequality. *Journal of Urban Economics*, 79(C):121–138.

Manning, A. (2004). We Can Work It Out: The Impact of Technological Change on the Demand for Low-Skill Workers. *Scottish Journal of Political Economy*, 51(5):581–608.

Mazzolari, F. and Ragusa, G. (2013). Spillovers from High-Skill Consumption to Low-Skill Labor Markets. *The Review of Economics and Statistics*, 95.

McGrattan, E. R., Rogerson, R., and Wright, R. (1997). An equilibrium model of the business cycle with household production and fiscal policy. *International Economic Review*, 38(2):267–90.

Moretti, E. (2012). *The New Geography of Jobs*. HMH Books.

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. (2011). Taxes, social subsidies, and the allocation of work time. *American Economic Journal: Macroeconomics*, 3(4):1–26.

Olivetti, C. (2006). Changes in women's hours of market work: The role of returns to experience. *Review of Economic Dynamics*, 9(4):557–587.

Ragan, K. S. (2013). Taxes and time use: Fiscal policy in a household production model. *American Economic Journal: Macroeconomics*, 5(1):168–92.

Rogerson, R. (2007). Taxation and market work: is Scandinavia an outlier? *Economic Theory*, 32(1):59–85.

Rupert, P., Rogerson, R., and Wright, R. (1995). Estimating substitution elasticities in household production models. *Economic Theory*, 6(1):179–193.

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 accross 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 5-percent 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 35hours per week and 40 weeks per year.[33] This restriction on the workers' sample is the same adopted by Autor and Dorn (2013).[34] 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 full-time workers considered is $1,674,247$ in 1980 and $533,021$ in 2008 while, when dealing with employment polarization, total observations rise to $3,093,320$ in 1980 and $705,536$ in 2008.

Skill distribution

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

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

and where

[33] Farmers activities and military have been excluded.

[34] In Appendix B we also report the results for different restrictions of the workers' sample.

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

To quantify this measure using individual wages w^{ik}, we need to provide values for the parameters $\alpha,\ \omega,\ A_h,\ \psi,\ \gamma$ 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 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.[35]

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

[35] The only parameter that cannot be independently calibrated in (5) is ψ.

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 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. To have an idea of how population is concentrated across cities in 1980, consider that 50% of the total population concentrates in the 23 largest cities (Phoenix with around 1.5 millions inhabitants being the median city), 33% of the total population concentrates within the top 10 cities (Miami-Hileah with around 2.6 millions inhabitants being the second tercile city), 25% of the total population concentrates in the top 6 cities (San Francisco-Oakland-Vallejo with around 3.2 millions being the third quartile city). The top 5 cities are Los Angeles-Long Beach (9.4 millions), New York-Northeastern (9.1 millions), Chicago-Gary-Lake (7.2 millions), Philadelphia (4.7 millions), Detroit (4.3 millions).

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

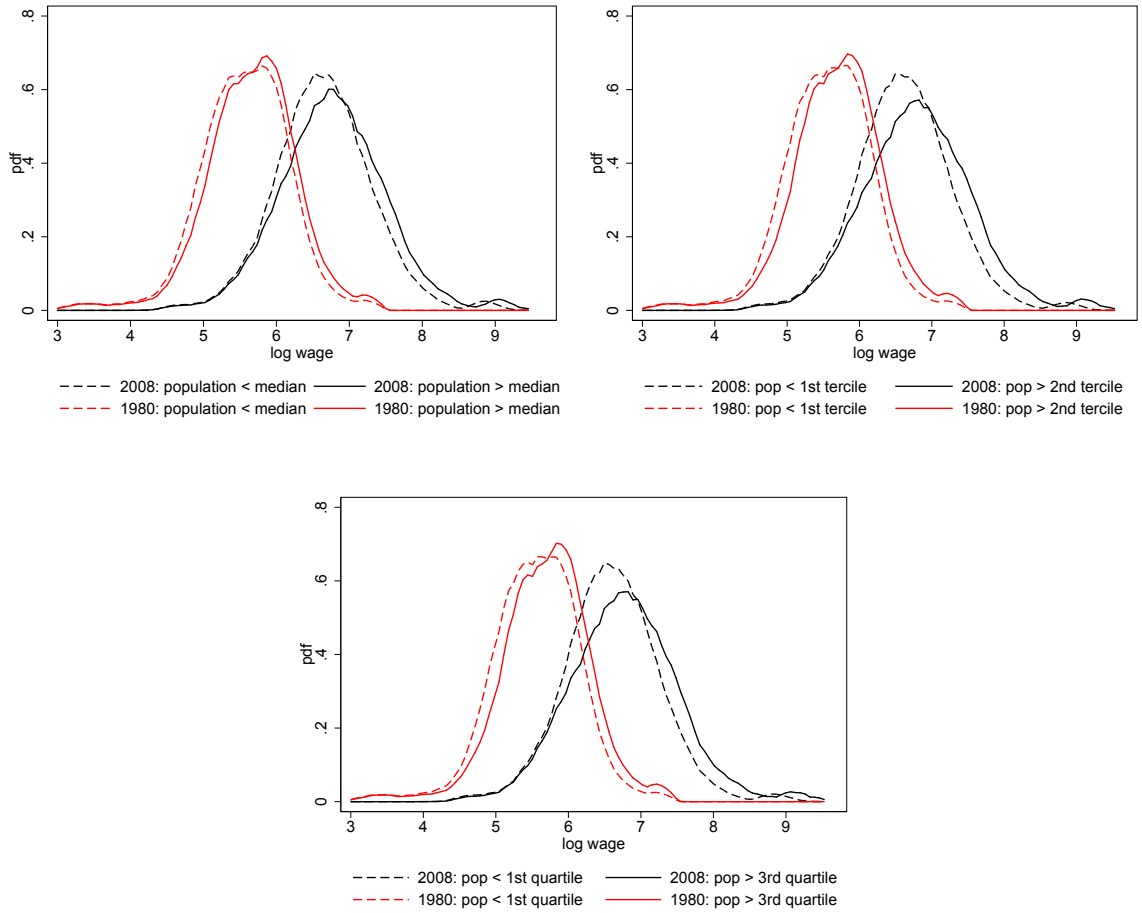

Figure 12: Wage distribution in 1980 (red) and 2008 (black) in small (dashed lines) and large (thick lines) cities. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel compares metropolitan areas with population above the 2nd and below the 1st tercile in 1980 and the bottom panel compares metropolitan areas with population above the 3rd and below the 1st quartile in 1980.

Figure 12 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 literature[36] 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

[36] Eeckhout et al. (2014) among the many.

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 12, 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.

Robustness analysis on the Quantile Regressions

In this section we show the results of quantile regressions by rescricting the workers' sample in different ways. Column "AD" of Tables 7 and 7 presents the results using the same restrictionused in Autor and Dorn (2013), which is the same adopted in the main text: it includes workers who worked at least 40 weeks and 35 hours per week. In column "EPS", we report the results for the sample restriction adopted in Eeckhout et al. (2014), which is slightly different: it includes workers who worked at least 36 usual hours per week and 48 weeks per year. The column "All" reports the results for all workers, without any restriction (the same sample used to compute the employment polarization graphs) while the column "25-65y" only imposes the age restriction from 25 to 65 years but no restriction on working time during the year or the week.[37] The main results of section 5.2 in the text are remarkably robust to changes in the sample restriction and the main message is confirmed: in 1980 urban premiums are homogeneous across the distribution of both wages and skills (being positive in the first case and close to zero in the second), while in 2008 they diverge, being significantly higher for top quantiles than for lower ones. In particular, table 8 shows how for 2008 the coefficients of the lower percentiles (10th and 25th) are always negative (with the exception of the 25th centile for the sample "All") and significantly lower than the respective coefficients in 1980, suggesting that low-skilled workers are disproportionally attracted in large cities in 2008 while the opposite happened in 1980. By contrast, the coefficients of the upper percentiles (75th and 90th) are clearly positive and higher with respect to 1980, when they were very similar to the coefficients of the lower percentiles. This represents additional evidence that fat tails in the skill distribution for large cities emerged in 2008 and is not a feature of the data in 1980.

[37] Notice that this restriction automatically excludes most of the part-time workers which are over-represented in the category below 25 years.

Table 7: Quantile regression coefficients for the wage distribution in 4 sample restrictions

Period	1980				2008			
Sample	AD	EPS	All	25-65y	AD	EPS	All	25-65y
10th centile	0.031	0.041	0.046	0.037	-0.000	-0.000	0.025	0.007
	(0.001[a])	(0.001[a])	(0.000[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.002[a])	(0.002)[a]
25th centile	0.035	0.043	0.047	0.034	0.035	0.038	0.030	0.026
	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])
50th centile	0.039	0.045	0.044	0.032	0.051	0.054	0.053	0.058
	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])
75th centile	0.033	0.038	0.034	0.035	0.072	0.077	0.075	0.074
	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])
90th centile	0.046	0.050	0.043	0.038	0.082	0.092	0.088	0.084
	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.002[a])	(0.001[a])	(0.001[a])	(0.001[a])
Nb obs	1674247	1568072	3093320	2117480	555761	531274	705636	568799
Weight	ind	ind	hours	hours	ind	ind	hours	hours
Pseudo R2	[0.003]	[0.003]	[0.004]	[0.003]	[0.009]	[0.009]	[0.009]	[0.009]

[a] denotes a significance at the 1% level.

Table 8: Quantile regression coefficients for the skill distribution in 4 sample restrictions

Period	1980				2008			
Sample	AD	EPS	All	25-65y	AD	EPS	All	25-65y
10th centile	0.006	0.010	0.017	0.009	−0.013	−0.011	−0.006	−0.014
	(0.000[a])	(0.000[a])	(0.000[a])	(0.000[a])	(0.001[a])	(0.001[a])	(0.000[a])	(0.001[a])
25th centile	0.009	0.011	0.018	0.010	−0.004	−0.001	0.002	−0.008
	(0.000[a])	(0.000[a])	(0.000[a])	(0.000[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])
50th centile	0.011	0.015	0.016	0.009	0.016	0.018	0.017	0.015
	(0.001[a])	(0.000[a])	(0.000[a])	(0.000[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])
75th centile	0.012	0.014	0.010	0.010	0.028	0.030	0.029	0.029
	(0.000[a])	(0.000[a])	(0.000[a])	(0.000[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])
90th centile	0.016	0.020	0.016	0.013	0.038	0.040	0.038	0.037
	(0.000[a])	(0.000[a])	(0.000[a])	(0.000[a])	(0.001[a])	(0.001[a])	(0.001[a])	(0.001[a])
Nb obs	1674247	1568072	3093320	2117480	555761	531274	705636	568799
Weight	ind	ind	hours	hours	ind	ind	hours	hours
Pseudo R2	[0.003]	[0.003]	[0.004]	[0.003]	[0.009]	[0.009]	[0.009]	[0.009]

[a] denotes a significance at the 1% level.

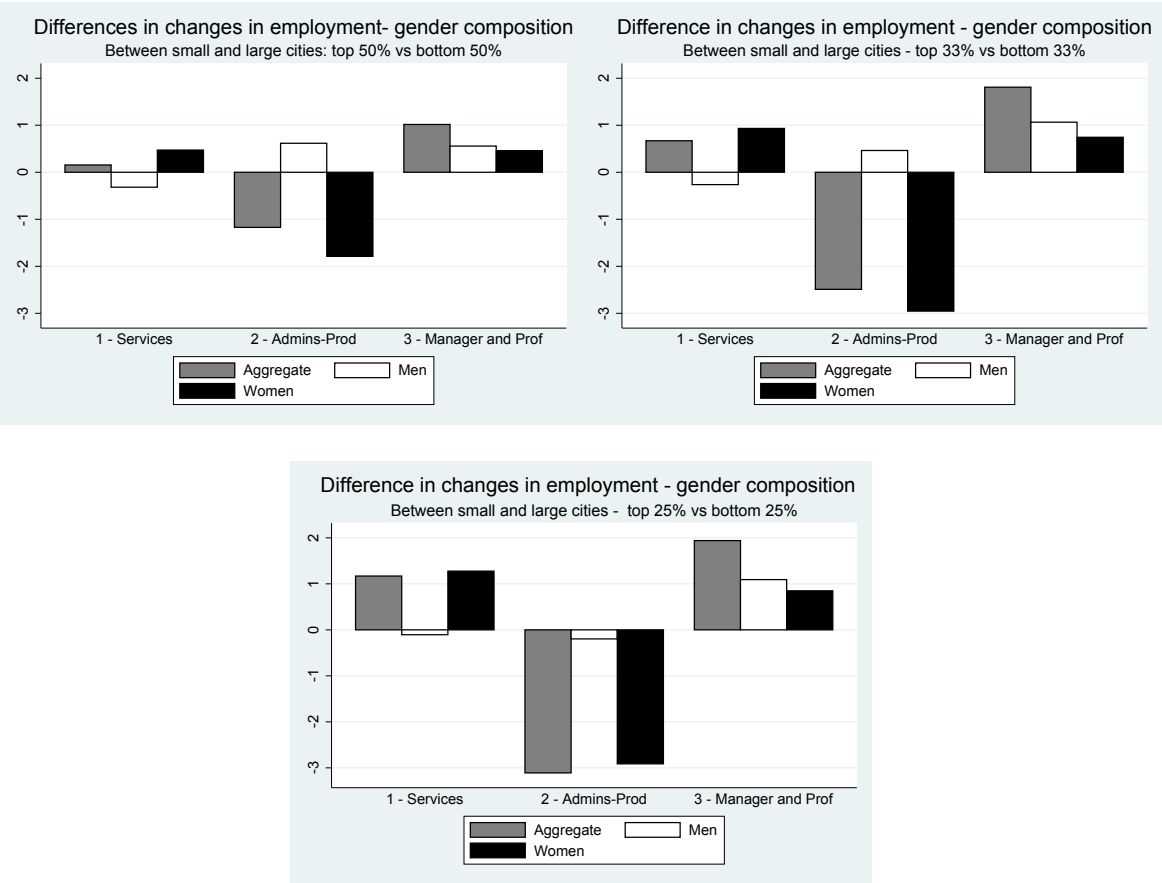

Figure 13: Difference in the change in employment shares between large and small cities in low-, middle- and high-skilled occupations across gender. The left panel compares metropolitan areas with population above and below the median in 1980, the right panel compares metropolitan areas with population 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.

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.

Figure 13 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, 2013b, 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.

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.unica.it

19/08 *Michele Battisti, Massimo Del Gatto, Christopher F. Parmeter,* "Skill Biased Technical Change and Misallocation: a Unified Framework"

19/07 *Fabrizio Fusillo, Francesco Quatraro, Stefano Usai,* "Going Green: Environmental Regulation, eco-innovation and technological alliances"

19/06 *Oliviero A. Carboni, Giuseppe Medda,* "External R&D Acquisition and Product Innovation"

19/05 *José J. Cao-Alvira, Luca G. Deidda,* "Development of Bank Microcredit"

19/04 *Vania Licio,* "When history leaves a mark: a new measure of Roman roads"

19/03 *Simone Franceschini, Gerardo Marletto,* "Reorganization of supply chains as a key for the envision of socio- technical transitions. The case of tourism"

19/02 *Elias Carroni, Dimitri Paolini,* "The business model of a streaming platform"

19/01 *Stefania Capecchi, Marta Meleddu, Manuela Pulina, Giuliana Solinas,* "Mixture models for consumers' preferences in healthcare"

18/13 *Adelaide Baronchelli, Teodora Erika Uberti,* "Exports and FDI: Comparing Networks in the New Millennium"

18/12 *Gabriele Cardullo, Maurizio Conti, Giovanni Sulis,* "Unions, Two-Tier Bargaining and Physical Capital Investment: Theory and Firm-Level Evidence from Italy"

18/11 *Jing Guan, J.D. Tena,* "Estimating the Effect of Physical Exercise on Juveniles' Health Status and Subjective Well-Being in China"

18/10 *Silvia Balia, Rinaldo Brau, Daniela Moro,* "Hospital choice with high long-distance mobility"

18/09 *Luca Deidda, Ettore Panetti,* "Banks' Liquidity Management and Financial Fragility"

18/08 *Gerardo Marletto, Cécile Sillig,* "Lost in mainstreaming? Agrifood and urban mobility grassroots innovations with multiple pathways and outcomes"

18/07 *Jing Guan, J.D. Tena,* "Do Social Medical Insurance Schemes Improve Children's Health in China?"

18/06 *Carmenr Aina, Daniela Sonedda,* "Investment in education and household consumption"

18/05 *Maria Gabriella Campolo, Antonino Di Pino, Edoardo Otranto*, "Reducing Bias in a Matching Estimation of Endogenous Treatment Effect"

18/04 *William Addessi, Bianca Biagi, Maria Giovanna Brandano,* "How tourist flows are affected by the introduction of the euro?"

18/03 *Luc Bauwens, Edoardo Otranto*, "Nonlinearities and Regimes in Conditional Correlations with Different Dynamics"

18/02 *Massimiliano Bratti, Maurizio Conti, Giovanni Sulis,* "Employment Protection, Temporary Contracts and Firm-provided Training: Evidence from Italy"

18/01 *Luca De Benedictis, Vania Licio, Anna Maria Pinna,* "The long-term effects of the historical Roman road network: Trade costs of Italian provinces"

17/11 *Massimo Del Gatto, Carlo S. Mastinu,* "A Huff model with heterogeneous retailers fits well in Southern Italy"

17/10 *Sara Calligaris, Massimo Del Gatto, Fadi Hassan, Gianmarco I.P. Ottaviano, Fabiano Schivardi,* "The Productivity Puzzle and Misallocation: an Italian Perspective"

www.crenos.unica.it

ISBN 9788868512347