NBER WORKING PAPER SERIES

LABOR MARKET POLARIZATION AND THE GREAT URBAN DIVERGENCE

Donald R. Davis
Eric Mengus
Tomasz K. Michalski

Working Paper 26955
http://www.nber.org/papers/w26955

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2020, Revised March 2024

Previously circulated as "Labor Market Polarization and the Great Divergence: Theory and Evidence." We would like to thank David Autor, Iain Bamford, Dominick Barthelme, Mehdi Benatiya Andaloussi, Leah Brooks, Guillaume Chapelle, Nicolas Coeurdacier, Wolfgang Dauth, Jorge De La Roca, Laurent Gobillon, James Harrigan, Reka Juhasz, Maxime Liegey, Bentley MacLeod, Alan Manning, David Nagy, Giacomo A. M. Ponzetto, and Howard Zihao Zhang, as well as seminar participants at BU, EIEF, IGC, Kraks Fond, Le Mans, OECD, PSU, Queens College, SMU, UAB, Wharton, University of Warsaw/WSE, UEA annual meeting 2018 (New York), Université Paris-Saclay, Joint French Macro workshop 2018, 7th workshop on "Structural Change and Macroeconomic Dynamics" (Cagliari), NBER ITI Spring 2019, 2019 European meetings of the UEA (Amsterdam), 2019 NBER Summer Institute ("Income Distribution and Macroeconomics" and "Urban Economics"), HEC Paris Workshop on Firm Location and Economic Geography, 2019 CURE conference, Workshop on "Job polarization and inequality" (Cergy) for helpful comments and discussions. This work was supported by Investissements d'Avenir [grant number ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047] and ANR grant ANR-22-CE26-0016-01. All remaining errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Labor Market Polarization and The Great Urban Divergence
Donald R. Davis, Eric Mengus, and Tomasz K. Michalski
NBER Working Paper No. 26955
April 2020, Revised March 2024
JEL No. J21,R12,R13

ABSTRACT

Labor market polarization is among the most important features in recent decades of advanced country labor markets. Yet key spatial aspects of this phenomenon remain under-explored. We develop four key facts that document the universality of polarization, a city-size difference in the shock magnitudes, a skew in the types of middle-paid jobs lost, and the role of polarization in the great urban divergence. Existing theories cannot account for these facts. Hence we develop a parsimonious theoretical account that does so by integrating elements from the literatures on labor market polarization and systems of cities with heterogeneous labor in spatial equilibrium.

Donald R. Davis
Department of Economics
Columbia University
1004 International Affairs Building
420 West 118th St.
New York, NY 10027
and NBER
drd28@columbia.edu

Tomasz K. Michalski
HEC Paris
Economics and Decisions Sciences Department
1 rue de la Libération
78350 Jouy en Josas
France
michalski@hec.fr

Eric Mengus
HEC Paris
Economics and Decisions Sciences Department
1 rue de la Libération
78350 Jouy en Josas
France
mengus@hec.fr

An online appendix is available at http://www.nber.org/data-appendix/w26955

1 Introduction

This paper examines, in a common framework, two of the most salient features in recent decades of labor markets in the United States and many European countries. The first of these is labor market polarization, the simultaneous loss of middle-paid jobs and growth of both low- and high-paid jobs. The second of these is the great urban divergence, the fact that initially more skilled, typically larger, cities have over time become even more skilled compared to initially less skilled, typically smaller, cities (see Austin et al. (2018), Autor (2019), Moretti (2012) for the US, Iammarino et al. (2018) for Europe as a whole and Guilluy (2010) on the diverging pattern of "France périphérique").

Labor market polarization and the great urban divergence arise in the same labor markets in the same time periods, yet their connection has not been previously explored. The theory in the labor market polarization literature has considered multiple cities (e.g. Autor and Dorn (2013)). However in the long run the theory does not even predict polarization in the aggregate, the primary fact motivating the literature, does not predict polarization in individual cities, and has no clear concept of city scale, so cannot make strong predictions about systematic patterns across cities of different sizes. The labor market polarization literature does include empirics focused specifically on heterogeneity across locations, but these have important limitations discussed below. From the other side, the two-skill setting for the literature on the great urban divergence is wholly inadequate for thinking about labor market polarization, either in the aggregate or in individual cities. The present paper improves our understanding through an examination of French labor markets over the period 1994-2015, replicating existing facts, developing new facts, and then articulating a theoretical framework consistent with all of these.

The segment of the labor market polarization literature focused on locations provides a natural starting point for thinking about which types of cities lose the most middle-paid jobs. The hypothesis developed is that middle-paid job loss will be highest in those locations which had the highest *initial exposure* to these jobs (cf. Autor et al., 2013; Acemoglu and Restrepo, 2020). Of course, whether this turns out to be correct depends on the economic structure connecting shocks and jobs, and perhaps as well on the specific lens through which one views the data. In our data, if we focus *only* on the middle-paid jobs *most* exposed to the posited shocks, then exposure is indeed a good predictor of this subset of middle-paid job loss. However, if we take a broader measure of middle-paid jobs, consistent with the heuristic of labor market polarization, then this result is reversed. This suggests the value of an inquiry that develops a richer set of facts to explain and that also provides a stronger link between theory and data.

Prominent research has examined the great urban divergence. Moretti (2004) described this and made it the central theme of his book, *The New Geography of Jobs* (Moretti, 2012). Unfortunately, there is an intellectual disconnect between the literatures on the great urban divergence and labor market polarization. The main text of a recent review article in the spirit of the great urban divergence does not even include the term "polarization" (Diamond and Gaubert, 2022). This disconnect is not an accident. The intellectual setting within which the great urban divergence literature developed relies on variants of the older two-skill models with skill-biased technical change strongly criticized in the labor market polarization literature (see Acemoglu and Autor, 2011, 2012). The two-skill models are inherently incapable of explaining labor market polarization due to the absence of a middle-paid job sector. Notably, this absence also means that this literature cannot explore any differences between large and small cities in the magnitudes of middle-paid job losses nor characterize the specific types of jobs lost in each. Our examination of labor market polarization and the great urban divergence in a common framework permits a richer, more textured understanding of the systematic, differential evolution of these labor markets.

Our inquiry proceeds in a few steps. We focus on developments in the French labor markets in the period 1994-2015. We begin by developing a set of four key facts characterizing the evolution of labor markets in this period. We relate these to facts known from the existing literature. Next we develop additional features of the French labor markets that should guide theorizing. We follow by documenting that existing models of labor market polarization cannot explain our key facts. Finally, we develop a theory to make sense of these facts and use simulations to examine the qualitative and quantitative relevance of our model.

Our first key fact, universal polarization, demonstrates that there is not only aggregate labor market polarization, but polarization in nearly all cities. This arises because all cities, in spite of their differences, share common features. Specifically, the shock of interest depresses returns in the middle-paid sector everywhere and, in our model with a continuum of skill types, each city finds that this releases labor at both a lower- and upper-margin of the middle-paid sector to the low- and high-paid sectors. Our second key fact is that the loss of middle-paid jobs is largest in the large cities where exposure to these jobs is initially lowest. This requires a focus on a more comprehensive set of middle-paid jobs most consistent with the labor market polarization heuristic. The fact relies on differences across cities of different sizes in both levels, since large cities have relatively fewer middle-paid jobs, and changes, as these jobs nonetheless decline more sharply there. Our third key fact, skewed middle-paid job loss, is that while large and small cities both lose middle-paid jobs, those in the large city are lost relatively in an upper rather than a lower tier. Our fourth key fact is the presence and strength of the great urban divergence in the French data.

We identify a set of additional features of the French data that both motivate and constrain our theorizing and that also provide inputs to simulations to follow. In brief, these are wage polarization; patterns of sectoral absolute and comparative advantage that are systematic across city sizes; and a characterization of the distribution of individual-level productivities and patterns of skill sorting across space.

We make progress on these issues through development of a model that encompasses these facts. The model builds on the foundational labor market polarization model of Autor and Dorn (2013) and the heterogeneous skill spatial equilibrium model of Davis and Dingel (2020). Autor and Dorn provide a basic framework with three labor tasks and a capital/offshoring good. We relax constraints so that our continuum of heterogeneous skills sorts endogenously across tasks and cities. To this we add elements of absolute and comparative advantage at two levels. At the individual level, absolute and comparative advantage combine to drive wage differences and sorting of individuals across tasks and cities. At the city level, absolute and comparative advantage jointly drive differences in city size, the shares of initial sectoral employment, and how cities of different sizes' employment shares respond to polarization shocks due to routinization or offshoring. In our setting, the same routinization or offshoring shocks that deliver labor market polarization also deliver the great urban divergence.

Our paper thus makes a number of contributions. First, we document for the case of France 1994-2015 four key facts about the data that concern city-level patterns of labor market polarization and the great urban divergence that contrast the experiences of larger, skilled and smaller, less skilled cities. Second, we develop a model that replicates these key features of the data. These include aggregate labor market polarization and our version of the great urban divergence. Our model goes beyond prior work, though, in providing an account for robust features of the data, particularly the contrasting evolution of middle-paid sectors in large and small cities, that heretofore have not been part of the discussion. These contributions both unify the literatures on labor market polarization and the great urban divergence and go beyond them to provide a theory that can account for these new facts.

One should care about these advances for a variety of reasons. Relative to the prior literature, we identify a richer set of key facts to understand. Motivated by these, we are able to propose a quite simple theoretical model that can account for the main qualitative and quantitative features. Our study also helps us to understand the underlying economic structure that translates a common shock to a systematic, but spatially heterogeneous, set of outcomes. Our focus is on a long difference over two decades, and so we work with a largely frictionless model.[1] Nevertheless, the insights from this work would do much to inform any future study that would take closer account of the many frictions that exist and that may shape the time path of adjustment. In this context, our focus on spatial heterogeneity of adjustment would

[1]In our model, the relative ranking of skills matters for the spatial and sectoral sorting. This permits accounting in the long-run equilibrium both for downgrading or upgrading of middle-paid sector workers who may choose different sectors after receiving shocks — consistent with the evidence on the transitions of previously middle-paid workers as e.g. in Cortes (2016), Cortes et al. (2017) or Keller and Utar (2023). But it also allows the model to capture changes due to the evolving composition of the workforce — for example the impact of young entrants in labor markets that might have more human capital in comparison to older cohorts who exit.

do much to inform the costs of these changes. Our study also helps to inform other questions focused squarely on the long run. Even in a world that is frictionless, important outcomes may be shaped by local characteristics central to our discussion. For example, we address how the skill and occupational structure evolve differently across cities of different sizes. Large cities become much more tilted to an upper tier, have a sharp decline in the prosperous middle class, and have some growth of lower tier jobs – overall a strongly growing class divide within large cities. Small cities have more modest changes, with some growth in the upper tier, a loss of lower middle class jobs, and stronger growth in low-paid jobs. Any effort to understand differences and changes in political economy across cities of different sizes will need to engage with these facts. Similarly, these distinct changes to who is in each type of city will affect differently the learning environments in each city both for local technological advance (Davis and Dingel, 2020) and for the different environments they provide for opportunity across generations (Chetty et al., 2014). We don't pursue these avenues here, but we do provide a structure in which the path from aggregate shocks to local effects can be understood.

Related Literature Our work builds on a number of literatures. Labor market polarization is documented in the United States in Acemoglu (1999), Autor et al. (2006), and Autor and Dorn (2013); and in European countries in Goos and Manning (2007) and Goos et al. (2009). Acemoglu and Autor (2011, 2012) provide an extended discussion of why recent periods should be investigated in frameworks consistent with labor market polarization. Autor and Dorn (2013) provide a foundational model incorporating what can be thought of as routinization or offshoring shocks. Cortes et al. (2017) expand this model and focus on transitions of workers from routine jobs to non-routine manual jobs and non-employment. Cortes (2016), following Jung and Mercenier (2014), introduces a continuum of labor types mobile across tasks, so is able to accommodate polarization at both the high and low margins, as well as to provide a rich model of the variability of wage shocks among those who remain in their initial task versus changes in tasks both up and down. An important set of papers has also examined the contributions of particular shocks to polarization, including, among others, Goos et al. (2014) on the relative contribution of offshoring and automation, Harrigan et al. (2016) on the role of techies within firms, Michaels et al. (2014) and Eeckhout et al. (2021) on the specific role of IT or Keller and Utar (2023) on the impact of international competition. An important literature has explored spatial dimensions of labor market polarization via the impact of shocks on local labor markets. Prominent examples include Autor et al. (2013) and Acemoglu and Restrepo (2020). These papers have focused on relating job loss to the *exposure* of these local labor markets to the most offshorable or routinizable occupations, or alternatively to robots. In his Ely lecture, Autor (2019) explores empirically how some impacts of labor market polarization have varied across areas of different densities in the United States. Our paper differs from these approaches by considering labor market polarization as a phenomenon affecting the entire skill distribution and taking into account the spatial dimension of labor markets. In particular, we do not focus on one specific shock contributing to polarization. As we are interested in the long run outcomes, we do not focus on the transitional dynamics of workers since large adjustments occur also through entry and exit in and out of labor markets. In our model, we allow agents with all skills to choose a location and a sector without frictions, thus allowing polarization at both the high and low skill margins. Finally, we focus on the important concept of local initial exposure to the posited shocks, considering both in theory and data whether this provides a robust indication of local vulnerability to these shocks.

The term "great divergence" was first applied to cities by Moretti (2012) and has been closely linked to models of skill-biased technical change.[2] The roots of this literature may be found in a seminal paper by Katz and Murphy (1992) and receives its fullest treatment in Goldin and Katz (2009). These works focus on the aggregate labor market and what they term the race between technology and education. In these settings, there is ongoing skill-biased technical change. In periods in which the relative supply of skills rises sufficiently rapidly, the matching of relative demand and supply shocks leaves the skill premium unaffected. When skill-biased technical change outpaces the rise in the relative supply of skill, the skill premium rises. Important contributions to research on the great urban divergence have included Moretti (2004), Diamond (2016), Eckert (2019), Ganong and Shoag (2017), Giannone (2017) and Cerina et al. (2022).[3] Harrigan et al. (2018) examine the role of Hicks neutral and skill-biased technical change on the level and composition of jobs at French firms for the period 2009-2013. Our framework seeks to bridge the literatures on labor market polarization and the great urban divergence, relying on the routinization and offshoring shocks in the former rather than the skill-biased technical change shocks common in the latter. In order to link up to the labor market polarization literature, we go beyond this skilled-unskilled labor dichotomy. We focus on a setting with three key tasks and a continuum of labor types, where routinization and offshoring activities substitute for middle-paid labor and complement low- and high-paid labor. This also allows us to have polarization at both high and low skill margins, and in our setting with many cities also to have polarization both in the aggregate and in all locations.

Our work also relates to a literature on heterogeneous labor and firms in spatial equilibrium across a system of cities such as Behrens et al. (2014), Davis and Dingel (2019, 2020), and Gaubert (2018). Some of these also build on Costinot (2009) and Costinot and Vogel (2010). Relative to these, we simplify in some dimensions and enrich in others, in order to focus on how polarization shocks translate into different evolutions in large and small cities. Where the prior

[2]We have modified this to be the great *urban* divergence to distinguish this from other uses of "great divergence," notably the historical separation in technology and incomes of parts of Europe from China.

[3]Relatedly, Michaels et al. (2018) study task-biased technical change and city comparative advantage.

literature focused on conditions for symmetry breaking, we take these as given in our baseline model to focus on new elements. We also emphasize the role of not only individual- but also city-level relative productivities across tasks. All of these considerations allow us to formulate hypotheses at the aggregate and city levels that we can compare to data.

2 Data description

We focus on a few key questions about the evolution of across- and within-city labor markets. We examine the characteristics and the evolution of labor market polarization in the aggregate and by metropolitan area. These require data on job characteristics (e.g. their routine or offshorable nature), hours worked, wages by occupation, and a measure of skills such as educational attainment. The data should be geographically detailed at the city level and comparable over time. French administrative data and the Censuses satisfy these requirements.

2.1 DADS-Postes data

Our main data source is DADS-Postes for the years 1994-2015, which is a part of the publicly available DADS ("Déclaration Annuelle des Données Sociales") data set.[4] This data is provided by INSEE, the French national statistical institute, and is based on mandatory annual reports by all French companies. It includes data about all legally held job positions, detailed at the plant level. The initial year 1994 is the first year of data that has comprehensive coverage of hours worked. For each worker, for a particular job position, the main reported data are the hours worked, remuneration (total compensation before taxes), occupation type, age and gender.[5] Establishment location is available at the *commune* level, the lowest administrative unit. There were 36,169 communes in mainland France as of January 1, 2015.

We use data only for privately incorporated companies in the period 1994-2015 for mainland France. We limit the sample to workers 25-64 years of age. We retain all positions where there were at least 120 hours worked in a year.

2.2 Occupations and their classification

In DADS-Postes the information on occupations is available at a 2-digit level according to the French occupation classification called PCS ("Nomenclature des professions et categories

[4]Detailed discussion (e.g. on the choice of DADS-Postes over the Labor Force Survey) and description of the constructed data set appears in the Online Appendix, F.1.

[5]In DADS-Postes, we cannot observe education data or job tenure, and it is not possible to aggregate incomes by individual workers at each year level (for 1994) nor to follow them through time. This is possible for a fraction of individuals (1/25 prior to 2001) in a companion data set, DADS-Panel, obtained from the same raw data. We use information from the latter for supplementary evidence.

socio-professionnelles"). French statistical authorities developed it to classify occupations according to their "socio-professional" status and there is not an exact correspondence at this level to other internationally used classifications such as e.g. the International Standard Classification of Occupations (ISCO). We will refer to these as "CS" codes.

The broad 1-digit codes represent CEOs or small-business owners (CS category "2"), "cadres" (high-paid professionals, code "3"), intermediate professions (codes starting with "4"), low-paid employees (code "5") and blue-collar workers (code "6"). The 2-digit categories provide more detail, allowing us to use 18 different CS 2-digit categories consistently between 1994 and 2015.[6] Because of data incompleteness, we exclude artisans (CS 21 and 22; many are unincorporated or not well recorded), agriculture-related (CS 10 and CS 69), and public-sector occupations (CS 33, 34, 42, 45, 52 and CS 53 employed in the public sector). In general data for farmers (CS 10) or all public sector workers are unavailable in DADS before the end of the 2000s.[7]

As a measure of *routinizability* we use the Routine Task Intensity (RTI) index employed by Autor and Dorn (2013) based on the RTI measure of Autor et al. (2003), classifying occupations according to the ease of their automation. For the measure of *offshorability*, we use the index developed by Goos et al. (2014) based on actual offshoring patterns, which identifies occupations readily substituted by imports. To map these indices and obtain exposure to automation and offshoring at the 2-digit CS level used in the French DADS data, we proceed as follows. We first merge the exposure classifications of Goos et al. (2014) (that include RTI in their dataset) based on 2-digit ISCO occupation classification into the 1994 French Labor Survey. Then we map their ISCO-based values into the 2-digit CS ones used in French data basing on ISCO occupations' hours shares into the 2-digit CS.[8] The CS category 54 (office workers) is the most routine and 67 (unskilled industrial workers) is the most offshorable. The top 4 highest-paid occupations are among the least routinizable and offshorable.

The list of 2-digit CS categories we use is in Table 1 along with a short description, their in-sample employment share, average wages in our sample of cities in 1994 and 2015, the routine occupation (RTI) and offshorability (OFF-GMS) ranking from Goos et al. (2014) and the relative wage changes over 1994-2015 (see Appendix F.2.1 for the methodology). The exact RTI and OFF-GMS index values for each category are given in the Appendix Table F.2.

[6]Table F.4 shows representative occupations within each category. We cannot use 4-digit categories over the period studied because (i) the classifications changed in 2003, preventing comparisons at such a level; and (ii) many firms did not file job descriptions with the required detail in the 1990s. See Caliendo et al. (2015) for the use of CS 1-digit categories to analyze firms' hierarchies.

[7]Using an alternative data set, the Labor Force Survey, we compare the aggregate patterns with and without such restrictions in the right hand panel of Figure F.7. The aggregate patterns are similar, taking into account that the bulk of occupational coverage — public sector jobs did not change in a similar fashion throughout the period. See a discussion on public sector employment in Section F.7. Pre-1994 developments based on Census data are discussed in Section F.6 therein.

[8]The ISCO and CS categories are both available only in the French Labor Survey and not directly in the DADS data. More details in the Online Appendix F.1.

Table 1: Sample statistics by 2 digit CS categories.

CS	Description	Employment Share in percent		Average City Wage (in 2015 euros)		Routine	Offshorable	Relative Wage change ranking
						(2015 vs. 1994)		
		1994	2015	1994	2015			
				high-paid occupations				
23	CEOs	1.0	0.9	42.81	59.20	16	17	2
37	**managers and professionals**	**6.2**	**10.2**	32.52	38.56	15	16	1
38	**engineers**	**5.1**	**9.0**	30.36	33.69	17	10	3
35	creative professionals	0.5	0.5	22.83	31.80	14	11	8
				middle-paid occupations				
48	**supervisors and foremen**	**4.1**	**2.7**	18.03	21.86	3	3	10
46	**mid-level professionals**	**12.3**	**7.6**	17.54	21.20	13	6	13
47	**technicians**	**5.7**	**6.3**	17.15	20.60	11	7	14
43	mid-level health professionals	0.8	1.5	15.05	18.05	10	13	9
62	**skilled industrial workers**	**14.1**	**9.3**	13.52	17.99	4	2	17
54	**office workers**	**11.8**	**11.2**	13.17	16.98	1	4	16
65	**transport and logistics**	**2.9**	**3.0**	11.96	16.00	5	5	15
63	**skilled manual workers**	**8.0**	**8.3**	11.90	15.50	7	8	11
64	**drivers**	**5.0**	**5.5**	11.50	14.46	18	18	12
67	**unskilled industrial workers**	**10.9**	**5.7**	11.02	14.72	2	1	18
				low-paid occupations				
53	security workers	0.7	1.4	10.60	14.60	9	12	7
55	**sales-related occupations**	**5.4**	**8.3**	10.44	13.74	6	15	5
56	personal service workers	2.2	4.8	9.97	12.63	12	14	4
68	**unskilled manual workers**	**3.3**	**3.8**	9.11	13.28	8	9	6

Notes: CS refers to the PCS 2-digit codes. In-sample values. Employment share for mainland France (excluding Corsica). "Routine" ranking based on the RTI measure of Autor et al. (2003) while "Offshorable" on the OFF-GMS measure from Goos et al. (2014), both mapped into PCS 2-digit employment categories from the ISCO classification used by Goos et al. (2014). The relative wage change ranking based on wage changes relative to the least paid CS 68 (value in 2015 compared to 1994). Occupations with employment shares above 2.5% in 1994 in bold. We borrow the translation of 2-digit CS categories from Harrigan et al. (2016).

2.2.1 Classification of occupations into wage groups

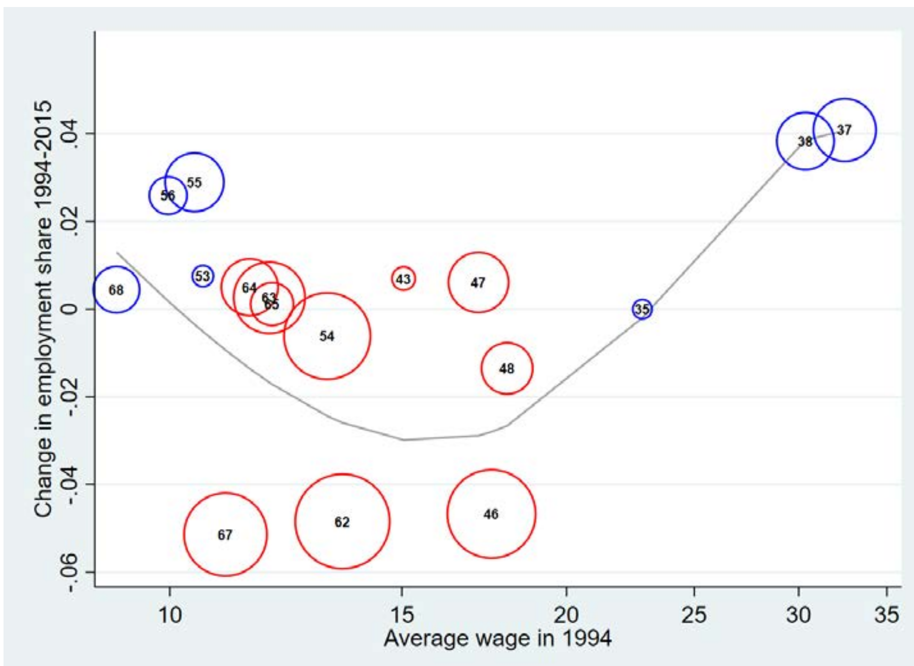
Consistent with the broad labor market polarization literature, we focus attention on three labor tasks with different levels of skill and pay. Matching data and theory thus necessitates mapping a richer set of occupations into three wage categories. Given that this tripartite division is just a heuristic for thinking about the data, there will no unique way to do this and any division will of necessity be imperfect. That said, the categorization may matter substantively for the empirics, so requires justification.

Labor market polarization is defined as the decline of middle-paid jobs along with the growth of high- and low-paid jobs. Following Autor and Dorn (2013) and Goos et al. (2014), we take the mean wage by occupation in the initial year of our data, here 1994, as a primitive ordering of our set of 18 occupations.[9] Forming the wage-occupation groups then requires criteria for the division into high-, medium-, and low-wage groups that respects this ordering.

[9]Table 1 shows that the ordering of average wages across jobs is almost perfectly stable through the period.

Figure 1 provides a visualization of occupational growth between 1994-2015. Only four occupations had declines of 1 percentage point or more of total employment in this period. These four are (from high to low wage) CS 48 supervisors and foremen, CS 46 mid-level professionals, CS 62 skilled industrial workers, and CS 67 unskilled industrial workers (the latter three declining approximately 5pp each). These four occupations each had their share of total employment decline by 34-47 percent in this period, with no other occupations with comparable declines in absolute or proportional terms. These alone provide a justification for including these four in a study of job declines in this period, hence making the relevant interval of middle-paid jobs from CS 48 supervisors and foremen to CS 67 unskilled industrial workers as those vulnerable to the posited shocks. Middle-paid jobs thus defined are highlighted in red in Figure 1.

Figure 1: Labor market polarization in France 1994-2015.



The figure shows the percentage point change in employment 1994-2015 of the considered 2-digit CS occupation categories plotted against their 1994 average wage in cities with >0.05m inhabitants as of 2015. Circle sizes correspond to employment shares in 1994. Middle-paid jobs are shown in red while high and low-paid ones in blue. The line shows a cubic relationship between the average wage and the percentage point change in employment shares, a similar U-relationship as in Autor and Dorn (2013). The CS category "23" - CEOs excluded in this Figure. It is a an "occupation" with a high-wage and stable but small employment share. Corresponding percentage changes and 100 × employment share changes by skill percentiles are shown in Figure F.7.

This definition of the middle-paid group thus also implicitly defines high- and low-paid groups. The implied high-wage group includes CS 23 CEOs and small business owners, as well as highly-paid CS 37 managers and professionals and CS 38 engineers ("cadres"), all of which have both high wages and a distinct social status There is a clear gap in terms of wages between these and the remainder of the occupations ("non-cadres"). The implied low-wage group includes all 2-digit occupations for which the Labor Survey of hours in 1994 had at least half of hours in what Goos et al. (2014) identified as low-wage occupations. These are comprised by CS 53 security workers, CS 55 sales-related occupations, CS 56 personal service workers, and CS 68 unskilled manual workers. All of the low wage occupations experienced growth in their share of employment, with the latter three by one percentage point or more of total employment. In proportional terms these grew by 15-118 percent.

We would like to make four points about this partition of occupations into wage groups.[10] First, from above, all occupations with one percentage point or more decrease in jobs' share are in the middle-wage group and all occupations with one percentage point or more increase in the share of jobs are in either the low- or high-wage groups. Moreover, all middle-paid jobs witness slower wage growth throughout the period in comparison to high- and low-paid jobs (see Table 1). These features are reassuring that our partition is broadly consistent with the labor market polarization heuristic. Second, this partition is governed entirely by an examination of *aggregate* changes in the distribution of occupations, whereas our novel results will concern changes in the cross-city patterns. This is reassuring that the partition into wage groups does not make use of the cross-city patterns that will appear as key facts. Third, while we do not use the measures of routinizability or offshorability to form our wage groups, it would be disturbing if the groups thus formed were grossly inconsistent with these hypotheses, which will be central to our theory. Table 1 provides rankings of the routinizability and offshorability indices for the CS occupations. For the middle wage occupations, the respective median ranks for routinizability and offshorability are (6, 5.5); for the low-wage occupations they are (8.5, 13); and for the high-wage occupations they are (15.5, 13.5). In short, by this measure the middle wage jobs are indeed more routinizable and offshorable, consistent with the labor market polarization hypothesis. Lastly, wages of high-paid occupations such as managers and engineers (CS 37 and 38) increased relatively more than those of low-paid workers (CS 53, 55, 56 and 68) which in turn increased more than *any* of the middle-paid jobs. But the most exposed to offshoring and automation categories such as CS 54, 62 and 67 enjoyed slowest wage growth over the time period 1994-2015 (respectively ranked 16, 17 and 18). Importantly, among our designated middle-paid occupations, even the most highly paid (CS 46, 47, and 48) also had slow wage growth in this period (respectively ranked 13, 14 and 10). In other words, our initial partition of occupations reveals patterns strongly indicative of aggregate job *and* wage polarization.[11]

[10]Throughout, we report robustness checks to show our key results hold (e.g. exhibited in Facts 1-4) when this division is altered. See for example Online Appendix section F.4.1 altering the assignment of the border categories CS 53 or CS 67 in respectively either the middle- or low-wage job category.

[11]Wage polarization that broadly matches job polarization enhances the plausibility of the labor market polarization hypothesis's focus on shocks to relative labor demand. Our finding that this extends to the occupational level underscores the appeal of our categorization of occupations into low-, medium-, and high-paid sectors. We examine this first in Table 1 in terms of raw wage changes. However we can also pursue this using the DADS-panel data set. We run within regressions (equation (30)) of individual wages on time-varying worker characteristics and fixed effects, time effects and an occupational component of the wage. We do this for two panels, 1993-1995 and 2013-2015, to see how these occupational components evolve (see Appendix F.2 for

2.2.2 Partitions of the middle-wage group

While much of our discussion treats the middle-paid sector as a composite, there are reasons to go beyond this. We do so in two ways. First, we divide middle-paid jobs simply as one tier of higher skill and pay versus a second tier of lower skill and pay, consistent with continuum of skill approaches such as Cortes (2016) and our own multi-city approach. The simplest version divides these middle-paid occupations at the median wage, with those above the median being CS 48, 46, 47, 43 and 62. The second way we divide these jobs connects with some of the prior work. While the early concern with labor market polarization was aggregate loss of middle-paid jobs, the hypotheses developed to explain this led later researchers to focus on a subset of these seen as most amenable to the posited factors. Consistent with this, we group CS 48, 54, 62 and 67 as the *most routine and offshorable* (MRO) jobs, as they have the lowest routinizability and offshorability rankings. They comprise 40.9% of hours worked in 1994 in our private-sector employment sample and span the entire wage distribution of middle-paid jobs. The complement to these we will term *other middle-paid* (OMP) jobs. These are frequently still quite routine or offshorable, especially in comparison to high-paid jobs. For example, CS 63 (skilled manual workers) or 65 (transport and logistics personnel) are both ranked as relatively routine/offshorable and CS 46 is 6th in terms of offshorability. While the routinizability and offshorability indices are helpful, they are also imperfect. In particular, they are not always well suited to grapple with the vertical structure of jobs, which could force them to be offshored jointly. Such jobs could also become more routinizable/offshorable with time (e.g. for CS 46: photographers, graphic designers, translators or secretaries; see also Appendix E).

2.3 Cities considered and final sample

We focus primarily on cross-city comparisons. We limit ourselves to data on jobs performed in cities (metropolitan areas) above 50,000 inhabitants as of 2015 unless otherwise noted. We aggregate commune-level data to the metropolitan area ("unité urbaine"), with city boundaries defined by INSEE as of 2010 unless otherwise indicated. There are 117 such cities in 2015 with the largest 55 above 100,000 inhabitants.[12]

These cities above 50,000 inhabitants encompass 54% of the total population of mainland France in 2015. In both 1994 and 2015, the jobs therein account for 73% of wages paid and

details). If we interpret occupations here as individual CS codes, then we can plot the change in occupational fixed effects as in Figure F.3, where wage polarization emerges strongly. Alternatively, we can do this while dividing these into low-, middle-, and high-paid sectors. Taking the low-paid sector as the base, we find that high-paid relative wages rise by 0.31 log points while middle-paid wages fall by a similar magnitude over this horizon. This result will serve as an input to our simulations in Section 5. In short, however we look at this, wage polarization emerges clearly. The estimation is discussed in more detail in Appendix F.2.1.

[12]They are shown in the Appendix: a map in Figure F.1, population data by city category in Table F.1 while the characteristics of the final sample in Table F.5.

68% of hours worked in the mainland in the non-farm private sector.[13][14] In 1994 and 2015 respectively, firms active in these cities for which we have data account for 396,637 and 633,851 firms. With the exclusions discussed above (on worker age or types of jobs), we retain as the main sample that accounts for 65% of total wages paid and 58% of hours worked in mainland France in 1994 and 2015.

We consider up to six major categories of cities for our analysis. Paris, given its size (10.7m inhabitants in the metropolitan area and 37.5% of jobs in 2015 in our final sample) is a category by itself. Then, we use 2 categories of cities above 0.5m: 0.5-0.75m and 0.75m and above (except Paris). Such a choice is warranted because there is a considerable size difference between the seventh largest metropolitan area – Bordeaux (904 thousand inhabitants) and the eighth – Nantes (634 thousand people). Moreover, cities with metropolitan areas of "0.75m and above" have also "urban areas" ("aires urbaines") as defined by INSEE of over 1m inhabitants. For other divisions we follow the ones of INSEE: 0.2-0.5m (size categories "71" and "72") , 0.1-0.2m (sizes "61" and "62") and 0.05-0.1m ("51" and "52"). We took the city size of 50,000 as a cutoff for our main discussion, although lowering this to 20,000 doesn't materially affect our results.

Throughout the studied period city populations increased by 9.9% on average. There is no significant differential growth in the sizes of cities, e.g. when one compares cities with population above 0.5m with the rest or smallest cities <0.1m.[15]

3 Four Facts on Polarization and Divergence

In this section we identify four key facts on labor market polarization and the great urban divergence based on French data for the period 1994-2015. First, labor market polarization is close to universal. Second, middle-paid job loss is strongest precisely in the large cities where initial exposure to them was small, not large. Third, there are marked differences between large and small cities in *which types* of middle-paid jobs were lost, with those lost concentrated relatively in an upper tier in large cities and a lower-tier in small cities. Fourth, we show that the great urban divergence is evident in the French data and we can provide a more textured

[13] In robustness checks, e.g. Online Appendix Table F.26 — a version of Table 3 we also consider urban areas ("airés urbaines") as defined by INSEE encompassing all communes in the metropolitan area ("unité urbaine") plus all communes where at least 40% of residents have employment in the same metropolitan area. These urban areas including the metropolitan areas that we consider account for 70% of the total population, 83% of wages paid and 79% of hours worked.

[14] We do not work with commuting zones defined by the INSEE with a partition of the entire country for the following reasons. First, we focus on jobs outside primary sectors exposed to polarization. Many commuting zones are rural and sparsely populated. Moreover, commuting zones are not consistently defined across French cities, in contrast to metropolitan areas. For example, to avoid obtaining a single commuting zone for the Paris agglomeration, INSEE applies different criteria there compared with the rest of France.

[15] Population data is from the INSEE for the Census years and 2015 and unavailble in 1994 at the commune and therefore city level. Hence, for weighting we use 1990 population from the Census.

account of its character when combined with the first three facts. We note additional relevant characteristics of cities on the wage evolution, productivity and skill sorting that theory should match. We conclude by arguing that existing models of labor market polarization and the great urban divergence fall short of accounting for these facts. Thus we motivate a unified approach to modeling them and develop this with related simulations in Section 4.

3.1 Universal labor market polarization

We first investigate how patterns of employment evolved in mainland France as a whole over the period 1994-2015. Labor market polarization is defined as a fall in the employment share of middle-paid occupations and a rise in the share both of high- and low-paid ones. Table 1 provides detail on this evolution for all jobs in mainland France[16]. The share of middle-paid jobs declined from 76% to 61% between 1994-2015. The bulk of job losses in this category occurred in what we term MRO jobs – the 4 most routine and offshorable occupations (supervisors and foremen (CS 48); office workers (CS 54); skilled (CS 62) and unskilled (CS 67) industrial workers), and their share in hours worked fell from 41% to 29%. The other middle-paid occupation experiencing a large overall employment share decline was mid-level associate professionals (CS 46), whose share of the labor force fell from 12% to 8%. Its rank in our offshorability and routinizability indices are 6th and 13th. At the same time, the overall shares of high-paid jobs increased from 13% to 21% and that of low-paid jobs from 12% to 18%.

The aggregate polarization patterns detailed at the 2-digit CS-level are exhibited in Figure 1 and confirm for France in the years 1994-2015 the U-shaped relationship studied by Autor et al. (2006) and Autor and Dorn (2013) for the U.S. and documented by Goos and Manning (2007), Goos et al. (2009) and Goos et al. (2014) for Europe.[17] They are also consistent with observations made by Harrigan et al. (2016) for France for the time period 1994-2007. Consistent with prior literature, this job polarization is closely paralleled by wage polarization (see Figure F.3), suggesting that the predominant shocks are to relative labor demand.

Figure 2 depicts labor market evolution at the individual city level for all cities in our sample. The horizontal axis measures the change in percentage points of all middle-paid jobs in the period 1994-2015, while the vertical axis provides the same information for high-paid jobs. In this space, a city experiences labor market polarization when it is in the third quadrant and below a ray from the origin with slope -1. Thus we see that at the individual city level, labor market polarization is close to ubiquitous. All of the 117 largest cities in France experienced some decline in employment of middle-paid jobs over the period 1994-2015. In 115 of these

[16] Table 2, last column, gives the corresponding statistics for cities in the sample.

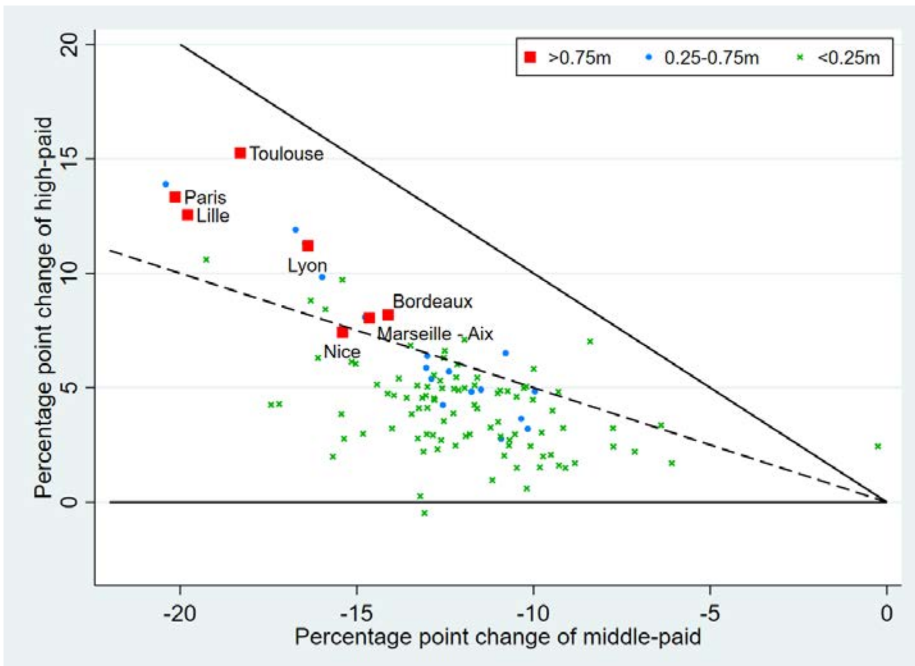
[17] We exclude here the category of CEOs - CS category 23. Firms typically report at most one CEO, if any. The CEO category is an outlier with highest average pay that has a rather constant population elasticity in sample and a share of 1% of total hours worked in 1994.

117 cities this was accompanied by a contemporaneous increase in the share of *both* low- and high-paid occupations at the city labor market level.[18]

In view of the above, we observe:

Fact 1 (Universal polarization)**.** *Over the period 1994-2015, French labor markets became more polarized in the aggregate and in nearly every individual city.*

Figure 2: Labor market polarization within cities and the great urban divergence.



This figure shows percentage point changes in employment shares of middle-paid against high-paid jobs for individual cities for the period 1994-2015. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.75m inhabitants), medium-sized (0.25-0.75m) or small (0.05-0.25m) city. Names of cities with more than 0.75m inhabitants are shown. N=117; 7 cities > 0.75m, 17 cities between 0.25-0.75m and 93 cities between 0.05-0.25m inhabitants in 2015.

3.2 Middle-paid job loss and initial exposure

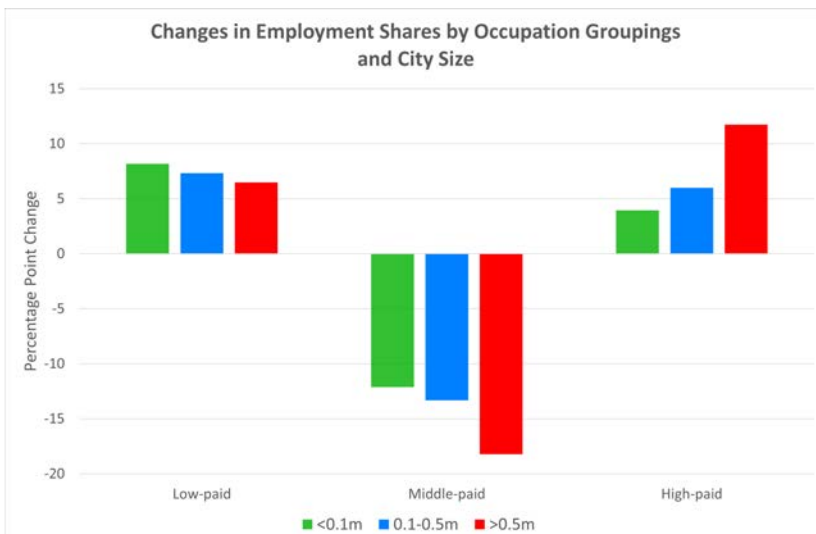
While labor market polarization was nearly universal in French cities, it was far from uniform. Figure 2 reveals a crucial role for city size in both the magnitude and composition of the shocks. Large cities (above 750,000) are represented by red squares; middle-size cities (250-750,000) by blue dots; and small cities (50-250,000) by green x's. Two key observations stand out. The first is that the typical large city has a much larger percentage point decline in

[18]The two exceptions are small cities below 60,000 inhabitants in 2015, Saint Cyprien and Salon de Provence. Given this pattern, it is unsurprising that labor market polarization is present for different groups of cities when we sum the hours worked in each job type. For example, for cities clustered into three categories: large (above >0.5m of inhabitants, 11 cities), medium (44 cities between 0.1-0.5m inhabitants) and small (62 cities between 0.05-0.1m), polarization for each group is depicted in Figure 3.

middle-paid jobs — reaching roughly 20 percentage points for Paris. The second is the nature of replacement jobs. In the figure, a city that has polarized and whose point is located above the dashed line with slope $-1/2$ has more than half of the replacement jobs in the high-paid sector (and the remainder in the low-paid sector), and vice versa if below the dashed line. Clearly the large cities have a much stronger propensity to lie above the dashed ray, and replacement jobs there are skewed toward high-paid jobs while in small cities toward low-paid jobs.

This contrast in the experience of cities of different sizes is summarized compactly in Figure 3. Cities of all sizes have large declines in the share of middle-paid jobs. But these losses are markedly stronger in large cities. Moreover, replacement jobs are primarily concentrated in high-paid jobs in the large cities and low-paid jobs in the small cities.

Figure 3: The great urban divergence and labor market polarization across three different city size groups, 1994-2015: 3 employment groups.



This figure shows percentage point changes in employment shares of high-, middle- and low-paid jobs with hours worked summed by the 3 job types and 3 city sizes: large (above >0.5m inhabitants), medium-sized (0.1-0.5m) and small (0.05-0.1m) in the period 1994-2015. Destruction of middle-paid jobs was the strongest in largest cities (18.2 pp) and weakest in smallest cities (12.1 pp). At the same time, the creation of high-paid jobs was strongest in largest agglomerations (11.7 pp) and weakest in smallest cities (3.9 pp). On the other hand, the strongest creation of low-paid jobs occurred in smallest cities (8.1 pp) while it was weakest in the cities above >0.5m (6.5 pp). The reallocation is clearly visible: nearly twice as many high-paid jobs as low-paid ones were created in the largest cities, while the reverse was true in the smallest ones.

Table 2 reveals a strong feature of the French data that may seem unexpected given prior literature. That table shows the evolution of the share of middle-paid jobs across six city sizes.[19] The second panel shows that in both 1994 and 2015 larger cities systematically had the lowest exposure to middle-paid jobs. This notwithstanding, the percentage point decline in employment shares of middle-paid occupations is greatest in the largest cities. In Paris, over the period 1994-2015, the middle-paid jobs share declined by 20 percentage points. In contrast,

[19]One can observe a similar pattern using 2-digit CS categories as shown by Figure F.10.

this decline was much lower in smaller cities, e.g. only 12 percentage points in metropolitan areas between 50 and 100 thousand inhabitants. If we consider this in proportional terms, the decline of middle-paid jobs in Paris was twice as large (31%) as in the smaller cities (15%). In short, the lower the initial exposure to middle-paid jobs, the greater the loss of those jobs.

Table 2: Share of high-, middle- and low-paid occupations in hours worked per metropolitan area size in 1994 and 2015.

High-paid							
Agglo.size	Paris	> .75m	.5-.75m	.2-.5m	.1-.2m	.05-.1m	All cities
1994	0.23	0.14	0.12	0.10	0.09	0.08	0.16
2015	0.37	0.25	0.21	0.16	0.14	0.12	0.25
change	0.13	0.11	0.09	0.06	0.05	0.04	0.10
growth in %	57	77	71	63	61	49	62
Middle-paid							
Agglo.size	Paris	> .75m	.5-.75m	.2-.5m	.1-.2m	.05-.1m	All cities
1994	0.65	0.74	0.75	0.77	0.79	0.79	0.72
2015	0.45	0.57	0.61	0.64	0.66	0.67	0.56
change	-0.20	-0.17	-0.15	-0.13	-0.13	-0.12	-0.17
growth in %	-31	-23	-19	-17	-17	-15	-23
Low-paid							
Agglo.size	Paris	> .75m	.5-.75m	.2-.5m	.1-.2m	.05-.1m	All cities
1994	0.12	0.12	0.12	0.13	0.12	0.13	0.12
2015	0.18	0.18	0.18	0.20	0.20	0.21	0.19
change	0.07	0.06	0.06	0.07	0.08	0.08	0.07
growth in %	59	48	48	55	68	64	57

This Table shows the means of shares of hours in total employment of different occupational groups in 1994 and 2015 for all 117 cities in our sample allocated in 6 bins according to city size (with Paris being a separate category), showing the percentage point changes and growth rates between 1994-2015. One observation per bin of the hours totals.

These points are underscored jointly in Tables F.15 and 3.[20] In the latter table, in the first three columns we compare the means of share changes across high- middle- and low-paid occupations in large and small cities. We observe a higher destruction overall of middle-paid jobs (by 18 percentage points on average) in cities of over 0.5m in comparison with the smallest cities (12 percentage points) even though the initial share of those jobs in total employment is lower in large cities (69% vs 78%).[21] Similar patterns are obtained for percentage changes shown in the lower panel of Table 3. This leads to the following fact:

Fact 2 (Middle-paid job loss and initial exposure)**.** *Labor market polarization led to greater*

[20] We obtain quantitatively and qualitatively similar results for these tables with other groupings of cities, for example opposing cities above 0.75m inhabitants and those below 0.25m.

[21] Appendix Table F.24 shows rank correlations between city size and middle-paid job loss are also statistically significant: Spearman's ρ and Kendall's τ between the city populations in 1990 and the percentage point changes of middle-paid jobs' employment shares over the 1994-2015 period are respectively -0.28 and -0.19.

destruction of middle-paid jobs in large relative to small cities, even though initial exposure to middle-paid jobs was lower in large cities.

This fact appears to be in tension with prominent results in the literature which focus on *initial exposure* to predict subsequent job loss, as in Autor et al. (2013). However the tension is only partial and focuses attention on the importance of precision in the specific question considered, which evolved over time. Early discussion of labor market polarization focused broadly on the loss of middle-paid jobs. Routinization and offshoring were then posited as potential mechanisms by which such jobs were affected. Most subsequent contributions thus focused on job loss among the *most* routinizable and offshorable jobs, with a notable exception of Cortes (2016). Indeed if we restrict attention to *only* the MRO jobs, we find the same pattern as Autor and Dorn (2013): MRO jobs decline more sharply where they are initially more present (see the discussion in Section F.5.1).

Table 3: Comparison of means of changes in employment shares of different occupations at the city level, cities >0.5m vs. 0.05-0.1m.

Item	high-paid	middle-paid	low-paid	middle-paid above median	middle-paid below median
Changes					
mean change, cities >0.5m	0.116	-0.181	0.065	-0.130	-0.051
mean change, cities 0.05-0.1m	0.037	-0.116	0.080	-0.073	-0.044
difference	0.079***	-0.065***	-0.015***	-0.057***	-0.008
Growth in percent					
mean growth, cities >0.5m	63.0	-26.5	54.4	-36.0	-16.0
mean growth, cities 0.05-0.1m	45.7	-14.9	62.2	-19.9	-10.2
difference in growth	17.2***	-11.6***	-7.8	-16.1***	-5.8***

Notes: 1990 population weighted, robust standard errors. N=73; 11 cities > 0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. The reported differences are coefficients in regressions of changes or growth of shares on a large city dummy. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels. Differences remain significant at least at the 1 % level without weighting or weighted by city population as of 2015 except for the difference in growth middle-paid below-median jobs for unweighted comparison. Group mean changes or growth rates are significantly different from zero at the 1 % level.

However we think it is important to return to the broader question of middle-paid job loss that has motivated the literature, as in Acemoglu and Autor (2011). As Figure 1 makes clear, there are occupations (e.g. CS 46 mid-level professionals) that are important to aggregate polarization yet do not fall into the MRO category. Two issues loom large here. One is that the routinization and offshoring scores are an index that doesn't provide a sharp characterization of how much more routinizable or offshorable are occupations with higher indices. A second issue is that vertical relations of occupations may imply that an occupation that may be very hard to offshore or routinize on its own may be pulled abroad or become unnecessary if a vertically related occupation moves or is routinized. In short, these factors suggest that even if offshoring and routinization are driving forces, we will want to move beyond looking only at

MRO occupations alone and consider the broader concern of middle-paid job loss.

3.3 Skewed middle-paid job loss

We investigate now the contrasting experience of large and small cities regarding middle-paid job loss. We split middle-paid occupations between those below versus above the median middle-paid wage. This will help us to understand if the contrasting experience of large and small cities arises due to job loss skewed toward one or the other margin in each.[22]

The results appear in Figure 4 with precise numbers and percentage changes in Table 3. This is identical to Figure 3 except for the new division of middle-paid occupations into two groups at the median wage. The new information is that the greater decline in middle-paid jobs in large cities is strongly concentrated in the upper tier of middle-paid jobs. And this provides a new perspective on the contrast in experience between large and small cities. For the largest cities, there is a 12 percentage point rise in high-paid jobs, which is only slightly less than the 13 percentage point decline in the upper-tier middle-paid jobs. By contrast, in the smallest cities, there is only a 3.7 percentage point growth of high-paid jobs, even as there is a 7.3 percentage point decline in these upper-tier middle-paid jobs.[23] The contrasts by city size are much more modest for the lower-tier of middle-paid jobs. Similar conclusions can be drawn when examining percentage changes across the different categories.

Thus, the following fact holds:

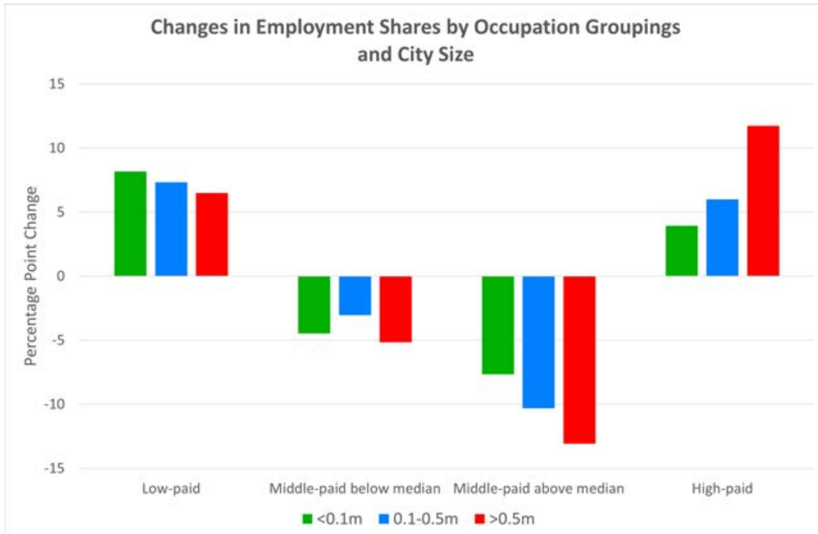
Fact 3 (Skewed middle-paid job loss)**.** *The greater loss of middle-paid jobs in large cities is due to the greater destruction of the upper tier of middle-paid jobs in comparison to small cities.*

Our conclusion is that an understanding of the variety of experience of cities of different sizes in the presence of offshoring or routinizability shocks driving labor market polarization needs to go beyond simple measures of exposure to the most offshorable or routinizable occupations. In particular, it is crucial to consider the heterogeneity of occupations within the middle-paid jobs. All cities lost jobs in roughly equal percentage points in the lower-paid, most routinizable and offshorable middle-paid jobs. But the contrasting differences across cities come in the upper tier of middle-paid jobs, where there are much sharper declines in larger cities.

[22] Connecting our work to prior discussions, we can divide the changes in middle-paid jobs into the most routine and offshorable (MRO) and other middle-paid (OMP) jobs as well. The most prominent feature of prior work (e.g. Autor and Dorn (2013)), that high MRO job loss occurs where initial exposure to MRO jobs is high, is also present in our data. At the same time, however, the initial exposure to MRO jobs is neither a good predictor of loss for other middle-paid (OMP) jobs nor for the set of middle-paid jobs taken as a whole. Neither is it true that aggregate exposure to middle-paid jobs correlates with a greater loss of all middle-paid jobs. Detailed discussions are in Sections F.5.1-F.5.2; see also Figures F.12 and Tables F.37-F.39 there.

[23] In Figure F.9, we provide further divisions of the middle-paid jobs depending on their wage. These regularities are also evident in rank correlations between city-level population and changes in occupational employment shares (Table F.24).

Figure 4: Labor market polarization and the great urban divergence across three different city size groups, 1994-2015: 4 employment groups



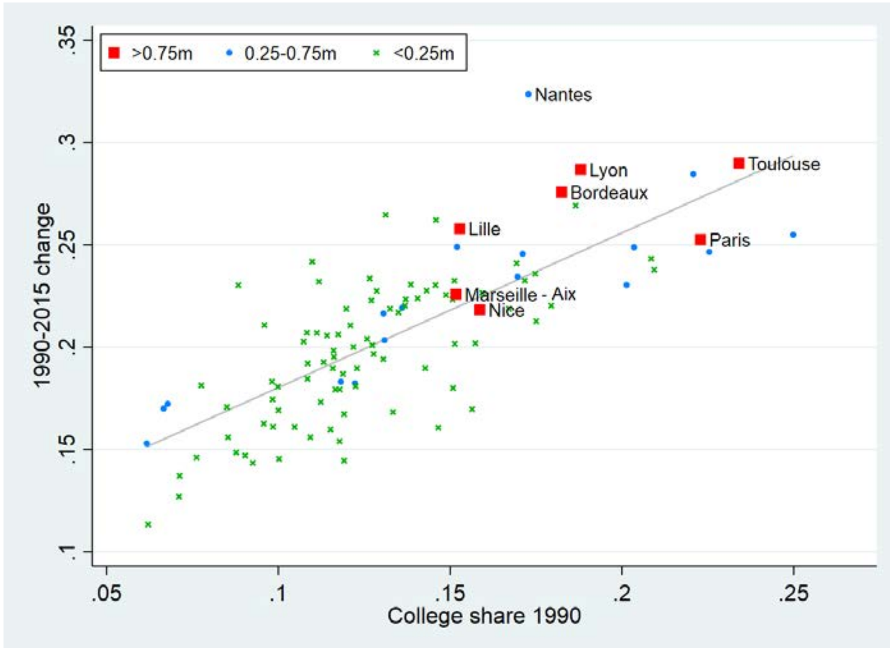
This figure shows percentage point changes in employment shares of high-, low- and different types of middle-paid jobs with hours worked summed by job types and 3 city sizes: large (above >0.5m inhabitants), medium-sized (0.1-0.5m) and small (0.05-0.1m) in the period 1994-2015. The bars for high- and low-paid jobs are exactly as in Figure 3. The division of middle-paid occupations is between the middle-paid jobs with average wages in 1994 above the median (CS 48, 46, 47, 43 and 62 in decreasing wage order) and those below the median (CS 54, 65, 63, 64, 67).

3.4 The Great Urban Divergence

We begin our documentation of the Great Urban Divergence by examining skill as educational attainment. Following the original formulation of Moretti (2004), Figure 5 plots a city's change in the share of college graduates against the initial college share, here for 117 French cities. Our data cover a longer period (1990-2015) than Moretti's data, but clearly confirm that the higher the initial college share, the higher the rise in that share. The figure also separates out cities of different sizes. While the ordering is not strict, this shows a strong relation between city size, initial skill share and the change in that share, whereby the large cities increasingly pull away from smaller cities. This evidence is consistent with our previous discussion on job share evolution as high-paid occupations and even many above median-paid occupations (e.g. CS 46 or CS 47) require college degrees. In short, by the traditional measure, the Great Urban Divergence is powerfully evident in the French data.

We can also examine this in our jobs data. This can be done either by comparing the evolution of high- versus low-paid jobs in large and small cities, or by considering all jobs divided at the median occupational wage. The contrasting evolution for high- vs. low-paid jobs is examined in the detailed data for six city groups in Table 2 (first and third panels respectively). This Table shows that the percentage point increase in high-paid jobs is monotonic in metropolitan area size, as is the initial share of high-paid jobs in total employment. In Paris and cities

Figure 5: The Great Urban Divergence in Skills in France.



The figure graphs the initial share of college graduates among the working age population in 1990 and the change in this share over the period 1990-2015 (both census years) at the individual city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.75m inhabitants), medium-sized (0.25-0.75m) or small (0.05-0.25m) city. Names of cities with more than 0.75m inhabitants are shown, as well as Nantes that is a city above 0.5m inhabitants and had the largest change of college graduates share. N=117; 7 cities > 0.75m, 17 cities between 0.25-0.75m and 93 cities between 0.05-0.25m inhabitants in 2015. All 117 of the largest cities in France experienced an increase in the share of college graduates over the period 1990-2015 by 10.5 pp on average. Linear slope of the relationship between the share of college graduates in 1990 and increases in the share of graduates is 0.76, statistically significant at 1%.

above 0.75m inhabitants the increase in such occupations is above 10 percentage points over the period 1994-2015. The smallest cities (0.05-0.1m of population) have the lowest gain, less than 4 percentage points. Although the variation is more modest, the percentage point increase in low-paid jobs is higher for smaller cities. These observations are reinforced by comparisons between large and small cities, as in Table 3 both for percentage point and percentage changes, and individual city evidence in rank correlation tests as reported in Online Appendix Table F.24. It can also be visualized directly in Figure 2. The large cities tend to have experienced a larger increase in the share of high-paid jobs, indicated by these cities' red squares primarily being above the dashed line.

When quantifying the difference between Paris and the smallest cities in Table 2 or large and small cities in Table 3, one would find that, in the large city, for every middle-paid job destroyed, 2/3 of the replacement jobs will be created in the high-paid sector and 1/3 in the low-paid sector. In small cities, the proportions are reversed, with 1/3 created in the high-paid sector and 2/3 in the low-paid sector.

Moreover, high-paid jobs increase by more in the large cities that feature an initially larger share of these jobs (and, at the other end of the spectrum, low-paid jobs increase by more in small cities that had them initially in higher proportion). Table 2 provides the basic relation. This is further jointly confirmed in Tables F.15 and 3 and summarized in Figure 3.[24]

The central contrast illustrated above would also be present if we had instead partitioned jobs into two groups divided according to the median occupational wage, between office workers (CS 54) and skilled industrial workers (CS 62). In this case, Figure 4 illustrates that the large compared to small cities grow relatively more in the above median-paid occupations.

Overall, all of this evidence points in the same direction. The creation of high-paid jobs is increasing with a city's size, even as they were initially more present there. These observations are consistent with the great urban divergence as described by Moretti (2012) and subsequent literature. We thus obtain:

Fact 4 (Great Urban Divergence). *New job growth is skewed in larger cities to high-paid jobs (where their share was already higher) and to low-paid ones in smaller cities.*

3.5 Can existing theories account for the four key facts?

Can Facts 1 to 4 be explained by existing theories of labor market polarization or the great urban divergence?

First, while the foundational model of Autor and Dorn (2013) is both an inspiration and an input for our work, it cannot explain our facts. Formally, even the aggregate polarization of jobs portion of Fact 1 is not possible in their setting since the total number of skilled/abstract workers is fixed in their model. Their model is geared rather to explain only a part of aggregate polarization, the shift of (middle-paid) routine workers to manual occupations. Their model does allow some cities to experience the growth in skilled jobs necessary for polarization, since it allows these workers to be mobile. However, with aggregate skilled workers fixed, if some cities have growth of skilled employment, others must have declining skilled employment. And in their long run all skilled workers move to a single city, so that all other cities are losing skilled jobs. That is, their model cannot account for the fact of near-universal expansion of skilled jobs, the other component of Fact 1. We run into problems again when we consider Facts 2 to 4. The Autor and Dorn (2013) model is scaleless and all of these facts require observing contrasts between large and small cities. In their model, the only fundamental source of variation across cities is in the Cobb-Douglas share of routine vs. skilled labor in goods production. A routinization drop in the price of computer capital would lead to a large drop in the number of routine jobs and to growth of skilled jobs precisely in those locations initially abundant in those routine jobs (in their model accounting for the entire set of middle-

[24] In Online Appendix F.3.4 we further discuss initial occupation shares across cities. See also Table F.35 we demonstrate from Census data that such a divergence also occurred in educational outcomes over the years 1990-2013.

skill jobs) rather than, as in the data, where these jobs were initially scarce per Fact 2.[25] In their model, all loss of routine middle-paid jobs in all cities is at the low-paid margin with routine jobs, so cannot account for Fact 3. In short, the foundational model in the literature accounts for none of our four key facts.

Second, there are existing models that predict aggregate labor market polarization, such as Cortes (2016). This allows them to capture the aggregate aspect of Fact 1. However they do not feature multiple cities, so fail to capture the city-level universal polarization of Fact 1 nor any of the characteristics of Facts 2 through 4, which rely on cross-city features.

Third, there is a model, in Cerina et al. (2022), that uses an extreme skill complementarity framework to contrast evolutions in a large and small city. However, in their approach the aggregate supply of each of the three skill types is fixed. As a result, they cannot account for aggregate polarization at all. Moreover, in their central two-city setting, if one city polarizes, then the other city must de-polarize, i.e. have the relative employment of high- and low-paid workers decline. The model accounts for none of our four facts.

Finally, existing theories of the great urban divergence focus on skill-biased technological change (SBTC). By design they explain Fact 4. However the very fact that there are no middle-paid jobs in these models means that the rich set of facts developed here is simply impossible to address in their framework. They cannot explain Fact 1, universal polarization, why nearly every single city experiences labor market polarization. They cannot explain Fact 2, middle-paid job loss and initial exposure, since these features can't exist in their setting. They cannot explain Fact 3, skewed labor market polarization. Since they can't discuss middle-paid jobs, there is no prospect of explaining *which* middle-paid jobs will be lost in large and small cities. Following the logic of Costinot and Vogel (2010), the skill-biased technical change central to the existing great urban divergence literature should lead to a pervasive growth in wage inequality rather than the wage polarization documented in Table 1 and discussed in Appendix F.2.1.[26]

[25]Other theories of the effects of technology on local labor markets (see Acemoglu and Restrepo, 2020, among others) also explain differences in the magnitude of middle-paid jobs losses as a (positive) function of initial exposures to these jobs. This fails to explain Fact 2.

[26]More recent contributions to the great urban divergence literature also have a role for e.g. automation shocks that eventually function like skill-biased technical change. For example, Eckert et al. (2022) build an interesting model in which a drop in ICT capital price leads to a demand for high-skilled labor due to a non-homothetic CES production function. As this stronger demand for high-skilled work is biased towards the more productive cities that attract such workers, the drop in the ICT capital price leads more productive cities to become even more high-skilled, consistent with Fact 4. But, like the other literature on the great urban divergence, this model does not feature middle-paid jobs, so cannot speak to the issue of polarization either in the aggregate or across cities.

4 Model of universal polarization

Why do middle-paid jobs disappear more in *initially less* exposed areas? Why is the decline in middle-paid jobs greater in large cities? Why are these losses skewed toward the upper-tier of middle-paid jobs in large relative to small cities? Is the Great Urban Divergence connected to Labor Market Polarization in the sense that a single shock may produce both? How are these facts connected with the observed patterns on wages?

To answer these questions, this section builds a model integrating the core framework of job polarization in Autor and Dorn (2013) with the system of cities model of Davis and Dingel (2020).[27] As in Autor and Dorn (2013), our model features middle-paid jobs that are relative substitutes with capital and/or offshored tasks. As in Davis and Dingel (2020), agents can decide where to live and in which sector they work. Our objective is to obtain as a spatial equilibrium outcome both the distribution of skills and jobs as well as their evolutions with respect to a polarization shock so as to match the facts that we have documented above. Overall, we find that, consistent with the data, spatial equilibrium does not necessarily lead to exposure-driven explanations of labor market evolutions and that shocks other than polarization are not necessary to generate the Great Urban Divergence.

More precisely, when the price of capital or offshored tasks decreases, we first find that polarization of the job market occurs both in the aggregate and in each city. Furthermore, when we match the model with productivity and sorting patterns (reported in Appendix F.3), we find that labor market polarization in the large city is biased in favor of high-paid jobs and leads to more destruction of middle-paid jobs despite an initially lower exposure to middle-paid jobs.

4.1 The environment

Let us consider an economy populated by households that provide heterogeneous labor, consume, and decide where to live and work. Households consume housing services and a final good that is produced using labor and a capital/offshoring good.

Locations. The set of cities is $c \in \{1, 2\}$.[28] In each city, there is a continuum of locations $\tau \in [0, \infty)$. τ denotes the distance from an ideal location inside a city. This can be interpreted in a variety of ways, including as commuting distance to a central business district or as remoteness from the core of a productive cluster with positive but spatially decaying spillovers. As will become clear, having multiple locations within a city allows us to introduce a trade-off

[27] The model features a long run full spatial equilibrium in which all workers freely choose a sector of production and a location. We thus look at comparative steady states rather than transition dynamics, which are beyond the scope of this paper.

[28] We extend our framework to N cities in Appendix B.3.

between living in a better location in a smaller and less productive city or in a worse location in a larger and more productive city.[29]

In each city c, we assume that the supply of locations $\{t|t \leq \tau\}$ is $S(\tau)$ with $S(0) = 0$, $S(.)$ strictly increasing and twice continuously differentiable.

Households. They consume a single final good and 1 unit of housing. Each household inelastically provides 1 unit of labor. Households have different innate skills that we denote by ω, where ω is distributed on $[\underline{\omega}, \overline{\omega}]$ with a pdf $n(.)$.[30]

Households freely choose where they live (the city c and the internal location $\tau \geq 0$). We denote the rental price of location (c, τ) by $r(c, \tau)$. We use the price of the final good as the numeraire and we normalize the price of unoccupied locations to 0 so that $r(c, \tau) \geq 0$. Locations are owned by absentee landlords who spend their rental income on the final good.

Households can also decide in which sector σ they work. Finally, we denote by $f(\omega, \sigma, c, \tau)$ the endogenous pdf of the distribution of households ω across sectors σ and locations $f(c, \tau)$.

Production. Production in this economy involves different sectors: final goods are produced out of intermediate goods $\{h, m, l, Z\}$. Goods $\{h, m, l\}$ are produced with labor and the capital/offshoring intermediate good Z is produced with, or traded for, the final good. All goods are traded with zero transportation costs except non-traded housing.

Final goods. They are produced by a continuum of identical competitive firms using intermediate goods $\{h, m, l, Z\}$. The production function of the representative firm is:

$$Q = \left(a(h)q(h)^{\zeta} + \left(a(m)q(m)^{\frac{1}{\theta}} + a(z)Z^{\frac{1}{\theta}}\right)^{\zeta\theta} + a(l)q(l)^{\zeta}\right)^{1/\zeta} \tag{1}$$

where $q(j)$ and $p(j)$, $j \in \{h, m, l\}$, are the quantity and the price of intermediate good j, p_z is the price of capital and/or an offshoring intermediate input with the rest being technological parameters that we assume to be fixed.

As in Autor and Dorn (2013), we assume that capital/offshoring goods Z are relative substitutes with intermediate goods produced by the middle-paid sector (m) but they are relative complements with the the intermediate goods produced by the high-paid (h) and low-paid sectors, that is $\zeta < \theta$.[31] In contrast with Autor and Dorn (2013), there is only one final good production function in the aggregate and no local ones: this implies that there are no local

[29] These locational choices will affect workers' productivity. For further interpretations of the location τ and the connection with other models of cities in the literature, see Davis and Dingel (2020).

[30] The model provides a description of long-run outcomes. From this perspective, innate skills are assumed to be time-invariant. This does not mean however that the distribution of actual skills/human capital is unchanged as change in sectoral composition may require different sets of skills – the comparison between long-run equilibrium allocations is then a relative one.

[31] This assumption implies no loss of generality as our results can be extended to any situation where a decline in the price of the capital goods leads to an increase in the relative prices of low- and high-paid sectors' inputs, $p(l)/p(m)$ and $p(h)/p(m)$. As we detailed in the previous section, our model also contrasts with Autor and Dorn (2013), as we eliminate immobility across locations, for unskilled labor, or sectors, for skilled labor.

complementarities either through production or demand between the low-paid sector and the rest of the economy of the city.

As we are using the final good as numeraire, the profits of the representative firms can be written as $\Pi = Q - p(h)q(h) - p(m)q(m) - p(l)q(l) - p_z Z$.

Intermediate goods. The intermediate goods $\{h, m, l\}$ are produced with a constant returns to scale technology using only labor. There is one sector to produce each of the $\{h, m, l\}$ goods. We label sectors by $\sigma \in \{h, m, l\}$ where h stands for high-paid, m for middle-paid and l for low-paid. We assume there is perfect competition in all three sectors, so that in each sector the wage per efficiency unit of labor equals the price of the intermediate good $p(\sigma)$.

A household with skill ω, living in city c and in a location τ has a productivity:

$$A(\sigma, c)H(\omega, \sigma)T(\tau) \tag{2}$$

That the location within city directly enters in productivity has several interpretations. The first interpretation follows traditional von Thunen models, which treat $T(\tau)$ directly as commuting time costs. The second one follows Arzaghi and Henderson (2008) who showed how locations offer heterogenous levels of productivity. The third one follows Xiao et al. (2021) that gets causal estimates of the productivity cost of commuting. They look at examples where a firm changes its location within a metro area. Since inventors have a location, some see their commutes made longer and others shorter, and they can see the impact on patenting, which is negative for longer commutes. We make the following assumptions on households' productivities:

Assumption 1 (Within-city productivity)**.** *$T(.)$ is a decreasing function, with $T(0) < \infty$, identifying the cost in productivity of being remote from the most productive location in a city.*

Assumption 2 (Absolute and comparative advantage of households)**.** *Higher-skilled households (with a high ω) have an absolute advantage in all sectors, i.e. $H(., \sigma)$ is increasing.*

Higher-skilled agents have a comparative advantage in higher-paid sectors, i.e. $H(\omega, \sigma)$ is log-supermodular in (ω, σ).

Assumption 3 (Absolute and comparative advantages of cities)**.** *City 1 has an absolute advantage in all sectors: $A(j, 1) > A(j, 2)$ for $j \in \{l, m, h\}$ and a comparative advantage in higher-paid sectors: $A(h, 1)/A(h, 2) > A(m, 1)/A(m, 2) > A(l, 1)/A(l, 2)$.*

Assumptions 2 and 3 are consistent with the evidence on productivities in Appendix F.3.[32]

[32]Nearly all urban models feature an absolute productivity advantage for larger cities, helping to explain the ubiquitous urban wage premium. A *comparative productivity advantage* of larger cities is a more novel element, but one that can be examined in the data. The relative productivities across large vs. small cities can be inferred from the sector-by-city fixed effects $\gamma_{p_o A_{co}}$ in within regressions (equation (32)) of individual wages on

Capital good/offshoring intermediate good. The intermediate good z is produced by transforming final goods using the following technology $Z = \frac{1}{\xi}q$, where q is the amount of final goods used and ξ is a technology parameter. Perfect competition implies $p_z = \xi$.

The intermediate good z has two interpretations. The first is that it is a capital good that substitutes for middle-paid labor as in Autor and Dorn (2013). Note that, as in Autor and Dorn (2013), this capital good would fully depreciate with production. With this view, ξ is a parameter that governs the efficiency of producing the capital good. The second interpretation is that Z is an imported intermediate and ξ is the terms of trade. As a result, a drop in p_z could be either due to routinization, a drop in the price of computer capital, or due to offshoring, a drop in the domestic price of the intermediate import due to technical progress abroad or the removal of trade barriers.

4.2 Household decisions

Let us first investigate location and sector decisions by agents and how these decisions depend on factor prices, $p(l)$, $p(m)$ and $p(h)$. The utility flow obtained by an agent with skill ω, location decisions (c, τ) and intermediate good sector σ is:

$$A(\sigma, c)H(\omega, \sigma)T(\tau)p(\sigma) - r(c, \tau) \tag{3}$$

We are interested in understanding in which city and in which sector a household with skill ω decides to work, that is, how the household maximizes (3) with respect to c, τ and σ.

Sectoral decisions. In each city c, we can define two thresholds $\omega(m, c)$ and $\omega(h, c)$:

$$A(m, c)H(\omega(m, c), m)p(m) = A(l, c)H(\omega(m, c), l)p(l) \tag{4}$$

$$A(h, c)H(\omega(h, c), h)p(h) = A(m, c)H(\omega(h, c), m)p(m) \tag{5}$$

The threshold $\omega(m, c)$ is such that a marginal household in a given city c is indifferent between working in the low- and middle-paid sectors. Similarly the threshold $\omega(h, c)$ is such that a marginal household in a city c is indifferent between the middle- and high-paid sectors.

time-varying worker characteristics and fixed effects, time effects and sector-by-city fixed effects in the DADS-Panel data for 1993-1995 (see Appendix F.3 for details). When taking ratios within sectors and across cities, price terms cancel out, so we obtain relative productivities. For example, for cities with population above 0.5m versus cities from 0.05-0.1m, the relative productivities for high-, medium-, and low-paid sectors respectively are (1.086, 1.059, 1.037). Estimated productivity values will inform our simulations in Section 5. More detail is in Appendix F.3. For our theory, we simplify by making absolute and relative producitivities exogenous. Davis and Dingel (2020, 2019) provide alternative approaches to endogenous absolute productivity differences across cities in a symmetry breaking setting. In the Appendix, we provide a way to obtain endogenous productivity differences consistent with the patterns of labor market polarization. Comparative productivity advantage by sector may arise simply when there is skill sorting and the relative supplies of skill types also affect sectoral productivity. Absolute and relative productivity differences could also arise from first nature differences.

These thresholds do not depend on the location τ as the productivity term $T(\tau)$ is separable.

The following proposition shows that these two thresholds are sufficient for characterizing sectoral decisions by households, breaking the skills into three intervals according to the intermediate sector those skills specialize in:

Proposition 1. *A household living in city c and with skill ω works in sector l when $\omega \leq \omega(m, c)$, in sector m when $\omega \in (\omega(m, c), \omega(h, c))$ and in sector h when $\omega \geq \omega(h, c)$.*

Across cities, these thresholds satisfy: $\omega(h, 1) < \omega(h, 2)$ and $\omega(m, 1) < \omega(m, 2)$.

Proof. See Appendix A.1

The ordering of these cutoffs in Proposition 1 will play a key role in discussions to follow regarding both levels and changes in the distribution of job types across cities, so it is important to grasp why these arise.

The differences in the thresholds across cities result from the comparative advantage of the larger cities in higher-paid sectors associated with the increasing importance of individual skills in higher-paid sectors. For given prices of intermediate goods, the same individual is relatively more productive in higher-paid sectors in the larger city and, thus, has more incentive to work in these sectors. Accordingly, in the larger city the least skilled worker in the high-paid sector is less skilled than the counterpart in the smaller city. A similar ranking holds for the least skilled worker in the middle-paid sector between the two cities. In the language of Costinot and Vogel (2010), the comparative advantage of the larger city in higher skill sectors leads to skill downgrading/task upgrading, the difference being that here this occurs with perfect factor mobility between locations.

Note that, in principle, it is possible that a sector does not exist in at least one of the two cities, even though the production function guarantees that this sector will exist in at least one city. This happens, for example, when $\omega(m, 1) \leq \underline{\omega}$. In this case, there is no low-paid sector in City 1. In what follows, we focus on situations where all three sectors are active in both cities.

In the end, Proposition 1 defines a function M such that $M(\omega, c)$ is the optimal sectoral decision for a household with skill ω in city c.

Location decisions. Let us now turn to location decisions. First note that a household with skill ω decides to work in city 1 and in location τ only if it is not better off working in the other city or in any other location τ', that is:

$$\max_{\sigma,\tau} A(\sigma, 1)H(\omega, \sigma)T(\tau)p(\sigma) - r(1, \tau) \geq \max_{\sigma',\tau'} A(\sigma', 2)H(\omega, \sigma')T(\tau')p(\sigma') - r(2, \tau'). \tag{6}$$

When this holds with equality the skill ω is present in the two cities. Using the results of Proposition 1, we can connect location decisions with the sectoral decisions and show that more skilled workers choose more attractive locations in each city.

Proposition 2 (Sorting within cities)**.** *In each city c, there exists $\overline{\tau}(h,c)$ and $\overline{\tau}(m,c)$ satisfying $\overline{\tau}(h,c) \leq \overline{\tau}(m,c) \leq \overline{\tau}(c)$ such that: if $\omega \geq \omega(h,c)$ then $\tau \leq \overline{\tau}(h,c)$, if $\omega \in [\omega(m,c), \omega(h,c)]$ then $\tau \in [\overline{\tau}(h,c), \overline{\tau}(m,c)]$, if $\omega \leq \omega(m,c)$ then $\tau \in [\overline{\tau}(m,c), \overline{\tau}(c)]$.*

In particular, $f(\omega, \sigma, c, \tau) = 0$ for all ω, σ, c and $\tau \geq \overline{\tau}(c)$, so that $\overline{\tau}(c)$ defines the limits of the occupied area of city c.

Proof. See Appendix A.3

Locations across cities. Let us now investigate how workers decide to locate between the two cities. We first show that, in equilibrium, locations occupied by the same skill ω have the same price $r(c,\tau)$.[33]

Due to productivity advantages, as per Assumption 3, there are locations in City 1 that are strictly more attractive than even the best location in City 2. Correspondingly, there will be a set of skills attracted to those locations in City 1 that will not locate at all in City 2. So long as these productivity advantages are not too large, City 2 will be occupied and it will have a maximum skill in the city $\bar{\omega}(2) < \bar{\omega}$.

Below the skill $\bar{\omega}(2)$, for each ω and for each τ, there exists $\tau' < \tau$ such that the productivities in City 1 and in City 2 are the same:

$$\begin{aligned} &A(M(\omega(\tau),1),1)H\left(\omega(\tau), M(\omega(\tau),1)\right)T(\tau)p\left(M(\omega(\tau),1)\right) = \cdots \\ &\cdots A(M(\omega(\tau),2),2)H\left(\omega(\tau), M(\omega(\tau),2)\right)T(\tau')p\left(M(\omega(\tau),2)\right). \end{aligned} \tag{7}$$

For $\omega \in [\underline{\omega}, \bar{\omega}(2)]$, households are indifferent between a less desirable location in the more productive and larger City 1 or a more desirable location in the less productive and smaller City 2. Figure 6 summarizes the distribution of skills across sectors and cities.[34]

Size and skill composition of cities. Finally, we show that location decisions associated with the assumptions on the productivity advantage of City 1 leads to:

Proposition 3. *City 1 is larger and contains a larger set of skills than City 2.*

Proof. See Appendix A.4.

As a consequence of this Proposition, we will refer to City 1 as the large city and City 2 as the small city, or respectively as the more-skilled city or less-skilled city.[35]

[33] See Appendix A.2 for more details on the intermediary steps on the results in this paragraph.

[34] Consistent with this, we show in Appendix Section F.3.5 that workers' transitions to larger cities are skewed to employment in better jobs, and vice versa for job transitions to smaller cities.

[35] The complementarity between skill and city productivity gives rise to a log-supermodular distribution of skills across cities. An implication of log-supermodularity is that the population elasticity with respect to city size is increasing in skills. Davis and Dingel (2020) find log-supermodularity of skills on US data. For our French data, the results are shown in Figure F.6 and discussed in Appendix F.3.3. There we see an ordering of the population elasticities of skills, with the two lowest skill groups having an elasticity statistically significantly below 1; two middle skill groups (with high school diplomas and some college) having an elasticity insignificantly different from 1; and a high skill category of workers with a graduate diploma that has a significant population elasticity of 1.18. In this sense, France's larger cities are relatively more skilled.

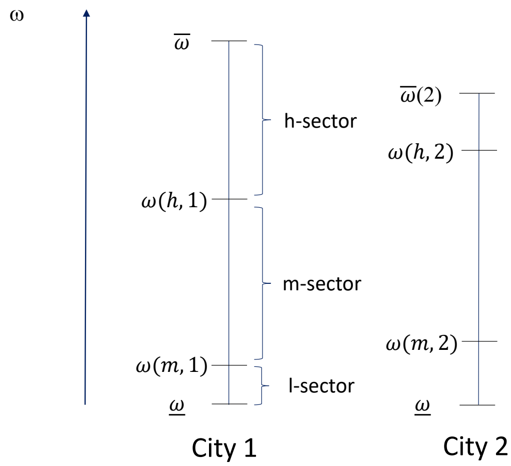


Figure 6: Skills, sectors and cities in equilibrium.

This figure depicts the equilibrium skill range and sectoral choice for individuals as a function of their skill in the large (City 1) and the small city when all sectors are present in both cities. All skill types are present in City 1. However, the small city lacks the most skillful agents with $\omega \in [\bar{\omega}(2), \bar{\omega}]$ who choose all to reside in City 1. In both cities, more able agents choose higher-paid sectors. Because of the assumptions about the absolute and comparative advantage of City 1 in higher-skill sectors, the skill thresholds for agents to choose the high- or middle-paid sectors are lower in the larger city: $\omega(h,1) < \omega(h,2)$ and $\omega(m,1) < \omega(m,2)$. If the absolute advantage of the high-skill sector in City 1 is large enough, the share of workers in the high- (middle-) skill sector will be higher (lower) in the larger city.

4.3 Universal polarization

We first show that a decline in the price of the capital/offshoring good p_z leads to polarization both in the aggregate and across cities. We will thus refer to this as a labor market polarization shock.

A relative price decline. We can investigate how a decrease of the price of the intermediate good z affects the distribution of jobs in our economy, as in Autor and Dorn (2013). To start, let us clarify the effect of a shock to the price of capital/offshoring intermediate goods on the relative prices of the middle-paid sector with the high- and low-paid sectors:

Proposition 4. *A decline in p_z leads to a decline of the relative prices of the middle-paid sector good relative to others, i.e. $p(m)/p(h)$ and $p(m)/p(l)$ fall.*

Proof. See Appendix A.5.

Using this pattern of relative prices, we can investigate how the shock to p_z affects the labor markets in the two cities.

Aggregate and Universal polarization. We now observe how this decline of the price of capital affects labor markets overall and in each city. As middle-paid jobs decline at both margins in both cities, we can infer the following proposition:

Proposition 5 (Aggregate and Universal polarization)**.** *A decline in p_z reduces the share of middle-paid jobs in the aggregate and in each city, while the shares of low- and high-paid jobs are at least weakly increasing.*

Proof. See Appendix A.6.

This result matches Fact 1 that documents such universal polarization for France from 1994 to 2015. As in Autor and Dorn (2013), a decline in the price of capital goods/offshoring intermediate goods leads firms to substitute middle-paid jobs by capital. Here this leads workers to reallocate, either to the high-paid or to the low-paid sectors, depending on workers' skills and, overall, the labor market becomes more polarized.[36] Proposition 5 implies in the context of our model also wage polarization in the aggregate (see Table 1 and Appendix F.2.1) and at the city level.

Importantly, this reallocation and the resulting polarization occur not only in the *aggregate* but also in *each* city. In addition, we obtain this conclusion in a spatial equilibrium context where all workers, no matter their skills, are free to move. Indeed, obtaining universal polarization in a spatial equilibrium setting is not obvious. In a model without labor mobility, polarization in any place immediately follows from polarization in the aggregate.[37] With free mobility of workers, as assumed in a spatial equilibrium context, the reallocation of workers in response to the shock could lead to polarization only in a subset of places. For example, in the spatial equilibrium model of Autor and Dorn (2013), only high-skilled workers can move and they migrate to the region where production is the most intensive in the routine task. This implies in their model that there is labor market polarization in only one city and a *decrease* in high-skill jobs everywhere else (see in particular pp.12 and 13 in Online Appendix F).

5 Patterns of polarization across cities in the calibrated model

We now study whether our model leads to patterns of polarization consistent, at least qualitatively, with the one we observe in the data. In our model, if any polarization shock should lead to universal polarization, the exact patterns of polarization depend on the relative productivities of households across sectors and cities.

In this section, we first match productivities in our model with data on wages in 1994. We then simulate the model in reaction to a relative decline of the price of the middle-paid good. Consistent with Facts 2 to 4, we show that, with such data-based calibration, 1) middle-paid jobs are both initially *relatively less abundant* and *decline relatively more sharply* in large

[36] Cortes (2016) studies labor market polarization in the aggregate economy in a model with occupational sorting driven by the comparative advantage of higher skilled in more complex tasks as in Gibbons et al. (2005). Three occupational groups ranked by ability (non-routine manual; routine and non-routine cognitive) are taken into account. As a result of increased automation, those with highest ability switch to non-routine cognitive jobs while those with low ability switch to non-routine manual jobs. These predictions are borne out in PSID data. Indeed those with highest skills switch into non-routine cognitive occupations the most.

[37] In this case, the intensity of polarization in a given place will stem from this place's exposure to the polarization shock. In Appendix D.1, we show that exposure is not necessarily the key predictor of labor market evolution in a spatial equilibrium context.

cities; 2) the middle-paid jobs disappearing in large cities are higher-skilled compared with those disappearing in small cities; and 3) the resulting creation of high-paid jobs between large and small cities leads to the great urban divergence, with large cities becoming relatively richer in high-paid jobs.

We stress that our results are developed as comparative statics in a fully frictionless model and are silent about the source of adjustments — whether coming from upgrading (downgrading) to higher (lower) paid jobs by workers initially in middle-paid or routine occupations (e.g. studied by Cortes, 2016; Cortes et al., 2017; Keller and Utar, 2023) or by entry of new and exit of old cohorts as well. This seems a valuable first approach given the long horizon, 1994-2015, we aim to understand.[38]

Matching data on wages with productivities in the model. We first match productivities in the model with data on wages (shown in Appendix F.2) and with the aggregate distribution of jobs across sectors in 1994 (see Appendix C for more details about this matching as well as on the algorithm to simulate the model).

The wage distribution is the first observable that we use. To connect this distribution to our model, we assume that the observed log hourly wage of an individual i is:

$$\log w_i = \log A(\sigma, c)p(\sigma) + \log H(\omega, \sigma) \tag{8}$$

where (ω, σ, c) are, respectively, the skill, the sector and the city of individual i.[39]From this specification, we can regress (equation (32) in the Appendix) log hourly wages on city/sector fixed effects. This allows us to obtain, for each $\sigma \in \{l, m, h\}$:

$$\log A(\sigma, 1)p(\sigma) - \log A(\sigma, 2)p(\sigma) = \gamma_{p_\sigma A_1 \sigma} - \gamma_{p_\sigma A_2 \sigma}$$

where $\gamma_{p_\sigma A_c \sigma}$ is the fixed effect of sector σ in city c. As sectoral prices $p(\sigma)$ cancel out, we can identify relative productivities $A(1, \sigma)/A(2, \sigma)$ from this equation. We contrast large cities of more than 500,000 inhabitants and small cities that are between 50,000 and 100,000 inhabitants. Productivity parameters $A(\sigma, c)$ consistent with data (reported in Table F.8). We then calibrate

[38]Of course, labor markets are not frictionless anywhere and certainly not in France. Unions, seniority, the timing and horizon for human capital investments are all important frictions. In the context of our model, this would suggest that the burdens and benefits of adjustment accrue to the young more than to the old. A simple exercise confirms this. Let the young be comprised of individuals aged 25-34 and the old individuals aged 55-64; let large cities be those above 0.5 million population and small cities be those with 50-100 thousand inhabitants. For these two groups, the ordering of the percentage points of growth or decline of our three types of jobs over the sample period are the same as the aggregates we have documented. But for every category, the absolute magnitude of the changes is larger for the young than the old (see Appendix Table F.33). This suggests that the model is valuable as a description of aggregate changes but that future work should also investigate in greater detail the transition path.

[39]Implicitly, this means that we do not take into account the term $T(\tau)$, consistently with the interpretation that $T(\tau)$ reflects commuting time costs.

the initial prices $p(\sigma)$, $\sigma \in \{l, m, h\}$ to match the aggregate shares in 1994 of the the low-, middle- and high-paid sectors in Table 2. In contrast, we put no constraints on the shares of these sectors at the city-level.[40]

We then recover individual fixed effects. We look at the shape of the individual worker fixed-effects curve within each city-size and across worker fixed-effects percentiles (Figure F.4). Notably, for each size group of cities, as we move from low to middle to high percentiles of individual fixed effects, the shape of the curve moves from concave to linear to convex, approximating a log-normal distribution.[41] We then assume that skills are distributed over Ω following a truncated normal distribution with a mean of 0 and standard deviation of 1 and that $H(\omega, \sigma) = \exp(\nu(\sigma)\omega + \mu(\sigma))$.[42] As a result, consistent with the observed patterns, the distribution of $H(\omega, \sigma)$ follows a truncated log-normal distribution with mean $\mu(\sigma)$ and standard deviation $\nu(\sigma)$. We then estimate μ and ν in the bottom, middle and high part of the wage distribution.

To put some values on the relevant parameters of $H(\omega, \sigma)$, we first revert to the ordering of occupations by wages in Table 1. In particular, we want to estimate the parameter values of the distribution of $H(\omega, \sigma)$ around the thresholds (implied by the employment shares) between low/medium-paid occupations and medium/high-paid occupations; and in the midpoint of the share of middle-paid jobs. Thus, in our benchmark calibration, we take the 12.5-22.5% range around the threshold between low- and middle-paid jobs to estimate the parameters for the low-paid sector. Similarly, we take the 67.5-77.5% range that is around the threshold between middle- and high-paid jobs to estimate the parameters in the high-paid sector and, finally, we take the 40-50% range for the middle-paid sector. The approach is to focus on parameters' values that are relevant for agents that are close to the thresholds. Finally, we assume that $\Omega = [-\overline{\omega}, \overline{\omega}]$ and, in the simulation, we take $\overline{\omega} = 10$, that is 10 times the standard deviation of ω. In the end, the distribution of wages that we obtain is close to what we obtain in the data (see Figure F.4). In addition, consistent with Assumption 2, we find that ν is increasing with sector σ, leading to a positive sorting of higher-skilled households to higher-paid sectors.

Following Davis and Dingel (2020), we parametrize the supply of locations $S(\tau) = \pi\tau^2$ and

[40] We conduct robustness checks with respect to the distribution of skills (Appendix C.4); different shocks (Appendix C.5); the presence of non-tradable low-skilled services (Appendix C.6); city sizes (cities > 200k and cities < 200k – Appendix C.7, cities > 750k and cities < 250k – Appendix C.8); and different intervals for productivity estimation (Appendix C.9).

[41] This is consistent, as well, with the evidence from the US presented by Song et al. (2018). The distribution of individuals worker-fixed effects in larger cities is positively skewed in comparison to smaller cities as in Davis and Dingel (2020) for the US. The relative strength of these forces in the upper tail would be consistent with models in which large cities have a superset of skills found in smaller cities, with the distinctive skills precisely in the upper tail.

[42] As we make clear in Appendix C, we cannot identify separately the distribution of skills from the mapping of skills to individual productivities, so we need to make an assumption on the former to identify the latter. Appendix C also reports the outcome of the model when $f(.)$ is uniform and we obtain similar qualitative outcomes.

Aggregate shares			
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.06	-0.20	+0.14
data	+0.07	-0.16	+0.09
Relative shares			
	$\Delta(s(l,1)-s(l,2))$	$\Delta(s(m,1)-s(m,2))$	$\Delta(s(h,1)-s(h,2))$
model	-0.10	-0.07	+0.17
data	-0.02	-0.06	+.08
Differences in initial exposure to middle-paid in 1994			
		$s(m,1)-s(m,2)$	
model		-0.04	
data		-0.11	

Table 4: Simulation-based sectoral distribution

the within-city productivity term as $T(\tau) = 1 - d_1\tau$. We set d_1 to match the relative size of inhabitants in cities larger than 500,000 inhabitants with employment in cities between 50,000 and 100,000 inhabitants.

Finally, in this setting, price $p(\sigma)$ and productivity $A(\sigma, c)$ changes are isomorphic and only relative changes matter. For expositional purposes, we will discuss this as holding fixed all productivity parameters $A(\sigma, c)$ as well as the price of the low-paid task $p(l)$ for 1994-2015. From the data on the evolution of relative value marginal products, (see Table F.6), we obtain a relative decrease of 3.1% of $p(m)$ and a relative increase of 3.1% of $p(h)$. Notice that such a price evolution is consistent with a polarization shock as spelled out by Proposition 4.

Implications. We now simulate the outcome of the model to a price change as described above. We report the results in Table 4. We then connect our findings to the facts that we document on the 1994-2015 period.

Let us first note that, as implied by Proposition 5, the labor market became more polarized in the model, as is the case in the data between 1994 and 2015. If anything, the model slightly overpredicts the fall in middle-paid job and the rise in high-paid jobs, as observed in the upper panel of Table 4.

Figure 7 shows aggregate and relative shares of each job type as prices change. The top panel plots the evolution of the aggregate share of (left to right) low-, middle-, and high-paid jobs. The bottom panel plots the difference in shares between large and small cities for each job type and how this changes in response to the price changes.

In all of these graphs, the shares are functions of a range of relative prices $p(m)/p(l)$ and $p(h)/p(l)$, consistent with the patterns exhibited in Table F.6, where $p(m)/p(l)$ decreases by 1% for each rise of 1% of $p(h)/p(l)$. Finally, we indicate by vertical red dashed lines the levels of relative prices consistent with the aggregate share of middle-paid jobs in 1994 and 2015.

Middle-paid job loss and initial exposure. The first observation is that the share of middle-

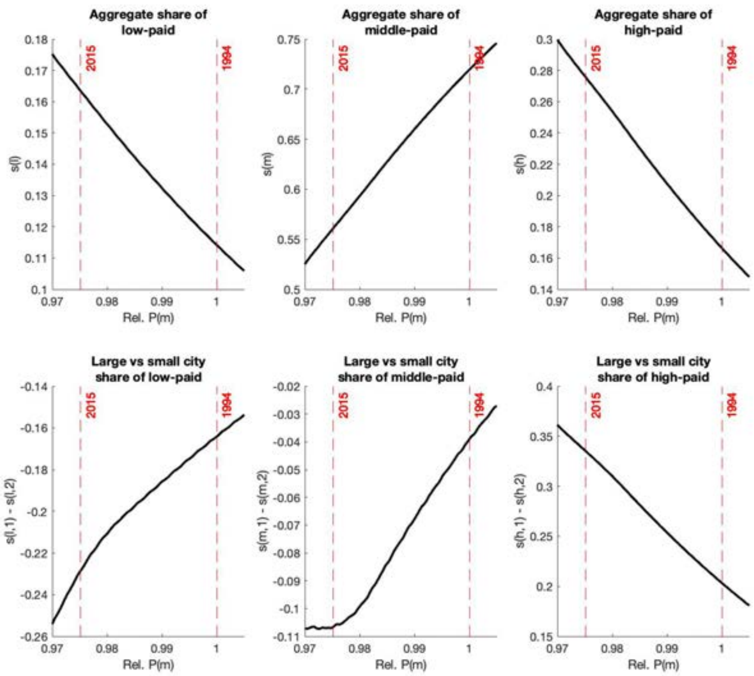


Figure 7: The effect of a decrease in the price of the middle-paid good

paid workers is declining by more in the large city compared with the small one: $s(1,m)-s(2,m)$ declines by 7 percentage points in the middle panel of Table 4 (see also the bottom left panel of Figure 7). This finding is consistent with Fact 2, where we find that large cities experienced a larger decline in middle-paid jobs. Quantitatively, the simulation predicts that the difference between the share of middle-paid workers in City 1 and in City 2 ($s(1,m) - s(2,m)$) falls by 7 percentage points. This number has to be compared with the actual decline in this difference that is close to 6 percentage points. The lower panel of Table 4 also shows that the share of middle-paid workers is initially lower in the large city than in the small city ($s(m,1) - s(m,2) < 0$), as in the data.

Taken together, these two findings correspond to Fact 2, where we find that large cities experienced a stronger decline in middle-paid jobs, even though they were initially less exposed to these jobs. Overall these findings mean that exposure to middle-paid jobs is not necessarily the key driver that explains the scale of their destruction in a particular location. Our interpretation is that technology or offshoring are necessary ingredients for the destruction of middle-paid jobs but they are not sufficient and one also needs to think about incentives to destroy these jobs. A direct implication of this finding is that we cannot instrument future job destruction only by city-level exposure or any other feasibility constraint for this destruction.

Key elements in the model that allow us to match Fact 2 are the productivity advantages of the large city. Absolute advantage under our assumptions implies that there is an interval of skills $(\bar{\omega}(2), \bar{\omega}]$ only in the large city and fully employed in the high-paid sector. Comparative advantage of the large city in the high-paid sector reinforces this advantage. These productivity advantages, paired with the assumption that agents can choose between the middle- and the high-paid sectors, explains both the lower initial exposure to middle-paid jobs and the stronger reallocation from middle- to high-paid jobs.[43]

On one hand, a sufficiently large comparative advantage for the high-paid sector in the large city leads to a lower threshold $\omega(h, 1)$ as implied by Proposition 1, and, thus, to a large share of employment in this sector. In turn, this leads the share of middle-paid jobs in the large city to become smaller relative to the share of these jobs in the small city.

On the other hand, the comparative advantage in the high sector associated with the margin of adjustment between middle- and high-paid sectors is also important for the evolution of middle-paid jobs. In our model, the incentive to destroy the upper tier of middle-paid jobs depends on city characteristics and the opportunity cost of keeping these jobs rather than creating new ones in other sectors, at this margin especially in the high-paid sector. The effects of a decline in the price of the middle-paid good then depend on how the thresholds $\omega(\sigma, c)$ evolve across cities and how many people are reallocated away from the middle-paid sector as a result of these variations in thresholds. This depends on the features of technology (individual productivity across sectors) and the distribution of skills, summarized by $H(\omega, M(\omega, c))$.

In our base calibration, our assumption that productivity is an exponential function of skill ω implies that thresholds move similarly across cities. The derivatives of the thresholds with respect to relative prices in this case do not depend on city-specific productivities:

$$\frac{\partial \omega(h,c)}{\partial \frac{p(m)}{p(h)}} = \frac{1}{\nu(h) - \nu(m)} \frac{p(h)}{p(m)} \text{ and } \frac{\partial \omega(m,c)}{\partial \frac{p(l)}{p(m)}} = \frac{1}{\nu(m) - \nu(l)} \frac{p(m)}{p(l)}$$

On the other hand, the skill distribution is normal and, given the relatively low share of high-paid jobs in either type of city initially (less than 20%), thresholds between middle- and high-paid sectors are on the right part of the normal distribution in both cities. However, the large city has a lower threshold $\omega(h, 1)$ due to the comparative advantage of the large city in the high-paid sector. This implies that the marginal high-paid worker in the large city is less skilled than in the small city, so that the same relative decline in the thresholds $\omega(h, c)$ leads to more reallocation of middle- to high-paid jobs in the large than the small city. On the other hand, for symmetric reasons, the small city experiences a stronger reallocation of middle-paid to low-paid jobs, as the small city has a comparative advantage in the low-paid sector. In this case, the low share of low-paid jobs initially leads thresholds between low- and middle-paid sectors to be on the left part of the normal distribution in both cities.

In our simulations, the reallocation towards high-paid jobs in the large city dominates, so overall middle-paid jobs decline more in the large city. A way to view this quantitative result

[43] This margin of adjustment is present in Cortes (2016) but absent in papers such as Autor and Dorn (2013).

is that the polarization shock leads to more reallocation to high-paid jobs overall, as in Table 4.[44] Finally, our results are robust to alternative assumptions on the distribution of skills. In Appendix C, we investigate an alternative with a uniform distribution of skills[45]

The Great Urban Divergence. In our setting, the Great Urban Divergence arises in case two features are present – the large city begins with a greater commitment to high- relative to low-paid jobs and the magnitude of this difference rises with the polarization shock. The difference in levels in our simulation is apparent in the lower panel of Figure 7, where the larger city has both a higher initial share of high-paid jobs and a lower initial share of low-paid jobs. Moreover, the same figure illustrates that the polarization shock that lowers the middle-paid price $P(m)$ increases the job share gaps between large and small cities, with large cities having a greater differential share of high-paid jobs and a more negative differential share of low-paid jobs. The middle panel of Table 4 establishes the magnitudes, where the high paid gap, $s(h,1) - s(h,2)$, increases by 17 percentage points, while the (negative) low paid gap $s(l,1) - s(l,2)$ grows in magnitude by 10 percentage points. In short, the model replicates the qualitative features of Fact 4, the Great Urban Divergence. The model tends to predict in the large city both a higher initial share of high-paid workers and a stronger reallocation towards this sector compared with what is observed in the data.[46]

Skewed middle-paid job loss. Finally, we can investigate which type of middle-paid jobs are destroyed in the two cities. Our findings, consistent with our previous discussions, are that in the large city middle-paid job loss is mainly about the upper tier of these jobs, while in the small city it is mainly about the lower tier.[47]

[44]In addition to data on the polarization shock from Table F.6, one reason for this is that agents' productivity is more sensitive to skill in the high-paid sector, leading to more adjustment at the top. Also, at the city-level, data on productivity from Table F.8 suggests that the comparative advantage in the high-paid sector of the large city is relatively stronger than the one of the small city in the low-paid sector.

[45]In this case, we need a productivity function with increasing convexity to match the distribution of individual fixed-effects of increasing convexity in ω, as in the data and as captured in our benchmark case by the combination of normal distribution of skills associated with exponential productivity. In this case, due to the comparative advantage of the large city in the high-paid sector, the threshold $\omega(h,1)$ declines by more than $\omega(h,2)$ as a result of the decline in the price of the middle-paid good and given the convexity of the production function. We then obtain qualitatively similar results in terms of the patterns of polarization. More generally, our understanding is that, given the $H(\omega,\sigma)$ functions that we approximate the $H(\omega, M(\omega, c))$ obtained from the data, we will obtain the skewed polarization result no matter how we split the distribution of individual fixed effects between the distribution of skills $f(\omega)$ and productivity $\omega \to H(\omega,\sigma)$ as long as the different $H(\omega,\sigma)$ functions take similar values in the vicinity of the studied thresholds.

[46]This likely reflects the model's simplifications that exclude many frictions. For example, the model does not include any zoning, social housing, or other policy interventions, frictions or lack of tradability of lower-skilled goods, that likely limit the specialization of the large city in high-paid activities. Even if quantitative patterns are not our primary objective in this section, we report in Appendix C.6 the results of the model with non-tradable goods. Introducing such non-tradable goods is one dimension along which quantitative results can be improved.

[47]We confirm this point in Figure C.1, we plot the difference across cities in the shares of middle-paid jobs when we split middle-paid jobs into those occupied by higher-skilled households and those occupied by lower-skilled households. As we can observe, in the large city, higher-skilled middle-paid jobs have disappeared at a faster pace while lower-skilled middle-paid jobs disappeared more quickly in the small city.

This finding replicates Fact 3 where we find that the destruction of middle-paid jobs concerns the upper tier of middle-paid jobs in large cities but not in small cities. In contrast to prior work, our theory places emphasis both on heterogeneity of middle-paid jobs by skill and how that translates, given a common shock, into distinct experiences in large and small cities. Our model emphasizes two margins of adjustment, as middle-paid jobs are substituted alternately by low- or high-paid jobs. And it stresses that the magnitudes of the middle-paid job losses, and the relative importance of each margin, will differ according to the size of the city. As discussed, our theory replicates the fact that the magnitude of loss of middle-paid jobs will be larger in large cities and that these cities will also see a relatively large loss of these jobs at the upper end of the middle-paid jobs, and vice versa for smaller cities.

6 Conclusions

Labor market polarization is a prominent feature in recent decades of many advanced economies. The defining loss of middle-paid jobs along with the growth of both low- and high-paid jobs appears in the United States and many European countries. Over the same time period, diverging fates of already-skilled, typically larger cities and less skilled, typically smaller cities, were observed. This second phenomenon is called the great urban divergence. This paper develops a set of facts that characterize the related aggregate and cross-city features in this area and builds a parsimonious theoretical model to account for these.

We identify four key facts that anchor our work. The first is what we term Universal Polarization. In our data covering the period 1994-2015, both France as a whole and 115 of 117 French cities in our data experience labor market polarization. The second key fact is that middle-paid job loss was greater in large relative to small cities, even though the initial exposure to these jobs was lower in large cities. The third key fact focuses on the *type* of middle-paid jobs lost. In large relative to small cities, the lost jobs are concentrated relatively in an upper tier of the middle-paid jobs. Finally, consistent with the Great Urban Divergence, job growth in large cities was concentrated relatively in the high skill segment in spite of the greater initial presence of high skill jobs in large cities, and vice versa in small cities.

We discuss existing theories of labor market polarization and the great urban divergence in order to demonstrate that they cannot account for these four key facts. We then develop a parsimonious theory that can account for these facts. In spite of the simplicity of the components of the theory, these yield a rich set of results. Building on the prior literature on labor market polarization, we consider three intermediate tasks that can be thought of as low-, middle-, and high-paid jobs. There is an input which is a relative substitute for the middle-paid job and a complement to the low-and high-paid jobs. This input can be thought of either as capital that allows routinization or an intermediate input that is offshored. To this standard setting for labor market polarization, we add elements of labor heterogeneity with individual-level comparative advantage, whereby individuals select into one or another of the three types of jobs. There is city-level absolute and comparative advantage across the jobs. And we add some intuitive structure on how technology and skill interact. Jointly these yield results consistent with the four key facts in our data.

In sum, we find that the period of study identifies two Frances. In the France of large cities, there is a dramatic change, as there is a sharp contraction of middle-paid jobs, particularly at the top end of these. However in the France of large cities, these middle-paid jobs are largely replaced by high-paid jobs, with a more modest expansion of low-paid jobs. Still there is very sharp polarization within these cities. In the France of small cities, there is a strong, yet more moderated, loss of middle-paid jobs. Some high-paid jobs are gained, but the lost middle-paid jobs are primarily replaced by low-paid jobs. Polarization of jobs in the aggregate and within cities is accompanied by a great urban divergence between the Frances of large and small cities. Our theory accounts for these facts.

References

Acemoglu, Daron, "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence," *American Economic Review*, December 1999, *89* (5), 1259–1278.

— **and David Autor**, "Skills, tasks and technologies: Implications for employment and earnings," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

— **and** — , "What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology," *Journal of Economic Literature*, June 2012, *50* (2), 426–63.

— **and Pascual Restrepo**, "Low-Skill and High-Skill Automation," *Journal of Human Capital*, 2018, *12* (2), 204–232.

— **and** — , "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

Arzaghi, Mohammad and J. Vernon Henderson, "Networking off Madison Avenue," *The Review of Economic Studies*, 2008, *75* (4), 1011–1038.

Austin, Benjamin, Edward Glaeser, and Lawrence Summers, "Jobs for the Heartland: Place-Based Policies in 21st-Century America.," *Brookings Papers on Economic Activity*, 2018.

Autor, David H., "Work of the Past, Work of the Future," *AEA Papers and Proceedings*, May 2019, *109*, 1–32.

— **and David Dorn**, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 2013, *103* (5), 1553–1597.

Autor, David H, David Dorn, and Gordon H Hanson, "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–68.

__ , Frank Levy, and Richard J Murnane, "The skill content of recent technological change: An empirical exploration," *Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

Autor, David H., Lawrence F Katz, and Melissa S Kearney, "The polarization of the US labor market," *American Economic Review*, 2006, *96* (2), 189–194.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

Bock, Sébastien, "Job Polarization and Unskilled Employment Losses in France," Technical Report 2020.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg, "The Anatomy of French Production Hierarchies," *Journal of Political Economy*, 2015, *123* (4), 809–852.

Cerina, Fabio, Elisa Dienesch, Alessio Moro, and Michelle Rendall, "Spatial Polarisation," *Economic Journal*, 2022, *133* (649), 30–69.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Cortes, Guido Matias, "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data," *Journal of Labor Economics*, 2016, *34* (1), 63–105.

__ , Nir Jaimovich, and Henry E. Siu, "Disappearing routine jobs: Who, how, and why?," *Journal of Monetary Economics*, 2017, *91*, 69–87.

Costinot, Arnaud, "An Elementary Theory of Comparative Advantage," *Econometrica*, 2009, *77* (4), 1165–1192.

__ and Jonathan Vogel, "Matching and Inequality in the World Economy," *Journal of Political Economy*, August 2010, *118* (4), 747–786.

Davis, Donald R. and Jonathan I. Dingel, "A Spatial Knowledge Economy," *American Economic Review*, January 2019, *109* (1), 153–70.

__ and __ , "The comparative advantage of cities," *Journal of International Economics*, 2020, *123*, 103291.

Diamond, Rebecca, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, March 2016, *106* (3), 479–524.

— **and Cecile Gaubert**, "Spatial Sorting and Inequality," *Annual Review of Economics*, 2022, *14* (1), 795–819.

Eckert, Fabian, "Growing Apart: Tradable Services and the Fragmentation of the U.S. Economy," 2019 Meeting Papers 307, Society for Economic Dynamics 2019.

— **, Sharat Ganapati, and Conor Walsh**, "Urban-Biased Growth: A Macroeconomic Analysis," NBER Working Papers 30515, National Bureau of Economic Research, Inc September 2022.

Eeckhout, Jan, Christoph Hedtrich, and Roberto Pinheiro, "IT and Urban Polarization," CEPR Discussion Papers 16540, C.E.P.R. Discussion Papers September 2021.

— **, Roberto Pinheiro, and Kurt Schmidheiny**, "Spatial Sorting," *Journal of Political Economy*, 2014, *122* (3), 554–620.

Ganong, Peter and Daniel Shoag, "Why has regional income convergence in the U.S. declined?," *Journal of Urban Economics*, 2017, *102* (C), 76–90.

Gaubert, Cecile, "Firm Sorting and Agglomeration," *American Economic Review*, November 2018, *108* (11), 3117–3153.

Giannone, Elisa, "Skill-Biased Technical Change and Regional Convergence," 2017 Meeting Papers 190, Society for Economic Dynamics 2017.

Gibbons, Robert, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent, "Comparative Advantage, Learning, and Sectoral Wage Determination," *Journal of Labor Economics*, 2005, *23* (4), 681–724.

Goldin, Claudia Dale and Lawrence F Katz, *The race between education and technology*, Harvard University Press, 2009.

Goos, Maarten, Alan Manning, and Anna Salomons, "Job Polarization in Europe," *American Economic Review*, May 2009, *99* (2), 58–63.

— **,** — **, and** — , "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, August 2014, *104* (8), 2509–2526.

— **and** — , "Lousy and lovely jobs: The rising polarization of work in Britain," *Review of Economics and Statistics*, 2007, *89* (1), 118–133.

Guilluy, Christophe, *Fractures Françaises*, François Bourin, Paris, 2010.

Harrigan, James, Ariell Reshef, and Farid Toubal, "The March of the Techies: Technology, Trade, and Job Polarization in France, 1994-2007," NBER Working Papers 22110 March 2016.

— **,** — **, and** — , "Techies, Trade, and Skill-Biased Productivity," NBER Working Papers 25295 November 2018.

Iammarino, Simona, Andres Rodriguez-Pose, and Michael Storper, "Regional inequality in Europe: evidence, theory and policy implications," *Journal of Economic Geography*, 2018, *19* (2), 273–298.

Jung, Jaewon and Jean Mercenier, "Routinization-biased technical change and globalization: understanding labor market polarization," *Economic Inquiry*, 2014, *52* (4), 1446–1465.

Katz, Lawrence F and Kevin M Murphy, "Changes in relative wages, 1963–1987: supply and demand factors," *Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Keller, Wolfgang and Hale Utar, "International trade and job polarization: Evidence at the worker level," *Journal of International Economics*, 2023, *145*, 103810.

Michaels, Guy, Ashwini Natraj, and John Van Reenen, "Has ICT Polarized Skill Demand? Evidence from Eleven Countries over Twenty-Five Years," *The Review of Economics and Statistics*, 03 2014, *96* (1), 60–77.

__ , Ferdinand Rauch, and Stephen J Redding, "Task Specialization in U.S. Cities from 1880 to 2000," *Journal of the European Economic Association*, 03 2018, *17* (3), 754–798.

Moretti, Enrico, "Chapter 51 - Human Capital Externalities in Cities," in J. Vernon Henderson and Jacques-François Thisse, eds., *Cities and Geography*, Vol. 4 of *Handbook of Regional and Urban Economics*, Elsevier, 2004, pp. 2243–2291.

__ , *The New Geography of Jobs*, Houghton Mifflin Harcourt, 2012.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality," *Quarterly Journal of Economics*, 10 2018, *134* (1), 1–50.

Xiao, Hongyu, Andy Wu, and Jaeho Kim, "Commuting and innovation: Are closer inventors more productive?," *Journal of Urban Economics*, 2021, *121*, 103300.

A Proofs of Propositions - for online publication only

A.1 Proof of Proposition 1

First, note that log-supermodularity of H implies that $\forall \omega > \omega'$ and $\forall \sigma > \sigma'$, $H(\omega, \sigma)H(\omega', \sigma') \geq H(\omega', \sigma)H(\omega, \sigma')$. Manipulating this expression, we obtain that $H(\omega, \sigma)/H(\omega, \sigma')$ with $\sigma > \sigma'$ is increasing in ω as $H(\omega, \sigma)/H(\omega, \sigma') > H(\omega', \sigma)/H(\omega', \sigma')$. Note then that $H(\omega(m, c), m)/H(\omega(m, c), l) = A(l, c)/A(m, c)p(l)/p(m)$. This latter ratio is larger in City 2 than in City 1 given Assumption 3 and thus $\omega(m, 1) < \omega(m, 2)$. In addition $\omega(h, c)$ solves the following equation:

$$\frac{A(m,c)}{A(h,c)}\frac{p(m)}{p(h)} = H(\omega(h,c),h)/H(\omega(h,c),m) \tag{9}$$

Given Assumption 3, the ratio $A(h, c)/A(m, c)$ is larger in City 1 than in City 2, thus implying $\omega(h, 1) < \omega(h, 2)$.

A.2 Intermediary results on location and sectoral decisions

We now prove a set of results that will be useful to prove our main implications from the model.

Location decisions. Let us start by describing the location decision within each city. Note that the set of locations occupied in city c is a bounded set. We denote by $\overline{\tau}(c)$ the maximum value of τ occupied in city c. More desirable locations have higher rental prices:

Lemma A.1. *Housing prices $r(c, \tau)$ are decreasing on $[0, \overline{\tau}(c)]$ and $r(c, \overline{\tau}(c)) = 0$. Finally, for all $\tau \in [0, \overline{\tau}(c)]$:*

$$S(\tau) = L \int_0^\tau \int_\sigma \int_\omega f(\omega, M(\omega, c), c, x) d\omega d\sigma dx \tag{10}$$

Proof. We closely follow here the proof of Lemma 1 in Davis and Dingel (2020). Let us now show that $r(c, \tau)$ is decreasing with τ. Suppose it is not. Then there exist τ' and τ'' satisfying $\tau' < \tau'' \leq \overline{\tau}(c)$ such that $r(c, \tau') \leq r(c, \tau'')$. Thus, $U(c, \tau', \sigma, \omega) > U(c, \tau'', \sigma, \omega)$ for all σ and all ω. This contradicts the fact that τ'' has to maximize utility for some individual with some skill ω and sectoral decision σ.

The continuity of $T(.)$ ensures that $r(c, \overline{\tau}(c)) = 0$. Indeed, suppose that $r(c, \overline{\tau}(c)) > 0$. Given that the location $\overline{\tau}(c)$ is populated, there exists ω such that, for any $\epsilon > 0$:

$$A(M(\omega, c), c)H(\omega, M(\omega, c))T(\overline{\tau}(c)) - r(c, \overline{\tau}(c)) \geq A(M(\omega, c), c)H(\omega, M(\omega, c))T(\overline{\tau}(c) + \epsilon).$$

However, this inequality contradicts the continuity of $T(.)$ when $r(c, \overline{\tau}(c)) > 0$. Thus, $r(c, \overline{\tau}(c)) = 0$.

Finally, suppose that there exists $\tau' < \overline{\tau}(c)$ so that

$$S(\tau') > L \int_0^{\tau'} \int_\sigma \int_\omega f(\omega, M(\omega, c), c, x) d\omega d\sigma dx$$

This implies that there exists $\tau \leq \tau'$ so that

$$S'(\tau) > L \int_\sigma \int_\omega f(\omega, M(\omega, c), c, \tau) d\omega d\sigma$$

This location is empty and so $r(c,\tau) = 0 = r(c, \overline{\tau}(c))$. However, as $\tau < \overline{\tau}(c)$, any ω located in $\overline{\tau}(c)$ is strictly better off selecting τ as a location, contradicting that $\overline{\tau}(c)$ maximizes utility for some agents. □

Furthermore, higher skill households occupy more desirable locations. We find this by obtaining a mapping between skill ω and location (c, τ):

Lemma A.2. *There exists a function K such that: $f(\omega, M(\omega, c), c, \tau) > 0 \Leftrightarrow K(c,\tau) = \omega$. The function $K(c,.)$ is continuous and strictly decreasing.*

In addition, when the low-paid sector exists in both cities, $\overline{\tau}(1)$ and $\overline{\tau}(2)$ are such that $K(2, \overline{\tau}(2)) = K(1, \overline{\tau}(1)) = \underline{\omega}$. Furthermore, $K(1,0) = \overline{\omega}(1) = \overline{\omega}$ and there exists $\overline{\omega}(2)$ such that $K(2,0) = \overline{\omega}(2)$.

Proof. Here, we follow Lemma 2 in Davis and Dingel (2020) and Lemma 1 in Costinot and Vogel (2010). Let us first define $f(\omega, c, \tau) = \int_\sigma f(\omega, c, \tau, \sigma) d\sigma$, $\Omega(\tau, c) = \{\omega \in \Omega, f(\omega, c, \tau) > 0\}$ and $\mathcal{T}(\omega, c) = \{\tau \in [0, \bar{\tau}(c)], f(c, \omega, \tau) > 0\}$. Using these objects, we obtain:

(i) $\Omega(\tau, c) \neq \emptyset$ for $0 \leq \tau \leq \bar{\tau}(c)$ and $\mathcal{T}(\omega, c) \neq \emptyset$ for at least one city as $f(\omega) > 0$.

(ii) $\Omega(\tau, c)$ is a non-empty interval for $0 \leq \tau \leq \bar{\tau}(c)$. If not, there exist $\omega < \omega' < \omega''$ such that $\omega, \omega'' \in \Omega(\tau)$ but not ω'. This means that there exists τ' such that $\omega' \in \Omega(\tau')$. Without loss of generality, suppose that $\tau' > \tau$. Utility maximization for both ω and ω' implies:

$$T(\tau')G(\omega', c) - r(c, \tau') \geq T(\tau)G(\omega', c) - r(c, \tau)$$
$$T(\tau)G(\omega, c) - r(c, \tau) \geq T(\tau')G(\omega, c) - r(c, \tau')$$

where $G(\omega, c) = A(M(\omega, c), c)H(\omega, M(\omega, c))$. These inequalities jointly imply that

$$\left(T(\tau') - T(\tau)\right)\left(G(\omega', c) - G(\omega, c)\right) \geq 0,$$

but this cannot be with $\tau' > \tau$ and $\omega' > \omega$. The same reasoning can be applied when $\tau' < \tau$. We can also conclude that for any $\tau < \tau'$, if $\omega \in \Omega(\tau)$ and $\omega' \in \Omega(\tau')$, then $\omega \geq \omega'$.

(iii) $\Omega(\tau, c)$ is a singleton for all but a countable subset of $[0, \bar{\tau}(c)]$. For any $\tau \in [0, \bar{\tau}(c)]$, $\Omega(\tau, c)$ is measurable as it a non-empty interval. Let $\mathcal{T}_0(c)$ denote the subset of locations τ such that $\mu(\Omega(\tau, c)) > 0$, μ being the Lebesgue measure over $\mathcal{R}$. Let us show that $\mathcal{T}_0(c)$ is a countable sets – any other $\Omega(\tau, c)$ where $\tau \notin \mathcal{T}_0(c)$ is a singleton as it is a interval with measure 0. For any $\tau \in \mathcal{T}_0(c)$, let us define $\underline{\omega}(\tau) \equiv \inf \Omega(\tau, c)$ and $\overline{\omega}(\tau) \equiv \sup \Omega(\tau, c)$. As $\mu(\Omega(\tau, c)) > 0$, $\underline{\omega}(\tau) < \overline{\omega}(\tau)$. Thus there exists a integer j such that $j\left(\overline{\omega}(\tau) - \underline{\omega}(\tau)\right) > (\bar{\omega}(c) - \underline{\omega})$. Given that $\mu(\Omega(\tau, c) \cap \Omega(\tau', c)) = 0$ for $\tau \neq \tau'$, for any j, we can then have at most j elements $\{\tau_1, ... \tau_j\} \equiv \mathcal{T}_j^0$

verifying $j\left(\overline{\omega}(\tau_i)-\underline{\omega}(\tau_i)\right) > (\bar{\omega}(c)-\underline{\omega})$. Thus $\mathcal{T}_j^0$ is countable. Given that $\mathcal{T}^0=\bigcup_{j=1}^{\infty}\mathcal{T}_j^0$ and that the countable union of countable sets is also countable, we conclude that $\mathcal{T}^0$ is countable.

(iv) $\mathcal{T}(\omega,c)$ is a singleton for all but a countable subset of Ω. As in Davis and Dingel (2020), we use the arguments as in steps 2 and 3.

(v) $\Omega(\tau,c)$ is a singleton for any $\tau\in[0,\bar{\tau}(c)]$. Suppose not: there exists $\tau\in[0,\bar{\tau}(c)]$ so that $\Omega(\tau,c)$ is not a singleton. Given step (ii), it is then an interval with strictly positive measure. Step (iv) implies that $\mathcal{T}(\omega,c)=\{\tau\}$ for almost all $\omega\in\Omega(\tau,c)$ Hence we obtain:

$$f(c,\omega,\tau)=f(\omega)\delta^{\text{Dirac}}\left(1-1_{\Omega(c,\tau)}\right)\text{ for almost all }\omega\in\Omega(c,\tau). \tag{11}$$

This contradicts assumptions on $S(\tau)$: integrating $f(c,.,\tau)$ over $\Omega(c,\tau)$ which has a strictly positive measure requires the supply of locations at τ to satisfy $S'(\tau)=\infty$, which cannot be for $\tau<\infty$.

In the end, in city c, for any $\tau\in[0,\bar{\tau}(c)]$, there exists a unique ω such that $\omega\in\Omega(c,\tau)$. This does defines a function K_c such that $K_c(\tau)=\omega$. This function is weakly decreasing as shown by step (ii) and even strictly decreasing as $\Omega(\tau,c)$ is a singleton almost everywhere, following step (iv). Furthermore, as $\Omega(\tau)\neq\emptyset$ for all $\tau\in[0,\bar{\tau}(c)]$, K_c is continuous and satisfies $K_c(0)=\overline{\omega}(c)$ and $K_c(\bar{\tau}(c))=\underline{\omega}$.

Indeed, the least skill agent, $\underline{\omega}$ is in both cities when the low-paid sector is in both cities. Suppose it is not the case. Let us denote by $\omega^*>\underline{\omega}$ the agent with the lowest skill that live in both cities. This agent is indifferent to live in both cities, that is:

$$\begin{aligned}A(M(\omega^*,1),1)H(\omega^*,M(\omega^*,1))T(\tau(1)^*)-r(1,\tau(1)^*)=\cdots\\ \cdots A(M(\omega^*,2),2)H(\omega^*,M(\omega^*,2))T(\tau(2)^*)-r(2,\tau(2)^*)\end{aligned}$$

Suppose then that every $\omega<\omega^*$ is not in City 1 – reciprocally, we may assume that such ω are not in City 2 and show a contradiction. This implies that $r(1,\tau(1)^*)=0$. In particular, this means that $\underline{\omega}$ is not in City 1. Yet, following Proposition 1, as the low-paid sector is in both cities, the least skilled agent that is in both cities has to work in the low-paid sector: as a result, agents with skill ω^* work in the low-paid sector. We then obtain:

$$A(M(\omega^*,1),1)H(\omega^*,M(\omega^*,1))T(\tau(1)^*)=H(\omega^*,M(\omega^*,2),2)T(\tau(2)^*)-r(2,\tau(2)^*)$$

Let us consider the least skilled agent ($\underline{\omega}$). First, as $\underline{\omega}<\omega^*$, agents with skill $\underline{\omega}$ work also in the low-paid sector. In addition, this agent is only is City 2 and, given that it is the least skilled in that city, this agent is located at the edge ($\bar{\tau}(2)$) with a 0 rent ($r(2,\bar{\tau}(2))=0$). These agents do not want to live in City 1 next to agents with skill ω^*, which implies:

$$A(M(\underline{\omega},2),2)H(\underline{\omega},M(\underline{\omega},2))T(\bar{\tau}(2))>A(M(\underline{\omega},1),1)H(\underline{\omega},M(\underline{\omega},1))T(\tau(1)^*).$$

Again, we can simplify this inequality, given that $M(\underline{\omega}, 2) = M(\underline{\omega}, 1) = l$ and dividing by $H(\underline{\omega}, M(\underline{\omega}, 2))$:

$$A(2, l)T(\bar{\tau}(2)) > A(1, l)T(\tau(1)^*).$$

However, by multiplying this inequality by $H(\omega^*, M(\omega^*, 1))$, we then obtain that:

$$A(M(\omega^*, 1), 1)H(\omega^*, M(\omega^*, 1))T(\tau(1)^*) < A(M(\omega^*, 2), 2)H(\omega^*, M(\omega^*, 2))T(\bar{\tau}(2))$$

which implies that agent with skill ω^* is better off moving to City 2 in location $T(\bar{\tau}(2))$ thus contradicting the definition of an equilibrium. □

Lemma A.2 implies that when the low-paid sector is in both cities the least skilled person $\underline{\omega}$ is also in both cities in location $\overline{\tau}(1)$ in City 1 and in location $\overline{\tau}(2)$ in City 2 so that City 1's set of skills is a strict superset of that in City 2.

Correspondance between locations in City 1 and City 2. In the end, for an $\omega \leq \bar{\omega}(2)$ and a τ, there exists a single τ' in City 2. This defines a function $\Gamma(\omega, \tau) = \tau'$, which identifies a location in City 2 at which factor ω has the same return as it would have in City 1 at τ. Spatial equilibrium thus implies that for all ω present in both cities:

$$A(M(\omega, 1), 1)H(\omega, M(\omega, 1))T(\tau) = A(M(\omega, 2), 2)H(\omega, M(\omega, 2))T(\Gamma(\omega, \tau), 2) \tag{12}$$

In equilibrium, in the location τ, if the agent with skill ω is the marginal buyer, we then have that $r(1, \tau) = r(2, \Gamma(\omega, \tau))$. In Davis and Dingel (2020), the function Γ would be constant with respect to ω, but, as the larger city has also a comparative advantage in higher-paid sector, we obtain the following result:

Lemma A.3. *For all ω, $\Gamma(\omega, .)$ is continuously increasing in τ and, for any τ, $\Gamma(., \tau)$ is continuous and weakly decreasing in ω.*

Proof. Let us consider the function Γ as defined by equation (12), that is:

$$A(M(\omega, 1), 1)H(\omega, M(\omega, 1))T(\tau) = A(M(\omega, 2), 2)H(\omega, M(\omega, 2))T(\Gamma(\omega, \tau))$$

$\Gamma(\omega, .)$ inherits the properties of the function T. For $\Gamma(., \tau)$, the function is continuous and either constant or decreasing in each segment defined by the thresholds $\omega(h, c)$ and $\omega(m, c)$. Given the definition of the thresholds, the function is continuous everywhere and, thus, given it is either constant or decreasing in each segment, it is globally weakly decreasing. □

Lemma A.4. *Households of skill ω occupying locations in the two cities select locations τ_1 in City 1 and τ_2 in City 2 that are such that $r(1, \tau_1) = r(2, \tau_2)$.*

Proof. Let us consider ω such that this skill is present in the two cities. Let us denote by τ_1 and τ_2 the locations occupied by this skill in City 1 and City 2 respectively. Suppose that $r(1, \tau_1) \neq r(2, \tau_2)$.

For example, $r(1,\tau_1) > r(2,\tau_2)$. The indifference condition between the two locations writes:

$$A(M(\omega,1),1)H(\omega,M(\omega,1))T(\tau_1) - r(1,\tau_1) = A(M(\omega,2),2)H(\omega,M(\omega,2))T(\tau_2) - r(2,\tau_2).$$

As $r(1,\tau_1) > r(2,\tau_2)$, we have:

$$A(M(\omega,1),1)H(\omega,M(\omega,1))T(\tau_1) > A(M(\omega,2),2)H(\omega,M(\omega,2))T(\tau_2).$$

On the other hand, there exists a location $\tau_1' > \tau_1$ such that productivities are equal:

$$A(M(\omega,1),1)H(\omega,M(\omega,1))T(\tau_1') = A(M(\omega,2),2)H(\omega,M(\omega,2))T(\tau_2).$$

The price of the location satisfies $r(1,\tau_1') < r(2,\tau_2)$: as this gives the same productivity as $(2,\tau_2)$, the household with skill ω would bid the same price $r(2,\tau_2)$ but Lemma A.2 implies that it is households with skill $\omega' < \omega$ that occupy location $(1,\tau_1')$ that bid a lower price. However, this contradicts the optimality of location $(2,\tau_2)$ as this location is strictly dominated by location $(1,\tau_1')$:

$$A(M(\omega,1),1)H(\omega,M(\omega,1))T(\tau_1') - r(1,\tau_1') > A(M(\omega,2),2)H(\omega,M(\omega,2))T(\tau_2) - r(2,\tau_2).$$

$\square$

A.3 Proof of Proposition 2

Let us consider the function $K(c,\tau)$ defined Lemma A.2. Following this lemma, this function is continuous and weakly decreasing with respect to τ. The results of the Lemma directly follows as the location of a individual with skill ω is $\{c,\tau\}$ such that $\omega = K(c,\tau)$. In particular, there exists a unique $\bar{\tau}(h,c)$ such that $K(c,\bar{\tau}(h,c)) = \omega(h,c)$ and a unique $\bar{\tau}(m,c) \geq \bar{\tau}(h,c)$ such that $K(c,\bar{\tau}(m,c)) = \omega(m,c)$.

A.4 Proof of Proposition 3

Given that City 1 has an absolute advantage compared City 2 in all sectors, a direct implication of the location decisions is that City 1 has a larger population that City 2, that is $\bar{\tau}(1) > \bar{\tau}(2)$ given that $S(\tau)$ is increasing and common for cities 1 and 2.

Indeed, suppose that this is not the case, that is $\bar{\tau}(1) \leq \bar{\tau}(2)$ in equilibrium. These two locations have a rental price of 0. Following Lemma A.2, the least skilled agent ($\underline{\omega}$) is in both cities and in location $\bar{\tau}(1)$ in City 1 and in location $\bar{\tau}(2)$ in City 2. Also, this agent works in the l sector. Overall, this implies that $A(1,l)T(\bar{\tau}(1)) > A(2,l)T(\bar{\tau}(2))$, thus implying that this agent is better off in City 1, thus contradicting the definition of an equilibrium.

A.5 Proof of Proposition 4

Lemma A.5. *A decline in $p(m)/p(h)$ implies that $\omega(h,c)$ declines in both cities $c \in \{1,2\}$. An increase in $p(l)/p(m)$ implies that $\omega(m,c)$ increases in both cities $c \in \{1,2\}$.*

Proof. From (4) and (5), we obtain that:

$$\frac{H(\omega(m,c),m)}{H(\omega(m,c),l)} = \frac{A(l,c)}{A(m,c)}\frac{p(l)}{p(m)} \text{ and } \frac{H(\omega(h,c),h)}{H(\omega(h,c),m)} = \frac{A(m,c)}{A(h,c)}\frac{p(m)}{p(h)}$$

As H is log-supermodular in (ω, c), the left hand term of the first equation is increasing in $\omega(m,c)$ and the left hand term of the second equation is increasing in $\omega(h,c)$. As $p(l)/p(m)$ is increasing and $p(m)/p(h)$ is decreasing, $\omega(m,c)$ is increasing and $\omega(h,c)$ is decreasing. $\square$

The problem solved by the representative firm is:

$$\max_{q(h),q(m),q(l),Z} Q - p(h)q(h) - p(m)q(m) - P_z Z - p(l)q(l)$$

Solving this problem allows to obtain:

$$\frac{q(l)}{q(m)} = \frac{a(l)}{a(m)}\left(\frac{p(m)}{p(l)}\right)^{\theta}\left(\frac{p(l)}{a(m)p(m)^{1-\theta}+a(z)P_z^{1-\theta}}\right)^{\theta-\zeta} \tag{13}$$

$$\frac{q(h)}{q(m)} = \frac{a(h)}{a(m)}\left(\frac{p(m)}{p(h)}\right)^{\theta}\left(\frac{p(h)}{a(m)p(m)^{1-\theta}+a(z)P_z^{1-\theta}}\right)^{\theta-\zeta} \tag{14}$$

Let us now consider a drop in the price p_z. Let us show that $p(m)/p(l)$ and $p(m)/p(h)$ both decline. Suppose, instead, that both $p(m)/p(l)$ and $p(m)/p(h)$ increase. As $a(m)p(m)^{1-\theta} + a(z)p_z^{1-\theta}$ declines and $\theta > \zeta$, the rightmost terms of both (13) and (14) both increase. If both $p(m)/p(l)$ and $p(m)/p(h)$ increase, we have that both $q(l)/q(m)$ and $q(h)/q(m)$ increase, which cannot be from Lemma A.5, as a rise in both ratios would require some agents to move out of the middle-paid sector – a contradiction. Suppose now that one of the relative price increases, say $p(m)/p(h)$. Our previous reasoning implies that $p(m)/p(l)$ has to decline. This implies that $p(l)/p(h)$ increases. On the other hand, dividing (13) by (14) allows to obtain that $p(l)/p(h)$ being increasing leads to $q(l)/q(h)$ to be decreasing. This cannot be as, from Lemma A.5, $p(m)/p(h)$ being increasing and $p(m)/p(l)$ decreasing leads to $q(l)/q(h)$ to be increasing – a contradiction. As a result, both relative prices have to fall.

A.6 Proof of Proposition 5

Following Propositions 1 and 2, we obtain that the share of middle-paid jobs in city 1 evolves as:

$$L(1)ds(m,1) = \underbrace{\left(S\left(T^{-1}\left(h(\omega(m,1),1)\right)\right)\right)' d\omega(m,1)}_{\text{Reallocation to the l-sector in city 1}} - \underbrace{\left(S\left(T^{-1}\left(h(\omega(h,1),1)\right)\right)\right)' d\omega(h,1)}_{\text{Reallocation to the h-sector in city 1}} + \cdots$$
$$\cdots \underbrace{\int_{\omega(m,1)}^{\omega(m,2)} \frac{\partial f(\omega,1)}{\partial p(m)/p(l)} d\omega \frac{dp(m)/p(l)}{p(m)/p(l)}}_{\text{Reallocation to the l-sector in city 2}}$$

The third term stems from the fact that a segment of households were initially indifferent between the middle-paid sector in city 1 and the low-paid sector in city 2. The share of middle-paid jobs in city 2 evolves as:

$$L(2)ds(m,2) = \underbrace{\left(S\left(T^{-1}\left(h(\omega(m,2),2)\right)\right)\right)' d\omega(m,2)}_{\text{Reallocation to the l-sector in city 2}} - \underbrace{\left(S\left(T^{-1}\left(h(\omega(h,2),2)\right)\right)\right)' d\omega(h,2)}_{\text{Reallocation to the h-sector in city 2}} + \cdots$$
$$\cdots \underbrace{\int_{\omega(h,1)}^{\omega(h,2)} \frac{\partial f(\omega,2)}{\partial p(m)/p(h)} d\omega \frac{dp(m)/p(h)}{p(m)/p(h)}}_{\text{Reallocation to the l-sector in city 2}}$$

The last term stems here from the fact a segment of households were initially indifferent between the high-paid sector in city 1 and the middle-paid sector in city 2.

From Lemma A.5, we obtain that, in both cities, $\omega(h,c)$ declines and $\omega(m,c)$ increases. As a result, the first two terms in the two expressions above are negative. In addition, as the relative price shock makes agents relatively better off in the other city – where they work in the other sector –, the third term is also negative. Overall, $ds(m,1)$ and $ds(m,2)$ are then both negative. From the same reasoning, we obtain that $ds(l,c) > 0$ and $ds(h,c) > 0$ in both cities. Finally, as polarization happens in both cities, it also declines overall.

B Extensions of the model - for online publication only

B.1 Thresholds in the normal distribution case

In this Appendix, we provide more intuition of skewed polarization in the normal distribution case by computing how thresholds move in this case as a function of relative prices. Also, we illustrate the sectoral choice made by households in this normal distribution case.

Thresholds. First, note that thresholds move in the same way across cities. Indeed, $\omega(h,c)$ and $\omega(m,c)$ satisfy:

$$A(h,c)\exp(\nu(h)\omega(h,c)+\mu(h))p(h) = A(m,c)\exp(\nu(m)\omega(h,c)+\mu(m))p(m)$$
$$A(m,c)\exp(\nu(m)\omega(m,c)+\mu(m))p(m) = A(l,c)\exp(\nu(l)\omega(m,c)+\mu(l))p(l)$$

As a result, we obtain that:

$$\omega(h,c) = \frac{1}{\nu(h)-\nu(m)}\left(\ln\left(\frac{A(m,c)}{A(h,c)}\right) + \ln\left(\frac{p(m)}{p(h)}\right) - \mu(h) + \mu(m)\right)$$
$$\omega(m,c) = \frac{1}{\nu(m)-\nu(l)}\left(\ln\left(\frac{A(l,c)}{A(m,c)}\right) + \ln\left(\frac{p(l)}{p(m)}\right) - \mu(m) + \mu(l)\right)$$

In particular, the derivative with respect to relative prices are:

$$\frac{\partial\omega(h,c)}{\partial\frac{p(m)}{p(h)}} = \frac{1}{\nu(h)-\nu(m)}\left(\frac{p(m)}{p(h)}\right)^{-1}$$
$$\frac{\partial\omega(m,c)}{\partial\frac{p(l)}{p(m)}} = \frac{1}{\nu(m)-\nu(l)}\left(\frac{p(l)}{p(m)}\right)^{-1}$$

which do not depend on city-specific productivities. As a result, thresholds move similarly across cities. If polarization is not uniform, this then depends on the distribution of skills across cities, as discussed in the main text.

Upper envelope. In Figure B.1, we plot the value marginal products in the different sectors implied by the exponential production functions. Agents choose sectors in order to reach the upper envelope of these value marginal products, as pictured in the figure.

B.2 Endogenous productivity

In the model, we treat productivity terms A as exogenous. In this subsection, we extend our results to the case where A is endogenous to the composition of labor in the city. More specifically, let us consider the following form for productivity terms:

$$A(c,j) = G\left(\int_{\omega\geq\omega(h,c)} g(\omega)f(\omega,c)d\omega, j\right)$$

where $G(.,j)$ is an increasing function for all $j \in \{l,m,h\}$ and $G(x,j)$ is log-supermodular in $\{x,j\}$. Both assumptions ensure that, in equilibrium, Assumption 3 is satisfied.

Under these conditions, we obtain that:

Proposition B.1. *A decline in p_z leads to*

(i) *an increase in the absolute advantage of city* 1*, i.e.* $A(j,1)/A(j,1)$ *increases for* $j \in \{l,m,h\}$,

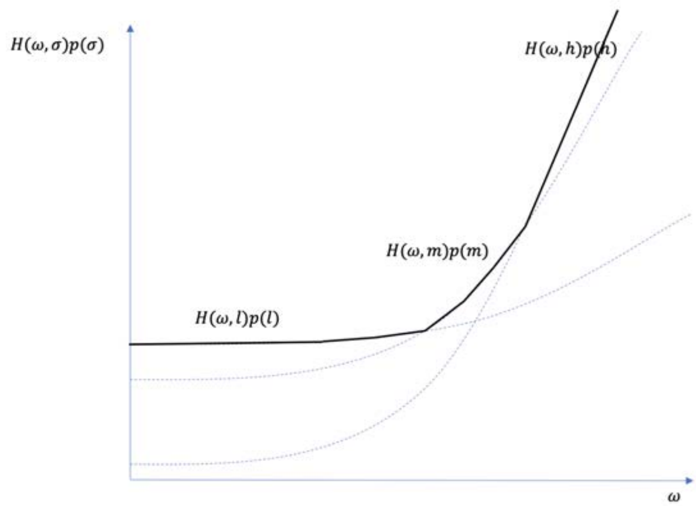


Figure B.1: Value marginal products in the different sectors.

This figure depicts the value marginal products as a function of the skill ω for the three sectors $\sigma \in \{h, m, l\}$. The value marginal productivity is the productivity function $H(\omega, \sigma)$ weighted by the price of the sector's output ($p(\sigma)$, $\sigma \in \{h, m, l\}$). The plain black line is the upper envelope of these value marginal products.

(ii) *an increase in the comparative advantage of City 1 in higher-paid activities, that is* $A(h,1)/A(h,2) - A(m,1)/A(m,2)$ *and* $A(m,1)/A(m,2) - A(l,1)/A(l,2)$ *are increasing.*

By increasing the share of the population in the high-paid sector, the polarization shock increases productivity. Yet, as shown in Proposition D.1, the increase in the share of the population in the high-paid sector is more important in the larger than in the smaller city. Thus, this increases productivity by more in the larger city, reinforcing the absolute advantage of this city. Given that we assume that productivity reacts by more for higher-paid occupations, this also leads to a reinforcement of the comparative advantage in the larger city for higher skill occupations.

Remark. A potential pitfall with endogenous productivity is that it can lead to multiple equilibria. For example, in our model, symmetric cities can also be an equilibrium outcome if productivity is endogenous. To extend our results to endogenous productivity would then require to maintain the assumption that we stay close to the selected equilibrium. We also refer the interested reader to Davis and Dingel (2020) for a discussion of the possibility of multiple equilibria in a related setting.

B.3 N cities

The benchmark model only considers two cities. We extend here the model to N cities.

Let us then index cities by $c \in \{1, 2, ..., N\}$. We order cities so that for any $i, j \in \{1, 2, ..., N\}$ so that if $i > j$, we assume that city i has an absolute advantage over city j in all occupations $A(i, \sigma) > A(j, \sigma)$ for $\sigma \in \{l, m, h\}$ and it has a comparative advantage in higher skill occupations: $A(i, h)/A(j, h) > A(i, m)/A(j, m) > A(i, l)/A(j, l)$.

As this can be observed, if $i > j > k$, the absolute advantage of i over j and the absolute advantage of j over k leads to an absolute advantage of i over k. Similarly, we obtain such a transitivity for the comparative advantage. To illustrate, the comparative advantage of city i in skill h with respect to city j and the same comparative advantage for city j with respect to city k leads to $A(i,h)/A(i,m) > A(j,h)/A(j,m) > A(k,h)/A(k,m)$, which implies $A(i,h)/A(k,h) > A(i,m)/A(k,m)$, that is that city i has a comparative advantage in skill h compared with city k.

Sectoral decisions As a result of these assumptions and extending Proposition 1, the thresholds $\omega(m,c)$ and $\omega(h,c)$ are decreasing with city size. Furthermore, we can extend Lemma D.1 and obtain that a change in p_z leads to a stronger decline in $\omega(h,c)$ in large cities and $\omega(m,c)$ increase by more in smaller cities.

Location decisions The description of location decision within a city as described in Lemmas A.1, A.2 and Proposition 2 given that these results apply for any city c. We then only need to describe how agents decide to choose locations between the different N cities.

To start with, for any $i \leq N-1$, there are locations in city $c \in \{1,...,i\}$ where the productivity of worker is strictly higher than what it could be in any city $c \geq i+1$. This happens for locations τ where productivity in city $c \leq i$ strictly exceeds what can be obtained in city $c > i$, even in the best location. More formally:

$$H(\omega(\tau(c)), M(\omega(\tau(c)), c), c)T(\tau(c)) > H(\omega(\tau(c)), M(\omega(\tau(c)), i+1), i+1)T(0) \tag{15}$$

where $\omega(\tau(c)) = K(c, \tau(c))$ is the value of ω occupying location $\tau(c)$ in city c. This defines a maximum value for the skill in city $i+1$, $\bar{\omega}(i+1)$ above which higher skills are only present in cities $\{1,...,i\}$. As a result, any agent with a skill higher than $\bar{\omega}(i+1)$ will decide to live only in cities $c \leq i$.

Below this threshold $\bar{\omega}(i+1)$, for each ω and for each τ, there exists $\tau' < \tau$ such that the productivities in City 1 and in City 2 are the same:

$$H(\omega(\tau), M(\omega(\tau, c), c), c)T(\tau) = H(\omega(\tau), M(\omega(\tau), i+1), i+1)T(\tau'). \tag{16}$$

which implies that this agent is indifferent in living between, at least, any city $c \leq i+1$. In the end, households are indifferent between a less desirable location in the more productive and larger city $c \leq i$ or a more desirable location in the less productive and smaller city $i+1$. Similarly, between two locations c and c' such that $c \leq c' \leq i$, households hesitate between more desirable locations in city c' and less desirable ones in city c.

Results. As for the 2-city case, labor market polarization will happen in the aggregate and across cities. This results from sectoral decisions. As for the 2-city case, the distribution of skills is going to be log-supermodular. Similarly, if, for any city $c \in \{1,...,N-1\}$ the comparative advantage of c over city $c-1$ in high skill occupations is sufficiently large, i.e. $A(i,h)/A(i-1,h)$ is sufficiently large compared with $A(i,m)/A(i-1,m)$ for all $i \leq N-1$, we also obtain the results of Proposition D.2 about initial exposures and of Proposition D.1.

B.4 Log-supermodularity of skills across cities

Let us first describe the allocation of skills and the exposures to different sectors across cities. Our main result is that the large city can have a smaller exposure to the middle-paid sector. Importantly, this result does not stem from large cities being poorer in middle-skill agents, as we show that this can happen even if the distribution of skills is log-supermodular.

Log-supermodularity Our second implication concerns the distribution of skills across the two cities. To this purpose, let us introduce the supply of locations within a city:

$$V(z) = -\frac{\partial}{\partial z} S\left(T^{-1}(z)\right)$$

This function indicates the number of locations within a city with $\tau = T^{-1}(z)$. Following Davis and Dingel (2020), we can now obtain the following proposition:

Proposition B.2. *Assume that the supply of locations in each city $V(z)$ has a sufficiently decreasing elasticity. Then, the distribution of skills $f(\omega,c)$ is strictly log-supermodular.*

Proof. To start with, let us derive the pdf of the distribution of agents across cities, $f(\omega,c)$. The population of individuals with skills between ω and $\omega + d\omega$ is:

$$L \int_{\omega}^{\omega+d\omega} f(x,c)dx = S\left(T^{-1}\left(h(\omega,c)\right)\right) - S\left(T^{-1}\left(h(\omega+d\omega,c)\right)\right) \tag{17}$$

again with $h(\omega,c)$ defined by $K(T^{-1}(h(\omega,c)),c) = \omega$. Taking the derivative with respect to $d\omega$ and taking $d\omega \to 0$ yield:

$$f(\omega,c) = -\frac{\partial}{\partial \omega} S\left(T^{-1}\left(h(\omega,c)\right)\right) = h'(\omega,c)V(h(\omega,c)) \tag{18}$$

with $V(.) = -\frac{\partial}{\partial \omega} S\left(T^{-1}(.)\right)$.

Let us first note that $f(\omega,c)$ is log-supermodular if and only if, for all $\omega > \omega'$ and $c > c'$, we have $f(\omega,c)f(\omega',c') > f(\omega',c)f(\omega,c')$. When $f(\omega,c')$ and $f(\omega',c')$ are different than 0, this condition amounts to verifying than $f(\omega,c)/f(\omega,c')$ is strictly increasing or, equivalently that:

$$f'(\omega,c)f(\omega,c') > f'(\omega,c')f(\omega,c). \tag{19}$$

Using the fact that $f(\omega,c) = h'(\omega,c)V(h(\omega,c))$, we can compute:

$$f'(\omega,c) = h''(\omega,c)V(h(\omega,c)) + (h'(\omega,c))^2 V'(h(\omega,c))$$

By denoting $\xi(V, h(\omega, c)) = h(\omega, c)V'(h(\omega, c))/V(h(\omega, c))$, we obtain that:

$$\begin{aligned} f'(\omega, c) =& h''(\omega, c)V(h(\omega, c)) + (h'(\omega, c))^2 \xi(V, h(\omega, c))V(h(\omega, c))/h(\omega, c) \\ =& f(\omega, c) \left(\frac{h''(\omega, c)}{h'(\omega, c)} + \xi(V, h(\omega, c)) \frac{h'(\omega, c)}{h(\omega, c)} \right) \end{aligned}$$

Replacing $f'(\omega, c)$ and $f'(\omega, c')$ by their values in (19), we then obtain the following condition:

$$\frac{h''(\omega, c)}{h'(\omega, c)} + \xi(V, h(\omega, c)) \frac{h'(\omega, c)}{h(\omega, c)} > \frac{h''(\omega, c')}{h'(\omega, c')} + \xi(V, h(\omega, c')) \frac{h'(\omega, c')}{h(\omega, c')} \tag{20}$$

A straightforward implication of this necessary and sufficient condition is the following.

Lemma B.1. (i) *If, for ω and ω' and for c and c', the occupation decisions are the same across cities, that is $M(\omega, c) = M(\omega, c')$ and $M(\omega', c) = M(\omega', c')$, then*

$$f(\omega, c)f(\omega', c') > f(\omega', c)f(\omega, c')$$

if and only if $\xi(V, x)$ is decreasing in x.

(ii) *If productivities are constant across occupations, $A(c, h) = A(c, m) = A(c, l)$ as in Davis and Dingel (2020), a necessary and sufficient condition for $f(\omega, c)$ to be log-supermodular is that $\xi(V, x)$ is decreasing in x.*

Proof. Suppose that $M(\omega, c) = M(\omega, c')$ and $M(\omega', c) = M(\omega', c')$, then, in equilibrium:

$$A(c, M(\omega, c))H(\omega, M(\omega, c))h(\omega, c) = A(c', M(\omega, c'))H(\omega, M(\omega, c'))h(\omega, c')$$

and thus $h(\omega, c) = h(\omega, c')$. By continuity, $M(\omega, c) = M(\omega, c')$ on a (right- or left-) neighborhood of ω and thus $h(., c) = h(., c')$ on this neighborhood, thus ensuring that locally $h''(., c) = h''(., c')$ and $h'(., c) = h'(., c')$ and in particular that $h''(\omega, c) = h''(\omega, c')$ and $h'(\omega, c) = h'(\omega, c')$. In the end, (20) simplifies into $\xi(V, h(\omega, c)) > \xi(V, h(\omega, c'))$, which is satisfied as long as V features decreasing elasticity.

$\square$

The conclusion of the second point is that, with V featuring decreasing elasticity, we obtain that f is log-supermodular on subsets where the occupation decisions are the same, that is $[\omega(h, 2), \overline{\omega}]$, $[\omega(m, 2), \omega(h, 1)]$ and $[\underline{\omega}, \omega(m, 1)]$.

Let us now turn to the segments $[\omega(h, 1), \omega(h, 2)]$ and $[\omega(m, 1), \omega(m, 2)]$, where households have different occupation choices depending on cities. Let us first show that it is sufficient to show that f is log-supermodular on each of these two segments to obtain log-supermodularity on $[\underline{\omega}, \overline{\omega}]$.

Lemma B.2. *Suppose that $f(x, c)$ is log-supermodular in $\{x, c\}$ on $[\underline{x}, \overline{x}]$ and $[\overline{x}, \overline{\overline{x}}]$, then $f(x, c)$ is log-supermodular in $\{x, c\}$ on $[\underline{x}, \overline{\overline{x}}]$*

Proof. Let us consider any x and x' in $[\underline{x}, \overline{\overline{x}}]$ such that $x > x'$. Let us also consider two cities c and c' such that $c > c'$. If x and x' are both in the same segment, either $[\underline{x}, \overline{x}]$ or $[\overline{x}, \overline{\overline{x}}]$, we already have log-supermodularity. So, let us consider the case where $x \geq \overline{x} \geq x'$.

Using log-supermodularity on $[\overline{x}, \overline{\overline{x}}]$, we have $\frac{f(x,c)}{f(x,c')} > \frac{f(\overline{x},c)}{f(\overline{x},c')}$. Using log-supermodularity on $[\underline{x}, \overline{x}]$, we have $\frac{f(\overline{x},c)}{f(\overline{x},c')} > \frac{f(x',c)}{f(x',c')}$. Combining these two equations, we obtain $\frac{f(x,c)}{f(x,c')} > \frac{f(x',c)}{f(x',c')}$. In the end, f is then log-supermodular on $[\underline{x}, \overline{\overline{x}}]$.

□

We now need to establish log-supermodularity on $[\omega(h,1), \omega(h,2)]$ and $[\omega(m,1), \omega(m,2)]$.

Let us start with some properties on the $h(\omega, c)$ function. The indifference condition between location implies that $\phi(\omega) = H(\omega, M(\omega,c), c)h(\omega,c) = H(\omega, M(\omega,c'), c')h(\omega,c')$.

Given that $M(\omega,c) \geq M(\omega,c')$ due to the comparative advantage of the large city and that $H(\omega, M(\omega,c), c) \geq H(\omega, M(\omega,c'), c')$, we have that $h(\omega,c) \leq h(\omega,c')$. Furthermore, given that $H(\omega, M(\omega,c), c)/H(\omega, M(\omega,c'), c')$ is an increasing function of ω, we obtain that $h(\omega,c')/h(\omega,c)$ is an increasing function of ω and $h'(\omega,c) \leq h'(\omega,c')$. Finally, $H(\omega, M(\omega,c), c)$ being log-supermodular, we obtain that $h(\omega,1)h(\omega',2) \leq h(\omega',1)h(\omega,2)$ and that

$$\frac{h'(\omega,1)}{h(\omega,1)} \leq \frac{h'(\omega,2)}{h(\omega,2)}$$

A first conclusion is then that when $\eta(V) \leq 0$ and decreasing, we obtain that:

$$\xi(V, h(\omega,1))\frac{h'(\omega,1)}{h(\omega,1)} > \xi(V, h(\omega,2))\frac{h'(\omega,2)}{h(\omega,2)}. \tag{21}$$

Second, note that (20) is invariant to equilibrium prices. In the end, when $\xi(V, h(\omega,1))$ is sufficiently decreasing, condition (20) is satisfied. □

Recall that a distribution is strictly log-supermodular when, for $c > c'$ (i.e. city c is larger than city c') and $\omega > \omega'$, $f(\omega,c)f(\omega',c') > f(\omega,c')f(\omega',c)$, which means that there are relatively more high skill workers in the larger city. Proposition B.2 extends Davis and Dingel (2020) to a situation where City 1 does not only have an absolute advantage over City 2 but has a comparative advantage in higher-skilled sectors.

Middle-paid jobs as a function of city size Jointly Proposition D.2 and B.2 lead to some important implications. Given our previous result in Proposition D.2 where we obtained conditions under which the share of middle-paid jobs is smaller in the larger city, we can also characterize the elasticity of the middle-paid jobs with respect to the size of the city:

Corollary 1. *Under the conditions of Proposition D.2, the elasticity for middle-paid jobs with respect to the size of the city is lower than* 1.

One implication of this result, associated with the fact that larger cities have a lower initial share of middle-paid workers as shown in Proposition D.2, is that occupations, in contrast to skills, do not need to be log-supermodular. The total number of jobs in the middle-paid occupations may be lower in the larger city compared with the smaller city.

More precisely, we obtain such a discrepancy between skills and sectors as a result of the endogenous sectoral decisions by households when the comparative advantage of City 1 is in the high-paid sector: interim-skilled (interim ω) residing in City 1 work less in the middle-paid sector and relatively more in the high-paid sector – formally, $\omega(h,1) < \omega(h,2)$. This is consistent with the idea that two similarly skilled individuals may not work in the same sector depending on the cities in which they live.

This result has to be contrasted for example with extreme-skill complementarity as put forward by Eeckhout et al. (2014): the implication of such complementarity would be that the smaller share of middle-paid jobs would stem from a smaller share of interim-skilled individuals.

Share of low-paid jobs across cities Given Proposition 1 and the log-supermodularity result, we can state:

Corollary 2. *The share of low-paid workers is lower in the larger city.*

C Simulating the model – for online publication only

In this Appendix, we provide further details on the algorithm that we use to simulate the model and the way we match it with data on wages. We also report results when we assume a uniform distribution for skills $f(.)$.

C.1 Algorithm

Stage 1. We first discretize the set of skills $[\underline{\omega}, \overline{\omega}]$ into N values, equally spaced. We then discretize the distribution $f(.)$ on this grid.

Stage 2. We compute the thresholds $\omega(m,c)$ and $\omega(h,c)$ in the two cities.

Stage 3. We determine $\bar{\omega}(2)$ in the following way: $\bar{\omega}(2)$ and the associated $\tau(2)$ is then defined by

$$A(1,h)H(\bar{\omega}(2), M(\bar{\omega}(2),1))T(\bar{\tau}(2)) = A(2,h)H(\bar{\omega}(2), M(\bar{\omega}(2),2))T(0)$$

$$\int_{\bar{\omega}(2)}^{\overline{\omega}} f(\omega)d\omega = S(\bar{\tau}(2))$$

Stage 4. We now allocate workers with skills in $[\underline{\omega}, \bar{\omega}(2)]$. Given initial $\tau_1 = \bar{\tau}_2$, $\tau_2 = 0$ and $\omega = \bar{\omega}(2)$, we iterate in the following way:

- Determine which sectors prevail in City 1 and City 2 by comparing ω and $\omega(m,c)$ and $\omega(h,c)$, in $c \in \{1,2\}$.

- Solve in $x \geq 0$ the following equation $L \times f(\omega) = S(x+\tau_1) - S(\tau_1) + S(\tau_2 + g(x)) - S(\tau_2)$ where $g(x)$ is defined as:

$$A(M(\omega,1),1)p(M(\omega,1))H(\omega,M(\omega,1))T(\tau_1+x) = \cdots$$
$$\cdots A(M(\omega,2),2)p(M(\omega,2))H(\omega,M(\omega,2))T(\tau_2+g(x))$$

- Iterate the algorithm with the next ω and $\tau_1' = \tau_1 + x$ and $\tau_2' = \tau_2' + g(x)$, until $\omega = \underline{\omega}$.

C.2 Connecting data to the model

We now use data to parametrize and match the model with data on wages.

Matching the wage distribution. The observed wage distribution is our first observable that we use. To connect this distribution to our model, we assume that the observed hourly wage of an individual i is:

$$w_i = A(\sigma,c)p(\sigma)H(\omega,\sigma) \tag{22}$$

where (ω,σ,c) are, respectively, the skill, the sector and the city of the individual i. Implicitly, this means that we do not take into account the term $T(\tau)$.

Remark. A concern with assuming away $T(\tau)$ can be that, in equilibrium, τ is correlated with ω, as higher skilled workers locate in better locations. The way to think about this assumption that $T(\tau)$ does not enter (22) is that $T(\tau)$ is a productivity loss that affects only the number of hours worked but not the hourly wage. For example, if workers have a fixed amount of time $\bar{l}$ to allocate between working and commuting and commuting time is $\bar{l} - T(\tau)$, the wage received by an individual i is $w_i T(\tau_i)$.

Taking the logarithm of Equation (22), we obtain $\log w_i = \log p(\sigma)A(\sigma,c) + \log H(\omega,\sigma)$. By running the regression:

$$\log w_i = C + \sum_{c,\sigma} \delta_{c,\sigma} + v_i \tag{23}$$

we obtain that, for each sector $\sigma \in \{l,m,h\}$:

$$\log A(\sigma,1)p(\sigma) - \log A(l,2)p(l) = \delta_{1,\sigma} - \delta_{2,l} \tag{24}$$

and individual effects $\log H(\omega,\sigma) = v_i$ that do not depend on city c, according to our model specification.

Let us now turn to function H that maps skills and sectors into productivity. Assuming a distri-

	Low-paid	Middle-paid	High-paid
City-sector productivities $A(c, \sigma)$			
$A(1, \sigma)/A(2, \sigma)$	1.037	1.059	1.086
Individual-sector productivities $H(\omega, \sigma)$			
$\mu(\sigma)$	-0.120	-0.065	-0.095
$\nu(\sigma)$	0.245	0.356	0.486

Table C.1: Model-based productivity estimates: base case

bution $f(.)$ for ω and denoting by g the distribution of v, we can then infer:

$$\int_{\underline{\omega}}^{\omega} H(x, \sigma) f(x) dx = \int_{\underline{v}}^{v} y g(y) dy$$

This relies on the result that higher skilled individuals have higher income. Notice, however, that this equation also means that H and f cannot be inferred independently.

Finally, as the share of the low-paid workers evolves between 12% and 19% between 1994 and 2015, we infer the values for the low-paid sector on the 15-20% range. For the high-paid, we take the 70-75% range, as it is where the high-paid sector starts to appear. Finally, We take data from the 20-70% range for the middle-paid sector. Table C.1 reports the estimates.

C.3 Upper and lower tier middle paid job loss

In Figure C.1, we plot the difference across cities in the shares of middle-paid jobs when we split middle-paid jobs into those occupied by higher-skilled households (right panel) and those occupied by lower-skilled households (left panel). As we can observe, in the large city, higher-skilled middle-paid jobs have disappeared at a faster pace while lower-skilled middle-paid jobs disappeared more quickly in the small city.

Figure C.1 also allows us to grasp the intuition on why we obtain skewed polarization that, in contrast to prior work, is inversely tied to initial exposure. The decline in the price of the middle-paid good corresponds, in both cities, to a decline in the threshold between high- and middle-paid jobs $\omega(h, c)$ and a rise in the threshold between low- and middle-paid jobs $\omega(m, c)$. However, these movements in thresholds do not correspond to similar outflows of middle-paid jobs across cities and, as Figure C.1 illustrates, this outflow is greater from middle- to high-paid jobs in the large city, sufficiently for middle-paid jobs to decline the most in the large city. Thus, the strength of labor market polarization and its direction does not depend simply on the *average* number of middle-paid workers but rather on the incentives and the numbers of *marginal* middle-paid workers that may shift to other sectors.

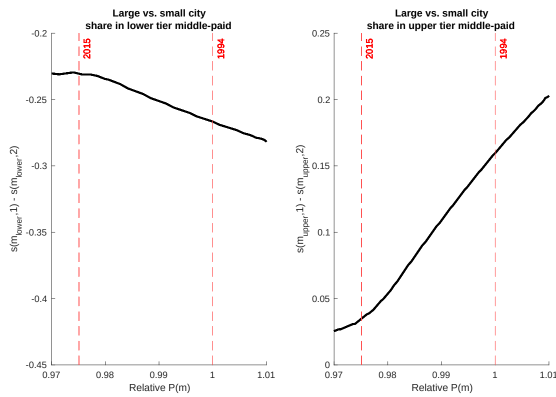


Figure C.1: Evolution across cities of upper-tier and lower-tier middle-paid jobs

In this figure, we compute the difference between the share of the upper-tier middle-paid jobs in large cities and the one in the small city (right panel – $s(m_{upper}, 1) - s(m_{upper}, 2)$) and the difference between the share of the lower-tier middle-paid jobs in large cities (left panel – $s(m_{lower}, 1) - s(m_{lower}, 2)$). We define upper-tier middle-paid jobs as jobs occupied by households with skill above average. The rest is classified as lower-tier middle-paid jobs.

Uniform distribution assumption. Our second approach is to assume that skills are uniformly distributed over Ω and the distribution of $H(\omega, \sigma)$ then derives from differences in the mapping between ω and $H(\omega, \sigma)$. To this purpose, we parametrize productivities as follows to capture the concave, linear and convex portions of H:

$$H(\omega, l) = \omega^{\phi},\ H(\omega, m) = \epsilon\omega \text{ and } H(\omega, h) = \exp \eta\omega - 1.$$

In addition, we assume that $\Omega = [0, \overline{\omega}]$. We obtain that $\phi = 0.135$, $\epsilon = 0.86$ and $\eta = 1.45$ – the parameters $A(\sigma, c)$ are unchanged.

Other parametrization and calibration. Following Davis and Dingel (2020), we parametrize $S(\tau) = \pi\tau^2$ and $T(\tau) = 1 - d_1\tau$. We calibrate d_1 to match the relative size of cities above 500k inhabitants and those between 50k and 100k inhabitants.

C.4 Results – uniform distribution

Let us investigate how a polarization shock affects the distribution of sectors across cities.

Aggregate shares			
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.005	-0.015	+0.01
data	+0.07	-0.16	+0.09
Relative shares			
	$\Delta(s(l,1) - s(l,2))$	$\Delta(s(m,1) - s(m,2))$	$\Delta(s(h,1) - s(h,2))$
model	-0.006	-0.005	+0.011
data	-0.02	-0.06	+.08
Initial exposure in 1994			
		$s(m,1) - s(m,2)$	
model		-0.15	
data		-0.11	

Table C.2: Simulation-based sectoral distribution – uniform distribution

Even if the quantitative effects are smaller, we observe that, qualitatively:

1. the initial share of the middle-paid sector is larger in the small city ($s(1,m) - s(2,m) < 0$),
2. a decline n the price of the middle-paid good leads to a decline in the middle-paid sector ($s(m)$ declines), consistent with labor market polarization in the aggregate,
3. this decline is, at least in the beginning stronger in the large city ($s(1,m) - s(2,m)$ declines),
4. the increase in the high-paid sector is stronger in the large city ($s(1,h) - s(2,h)$ is positive and increases),
5. the increase in the low-paid sector is stronger in the small city ($s(1,l) - s(2,l)$ is negative and decreases),

The intuition why, with a uniform distribution, we still get skewed polarization requires some further analysis. Indeed, in contrast with the normal distribution, the number of agents concerned by a change of sector does not depend on the initial level of thresholds. In this case, skewed polarization results from different reactions of thresholds to the same shock. We clarify this point in Appendix D with the functional form assumed with the uniform distribution.

C.5 Normal distribution – robustness – shock

In this subsection, we show two robustness exercices for different shock structures. First, we show the case in which the relative price $p(h)/p(l)$ is kept constant and the price of the middle-paid good decreases. This case is reported in Table C.3. Second, we show the case in which the price $p(h)$ increases while the prices $p(m)$ and $p(l)$ are constant. This case is reported in Table C.4.

The overall picture of Table C.3 is not so different compared to what we obtained in the benchmark case, at least qualitatively. Quantitatively, the effect of polarization on middle-paid jobs in large cities is much milder. In the absence of an increase in the price of the high-paid good, the incentives to shift to the high-paid sector are lower. This then limits polarization especially in the large city in which the polarization is titled to this high-paid sector. The increase is in the price of the high-paid good is then useful to have a quantitatively meaningful skewed polarization but it is not necessary to qualitatively obtain the skewed polarization.

	Aggregate shares		
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.07	-0.14	+0.07
data	+0.07	-0.16	+0.09
	Relative shares		
	$\Delta(s(l,1)-s(l,2))$	$\Delta(s(m,1)-s(m,2))$	$\Delta(s(h,1)-s(h,2))$
model	-0.08	0.00	+0.08
data	-0.02	-0.06	+.08
	Initial exposure in 1994		
		$s(m,1)-s(m,2)$	
model		-0.04	
data		-0.11	

Table C.3: Simulation-based sectoral distribution – no variation of the price of the high-paid good.

Table C.4 conveys a similar picture. A shock only to $p(h)$ increases the overall share of high-paid jobs, at the expense of middle-paid jobs. Given the comparative advantage of the large city for these jobs, the large becomes richer in high-paid jobs and loses more middle-paid jobs. However, the effects are also quantitatively small even if this leads qualitatively to skewed polarization. Notice also that low-paid jobs do not gain any importance in this case and they remain a constant share of the workforce, overall and in each individual city.

Aggregate shares			
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.00	-0.07	+0.07
data	+0.07	-0.16	+0.09
Relative shares			
	$\Delta(s(l,1)-s(l,2))$	$\Delta(s(m,1)-s(m,2))$	$\Delta(s(h,1)-s(h,2))$
model	-0.01	-0.07	+0.08
data	-0.02	-0.06	+.08
Initial exposure in 1994			
		$s(m,1)-s(m,2)$	
model		-0.04	
data		-0.11	

Table C.4: Simulation-based sectoral distribution – no variation of the price of the middle-paid good

C.6 Normal distribution – robustness – non-tradable services

In this appendix, we report the results of a simulation in which a low-paid jobs are working in a non-tradable sector. Following Davis and Dingel (2019), this sector is proportional in size to the total population in the city – e.g., there is a demand of one unit of non-tradable good per inhabitant. Also, we assume that the productivity in the low-paid sector does not depend on localisation τ. We calibrate the total share of low-paid in the non-tradable sector at 7% of total workforce to match the relative decline in middle-paid jobs in the large city.

	Aggregate shares		
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.04	-0.17	+0.13
data	+0.07	-0.16	+0.09
	Relative shares		
	$\Delta(s(l,1)-s(l,2))$	$\Delta(s(m,1)-s(m,2))$	$\Delta(s(h,1)-s(h,2))$
model	-0.09	-0.06	+0.15
data	-0.02	-0.06	+.08
	Initial exposure in 1994		
		$s(m,1)-s(m,2)$	
model		-0.13	
data		-0.11	

Table C.5: Simulation-based sectoral distribution – non-tradable services

C.7 Normal distribution – robustness – cities $<$ 200k and cities $>$ 200k

In this appendix, instead of comparing cities below 500k and cities below 50k-100k, we compare cities below 200k inhabitants and cities above 200k inhabitants.

	Aggregate shares		
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.06	-0.19	+0.13
data	+0.07	-0.16	+0.09
	Relative shares		
	$\Delta(s(l,1)-s(l,2))$	$\Delta(s(m,1)-s(m,2))$	$\Delta(s(h,1)-s(h,2))$
model	-0.05	-0.04	+0.08
data	-0.02	-0.05	+0.06
	Initial exposure in 1994		
		$s(m,1)-s(m,2)$	
model		-0.09	
data		-0.09	

Table C.6: Simulation-based sectoral distribution – cities $<$ 200k and cities $>$ 200k

C.8 Normal distribution – robustness – cities < 250k and cities > 750k

In this appendix, instead of comparing cities below 500k and cities below 50k-100k, we compare cities below 250k inhabitants and cities above 750k inhabitants.

Aggregate shares			
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.07	-0.20	+0.13
data	+0.07	-0.16	+0.09
Relative shares			
	$\Delta(s(l,1)-s(l,2))$	$\Delta(s(m,1)-s(m,2))$	$\Delta(s(h,1)-s(h,2))$
model	-0.05	-0.07	+0.11
data	-0.01	-0.06	+0.07
Initial exposure in 1994			
		$s(m,1)-s(m,2)$	
model		-0.09	
data		-0.09	

Table C.7: Simulation-based sectoral distribution – cities < 250k and cities > 750k

C.9 Normal distribution – robustness – intervals for productivity estimation

In this Appendix, we show two robustness exercises on parameters used in the base scenario shown in Table C.1 and Figure 7 for different estimates of productivity. In Table C.9, we take estimates — reported in Table C.8 — on the range $[0, 20]$ for the low-paid sector, $[20, 70]$ for the middle-paid sector and $[70, 100]$ for the high-paid sector. In Table C.11, we take estimates — reported in Table C.10 on the range $[15, 20]$ for the low-paid sector, $[42.5, 47.5]$ for the middle-paid sector and $[70, 75]$ for the high-paid sector.

Overall, we still obtain skewed labor market polarization, but with quantitative variations compared with our benchmark case.

	Low-paid	Middle-paid	High-paid
Individual-sector productivities $H(\omega,\sigma)$			
$\mu(\sigma)$	-0.174	-0.053	-0.257
$\nu(\sigma)$	0.195	0.355	0.685

Table C.8: Model-based productivity estimates

Aggregate shares			
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.04	-0.09	+0.05
data	+0.07	-0.16	+0.09
Relative shares			
	$\Delta(s(l,1) - s(l,2))$	$\Delta(s(m,1) - s(m,2))$	$\Delta(s(h,1) - s(h,2))$
model	-0.05	-0.02	+0.06
data	-0.01	-0.06	+0.07
Initial exposure in 1994			
		$s(m,1) - s(m,2)$	
model		-0.04	
data		-0.09	

Table C.9: Simulation-based sectoral distribution – alternative estimates 1/2

	Low-paid	Middle-paid	High-paid
Individual-sector productivities $H(\omega, \sigma)$			
$\mu(\sigma)$	-0.118	-0.065	-0.094
$\nu(\sigma)$	0.247	0.359	0.484

Table C.10: Model-based productivity estimates

Aggregate shares			
	$\Delta s(l)$	$\Delta s(m)$	$\Delta s(h)$
model	+0.08	-0.21	+0.13
data	+0.07	-0.16	+0.09
Relative shares			
	$\Delta(s(l,1)-s(l,2))$	$\Delta(s(m,1)-s(m,2))$	$\Delta(s(h,1)-s(h,2))$
model	-0.09	-0.08	+0.16
data	-0.01	-0.06	+0.07
Initial exposure in 1994			
		$s(m,1)-s(m,2)$	
model		-0.09	
data		-0.09	

Table C.11: Simulation-based sectoral distribution – alternative estimates 2/2

D Additional theoretical results – for online publication only

In this Appendix, we provide some additional theoretical results on skewed labor market polarization and the great urban divergence. To obtain these results, we rely on the productivity functions as in the calibration in the case of a uniform distribution (Appendix C.4). Such productivity forms feature increasing convexity. In this case, we show that thresholds between sectors move differently across cities as a function of cities' comparative advantage and may lead to skewed polarization – at least in the case in which the large city concentrates all the high-paid jobs.

D.1 Middle-paid job loss and initial exposure

We now investigate the evolution of high-, middle- and low-paid jobs across cities and we connect this evolution to the initial exposure to middle-paid jobs. We find conditions under which middle-paid jobs decrease by *more* in large cities, despite a *smaller* initial exposure of these cities to middle-paid jobs.

The evolution of jobs. Let us characterize the evolution of high-, middle- and low-paid jobs. To this end, we first observe what the key drivers of such evolutions, for example, using the evolution of the share of high-paid jobs. In city c, in response to a change $dp < 0$ in the relative price of middle-paid to high-paid goods, this share evolves as:

$$ds(h,c) = \underbrace{-\frac{f(\omega(h,c),c)}{L(c)}\frac{1}{\Theta(\omega(h,c),c,h)}\frac{dp}{p}}_{\text{Within-city reallocation}} + \underbrace{\int_{\omega(h,1)}^{\omega(h,2)}\frac{\partial f(\omega,c)}{\partial p}d\omega\frac{dp}{p}}_{\text{Across-city reallocation}}. \tag{25}$$

The evolution of this share depends on two factors. On the one hand, this share increases in each city due to the reallocation of middle-paid workers that already were in city c to the high-paid sector. The strength of this first channel depends on the density of middle-paid jobs close to the threshold, $f(\omega(h,c),c)/L(c)$, and the variation of the threshold due to the relative price change. This latter variation is $d\omega(h,c) = 1/\Theta(\omega(h,c),c,h)dp/p$ and results from the indifference condition (Equation (5)) between the high- and the middle-paid sectors. On the other hand, the relative price decline makes city 1 more attractive for households with productivity between $\omega(h,1)$ and $\omega(h,2)$: these households are initially indifferent between the high-paid sector in City 1 and the middle-paid in City 2. For this range of skills, a decline in the relative price of the good produced by the middle-paid sector then reduces the attractiveness of City 2 compared with City 1, thus leading to a reallocation of workers to City 1 – the term of across-city reallocation is then positive. In the end, when the comparative advantage of the large city in the high-paid sector is sufficiently large, the share of high-paid workers increases only in the large city, proving the result.

A simple but important observation also emerges from (25): what drives the evolution of high-paid jobs is how the relative price change affects the incentives of workers close to the threshold $\omega(h,c)$ and these incentives are influenced by the the distribution of such workers ($f(\omega(h,c),c)$ as well as by the *local* patterns of workers productivity, that affects $\Theta(\omega(h,c),c,h)$. In particular, this implies that such evolutions are not related to exposures, inter alia the exposure to middle-paid jobs, which would correspond to an *average* on a large set of skills (ω). Note that a similar observation can be made on the evolution of the share of low-paid jobs and, as by construction $ds(m,c) = -ds(h,c) - ds(l,c)$, on the evolution of the share of middle-paid jobs.

Skewed polarization. Let us now find conditions under which, consistent with the patterns observed in the data, polarization is tilted towards high-paid jobs and middle-paid jobs decrease by more in the large city. Of course, as these conditions are not necessary ones, one can obtain a stronger shock on middle-paid jobs in the large city under milder assumptions.

To obtain some of these results, let us consider the following functional form:

Assumption D.1 (Functional form of productivity)**.** *In city c, the productivity of an agent with productivity ω where $1 \leq \underline{\omega} \leq \omega \leq \overline{\omega}$ is:*

$$H(\omega,l) = \omega^{\phi} \textit{ with } \phi \in (0,1),\ H(\omega,m) = \omega \textit{ and } \ H(\omega,h) = e^{\eta\omega} - 1 \textit{ with } \eta = 1/\underline{\omega}.$$

This form of productivity is the one we use in the case of a uniform distribution of skills in

Appendix C.4. It helps to capture the increasing convexity of individual fixed effects that we obtain in the data.

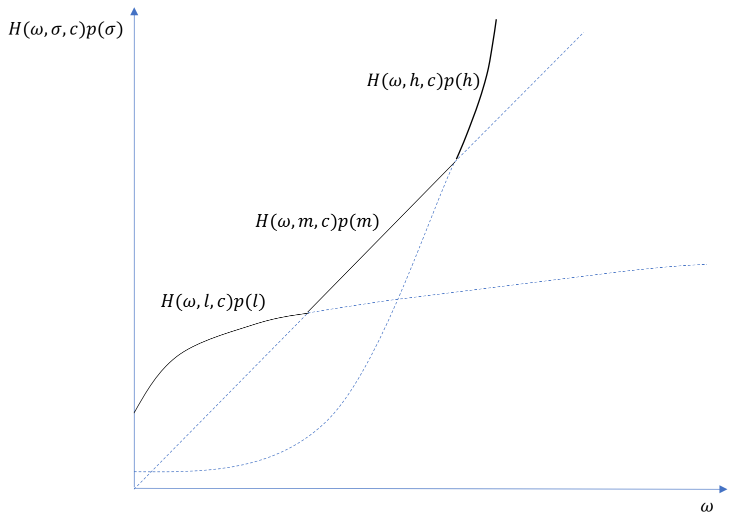


Figure D.1: Value marginal products in the different sectors.

This figure depicts the value marginal products as a function of the skill ω for the three sectors $\sigma \in \{h, m, l\}$. The value marginal productivity is the productivity function $H(\omega, \sigma, c)$ weighted by the price of the sector's output ($p(\sigma)$, $\sigma \in \{h, m, l\}$). The plain black line is the upper envelope of these value marginal products.

This functional form is log-convex for the high-paid sector. Figure D.1 plots the value marginal products resulting from this Assumption. We then obtain:

Proposition D.1 (Middle-paid job loss)**.** *When $\frac{A(h,1)}{A(h,2)}$ is sufficiently large relative to $\frac{A(m,1)}{A(m,2)}$, i.e. comparative advantage in these sectors is sufficiently strong, then a decline in p_z implies that in the large relative to the small city the increase in high-paid jobs is larger in percentage points.*

Furthermore, when $\frac{A(m,1)}{A(m,2)}$ is also sufficiently large relative to $\frac{A(l,1)}{A(l,2)}$ and under Assumption D.1, in the large relative to the small city:

(i) *The decline in middle-paid jobs is larger in percentage points.*

(ii) *The increase in low-paid jobs is smaller in percentage points.*

Proof. See Appendix D.4.2. $\square$

Let us first note that, here we obtain skewed polarization in the extreme case in which the comparative advantage of the large city in the high-paid sector is sufficiently large so that this sector is only in the large city. As this condition is not satisfied in the data, we see this result as being inferior to the simulations and more general formal conditions for skewed polarization seems beyond reach.

Let us give some intuition on the proof of Proposition D.1. The evolution of the share of high-paid jobs is described by (25). In particular, note that when the comparative advantage of the large city in the high-paid sector is sufficiently large, the high-paid sector is absent in City 2, thus leading this first term describing the term of the within-city reallocation to be 0 in City 2 and strictly positive in City 1. On the other hand, the relative price decline in the middle-paid sector makes city 1 more attractive for households with productivity between $\omega(h,1)$ and $\omega(h,2)$: these households are initially indifferent between the high-paid sector in City 1 and the middle-paid in City 2. Intuitively, such a term is positive in City 1 but negative in City 2. In the end, when the comparative advantage of the large city in the high-paid sector is sufficiently large, the share of high-paid workers increases only in the large city, proving the result.

The decrease in the price of capital/offshored goods corresponds to a negative demand shock for middle-paid jobs, but its effects across location depends on local comparative advantages. Under Assumption D.1 such a decline leads, on the one hand, to a stronger decrease in the threshold $\omega(h,1)$ than in $\omega(h,2)$ ($\Theta(\omega(h,c),c,h)$ is smaller in the large city): households have a stronger incentive to shift to the high-paid sector in the large city than in the small city due to the comparative advantage of the large city in the high-paid sector. On the other hand, the threshold $\omega(m,1)$ increases by less than $\omega(m,2)$: the incentive to shift to the low-skill sector increases by more in the small city. We illustrate this point in Figure D.2 in the special case where $A(m,1) = A(m,2)$. In the case where the comparative advantage of the large city in the high-paid sector is large enough, i.e., $A(h,1)/A(h,2)$ is sufficiently large relative to $A(m,1)/A(m,2)$ and, by transitivity, to $A(l,1)/A(l,2)$, the first effect dominates. Not only are more high-paid jobs created in the large city, but also more middle-paid jobs are destroyed there. There is a rise in low-paid jobs in the large city, but of a lesser magnitude than of high-paid jobs. Figure D.3 summarizes all these results.[48]

Result (ii) in Proposition D.1 implies that the increase in low-paid jobs is smaller in larger cities. The logic is as discussed earlier. The comparative advantage of the large relative to small city in the middle- relative to low-paid occupations implies that the cutoff for middle- relative to low-paid occupations is lower in the large city. However at this margin, this implies that the large city adjusts less elastically. The contrast between the cities will be stronger when either the concavity of productivity with respect to skill in the low-paid sector is strong or the comparative advantage of the large city in the middle- to low-paid sector is strong.

Initial exposure to middle-paid jobs. We now investigate the initial exposure to middle-paid jobs as a function of city size under our assumptions on technologies. Our conclusion is that large cities are the less exposed to middle-paid jobs. Combined with the results of Proposition D.1, this implies that destruction of middle-paid jobs is the strongest where the exposure is initially the *smallest*.

Proposition D.2 (Middle-paid job loss and initial exposure)**.** *When* $\frac{A(h,1)}{A(h,2)}$ *is sufficiently large relative*

[48]We have obtained proposition D.1 as an asymptotic result on the comparative advantage of the large city for the high-paid jobs in a context where productivities are relatively more elastic to skills for these high-paid jobs. In this way, our result does not depend on the skill distribution ($n(.)$).

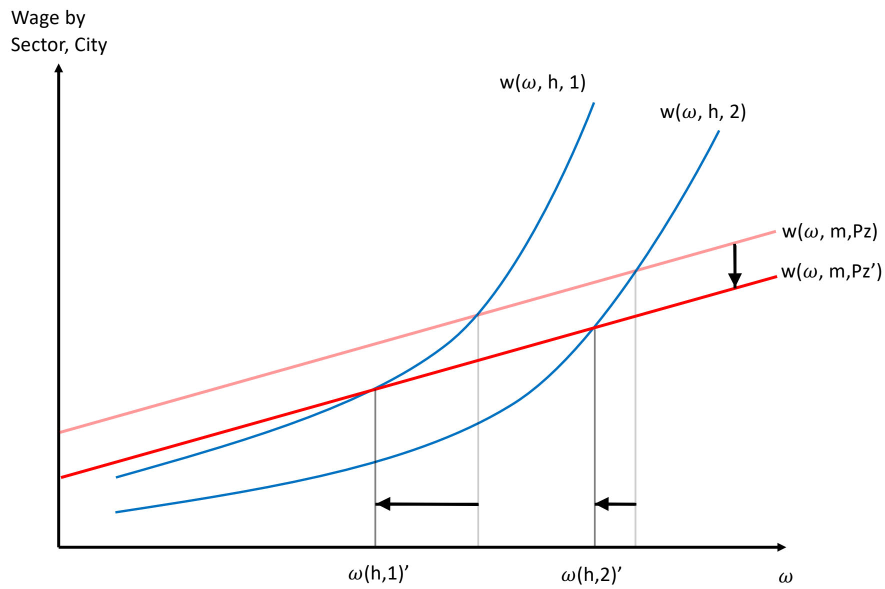


Figure D.2: Effect of a decline of the price of the capital/offshoring good on sector decisions – special case of equal productivities of cities in the m sector

This graph plots the curves for the wages in sector m – in red – and sector h – in blue – as a function of skill ω for both City 1 and City 2 and for two levels of prices for the capital/offshoring good ($P_z > P_{z'}$). Because of the increasing convexity of the wage when shifting from the m to the h sector, a decline in P_z leads to a stronger decline in the threshold $\omega(h,1)$ than in $\omega(h,2)$.

to $\frac{A(m,1)}{A(m,2)}$*, the share of middle skill sector jobs is smaller in the larger city.*

Under the conditions of Proposition D.1, the destruction of middle-paid jobs is the largest in percentage points in the large city where there is, initially, the lower share of middle-paid jobs.

Proof. See Appendix D.4.3. □

The intuition behind Proposition D.2 is simple: a sufficiently large comparative advantage for the high-paid sector in the large city leads to a large share of employment in this sector. In turn, this leads the share of middle-paid jobs to become smaller relatively to the share of these jobs in the small city.

D.2 The heterogeneity among middle-paid jobs across cities

We now look more closely at heterogeneity in adjustments within the middle-paid sector in each city. As in the low- and high-paid sectors, workers occupying middle-paid jobs are heterogeneous with respect to their skills – they have different values for ω. It is then possible to further analyze how labor market polarization affects middle-paid jobs across cities depending on workers' skills. Here, we show that, in large cities, it is the most skilled (i.e. with the highest ω or, equivalently, with the highest wage) middle-paid jobs that are destroyed and replaced by high-paid jobs. In small cities, it is mainly the least-skilled middle-paid jobs that are destroyed and replaced by low-paid jobs.[49]

[49] Our partition of occupations into low-, middle-, and high-paid sectors in Table 1 emphasized vulnerabilities to our posited shocks. Heterogeneity within groups notwithstanding, the middle-paid sector really is different.

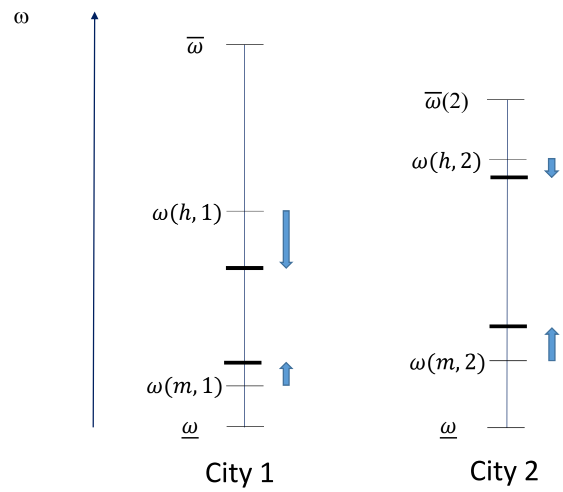


Figure D.3: The effects of a decline in the price of capital/offshored goods.

This figure shows the change in the equilibrium sectoral choices of agents as a result of the decline of the price of the capital/offshorable good p_z. Since the wages obtained by agents active in the middle-skill sector decline, agents with the highest opportunity costs working in this sector prior to the shock switch to high- and low-skill sectors in both cities. Given the technological assumptions, the decline (increase) in the skill threshold for agents to choose employment in the high- (low-) skill sector is larger (smaller) in City 1. As a result, there is a higher decrease in the middle-sector employment (despite a lower pre-shock share of employment in this sector as compared to City 2) and a higher (lower) increase in high- (low-) skill employment in the larger City 1.

To this end, we split middle-paid jobs into high-wage and low-wage tiers. Given that nominal wages in the model are functions of the skill ω, it is equivalent to splitting middle-paid jobs into higher-middle-paid and lower-middle-paid jobs. Let $\hat{\omega}$ be the threshold between these two categories. To avoid having empty sets, $\hat{\omega} \in [\omega(m,c), \omega(h,c)]$ for $c \in \{1,2\}$. With this threshold in hand, we can define higher wage middle-paid workers as the workers working in the m sector with a skill higher than $\hat{\omega}$ and the ones with a skill lower than $\hat{\omega}$ are lower skilled middle-paid workers.

Proposition D.3 (Heterogenous middle-paid job losses)**.** *The share of higher wage middle-paid jobs decreases by more in percentage points in the large city.*

The share of lower wage middle-paid jobs decreases by more in percentage points in the small city.

Proof. By definition of higher wage middle-paid jobs, the evolution of their share $s(m,c)^h$ is $ds(m,c)^h = -ds(h,c)$ and $ds(m,c)^l = -ds(l,c)$. The results of the proposition then follow from Proposition D.1. □

In terms of the model's notation, what stands behind this proposition is the relative behavior of the thresholds $\omega(h,c)$ and $\omega(m,c)$ across cities. These thresholds correspond to the indifference condition between, respectively, the high- and the middle-paid sectors and the middle- and the low-paid sectors.

All eight of the most offshorable occupations and six of eight of the most routinizable occupations are in the middle-paid sector. The lowest-paid of these, CS 67 unskilled industrial workers is the most offshorable occupation and the second most routinizable one. At the other end, the highest-paid middle-wage occupation, CS 48 supervisors and foremen is third both in offshorability and routinizability. These underscore the value in our framework of examining margins with both low- and high-paid jobs

Under Assumption D.1, the upper threshold $\omega(h, c)$ decreases by more in the large city (City 1) and the lower threshold $\omega(m, c)$ increases by more in the small city (City 2). As a result, more higher wage middle-paid jobs are destroyed in the large city following a labor market polarization shock while more lower wage middle-paid jobs are destroyed in the small city.

D.3 The effects on high-paid jobs and the great urban divergence

We now turn to what happens to high-paid jobs. Our main conclusion is that the polarization shock leads to the great urban divergence across cities, when large cities have a comparative advantage in high-paid jobs. More precisely, we obtain:

Proposition D.4 (The Great Urban Divergence)**.** *Under the condition of Proposition D.1, the share of high-paid jobs increases by more in the larger cities which already had an initially larger share of high-paid jobs.*

Proof. This proposition first results from Proposition D.1, which shows that the large city experiences a larger increase in high-paid jobs. We obtain that the share of high-paid workers is larger in the large city by combining Corollary 2 in Appendix B.4 and Proposition D.2, which show that the share of low-paid and middle-paid are both smaller in the large city. □

D.4 Proofs

D.4.1 Sectoral decisions and factor prices.

As this can be observed from equations (4) and (5), the two thresholds are functions of intermediate good prices $p(l)$, $p(m)$ and $p(h)$. The following lemma clarifies how these thresholds move as a function of the relative prices $p(m)/p(h)$ and $p(l)/p(m)$ when Assumption D.1 holds.

Lemma D.1. *Suppose that Assumption D.1 holds.*

A decline in $p(m)/p(h)$ implies a relatively larger decline for $\omega(h, 1)$ than for $\omega(h, 2)$.

An increase in $p(l)/p(m)$ implies a relatively larger increase in $\omega(m, 2)$ than for $\omega(m, 1)$.

Proof. Let us define $\tilde{H}(\omega, \sigma, c) = A(\sigma, c)H(\omega, \sigma)$.

Let us now compute how a change in price of intermediate goods modifies the thresholds. By rewriting the indifference condition as $\frac{\tilde{H}(\omega(h,c),h,c)}{\tilde{H}(\omega(h,c),m,c)} = \frac{p(m)}{p(h)}$, we obtain, by differentiating both the right and the left hand terms:

$$\frac{d\left(\frac{\tilde{H}(\omega(h,c),h,c)}{\tilde{H}(\omega(h,c),m,c)}\right)}{\frac{\tilde{H}(\omega(h,c),h,c)}{\tilde{H}(\omega(h,c),m,c)}} = \frac{d\left(\frac{p(m)}{p(h)}\right)}{\frac{p(m)}{p(h)}}$$

Let us compute the different terms separately:

$$\frac{d\left(\frac{\tilde{H}(\omega(h,c),h)}{\tilde{H}(\omega(h,c),m)}\right)}{\frac{\tilde{H}(\omega(h,c),h)}{\tilde{H}(\omega(h,c),m)}} = \left(\frac{\tilde{H}_\omega(\omega(h,c),h)}{\tilde{H}(\omega(h,c),h)} - \frac{\tilde{H}_\omega(\omega(h,c),m)}{\tilde{H}(\omega(h,c),m)}\right) d\omega(h,c)$$

As a result, the effect of a relative decline in prices is such that:

$$d\omega(h,c) = \frac{1}{\Theta(\omega(h,c),c)} \frac{d\left(\frac{p(m)}{p(h)}\right)}{\frac{p(m)}{p(h)}} \tag{26}$$

with $\Theta(\omega(h,c),c) = \frac{\tilde{H}_\omega(\omega(h,c),h,c)}{\tilde{H}(\omega(h,c),h,c)} - \frac{\tilde{H}_\omega(\omega(h,c),m,c)}{\tilde{H}(\omega(h,c),m,c)} > 0$. This term is positive given that $\tilde{H}$ is log-supermodular in (ω,σ). As a result, a decline in $p(m)/p(h)$ then leads to a decline in $\omega(h,c)$. Similarly, we obtain:

$$d\omega(m,c) = \frac{1}{\Theta(\omega(m,c),c)} \frac{d\left(\frac{p(l)}{p(m)}\right)}{\frac{p(l)}{p(m)}} \tag{27}$$

where $\Theta(\omega(m,c),c) > 0$ due to log-supermodularity. As a result, an increase in $p(l)/p(m)$ then leads to an increase in $\omega(m,c)$.

We now want to know where the decline in $\omega(h,c)$ and the increase in $\omega(m,c)$ are the strongest.

For the first point, this amounts to comparing $\Theta(\omega(h,1),1)$ and $\Theta(\omega(h,2),2)$, that is to determine the sign of:

$$\frac{\tilde{H}_\omega(\omega(h,1),h,1)}{\tilde{H}(\omega(h,1),h,1)} - \frac{\tilde{H}_\omega(\omega(h,1),m,1)}{\tilde{H}(\omega(h,1),m,1)} - \frac{\tilde{H}_\omega(\omega(h,2),h,2)}{\tilde{H}(\omega(h,2),h,2)} + \frac{\tilde{H}_\omega(\omega(h,2),m,2)}{\tilde{H}(\omega(h,2),m,2)}$$

For the second point, this amounts to comparing $\Theta(\omega(m,1),1)$ and $\Theta(\omega(m,2),2)$, that is to determine the sign of:

$$\frac{\tilde{H}_\omega(\omega(m,1),m,1)}{\tilde{H}(\omega(m,1),m,1)} - \frac{\tilde{H}_\omega(\omega(m,1),l,1)}{\tilde{H}(\omega(m,1),l,1)} - \frac{\tilde{H}_\omega(\omega(m,2),m,2)}{\tilde{H}(\omega(m,2),m,2)} + \frac{\tilde{H}_\omega(\omega(m,2),l,2)}{\tilde{H}(\omega(m,2),l,2)}$$

Using our assumption on the function H, this simplifies the two expressions into $\frac{1-\phi}{\omega(m,1)} - \frac{1-\phi}{\omega(m,2)} \geq 0$ which is positive as $\omega(m,1) \leq \omega(m,2)$ and:

$$\frac{\tilde{H}_\omega(\omega(h,1),h,1)}{\tilde{H}(\omega(h,1),h,1)} - \frac{1}{\omega(h,1)} - \frac{\tilde{H}_\omega(\omega(h,2),h,2)}{\tilde{H}(\omega(h,2),h,2)} + \frac{1}{\omega(h,2)}$$

Let us investigate the sign of this expression. Note that it is negative as long as:

$$\frac{\tilde{H}_\omega(\omega(h,1),h,1)}{\tilde{H}(\omega(h,1),h,1)} - \frac{\tilde{H}_\omega(\omega(h,2),h,2)}{\tilde{H}(\omega(h,2),h,2)} \leq \frac{1}{\omega(h,1)} - \frac{1}{\omega(h,2)}$$

which is satisfied. □

A crucial assumption to obtain Lemma D.1 is the one of the relative convexity of $\tilde{H}(\omega, l, c)$, $\tilde{H}(\omega, m, c)$ and $\tilde{H}(\omega, h, c)$. Let us explain why by focusing on the threshold between high-paid and middle-paid jobs, $\omega(h, c)$. The comparative advantage of the large city in the high-paid sector leads this threshold to be smaller in the large city as shown by Proposition 1 ($\omega(h, 1) < \omega(h, 2)$). Actually, the larger is this comparative advantage, the lower will be the threshold $\omega(h, 1)$ compared with $\omega(h, 2)$, as $\tilde{H}(\omega, h, c)/\tilde{H}(\omega, m, c)$ is an increasing function of ω.

Through our assumptions, the relative productivity between the high- and the middle-paid sectors ($\tilde{H}(\omega, h, c)/\tilde{H}(\omega, m, c)$) is a convex function of the skill ω. As a result, a decline of the relative price of middle-paid sector's good leads to a stronger decline of $\omega(h, 1)$ than of $\omega(h, 2)$ as the former is a region where the relative productivity is flatter. Using other words: the incentive for a middle-paid worker to become a high-paid worker increases for a larger set of skills in the large than in the small city.

In the end, the comparative advantage of the large city in the high-paid sector leads both to a lower threshold $\omega(h, c)$ in that city but also to a stronger decline of this threshold in the case of a decline of the price of the capital/offshored good.

Similarly, the incentive for a middle-paid worker to become a low skill worker also increases in both cities. $\tilde{H}(\omega, m, c)/\tilde{H}(\omega, l, c)$ being an increasing function of ω, Proposition 1 shows that $\omega(m, 1) < \omega(m, 2)$. However, $H(\omega, m, c)/H(\omega, l, c)$ is a concave function of ω, thus leading $\omega(m, 2)$ to increase by more than $\omega(m, 1)$ for the same variation of the price of the capital/offshored good.

Figure D.3 summarizes these findings.

Evolution with respect to comparative advantage and convexity. Let us have a few words about how the results of Lemma D.1 evolve as a function of the comparative advantage of the two cities in the different sectors and as a function of the convexity assumptions on productivities.

Let us start with the threshold between the high-paid and the middle-paid sectors. In some of our results, we are going to focus on situations where the productivity in the high-paid sector in City 1 ($A(h, 1)$) is large. The relative stronger decline of $\omega(h, 1)$ compared with $\omega(h, 2)$ is more pronounced when the slope of productivity in the high-paid sector (η) is larger and/or when the comparative advantage in the high-paid sector in the large city is stronger ($A(h, 1)/A(h, 2)$ as compared with $A(m, 1)/A(m, 2)$). As a result, a high productivity $A(h, 1)$ results in a lower threshold $\omega(h, 1)$ so that this threshold ends up in a region that is even flatter. An implication is then that a decrease in the price of the middle-paid sector input has even more stronger downward effect on $\omega(h, 1)$ when $A(h, 1)$ is large.

Conversely, the relative increase of $\omega(m, 2)$ compared with $\omega(m, 1)$ is more pronounced when the slope of productivity in the low-paid sector (ϕ) is greater and/or when the comparative advantage in the low-paid sector in the small city is stronger ($A(l, 2)/A(l, 1)$ as compared with $A(m, 2)/A(m, 1)$). In particular, when $A(l, 2)/A(l, 1)$ is very close to $A(m, 2)/A(m, 1)$, $\omega(m, 2)$ behaves as $\omega(m, 1)$.

D.4.2 Proof of Proposition D.1

The share of high-paid jobs Let us first start with the share of high-paid jobs. As written in the text, the evolution of the share of high-paid jobs is

$$ds(h,c) = -\frac{f(\omega(h,c),c)}{L(c)}\frac{1}{\Theta(\omega(h,c),c,h)}\frac{dp}{p} + \int_{\omega(h,1)}^{\omega(h,2)} \frac{\partial f(\omega,c)}{\partial p} d\omega \frac{dp}{p}.$$

The second term is positive for city 1 and negative for city 2 at the first order. Between the thresholds $\omega(h,1)$ and $\omega(h,2)$, households hesitate between working in City 1 in the high-paid sector and City 2 in the middle-paid sector. As this second option becomes relatively less valuable, if any change happens in the distribution of households of skill $\omega \in [\omega(h,1), \omega(h,2)]$, this leads to increasing $f(\omega,1)$ and to decreasing $f(\omega,2)$. As a result, a sufficient condition for $ds(h,1) > ds(h,2)$ as a result of a price decline ($dp < 0$) is that $\frac{f(\omega(h,c),c)}{L(c)}\frac{1}{\Theta(\omega(h,c),c,h)}$ is larger in City 1.

Let us then fix the initial level of the relative price of the middle-paid sector good p. When $A(1,h)$ is sufficiently large compared with $A(2,h)$, the high-paid sector is present only in City 1. More precisely, this happens when $\omega(h,2) > \omega\bar{(}2)$. To recall, $\omega\bar{(}2)$ is the highest skill present in City 2.

Middle- and low-paid jobs As $\omega(m,1)$ converges to $\underline{\omega}$, we obtain that $s(l,1) = 0$ whatever the price of the capital/offshoring good p. As a result, $ds(l,1) = 0$ as well. In City 2, the evolution of the share of low-paid jobs is, in percentage points:

$$ds(l,2) = -\frac{\left(\frac{V(h(\omega(m,2),2))}{\Theta(\omega(m,2),2,m)} h'(\omega(m,2),2)\right)}{S\left(T^{-1}\left(h(\underline{\omega},2)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(2),2)\right)\right)}\frac{dp}{p} + \int_{\omega(m,1)}^{\omega(m,2)} \frac{\partial f(\omega,2)}{\partial p} d\omega \frac{dp}{p} \geq 0$$

when $dp < 0$. In the end, $ds(l,2) \geq ds(l,1) = 0$.

Finally, let us prove that the fall in the share of middle-paid workers is stronger in City 1. Note that $ds(m,c) = -ds(l,c) - ds(h,c)$. Given previous results, we have then to compare $ds(h,1)$ with $ds(l,2)$. Let us suppose that Assumption D.1 holds. Note that following Lemma D.1, when $A(1,h)/A(2,h)$ is arbitrarily large, $\omega(h,1)$ converges to $\underline{\omega}$ and $\Theta(\omega(h,1),1,h)^{-1} = \omega(h,1)/(\eta\omega(h,1) - 1)$ diverges to ∞. This implies that there exists a level for $A(h,1)$ so that the threshold $\omega(h,1)$ falls below $\underline{\omega}$ for an arbitrarily small variation of the relative price dp. In particular, we can select a $A(h,1)$ and a level of relative price p so that $ds(h,1) = s(m,1)$ – the marginal increase in price squeezes the middle-paid sector – and $s(m,1) > ds(l,2)$.

D.4.3 Proof of Proposition D.2

In city c, individuals in the middle-paid sector have a skill ω is between $\omega(h,c)$ and $\omega(m,c)$. As a result, the population of such individuals is:

$$L\int_{\omega(m,c)}^{\omega(h,c)} f(x,c)dx = S\left(T^{-1}\left(h(\omega(m,c),c)\right)\right) - S\left(T^{-1}\left(h(\omega(h,c),c)\right)\right) \tag{28}$$

where $K(T^{-1}(h(\omega, c)), c) = \omega$. The share of agents in the middle-paid sector in city c is then:

$$s(m,c) = \frac{\int_{\omega(m,c)}^{\omega(h,c)} f(x,c)dx}{\int_{\underline{\omega}}^{\overline{\omega}(c)} f(x,c)dx} = \frac{S\left(T^{-1}\left(h(\omega(m,c),c)\right)\right) - S\left(T^{-1}\left(h(\omega(h,c),c)\right)\right)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{29}$$

Using the continuity of the different functions and given that $\omega(h,c)$ is decreasing in $A(c,h)/A(c,m)$ and, thus, $s(m,c) = 0$ when $A(c,h)/A(c,m) \to \infty$, we obtain that, when $A(h,1)/A(m,1)$ is sufficiently large compared with $A(h,2)/A(m,2)$, shares satisfy $s(m,1) \leq s(m,2)$.

E A simplified model of the middle-paid sector - for online publication only

In this section, we further investigate the heterogeneity across middle-paid jobs. In particular, we consider a model where there are two types of middle-paid jobs, a first type that is lower-skilled and more routinizable (as MRO jobs) and a second type that is higher-skilled and less routinizable (as OMP jobs). We interpret routinizability here explicitly as a cost in units of capital to replace a middle-paid job. We first show that the large city can be relatively specialized in the higher-skilled type of middle-paid jobs and the small city in the lower-skilled type of middle-paid jobs. Second, we show that, despite being less routinizable, higher-skilled middle-paid jobs can be destroyed in the large city before lower-skilled middle-paid jobs in smaller cities, consistent with Fact 3.

To this purpose, let us consider the following simplified version of our model that is zoomed in to focus only on the middle skill workers.

Production using middle skill jobs. We now split middle skill jobs into high and low wage occupations. As in Acemoglu and Restrepo (2018), these two occupations are differently substitutable with capital: capital is less effective to replace the high wage middle-paid jobs than the low wage middle-paid jobs. To simplify, we assume that jobs and capital are perfect substitutes.

We assume that the production function to produce the low-wage middle-paid sector's input is $q(m_l) + k_l$, with $q(m_l)$ the quantity of efficient labor used for production and k_l the amount of capital. The production function to produce the high-wage middle-paid sector's input is $q(m_h) + \gamma_k k_h$, with $q(m_h)$ the quantity of efficient labor used for production and k_h the amount of capital. $\gamma_k < 1$ is a technological parameter – it is less than one as capital is less productive to replace high-wage middle-paid jobs. The price for the first type of input is $p(m,l)$ and the price for the second type of input is $p(m,h)$. Capital is still provided using an exogenous production function so that the price of capital is ξ.

The households. We denote by $\omega \in \{\underline{\omega}, \overline{\omega}\}$ the skill of an agent, with $\underline{\omega} < \overline{\omega}$. To also streamline the model in terms of location decisions, we assume that the cost to live in city c is $r(c)$. In each city, the two middle-paid types hesitate to work in different sectors. This leads to the reservation wage

$\bar{w}(c, \omega)$.

Equilibrium. Let us investigate the choice between labor and capital to produce the two intermediate goods. In equilibrium, we have that: $\bar{w}(c, \omega) = w(c, \omega)$. Labor from city c is predominantly used in the low wage middle-paid sector when $w(c, \underline{\omega}) \leq \xi$ and labor from city c is predominantly used in the high wage middle-paid sector when $w(c, \overline{\omega}) \leq \xi + 1 - \gamma_k$.

Given that nominal wages are always higher in the large city, $\bar{w}(1, \omega) > \bar{w}(2, \sigma)$ for $\omega \in \{\underline{\omega}, \overline{\omega}\}$, we directly obtain the following lemma:

Lemma E.1. *For both high- and low-wage middle skilled jobs, automation takes place first in the large city and then in the small city.*

In a given city, the automation of the low-wage middle-paid workers takes place before the automation of the high-wage middle-paid workers when

$$w(c, \underline{\omega}) \geq w(c, \overline{\omega}) + \gamma_k - 1$$

The incentive to first replace lower-wage middle skilled jobs with capital is the balance of two forces. On the one hand, these jobs are relatively cheaper ($w(c, \overline{\omega}) > w(c, \underline{\omega})$) but, on the other hand, they are more efficiently replaced by capital compared to high-wage middle-paid jobs (as measured by the parameter γ_k). This leads to the following lemma:

Lemma E.2. *When γ_k is sufficiently low, in both cities, automation of low-wage middle-paid jobs takes place before automation of high-wage middle-paid jobs.*

In city c, when the wage of high-skilled middle-paid jobs is sufficiently large ($w(c, \overline{\omega})$ compared with $w(c, \underline{\omega})$), automation of high-wage middle-paid jobs takes place before automation of low-wage middle-paid jobs.

Less-routinizable jobs can be automated before more-routinizable ones. Let us now investigate which jobs are more likely to be automated. Given the linearity of our environment, our criterion is to check which jobs are automated first, that is the ones for which a smaller decrease in the price of capital is sufficient for capital to replace them.

The following proposition establishes that, when the opportunities for high-skilled middle-paid jobs are important enough in the large cities, this is sufficient to lead these jobs to be automated first:

Proposition E.1. *There exists a reservation wage in City 1 $w(1, \overline{\omega})$ sufficiently large such that the automation of high-skilled middle-paid jobs located in City 1 takes place before the automation of high-skilled middle-paid jobs located in City 2 and low-skilled middle-paid jobs in cities 1 and 2.*

The decision to live in City 1 or City 2 for households having high-skilled middle-paid jobs amounts to comparing $\bar{w}(1, \overline{\omega}) - r(1)$ and $\bar{w}(2, \overline{\omega}) - r(2)$. When $\bar{w}(1, \overline{\omega})$ is sufficiently large, these households move to City 1.

Replacing high-skilled middle-paid jobs located in City 1 by capital happens when the price of capital is lower than $w(1, \overline{\omega}) + \gamma_k - 1$. Replacing high-skilled middle-paid jobs located in City 2 by capital happens when the price of capital is lower than $w(2, \overline{\omega}) + \gamma_k - 1$. When $w(1, \overline{\omega})$ is sufficiently large, the incentives to replace high-skilled middle-paid jobs located in City 1 is larger than the incentives to replace high-skilled middle-paid jobs located in City 2. In addition, when $w(1, \overline{\omega})$ is sufficiently large, some high-skilled middle-paid jobs are indeed located in City 1 ($w(1, \overline{\omega}) - r(1) \geq w(2, \overline{\omega}) - r(2)$).

Low-skilled middle-paid jobs are replaced by capital when the price of capital is lower than $w(1, \underline{\omega})$ in City 1 and $w(2, \underline{\omega})$ in City 2. In the end, when the wage of high-skilled middle-paid jobs is large enough, that is:

$$\max_c w(c, \underline{\omega}) \geq w(1, \overline{\omega}) + \gamma_k - 1,$$

automating high-skill middle-paid jobs requires a smaller fall in the price of capital compared with automating low-skill middle-paid jobs. This happens because the incentive to replace these jobs is sufficiently strong and despite that the cost of automation is higher for these high-skill middle-paid jobs (capital is less efficient to replace these jobs).

Mapping with the large model. Let us connect this simple model with our benchmark model.

In the benchmark model, all middle-paid jobs had the same degree of substitutability with capital. Yet, the incentives to replace them were different depending on the skills of agents and the location of the jobs: in the large city, higher-skilled middle-paid workers have a strong incentive to shift to the high-paid sector.

In the simple model presented in this section, this incentive is captured through the reservation wage for high-skilled middle-paid jobs $\bar{w}(1, \overline{\omega})$. We then show that when this reservation wage is sufficiently strong, the incentive to replace higher skilled jobs can dominate the potential higher cost of routinizability of these jobs.

F Additional empirical results, Figures and Tables - for online publication only

F.1 More detailed data description

Figure F.1: Map of France with largest metropolitan areas in 2015.

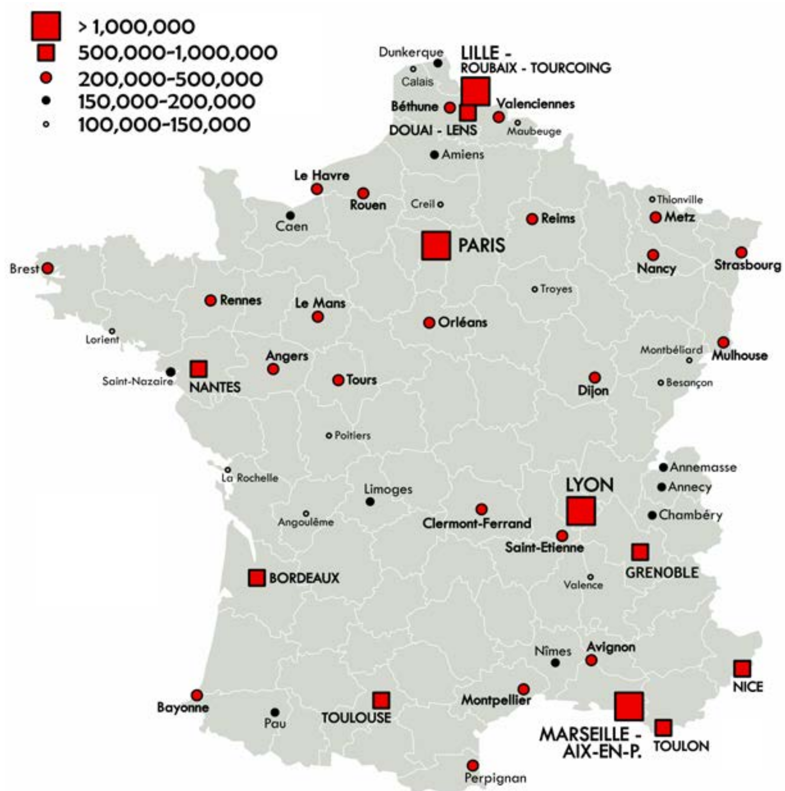


Table F.1: City size categories, number of cities, population and the share of hours worked in 2015.

city size	number	total population	share of hours worked
>2,000,000	1	10,706,072	.375
750,000-2,000,000	6	7,060,599	.206
500,000-750,000	4	2,219,618	.055
200,000-500,000	22	6,691,222	.169
100,000-200,000	22	3,245,887	.083
50,000-100,000	62	4,414,317	.112
Total	117	34,337,715	

Notes on dataset construction. In dataset construction, we include firms that are incorporated and have the legal category starting with "5" in the INSEE classifications, excluding privatized firms or those that changed status from public to private incorporation (which affects for example the public or private law under which labor contracts are offered). Data on self-employed are not reported prior to the late 2000s.

Some firms in the finance, insurance and real estate sectors reported pre-2001 their employees from branches at few establishments for example at the department level. This represents a small fraction of employment in these sectors based on INSEE assessments on 2001-2003 data and may introduce minor errors given the scale of the problem when we use metropolitan area-level data. Excluding these sectors from our analysis does not change our results considerably and does not impact our conclusions. We include Table F.32 without these sectors as a replication of Table 3 in a robustness test.

DADS-Postes has data on public sector employment only since the end of the 2000s. The evolution of public sector employment based on Census data for years 1990-2015 is in the Online Appendix F.7. Public sector employment increased in mainland France by 0.5 percentage points in the period and its evolution does not reveal systematic spatial nor high- middle- or low-skill patterns confounding our analysis.

We retain all positions where there were at least 120 hours worked in a year to minimize erroneous entries — that would result in e.g. abnormally high average wages. We do not observe, however, a material difference in our results if no filtering is applied or filtering based on end-of-year presence with at least 30 days in the firm. The INSEE provides filtering in the DADS data set, but it is not consistent between 1994 and 2015.

We use the 2-digit occupation codes because of data availability. Firms should report their data to the INSEE using much finer 4-digit codes, but many failed to do so especially before 2003. Morevoer, during the 2003 revision of codes many 4-digit codes were changed without an onto mapping between codes in either direction. One can perform an imperfect matching based on 2003 employment shares of 4-digit occupations between the 1982 and 2003 classifications, but this comes with a strong assumption that the shares of jobs according to 1982 classification in 4-digit occupations defined in 2003 remain fixed until 2015. A mapping at the 2-digit CS level is, however, possible and hence we can obtain a consistent data series in the period 1994-2015. In 2003 a new 2-digit category, CS 31 was created, encompassing "liberal" professionals such as lawyers previously included in CS 37. In all our data we merge CS 31 and CS 37 together without loss of generality, as these are high-paid occupations requiring high skills.

There are two alternative sources of data to DADS-Postes on hours and wages by occupation in the studied period. The DADS-panel data set is constructed by the INSEE on the basis of DADS-Postes, retaining a fraction (1/25 of total pre-2001) of individuals in the main data set. It shares the advantages and limitation of DADS-Postes; therefore for our purposes of analyzing city labor markets at two distant dates it offers no advantages at the cost of lower precision.

Another data set is the French Labor Survey is available since 1982. For early years it has approximately 60,000 observations per year with occupations available at the 4-digit PCS level. This allows to document general facts about labor market polarization (see e.g. right-hand panel of Figure

F.7 for 1994-2015 changes), but disallows constructing precise changes in labor market evolutions at the city level even for 2-digit CS occupations. This, and also given the break in the 4-digit PCS classification in 2003 renders the Survey data impractical for our purposes. Moreover, the survey suffers from all the shortcomings of such data: the hours and wages are self-reported, the non-response bias, especially at the top is not known. In contrast, the DADS-Postes data is exhaustive, the data submitted by firms is mandatory and it gives inter alia more geographical details.

Classification of jobs. Here we provide more details and discussion about how we classify occupations into high-, middle- and low-paid and obtain their exposure to automation and offshoring, complementing Sections 2.2 and 2.2.1.

In the first step, we obtained data from the data Appendix of Goos et al. (2014) on exposure classifications in the file "task.dta". We use the series "RTI_alm_isco_77" for measure of routinizability and "OFF_gms" for offshoring. The data is available for the 2-digit ISCO occupation classification.

There is no official passage between the 2-digit PCS classification used by the INSEE and available in DADS-Postes and the 2-digit ISCO for the first years when DADS-Postes data is available. Both classifications, however, are available in the French Labor Surveys and we use the 1994 vintage to perform a mapping between the ISCO classifications and the PCS. We used hours worked in 1994 available in the Survey as weights to construct characteristics of 2-digit CS categories (measures of routinizability and offshorability) inherited from the properties of occupations of the 2-digit ISCO classification. Using different periods from the Survey - e.g. the entire 1982-1994 (the Survey started in 1982) yields similar results.

A plot of resulting routinizability and offshorability measures for the considered occupations is shown in Figure F.2 while Table 1 with these values instead of ranks is reproduced as Table F.2.

For grouping occupations as done in Section 2.2.1, an alternative to consider is whether an ordering based on the PCS 1-digit codes might make a reasonable partition into the wage groups. The codes CS 2 for CEOs and small business owners and CS 3 for high-paid professionals, if combined into the high-paid sector, would indeed yield the same boundary between high- and middle-paid occupations as the one used. Adding CS 4 occupations to the high-paid group would be consistent with a single cut between high-paid and other occupations, but it would have two downsides. First, it would require bridging a clear 21 percent wage gap at the boundary of the 2-digit CS 3 and CS 4 occupations (see Table 1). Moreover, that would put in the high wage group two of the occupations (CS 46 and 48) whose jobs declined most sharply in absolute and relative terms in our period of study, hence be inconsistent with the spirit of the labor market polarization approach. In sum, using the 1-digit PCS codes suggests the same boundary between high- and middle-paid occupations as our simple visualization exercise.

Trying to use the remaining 1-digit CS codes 4, 5, and 6 to define a boundary between middle- and low-paid groups immediately runs into problems. The 2-digit CS 5 and CS 6 occupations have no clear ordering by initial mean wage, so cannot be sensibly separated. This is also by construction according to the PCS classification as CS 5 ("employees", typically service workers) and 6 ("workers", typically blue-collar workers working in industry and various artisans) need not differ much in skills.[50] If they are combined as the low-paid sector, then this would include two of the sectors with the largest absolute and relative job declines, CS 62 and 67, in the low-paid sector. Again, this is against the spirit of the labor market polarization approach.

Taken together with our prior observations, this suggests that our initial approach in Section 2.2.1, focusing on encompassing within the middle-paid group those occupations with the largest absolute and proportional job declines is likely to be the best to define our occupation-wage groups. Moreover, our main results are not materially affected when we move the border between low- or middle-paid occupations. For example, such robustness checks for the comparison of means of changes in occupation shares between small and large cities of Table 3 – are shown in Online Appendix Tables F.28 and F.29.

[50]The algorithm used to classify occupations here is complex; e.g. cooks, depending on seniority, can be classified either as CS 56 (unskilled) or 63 (skilled). Occupations with very low skills would also be a part of either main category – for example janitors or cleaners are coded into CS 56 while unskilled garbage collectors into CS 68.

Table F.2: Basic statistics by 2 digit CS categories: Full table.

CS 2-digit	description	employment share in %		average city wage (in 2015 euros)		Routine	Offshorable
						(index values)	
		1994	2015	1994	2015		
		high-paid occupations					
23	CEOs	1.0	0.9	42.81	59.20	-0.75	-0.59
37	**managers and professionals**	**6.2**	**10.2**	32.52	38.56	-0.75	-0.59
38	**engineers**	**5.1**	**9.0**	30.36	33.69	-0.82	-0.39
35	creative professionals	0.5	0.5	22.83	31.80	-0.72	-0.49
		middle-paid occupations					
48	**supervisors and foremen**	**4.1**	**2.7**	18.03	21.86	0.42	1.23
46	**mid-level professionals**	**12.3**	**7.6**	17.54	21.20	-0.48	-0.16
47	**technicians**	**5.7**	**6.3**	17.15	20.60	-0.40	-0.29
43	mid-level health professionals	0.8	1.5	15.05	18.05	-0.35	-0.57
62	**skilled industrial workers**	**14.1**	**9.3**	13.52	17.99	0.38	1.24
54	**office workers**	**11.8**	**11.2**	13.17	16.98	2.03	0.87
65	**transport and logistics personnel**	**2.9**	**3.0**	11.96	16.00	0.33	0.27
63	**skilled manual workers**	**8.0**	**8.3**	11.90	15.50	0.17	-0.33
64	**drivers**	**5.0**	**5.5**	11.50	14.46	-1.50	-0.63
67	**unskilled industrial workers**	**10.9**	**5.7**	11.02	14.72	0.45	2.09
		low-paid occupations					
53	security workers	0.7	1.4	10.60	14.60	-0.28	-0.51
55	**sales-related occupations**	**5.4**	**8.3**	10.44	13.74	0.30	-0.57
56	personal service workers	2.2	4.8	9.97	12.63	-0.43	-0.57
68	**unskilled manual workers**	**3.3**	**3.8**	9.11	13.27	0.06	-0.36

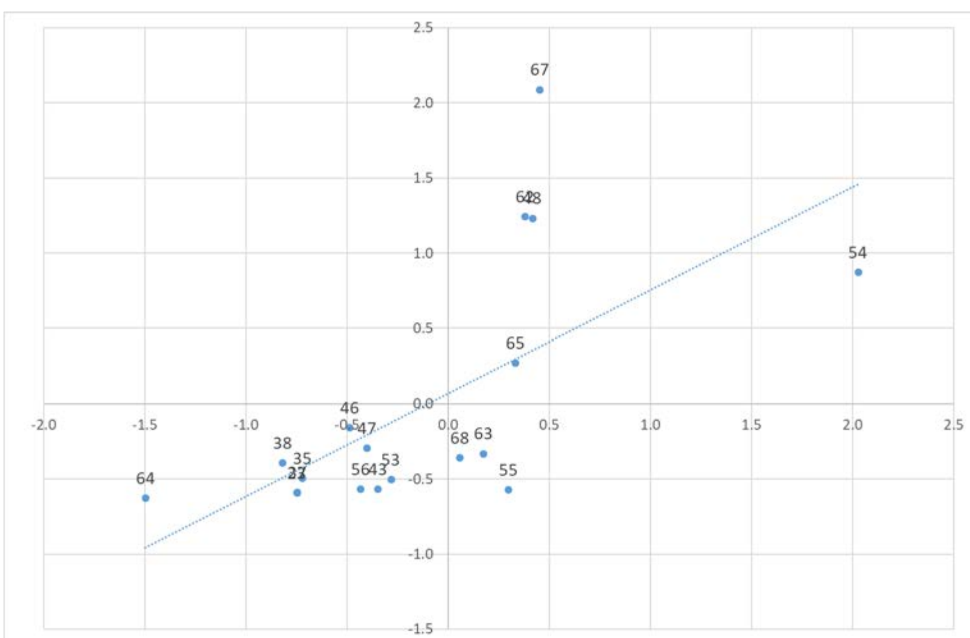
Notes: In-sample values. Employment share for mainland France. Average city wages in constant 2015 euros. Categories in bold are those with employment shares above 2.5% in 1994 in sample. Translation from French of category names other than PCS 23, 35, 43 and 53 taken from Table 2 of Harrigan et al. (2016).

Table F.3: RTI and OFF-GMS indexes for different types of jobs

Job type	Routine index values	Offshorable
3 types		
high-paid	-0.77	-0.51
middle-paid	0.31	0.47
low-paid	-0.10	-0.52
4 types		
high-paid	-0.77	-0.51
middle-paid above median	-0.08	0.35
middle-paid below median	0.69	0.58
low-paid	-0.10	-0.52

"Routine" index based on the RTI measure of Autor et al. (2003) while "Offshorable" on the OFF-GMS measure from Goos et al. (2014), both mapped into PCS 2-digit employment categories from the ISCO classification used by Goos et al. (2014).
This Table gives the employment-share weighted values of the routiness and offshorability indexes of the main occupation groupings considered in the paper.
The middle-paid jobs are the most exposed to automation and offshoring shocks while high-paid the least.

Figure F.2: RTI and OFF-GMS values for different PCS 2-digit occupations.



Note: Dashed line shows linear fit between the two measures.

Table F.4: 2-digit PCS categories and representative 4-digit PCS categories (private sector)

23	**CEOs of firms above 9 employees**	54	**Office workers**
35	**Creative professionals**		Receptionists, secretaries
	Journalists and writers		Administrative/clerical workers, various sectors
	Media, publishing houses and performing arts managers		Computer operators
	Artists		Bus/train conductors, etc
37	**Top managers, professionals and liberal professions (PCS 31)***	55	**Retail workers**
	Managers of large businesses		Retail employees, various establishments
	Finance, accounting, sales, and advertising managers		Cashiers
	Other administrative managers		Service station attendants
	Doctors and pharmacists	56	**Personal service workers**
	Legal and technical liberal professions (lawyers, architects)		Restaurant servers, food prep workers
38	**Technical managers and engineers**		Hotel employees: front desk, cleaning, other
	Technical managers for large companies		Barbers, hair stylists, and beauty shop employees
	Engineers and R&D managers		Child care providers, home health aids
	Eletrical, mechanical, materials and chemical engineers		Residential building janitors, caretakers
	Purchasing, planning, quality control, and production managers	62	**Skilled industrial workers**
	Information technology R&D engineers and managers		Skilled construction workers
	Information technology support engineers and managers		Skilled metalworkers, pipe?tters, welders
	Telecommunications engineers and specialists		Skilled heavy and electrical machinery operators
43	**Mid-level health professionals and social workers**		Skilled operators of electrical and electronic equipment
	Nurses		Skilled workers in various industries
	Masseurs and therapists	63	**Skilled manual laborers**
	Medical technicians		Gardeners
	Specialized educators		Master electricians, bricklayers, carpenters, etc
	Leisure and cultural activity organizers		Skilled electrical and electronice service technicians
46	**Mid-level professionals**		Skilled autobody and autorepair workers
	Mid-level professionals, various industries		Master cooks, bakers, butchers
	Supervisors in financial, legal, and other services		Skilled artisans (jewelers, potters, etc)
	Store, hotel, and food service managers	64	**Drivers**
	Sales and PR representatives		Truck, taxi, and delivery drivers
47	**Technicians**	65	**Skilled transport workers**
	Designers of electrical, electronic, and mechanical equipment		Heavy crane and vehicle operators
	R&D technicians, general and IT		Warehouse truck and forklift drivers
	Installation and maintenance of non-IT equipment		Other skilled warehouse workers
	Installation and maintenance of IT equipment	67	**Low skill industrial workers**
	Telecommunications and computer network technicians		Low skill construction workers
	Computer operation, installation and maintenance technicians		low skill electrical, metalworking, and mechanical workers
48	**Foremen, Supervisors**		low skill shipping, moving, and warehouse workers
	Foremen: construction and other		low skill transport industry workers
	Supervisors: various manufacturing sectors		Low skill production workers in various industries
	Supervisors: maintenance and installation of machinery	68	**Low skill manual laborers**
	Warehouse and shipping managers		Low skill mechanics, locksmiths, etc
	Food service supervisors Other		Apprentice bakers, butchers
53	**Security workers**		Building cleaners, street cleaners, sanitation workers
	Guards, bodyguards		Various low skill manual laborers

Notes: Translation of categories other than PCS 23, 35, 43 and 53 taken from Table 2 of Harrigan et al. (2016). (*): The PCS 31 – liberal professions category was created after the 2003 revision. Since we work with data from earlier years, we merge 37 and 31 together.

Table F.5: Summary statistics at the city level.

Item		year	mean	stdev	min	max
population		2015	293,485	1,007,302	50,571	10,706,072
number of firms with jobs in the city		1994	3,523	13,284	529	141,932
		2015	5,728	20,903	881	222,249
employment share	high-paid jobs	1994	0.090	0.027	0.052	0.233
		2015	0.139	0.049	0.080	0.367
	middle-paid jobs	1994	0.775	0.050	0.602	0.872
		2015	0.651	0.059	0.449	0.799
	low-paid jobs	1994	0.135	0.041	0.063	0.312
		2015	0.210	0.051	0.084	0.441
	MRO jobs	1994	0.419	0.078	0.240	0.625
		2015	0.307	0.055	0.176	0.489
	OMP jobs	1994	0.356	0.038	0.247	0.438
		2015	0.344	0.027	0.261	0.394
	middle-paid with wages above median	1994	0.374	0.059	0.231	0.614
		2015	0.289	0.051	0.137	0.463
	middle-paid with wages below median	1994	0.401	0.046	0.259	0.513
		2015	0.362	0.044	0.245	0.498
employment share percentage change 1994-2015	high-paid jobs		0.048	0.029	-0.005	0.153
	middle-paid jobs		-0.124	0.030	-0.204	-0.003
	low-paid jobs		0.076	0.024	-0.022	0.137
	MRO jobs		-0.112	0.041	-0.255	0.022
	OMP jobs		-0.012	0.036	-0.097	0.123
	middle-paid with wages above median		-0.085	0.040	-0.263	0.040
	middle-paid with wages below median		-0.039	0.035	-0.126	0.109

F.2 Evidence on SBTC and LMP

We begin by investigating wage polarization between our low-, middle-, and high-paid sectors over the period of interest. To do so, we run the following within regression of the individual logarithm of wages in yearly worker panel data for men in their prime working age:

$$ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_o} + \delta_t + v_i + e_{it} \tag{30}$$

where X_{it} are time-varying worker characteristics — age, tenure at current firm (both up to their quartic terms) or an indicator of obtaining a new job; $\gamma_{p_o A_o}$ are occupation (high-, middle- and low- paid) fixed effects composed of sectoral prices p_o and productivities A_o that we cannot identify separately; δ_t are year-fixed effects; and v_i are worker-fixed effects.[51] We use DADS-panel data for 1993-1995 to exploit the panel data dimension around 1994 and (in a separate regression) 2014-2016 around 2015. The low-paid sector is treated as the base sector. Sectoral price × productivity ratios inferred from occupation fixed effects are exhibited in Table F.6.[52]

Table F.6: Ratios of price×sector productivity fixed effects relative to the low-paid sector

Ratios	1993-1995	2014-2016	log-change
$\frac{p_m \times A_m}{p_l \times A_l}$	1.069	1.036	-0.031
$\frac{p_h \times A_h}{p_l \times A_l}$	1.154	1.191	0.031

Table F.6 shows that sectoral wages polarize. Taking the sector-component of the wage in the low-paid sector as a base, middle-paid wages decline by 0.31 log points, while the high-paid sector component rises in an equal magnitude (a symmetry feature we employ in Section 5).

We adapt the definition of skill-biased technical change (SBTC) from the model of Costinot and Vogel (2010) to our exercise. Their model features a continuum of tasks. In our empirical approach we order jobs by skill into three groups of low-, middle- and high-paid occupations. SBTC in the definition of Costinot and Vogel (2010) involves changes in relative factor demand biased towards higher skill workers. This translates in our setting into higher relative demand for workers in better-paid jobs (groups of tasks). Denote by $'$ later period values occurring after the change in demand. When the elasticity of substitution between intermediate goods in final good production is greater than 1, SBTC can be defined as $\left(\frac{A'_h}{A_h}\right) \geq \left(\frac{A'_m}{A_m}\right) \geq \left(\frac{A'_l}{A_l}\right)$, leading to a monotonic change in the wages (marginal value products)

$$\left(\frac{p'_h A'_h}{p_h A_h}\right) \geq \left(\frac{p'_m A'_m}{p_m A_m}\right) \geq \left(\frac{p'_l A'_l}{p_l A_l}\right) \tag{31}$$

Thus defined SBTC cannot explain the data patterns of Table F.6 by itself because the set of inequalities in (31) does not hold even if $\frac{p_h \times A_h}{p_l \times A_l}$ increases over time. This calls into question whether it is among the most important drivers of the Great Urban Divergence (Diamond and Gaubert, 2022).

[51] Without further data, we cannot separate at the sectoral level the changes in prices p_o from A_o.
[52] The differences between parameter values (within year/across time) are statistically significant.

However, automation and offshoring shocks affecting middle-paid jobs and inducing labor market polarization would reveal themselves through the following relative sectoral price evolution: $\left(\frac{p'_m}{p'_h}\right) < \left(\frac{p_m}{p_h}\right)$ and $\left(\frac{p'_m}{p'_l}\right) < \left(\frac{p_m}{p_l}\right)$. With such changes in relative sectoral prices alone we are able to rationalize the observed patterns absent of sectoral productivity changes[53].

As further evidence, we can expand the considered sectors and split the middle-paid sector into MRO and OMP tasks as defined in Section 2.2.2. These middle-paid jobs require similar skills but should be differentially affected by automation and offshoring, irrespectively of any SBTC shocks. We rerun equation (30) splitting the middle-paid jobs along MRO/OMP lines with results in Table F.7.

Table F.7: Ratios of price×sector productivity fixed effects across time relative to the low-paid sector

Ratios	1993-1995	2014-2016	log-change
$\frac{p_{MRO}\times A_{MRO}}{p_l\times A_l}$	1.068	1.038	-0.028
$\frac{p_{OMP}\times A_{OMP}}{p_l\times A_l}$	1.059	1.037	-0.021
$\frac{p_h\times A_h}{p_l\times A_l}$	1.147	1.192	0.039

The decline in $\frac{p_{MRO}\times A_{MRO}}{p_l\times A_l}$ is higher than that of $\frac{p_{OMP}\times A_{OMP}}{p_l\times A_l}$, indicative of an automation or offshoring shock affecting MRO jobs in particular — associated with labor-market polarization (though we argue in Appendix E that OMP jobs should be affected as well). In the following section, we explore the relative wage evolution in all 2-digit CS categories obtaining wage polarization. As above, this is evidence in favor of automation or offshoring shocks. A purely SBTC shock cannot explain the wage evolution observed in data.

F.2.1 Evolution of wages 1994-2015

We conduct similar exercises to the above ones using all 2-digit job categories, and we obtain wage polarization, shown in Table 1 and Figure F.3. Aggregate data strongly points to the presence of labor market polarization both in quantities and wages.

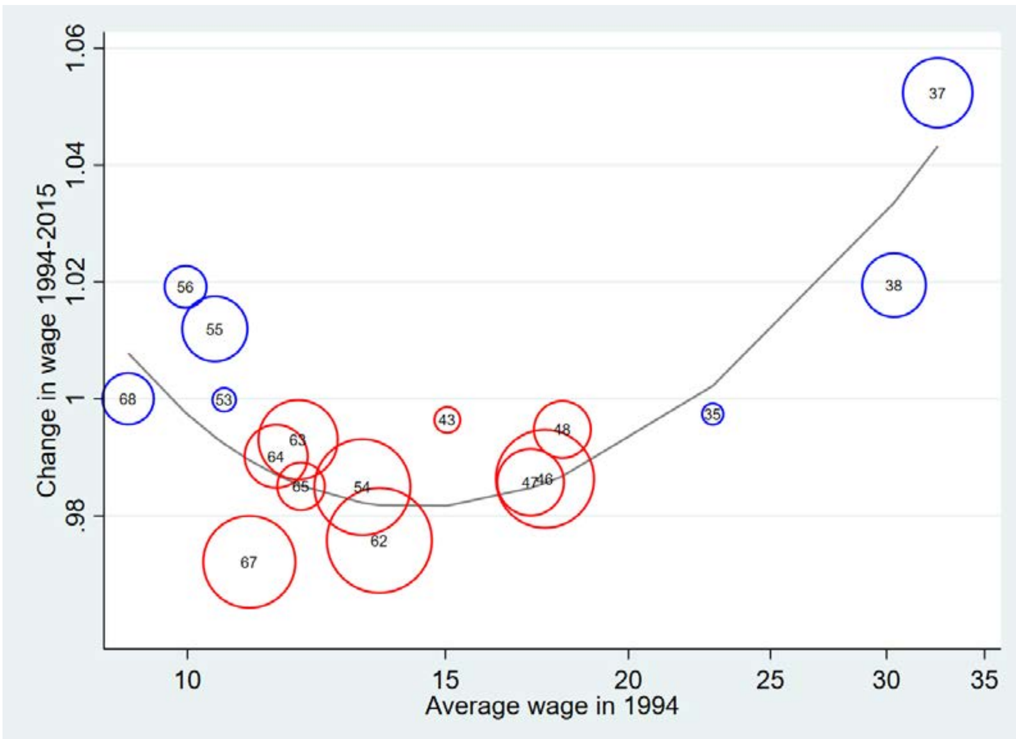
We estimate equation (30) using all 18 2-digit CS job categories instead of 3 sectors.

We focus on the estimates of $\gamma_{p_o A_o}$ for each occupation and period. These represent average value marginal products in each occupation — conditional wages — after accounting for worker observables and individual worker fixed effects. The lowest-paid category CS 68 (unskilled manual workers) is the

[53]We can further explore the extent of the labor market polarization implied by data constrained by different assumptions on the evolution of technology. Suppose that $\left(\frac{p'_h A'_h}{p_h A_h} / \frac{p'_l A'_l}{p_l A_l}\right) = \left(\frac{A'_h}{A_h} / \frac{A'_l}{A_l}\right)$ gives the extent of SBTC (meaning the relative p_h/p_l remain constant). Whenever $\left(\frac{A'_m}{A_m} / \frac{A'_L}{A_l}\right) \geq 0$ this implies given data $\ln\left(\frac{p'_m}{p'_h}\right) - \ln\left(\frac{p_m}{p_h}\right) = \ln\left(\frac{p'_m}{p'_l}\right) - \ln\left(\frac{p_m}{p_l}\right) < 0$ or labor market polarization. In other words, with our estimates, this implies that if there is SBTC in data, it *exacerbates* the needed relative decline in middle-paid sector prices to match the obtained estimates. In particular, when $\frac{A'_h}{A_h} \geq \frac{A'_m}{A_m}$ that the fall in $\frac{p_m}{p_h}$ and $\frac{p_m}{p_l}$ is restricted to the range $[-0.031, -0.063]$ required by our model simulation based on the Normal skill distribution (see Section 5) to obtain skewed polarization.

base. We calculate the ratios of estimated value marginal products (conditional wages) of individual CS occupations relative to that ratio for the base category CS 68 wage for the period 2014-2016 relative to 1993-1995. We can find thus the growth rates of conditional wages by occupation relative to CS 68 over this period. A value above 1 indicates that the relative wage of the CS job category increased in comparison to CS 68 (for which the normalization is 1). We plot the results in Figure F.3. We can also construct rankings of occupational wage growth over the 1994-2015 period, shown in Table 1.

Figure F.3: Changes in conditional wages 1994-2015 by 2-digit CS relative to CS 68.



The figure shows the change in the ratios of marginal value products (conditional wages after accounting for worker observables and individual fixed effects) of the considered 2-digit CS occupations relative to CS 68 in each year plotted against their 1994 average wage. The change in the CS 68 wage is normalized to 1. Circle sizes correspond to employment shares in 1994. Middle-paid jobs are shown in red while high and low-paid ones in blue. The line shows a cubic relationship between the average wage in 1994 and the relative wage change. The CS category "23" - CEOs excluded.

Figure F.3 documents aggregate wage polarization that occurred between 1994-2015 in mainland France and complements the job polarization exhibited in Figure 1.

The wages of the most skilled, best-paid occupations in 1994 such as managers and professionals (CS 37) and engineers (CS 38) increased the most relative to the least-paid CS 68 group. At the other end of the income distribution, some low-paid occupations' wages (CS 55 or 56) increased as well in relative terms. In contrast, the wages of *all* middle-paid occupations fell relative to those of CS 68, and had the slowest increases over 1994-2015, below any high- or low-paid occupations. In particular, the wages of occupations most exposed to automation and offshoring: unskilled (CS 67) and skilled (CS 62) industrial workers or office workers (CS 54) increased the least, ranking respectively 18th, 17th and 16th in terms of growth (see Table 1 for full ranking). The fitted cubic curve weighted by 1994 employment shares shows a similar U-relationship between initial average wages and relative wage growth as in Autor and Dorn (2013). Very similar patterns are obtained for quantity (employment share) changes in Figure 1. Thus, wage growth and employment shares changes go hand in hand in data for France between 1994-2015. Labor market polarization in our data is revealed as both job and wage polarization.

F.3 Model validation with data

In this section we discuss important features of the French data. Some, such as productivity across sectors and cities are those that we include in our model assumptions. Other features — such as log-supermodularity, composition of job shares across cities of different sizes or job-switching across cities and occupations are also implied by our model (Propositions D.2 and B.2 with Corollaries 1-2). This provides a cross-sectional model validation.

F.3.1 Wages and productivity across cities and occupations

Traditional urban models have focused on differences in city total factor productivity as a fundamental element in explaining city size differences and recent models have suggested the potential relevance of city-size-sector comparative advantage as well as skill sorting across cities. Our panel data allows us to examine these in this and the following subsection. We run within regressions (32) on DADS-panel data for 1993-1995 as in equation (30) including city size category × occupation fixed effects ($\gamma_{p_o A_{co}}$) instead of occupation ($\gamma_{p_o A_o}$) fixed effects only:

$$ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_{co}} + \delta_t + v_i + e_{it} \tag{32}$$

where X_{it} are time-varying worker characteristics — age, tenure at current firm (both up to their quartic terms) and an indicator of a obtaining a new job; p_o are sectoral prices (high-, middle- and low-paid tasks) while A_{co} are city size category × occupation fixed effects that cannot be identified separately; δ_t are year-fixed effects; v_i are worker-fixed effects; and e_{it} are error terms.

We focus on 1993-1995 to exploit the panel data dimension around 1994, the first year for which we have the exhaustive the DADS-Postes data used in the main study at the city level. The relative productivities between large and small cities by sector are shown in this Appendix Table F.8. For such comparisons the sectoral prices p_o cancel out. For example, in our leading grouping where we compare cities >0.5m with the smallest ones between 0.05-0.1m inhabitants (column 4), largest cities have a 1.086 times higher productivity in high-paid sector than smaller cities.

Whatever the grouping of large vs. small cities that is used, larger cities exhibit larger absolute productivities in all sectors and comparative advantage in high- relative to middle-paid sectors, as well as middle- relative to low-paid sectors. That is, large cities have a comparative advantage in more skilled sectors.[54] This data further justifies Assumption 3 made in Section 4.

[54]The estimated absolute and relative comparative advantages of larger cities are probably lower bounds.

Table F.8: Relative productivity across sectors and cities

Compared cities	Paris vs 50-100k	Paris, Lyon, Marseille vs 50-100k	Cities >1m vs 50-100k	Cities >500k vs 50-100k	Cities >200k vs <200k
high-paid sector	1.114 ***	1.093 ***	1.084 ***	1.086 ***	1.053 ***
middle-paid sector	1.092 ***	1.071 ***	1.058 ***	1.059 ***	1.045 ***
low-paid sector	1.060 ***	1.039 **	1.035 **	1.037 **	1.029 **

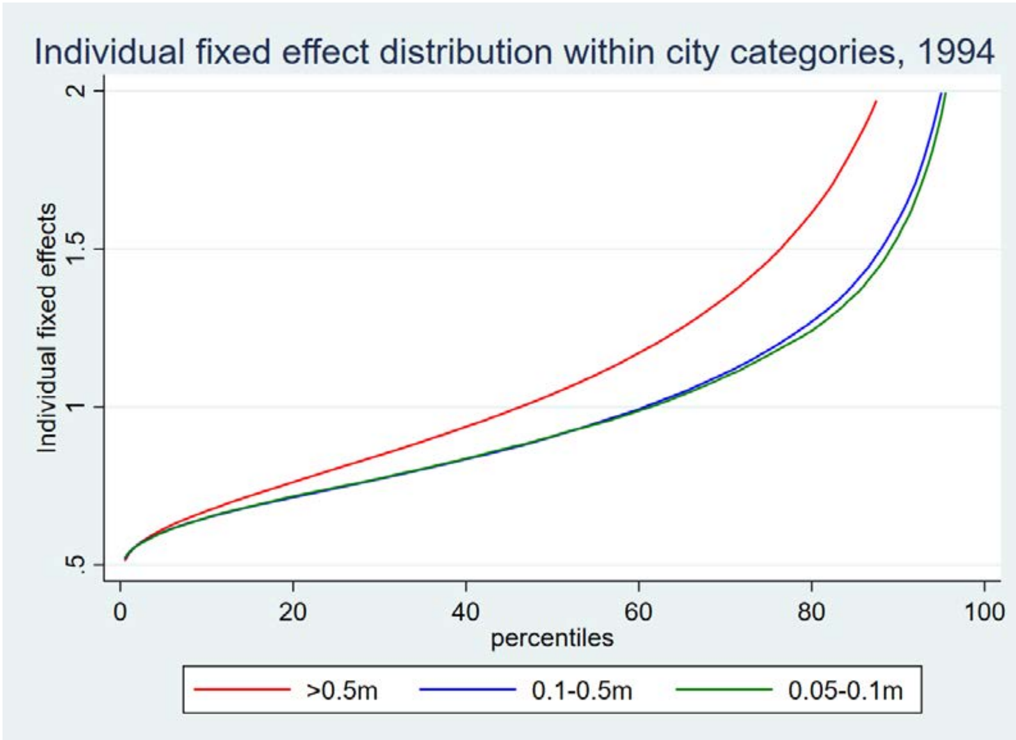
This Table presents the relative productivities across different groups of largest vs. smallest cities inferred from the terms $\gamma_{p_o A_{co}}$ in within regressions $ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_{co}} + \delta_t + v_i + e_{it}$ on DADS-Panel data for 1993-1995 using different city size groupings. In all cases we observe absolute productivity advantages of large versus small cities that are increasing in the average wage of the sector. Robust standard errors. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively of tests of hypotheses that the productivity coefficients across cities are equal.

F.3.2 Individual fixed effects from wage data

Figure F.4 shows the distribution of individual fixed effects obtained from regression (32) truncating their values at 2 for readability, while Figure F.5 gives the full exposition.

Our data does not contain non-wage compensation such as stock options that would typically figure more prominently in the compensation of top high-earners working at firm headquarters located predominantly in the largest cities.

Figure F.4: Individual fixed effects distribution within cities, value capped at 2.



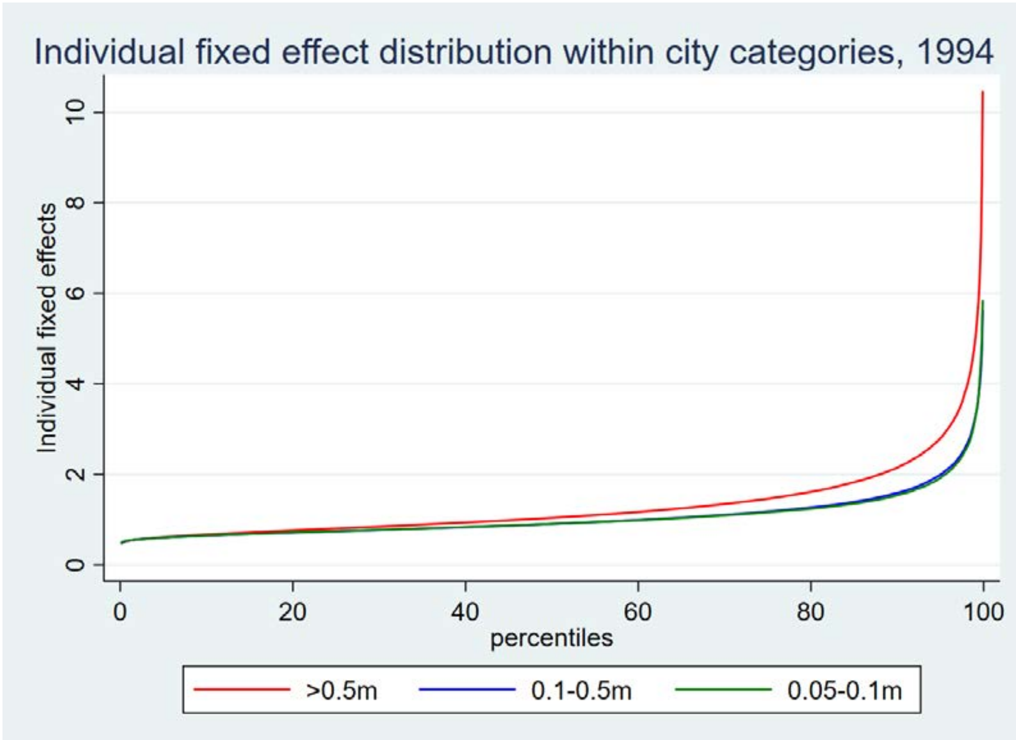
The figure shows the distribution of recovered individual fixed effects from a within regression (32) on a yearly worker DADS-panel data for men in 1993-1995 for cities above 0.5m, between 0.1-0.5m and 0.05-0.1m inhabitants:

$$ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_{co}} + \delta_t + v_i + e_{it} \tag{33}$$

where X_{it} are time-varying worker characteristics — age, tenure at current firm (both up to their quartic terms) or an indicator of a obtaining a new job, p_o are sectoral prices (high-, middle- and low-paid tasks) while A_{co} are city size category $\times$ occupation productivity, $\gamma_{p_o A_{co}}$ are city $\times$ sector fixed effects that cannot be identified separately, δ_t are year-fixed effects and v_i are worker-fixed effects.

Individual worker-fixed effects are recovered by calculating the mean prediction error. Given that the distribution of the fixed effects is positively skewed, we plot thus obtained worker fixed-effects truncating the individual fixed effect values at 2.

Figure F.5: Individual fixed effects distribution within different types of cities.



The figure shows the distribution of recovered individual fixed effects from a within regression (32) on a yearly worker DADS-panel data for men in 1993-1995:

$$ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_{co}} + \delta_t + v_i + e_{it} \tag{34}$$

where X_{it} are time-varying worker characteristics — age, tenure at current firm (both up to their quartic terms) and an indicator of a obtaining a new job; A_{co} are city size category $\times$ occupation (high-, middle- and low-paid) fixed effects; δ_t are year-fixed effects; v_i are worker-fixed effects; and e_{it} are error terms.
Individual worker-fixed effects are recovered by calculating the mean prediction error..

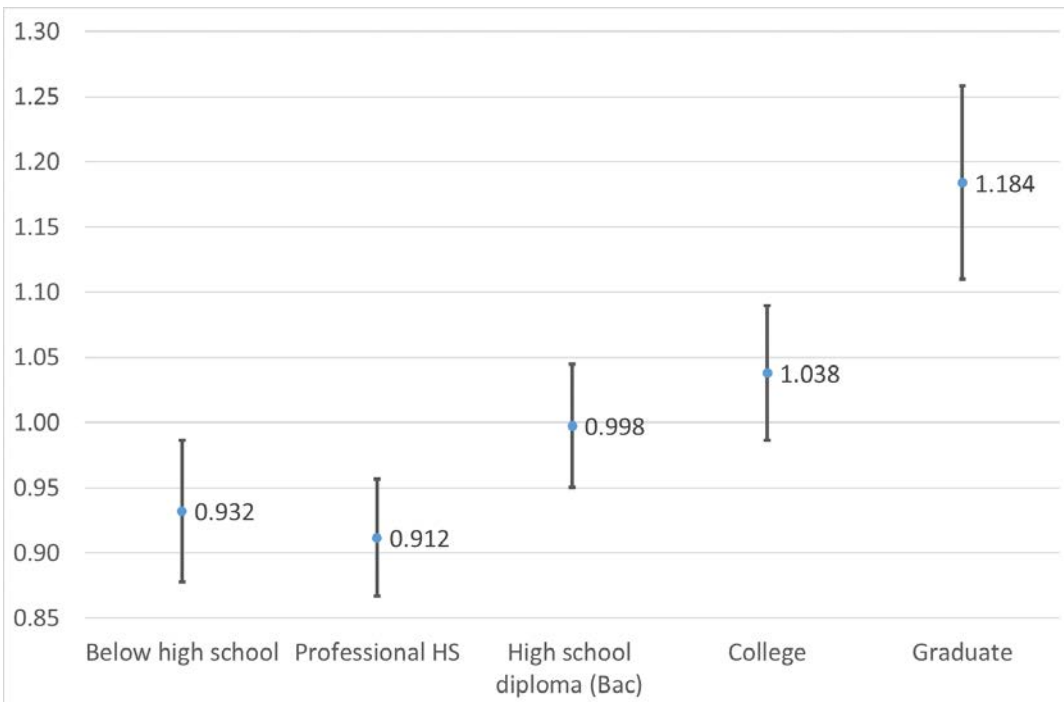
F.3.3 Log-supermodularity in data

Proposition B.2 provides conditions, as in Davis and Dingel (2020), under which the distribution of skills $f(\omega, c)$ is log-supermodular in city size. To obtain a measure of skills we turn to the 1999 Census, which has the best data on both diplomas and commune of residence among the Censuses spanning our time period.[55] We measure skills by the highest diploma received by individuals. The results are illustrated in Figure F.6 with further results in Tables F.9-F.14). As expected, Figure F.6 shows there is an ordering of the population elasticities of skills, with the two lowest skill groups having an elasticity statistically significantly below 1; two middle skill groups (with high school diplomas and some college) having an elasticity insignificantly different from 1; and a high skill category of workers with a graduate diploma that has a significant population elasticity of 1.18. These observations carry

[55] It spans 5% of population; provides data on education, nationality of respondents, and allows us to identify their location at the commune level. We also use the less-detailed 1990 and 2013 Censuses to document the evolution of e.g. educational attainment across cities.

over when we consider only French-born individuals; the presence of low-paid immigrants does not change these patterns. We also confirm these results using our classification of high, middle and low-paid jobs and the broad 1-digit CS categories in Tables F.12-F.13. It is not a coincidence that the population elasticity coefficients for high-paid jobs and “cadres” (respectively 1.14 and 1.16) are similar: “cadres” perform the bulk of high-paid occupations. The coefficients on middle- and low-paid jobs that are statistically significantly below 1 show that larger French cities have not only fewer low-paid jobs, but also fewer middle-paid jobs. This conforms with Corollary 1. Similar patterns in terms of population elasticities for different diploma categories can be obtained from the 1990 and 2013 Censuses (not shown), confirming the notion that log-supermodularity of skills holds for French cities over the entire studied period.

Figure F.6: Population elasticities by diploma (5 categories) in the 1999 Census data.



Notes: This sample contains 112 cities with > 0.05m inhabitants defined by INSEE as of 1999 with population figures as of 1999. Data on diplomas and residency is from the 1999 Census. Exclusions in terms of 2-digit CS and age as for the main DADS data used in the paper. 95% confidence intervals shown.

This Figure shows coefficients from regressions of the logarithm of the number of workers by five educational categories on the logarithm of city size. We observe log-supermodularity of skill distribution in city size as in Davis and Dingel (2020). The population elasticity for workers with graduate education (Master degrees and beyond) is 1.184 (significantly different from one at the 1% level) while for those with college (undergraduate) is 1.038. This means that larger cities have on average relatively more educated workers. At the same time, the least skilled (those with no diploma/a diploma below the general high school one or vocational – professional high school diplomas) are more likely to reside in smaller cities: the population elasticity estimates are significantly below one. The patterns do not qualitatively differ depending on whether we consider only the French-born fraction of the population. Table F.9 follows with more details.

Table F.9: Log-supermodularity, population elasticities by diploma (5 categories) in the 1999 Census data.

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
Below high school x ln pop	0.932**	0.914***	0.24	0.84
	(0.0274)	(0.0227)		
High school professional diploma (CAP, BEP) X ln pop	0.912***	0.907***	0.31	0.96
	(0.0226)	(0.0222)		
End of high school diploma (Bac) X ln pop	0.998	0.993	0.15	0.95
	(0.0238)	(0.0232)		
Undergraduate studies X ln pop	1.038***	1.034***	0.15	0.97
	(0.0262)	(0.0261)		
Graduate studies X ln pop	1.184***	1.18***	0.14	0.94
	(0.0374)	(0.0365)		

Notes: 112 cities > 0.05m inhabitants defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of 2-digit CS and age as in main sample. Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.
In this table the names of diplomas pertain to the following. CAP or *Certificat d'aptitude professionnelle* is obtained at the age of 16, the BEP or *Brevet d'études professionnelles* is also obtained at the age of 16 but is a prerequisite for obtaining the more advanced *bac professionnel* at the age of 18 that is included here with the general high school diploma (Bac).

Table F.10: Log-supermodularity, population elasticities by diploma (9 categories) in the 1999 Census data.

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
No diploma X ln pop	0.94*	0.91***	.12	.75
	(0.032)	(0.0254)		
End of primary school X ln pop	0.89***	0.88***	.08	.89
	(0.033)	(0.029)		
End of middle school (collège) X ln pop	0.98	0.97	.07	.94
	(0.024)	(0.023)		
Vocational school diploma (CAP) X ln pop	0.91***	0.90***	.20	.96
	(0.025)	(0.024)		
Vocational high school intermediate diploma (BEP) X ln pop	0.92***	0.92***	.10	.96
	(0.022)	(0.022)		
High school vocational diploma (bac technologique or professionnel) X ln pop	0.97	0.97	.09	.97
	(0.026)	(0.026)		
General high school diploma (Bac) X ln pop	1.04	1.04	.06	.93
	(0.029)	(0.028)		
Undergraduate studies X ln pop	1.04	1.03	.15	.97
	(0.026)	(0.026)		
Graduate studies X ln pop	1.18***	1.18***	.14	.94
	(0.037)	(0.037)		

Notes: 112 cities > 0.05m inhabitants defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as in main sample. Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.
In this table the names of diplomas pertain to the following. CAP or *Certificat d'aptitude professionnelle* is obtained at the age of 16, the BEP or *Brevet d'études professionnelles* is also obtained at the age of 16 but is a prerequisite for obtaining the more advanced *bac professionnel* at the age of 18.

Table F.11: Population elasticities by diploma (3 categories) in the 1999 Census data.

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
Below high school × ln pop	0.93**	0.91***	0.27	0.84
	(0.027)	(0.023)		
High school professional or general diploma × ln pop	0.94***	0.93***	0.44	0.96
	(0.022)	(0.022)		
Higher education × ln pop	1.10***	1.09***	0.29	0.96
	(0.030)	(0.030)		

Notes: Data from the 1999 Census. 112 cities > 0.05m inhabitants defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as for the main DADS data used in the paper. Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

Table F.12: Population elasticities by high-, middle- and low-paid categories in the 1999 Census data.

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
High-paid X ln pop	1.14***	1.14***	.17	.96
	(0.037)	(0.036)		
Middle-paid X ln pop	0.95*	0.95**	.64	.94
	(0.025)	(0.024)		
Low-paid X ln pop	0.94***	0.92***	.19	.85
	(0.018)	(0.015)		

Notes: 112 cities above 0.05m inhabitants as defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as in main sample. Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

Table F.13: Population elasticities by 1-digit CS categories in the 1999 Census data.

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
Cadres (CS 3) X ln pop	1.16***	1.15***	.17	.96
	(0.038)	(0.037)		
Intermediate professionals (CS 4) X ln pop	1.02	1.02	.28	.97
	(0.026)	(0.026)		
Low-skill employees (CS 5) X ln pop	0.97	0.96**	.27	.92
	(0.021)	(0.019)		
Blue-collar workers (CS 6) X ln pop	0.88***	0.86***	.28	.86
	(0.027)	(0.026)		

Notes: 112 cities above 0.05m inhabitants as defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as in main sample. CS 23 category – CEOs – not included in the category "cadres". Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

Table F.14: Population elasticities by 2-digit CS categories in the 1999 Census data.

CS code	description	All workers	French born	Population share	French born share
high-paid occupations					
23	CEOs	1.02	1.01	0.011	0.95
37	managers and professionals	1.12***	1.12***	0.090	0.97
38	engineers	1.25***	1.25***	0.059	0.96
35	creative professionals	1.23***	1.21***	0.016	0.93
middle-paid occupations					
48	supervisors and foremen	0.97	0.96	0.034	0.95
46	mid-level associate professionals	1.07**	1.06**	0.120	0.97
47	technicians	1.06	1.06	0.061	0.97
43	mid-level health professionals	0.96	0.96	0.064	0.98
62	skilled industrial workers	0.86***	0.84***	0.067	0.91
54	office workers	1	1	0.122	0.97
65	transport and logistics personnel	0.95	0.93*	0.020	0.92
63	skilled manual workers	0.94***	0.92***	0.066	0.85
64	drivers	0.95*	0.94**	0.031	0.92
67	unskilled industrial workers	0.79***	0.77***	0.053	0.86
low-paid occupations					
53	security workers	1.01	1	0.031	0.96
55	sales-related occupations	0.94***	0.93***	0.047	0.94
56	personal service workers	0.95*	0.92***	0.070	0.81
68	unskilled manual workers	0.92***	0.88***	0.038	0.74

Notes: This Table shows coefficients from regressions of the logarithm of the number of workers by occupational categories on the logarithm of city population (columns 3 and 4 for all and no foreign born population respectively). Sample includes 112 cities above 0.05m inhabitants as defined by INSEE as of 1999. Exclusions in terms of CS and age as in main sample. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

F.3.4 Occupation shares across cities

We now turn to occupation share patterns that can be obtained from the detailed DADS data on hours worked. As indicated in the theory section, people with the same skills may perform different occupations depending on the city where they choose to reside, and this has implications on the observable employment shares across cities.[56]

In Table F.15 we compare the means of occupational shares among the 11 largest cities in our sample (>0.5m inhabitants) and 62 smallest cities (0.05-0.1m inhabitants). Observing the first three columns of that Table, it is clear that the differences in the shares of high-, medium- and low-paid jobs across cities of different sizes are statistically different from one another. Larger cities have higher shares of high-paid jobs and lower shares of middle- and low-paid jobs than small cities both in 1994 and 2015. Our theory can account for these patterns.

Table F.15: Comparison of means of employment shares of different occupations, cities >0.5m vs. 0.05-0.1m.

Item	high-paid	middle-paid	low-paid	middle-paid above median	middle-paid below median
1994					
mean, cities >0.5m	0.188	0.690	0.123	0.363	0.327
mean, cities 0.05-0.1m	0.081	0.780	0.140	0.360	0.420
difference	0.107***	-0.09***	-0.017*	0.003	-0.094***
2015					
mean, cities >0.5m	0.303	0.509	0.188	0.233	0.276
mean, cities 0.05-0.1m	0.117	0.664	0.219	0.287	0.377
difference	0.186***	-0.155***	-0.031***	-0.054***	-0.101***

Notes: 1990 population weighted, robust standard errors. N=73; 11 cities > 0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

The Table shows the means of hours shares in total employment of different occupational groups for cities with >0.5m (large cities) and 0.05m-0.1m (small cities) inhabitants, and the comparison between the two types of cities. The reported difference in the means is a coefficient in the regression of shares on a large city dummy. Values are population weighted at the city level. The average share of high-paid jobs is higher while those of middle- or low-paid ones lower in larger cities both in 1994 and 2015, with the differences being significant at least at a 10% level. The discrepancies appear to grow with time (cf. Table 3 for tests). The difference in middle-paid jobs patterns across cities in 1994 and also in 2015 comes from the shares middle-paid job categories with wages below the median average wage that are less prevalent in large cities. There is no statistically significant difference between the average shares of middle-paid jobs with wages above the median average wage between the large and small cities in 1994. However, such a difference appears in 2015, and large cities have on average fewer middle-paid jobs in all categories.

Furthermore, in Table 2 we can observe the share of high-, middle-, and low-paid occupations in total employment across our six categories of cities in 1994 and 2015. The share of high-paid occupations in total employment increases monotonically with city size in both years. This is implied by our Corollaries 1 and 2. The differences are sizeable, especially when comparing the extremes – the Paris metropolitan area and cities with population between 0.05-0.1m. In both 1994 and 2015, the fraction of high skill jobs in Paris was roughly three times as high as in cities of 0.05-0.1m population. Given the overall rise in skilled jobs, this gap rose from 15 percentage points to 25 percentage points.

[56] In Table F.17 we show the joint distribution of diplomas and occupation categories in 1990 in the Census data. The distribution of higher-skill requiring diplomas is correlated with occupations ranked by wages.

The share of the middle-paid jobs monotonically declines with city size – accounted for by Proposition D.2 in both 1994 and 2015. The share of the lowest-paid occupations is highest in the smallest cities in either of the years, although the cross-city variation is modest.[57]

Rank-correlation statistics confirming these patterns are in Table F.16.

We conclude that in larger cities, the share of high-paid jobs is larger and the share of middle- and low-paid occupations is smaller in both 1994 or 2015, and our theory can capture these features of data.

Table F.16: Rank correlation statistics between city-level population in 1990 and mean shares of different occupation categories in 1994 and 2015

Occupation category	1994		2015	
	Spearman's ρ	Kendall's τ	Spearman's ρ	Kendall's τ
high-paid	0.50***	0.36***	0.58***	0.42***
middle-paid	-0.20**	-0.15**	-0.29***	-0.21***
low-paid	-0.12	-0.08	-0.22**	-0.15**
MRO	-0.18*	-0.13**	-0.29***	-0.21***
OMP	0.08	0.06	-0.14#	-0.10#
top 3 middle-paid with highest wages	0.56***	0.41***	0.41***	0.29***
least-well-paid middle-paid	-0.38***	-0.28***	-0.44***	-0.32***
intermediate professions	0.54***	0.39***	0.42***	0.30***
employees and blue-collar workers	-0.35***	-0.26***	-0.42***	-0.30***
middle – wages above median	0.26***	0.18***	-0.01	-0.02
middle – wages below median	-0.53***	-0.37***	-0.43***	-0.30***

Notes: 1990 population weighted, robust standard errors. 117 cities with > 50,000 inhabitants as of 2015. ***, **, * and # denote statistical significance at the 1 %, 5 %, 10 % and 15 % levels. Top 3 middle-paid with highest wages: CS 46, 47, 48. Least-well-paid middle-paid: CS 43, 54, 62, 63, 64, 65, 67. Intermediate professions: CS 43, 46, 47, 48. Employees and blue-collar workers: CS 54, 62, 63, 64, 65, 67.

[57] The decline of low-paid occupation shares with city size is, however, clear when one measures the share of hours worked for the three lowest-paid jobs (sales-related occupations, personal service workers and unskilled manual workers; see Table F.18).

Table F.17: Joint distribution of education levels and job types in 1990

	high-paid	middle-paid above the median	middle-paid below the median	low-paid	Row total in pct
none or at most middle school	1.47	7.03	15.98	9.97	34.45
CAP, BEP	1.74	9.08	12.01	4.99	27.83
High school (general, vocational)	3.12	7.66	6.63	2.07	19.48
At least college	9.35	6.99	1.55	0.36	18.25
Column total in pct	15.68	30.76	36.17	17.39	100

Note: Data from the 1990 Census for workers aged 25-64 years in cities >0.05m employed in the private sector, and within the 18 CS categories considered in the paper.

Table F.18: Shares of hours worked for middle- and low-paid jobs across agglomerations when CS 53 "Security workers" included in middle-paid jobs

Middle-paid						
Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1994	0.66	0.74	0.76	0.78	0.80	0.80
2015	0.47	0.59	0.62	0.65	0.67	0.68
change	-0.19	-0.16	-0.14	-0.13	-0.13	-0.12
growth in %	-28	-21	-18	-16	-16	-15
Low-paid						
Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1994	0.11	0.12	0.12	0.12	0.11	0.12
2015	0.16	0.16	0.17	0.18	0.19	0.20
change	0.05	0.05	0.05	0.06	0.08	0.08
growth in %	51	42	42	53	69	64

Table F.19: Shares of hours worked for middle- and low-paid jobs across agglomerations when CS 67 "Low-skilled industrial workers" included in low-paid jobs

Middle-paid						
Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1994	0.61	0.66	0.67	0.68	0.69	0.66
2015	0.42	0.53	0.56	0.59	0.60	0.60
change	-0.18	-0.13	-0.11	-0.09	-0.09	-0.06
growth in %	-30	-19	-16	-14	-13	-9
Low-paid						
Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1994	0.16	0.20	0.21	0.22	0.22	0.26
2015	0.21	0.22	0.23	0.25	0.26	0.28
change	0.05	0.02	0.02	0.03	0.04	0.02
growth in %	30	8	10	13	16	7

Table F.20: Share of 6 most-offshorable occupations per metropolitan area size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.49	0.54	0.54	0.56	0.58	0.59
2015	0.30	0.36	0.38	0.41	0.42	0.43
change	-0.19	-0.17	-0.16	-0.16	-0.16	-0.16
growth in %	-39	-33	-29	-28	-27	-27

F.3.5 Job and city transitions

One of the implications of our model is that individuals could hold a better-paid job in a larger city than in a small one if their skill happens to be just above a threshold in the larger city. We provide simple statistics of the patterns of job and city transitions.

We use 1993-1995 DADS-Panel data on working age men using 6 city groups as in Table 2.[58] We first show in Table F.21 cross-tabulations of categorical variables carrying information on occupation and city transitions. We code the occupation change variable as -1, 0, +1 for a change to a worse-paid job category, no change (of a job or category) or an upgrade to a better-paid job respectively within the 3 year window around 1994. Similarly, we code the variable capturing city change as -1 or +1 for moving into smaller/larger cities and 0 for no change in city size or no move). In this data we keep workers that did not change their place of work nor a job in the studied time period.

[58] The patterns are qualitatively similar using fewer city groups, e.g. >500k, 100-500k and 50-100k but with fewer observed moves.

Table F.21: Job type and city size transitions

city change	occupation change			
	-1	0	1	Total
-1	272	2,303	279	2,854
0	4,219	184,145	5,329	193,693
1	224	2,262	340	2,826
Total	4,715	188,710	5,948	199,373

Relatively more people upgrade their job upon moving to a larger city (340 upgrades versus 224 downgrades, a ratio of 1.51) than when they move to a smaller city (279/272, a ratio of 1.02).

To further explore these relations, we run an ordered logit regression of occupation change on city change, age up to a quartic term (to account for the likelihood of moving connected with age) and recovered individual fixed effects (to account for possible ongoing sorting) from a regression as in equation (32) using 6 city groups[59]:

$$ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_{co}} + \delta_t + v_i + e_{it} \tag{35}$$

where X_{it} are time-varying worker characteristics — age, tenure at current firm (both up to their quartic terms) or an indicator of a obtaining a new job, p_o are sectoral prices (high-, middle- and low-paid tasks) while A_{co} are city size category $\times$ occupation fixed effects that cannot be identified separately, δ_t are year-fixed effects and v_i are worker-fixed effects.

The ordered logit regression results are shown in Table F.22.

[59]Simpler specifications yield quantitatively similar results to those reported.

Table F.22: Ordered Logit Results

	occupation change
city change	0.388
	(0.113)
age	0.774
	(0.419)
age^2	-0.026
	(0.015)
age^3	-0.000
	(0.000)
age^4	(0.000)
	(0.000
individual FE	0.14
	(0.022)
cutpoint $(-1 \rightarrow 0)$	-12.49
cutpoint $(0 \rightarrow 1)$	-5.26

Ordered logistic regression. Robust standard errors. Std. errors in parentheses.

On the basis of the estimated model and the cutpoints we can calculate then the implied frequencies of changing to a particular type of a job as there is a transition to a different city size exhibited in Table F.23. We calculate the values for an individual of 38 years of age (the mean in data) with an average individual fixed effect.

Table F.23: Observed job type change percentages by city size and inter-city migration

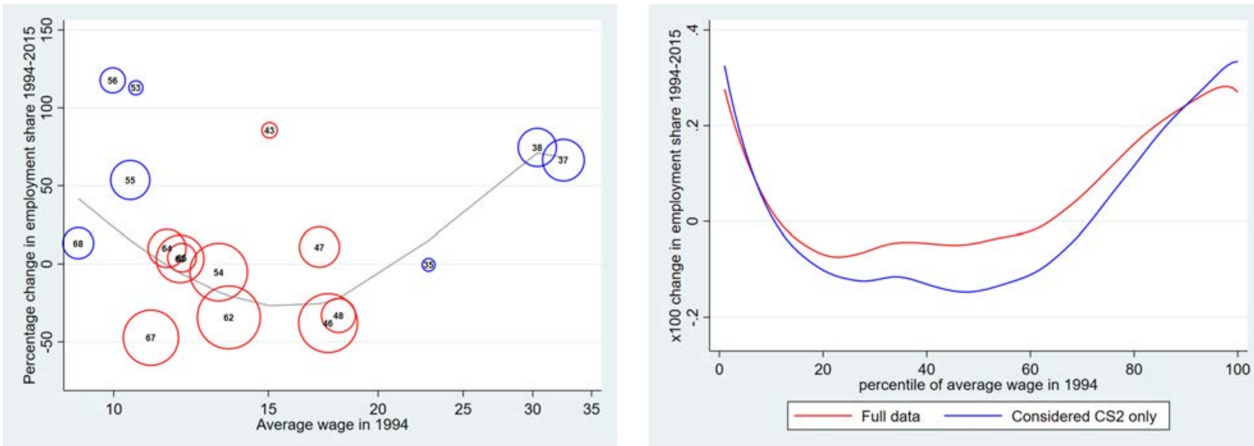
Change to:	worse-paid occupation	better-paid occupation
move to smaller city	0.034	0.020
no change in city size or no move	0.023	0.030
move to a larger city	0.016	0.043

The complement of the exhibited percentages is the one of no change in the job type.

Our results imply that within our timeframe upon moving to a larger city in 4.3% of cases workers should get a better job than held previously. This can be compared with 3% of cases when no move happened at all or to a city of different size or only 2.0% if the move occurred to a smaller city. On the other hand, upon moving to a smaller city workers were more likely to land a worse job (3.4%) than on average if not moving (2.3%) or migrating to a larger city (1.6%).

F.4 Labor Market Polarization: Additional Results

Figure F.7: Labor market polarization in France 1994-2015, percentage changes.



Left-hand figure: This figure, a counterpart to Figure 1 shows the percentage changes in employment 1994-2015 of the considered 2-digit CS occupation categories plotted against their 1994 average wage in cities with >0.05m inhabitants as of 2015. Circle sizes correspond to employment shares in 1994. Middle-paid jobs are shown in red while high and low-paid ones in blue. The line shows a cubic relationship between the average wage and the percentage point change in employment shares, a similar U-relationship as in Autor and Dorn (2013). The CS category "23" - CEOs excluded in this Figure. It is a an "occupation" with a high-wage and stable but small employment share.

Right-hand figure: This figure uses Labor Force Survey data. It shows the 100 × employment share changes between 1994 and 2015 by skill percentiles (as in Autor and Dorn (2013)) based on the average wage ranking of (i) all 348 4-digit occupations in 1994 with usable data (red line) (ii) 296 4-digit occupations in 1994 encompassed by 2 digit CS categories used in the paper without the CS category "23" - CEOs (blue line). Sample includes workers in the age 15-64. Changes are smoothed by using a locally weighted scatterplot smoothing model (LOWESS) with a .5 bandwidth. The merging of 4-digit PCS categories according to the 1982 PCS classification and the 2003 one done as in Bock (2020).
Both lines (on all or the CS2 categories used in the main paper) show polarization, albeit stronger for the sample used in our paper. The reason is that the wider category contains predominantly public sector workers for which no polarization occurs (see Section F.7).

F.4.1 Robustness: means of changes in shares of different occupations

In this Section we present different versions of Table 3 in the main body of the paper using different samples to show the robustness of the obtained patterns.

First, it is important to scrutinize the patterns once we drop Paris from the sample given the preponderance of this city in the French population and hours worked (Table F.25).

Next, in Table F.26 we change the definition of the city from "unite urbaine" to "aire urbaine" as defined by the INSEE. An "aire urbaine" constitutes a "unite urbaine" plus all communes where at least 40% of the resident working population has employment within the core "unite urbaine". The composition of smallest aires urbaines is slightly different from "unite urbaines" and there is more of them in the 0.05-0.1m category (65 vs. 62).

Then in Table F.27 we compare 11 largest cities with >0.5m inhabitants with 133 cities between 0.02-0.05m of inhabitants as of 2015.

In Table F.28 we assign PCS 53 "Security workers" as middle-paid jobs (belonging to those paid below the median) instead of low-paid jobs while in Table F.29 we assign PCS 67 "Low skill industrial

Table F.24: Rank correlation statistics between city-level population in 1990 and **percentage point changes** in employment shares of different occupation categories at the city level in the period 1994-2015.

Occupation category	Spearman's ρ	Kendall's τ
high-paid	0.49***	0.34***
middle-paid	-0.28***	-0.19***
low-paid	-0.30***	-0.21***
MRO	-0.07	-0.04
OMP	-0.20**	-0.14**
top 3 middle-paid with highest wages (CS 46, 47, 48)	-0.25***	-0.18***
least-well-paid middle-paid (CS 43, 54, 62, 63, 64, 65, 67)	-0.06	-0.04
intermediate professions (CS 43, 46, 47, 48)	-0.27***	-0.19***
employees and blue-collar workers	-0.05	-0.03
middle-paid with wages above median	-0.38***	-0.25***
middle-paid with wages below median	0.10	0.07

Notes: 117 cities with >0.05m inhabitants as of 2015. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

The Table shows Spearmans' ρ and Kendall's τ rank correlation statistics between city population ranks and percentage point changes in hours' shares of different occupational categories over the period 1994-2015.
This provides evidence that city sizes matter for the diverging patterns of labor market polarization, both in terms of magnitude and reallocation. Middle-paid jobs are destroyed the most in largest cities. There is a stronger creation of high-paid jobs in more populous cities. In contrast, there is a weaker growth of low-paid jobs in larger cities over the period (given the positive average share change at the city level in the period). Comparing changes among different groupings of middle-paid jobs one observes no significant correlation between city size and changes in the employment shares of MRO, seven least-well-paid middle-paid occupations, employees and blue-collar workers, and middle-paid with wages below median average wage. But OMP or the better-paid middle-paid jobs (top 3-paid; "intermediate" professions; jobs with wages above the median average wage) appear to be destroyed by more in larger cities over the period 1994-2015.

workers" as low-paid jobs instead of middle-paid ones. The latter assignment is counter to the spirit of our base classification as they are the second-most routine occupation and the most offshorable in our data.

In Table F.30 we show the results without restricting the hours worked to above 120 per payslip while in Table F.31 we give the patterns not weighting observations by city population.

Table F.33 presents the patterns comparing the evolution of employment shares among respectively the young (age 25-34) and old workers (age 55-64) in each studied year.

Table F.25: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample without Paris.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.095	-0.156	0.062	-0.111	-0.045	-0.110	-0.046
mean change, cities 0.05-0.1m	0.037	-0.116	0.080	-0.111	-0.006	-0.073	-0.044
difference	0.058***	-0.04***	-0.018***	0.000	-0.040***	-0.037***	-0.003
Growth in percent							
mean growth, cities >0.5m	70.0	-21.3	49.0	-29.9	-11.8	-29.1	-13.0
mean growth, cities 0.05-0.1m	45.7	-14.9	62.2	-25.2	-0.6	-19.9	-10.2
difference in growth	24.3***	-6.3***	-13.2**	-4.7**	-11.3***	-9.2***	-2.9

Notes: 1990 population weighted, robust standard errors. N=72 (10 cities > 0.5m as of 2015, without Paris). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.26: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample using "aires urbaines" as the definition of the city.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.104	-0.168	0.064	-0.109	-0.059	-0.120	-0.048
mean change, cities 0.05-0.1m	0.030	-0.109	0.080	-0.110	0.001	-0.067	-0.042
difference	0.074***	-0.059***	-0.015***	0.002	-0.060***	-0.052***	-0.007
Growth in percent							
mean growth, cities >0.5m	62.0	-24.1	54.3	-31.7	-16.2	-33.0	-14.4
mean growth, cities 0.05-0.1m	41.1	-14.0	61.5	-24.6	1.4	-19.2	-9.5
difference in growth	20.8***	-10.1***	-7.2	-7.1***	-17.7***	-13.8***	-4.9***

Notes: 1990 population weighted, robust standard errors. N=76 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.27: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.02-0.05m.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.116	-0.181	0.065	-0.108	-0.073	-0.130	-0.051
mean change, cities 0.05-0.1m	0.033	-0.116	0.083	-0.113	-0.003	-0.073	-0.043
difference	0.082***	-0.065***	-0.017***	0.005	-0.07***	-0.057***	-0.008
Growth in percent							
mean growth, cities >0.5m	63.0	-26.5	54.4	-33.1	-20.1	-35.9	-16.0
mean growth, cities 0.05-0.1m	44.4	-14.8	68.0	-24.3	0.5	-19.8	-9.6
difference in growth	18.6***	-11.8***	-13.7***	-8.8***	-20.6***	-16.1***	-6.3***

Notes: 1990 population weighted, robust standard errors. N=144 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.28: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample with PCS 53 "Security workers" counted towards middle-paid jobs (below the median).

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.116	-0.169	0.053	-0.108	-0.061	-0.130	-0.039
mean change, cities 0.05-0.1m	0.037	-0.113	0.076	-0.111	-0.002	-0.073	-0.040
difference	0.079***	-0.056***	-0.023***	0.003	-0.059***	-0.057***	0.001
Growth in percent							
mean growth, cities >0.5m	63.0	-24.4	47.7	-33.1	-16.3	-36.0	-11.8
mean growth, cities 0.05-0.1m	45.7	-14.3	62.3	-25.2	0.4	-19.9	-9.1
difference in growth	17.2***	-10.1***	-14.7***	-7.9***	-16.7***	-16.0***	-2.7*

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.29: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample with PCS 67 "Low skill industrial workers" counted towards low-paid jobs.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.116	-0.153	0.037	-0.108	-0.073	-0.130	-0.023
mean change, cities 0.05-0.1m	0.037	-0.059	0.023	-0.111	-0.006	-0.073	0.013
difference	0.079***	-0.094***	0.015	0.003	-0.068***	-0.057***	-0.036***
Growth in percent							
mean growth, cities >0.5m	63.0	-24.5	21.2	-33.1	-20.1	-36.0	-8.9
mean growth, cities 0.05-0.1m	45.7	-8.9	10.2	-25.2	-0.6	-19.9	5.3
difference in growth	17.2***	-15.7***	11.0*	-7.9***	-19.5***	-16.0***	-14.2***

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.30: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample keeping hours worked <120 hours / year (no "filtering" of observations by hours).

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.115	-0.181	0.066	-0.108	-0.073	-0.130	-0.051
mean change, cities 0.05-0.1m	0.036	-0.117	0.080	-0.111	-0.006	-0.073	-0.044
difference	0.079***	-0.064***	-0.014***	0.003	-0.067***	-0.057***	-0.008
Growth in percent							
mean growth, cities >0.5m	62.7	-26.6	54.7	-33.2	-20.0	-36.0	-15.9
mean growth, cities 0.05-0.1m	45.4	-15.0	62.4	-25.3	-0.6	-20.0	-10.1
difference in growth	17.3***	-11.6***	-7.6	-7.9***	-19.5***	-16.0***	-5.8***

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.31: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample without weighting the results by city population in 1990.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.096	-0.157	0.062	-0.111	-0.047	-0.113	-0.045
mean change, cities 0.05-0.1m	0.036	-0.116	0.079	-0.109	-0.006	-0.072	-0.044
difference	0.059***	-0.042***	-0.018***	-0.001	-0.040***	-0.041***	-0.001
Growth in percent							
mean growth, cities >0.5m	67.5	-21.8	48.9	-30.4	-12.2	-30.2	-12.9
mean growth, cities 0.05-0.1m	45.9	-14.9	61.5	-24.9	-0.8	-20.0	-10.1
difference in growth	21.6***	-6.9***	-12.6**	-5.5***	-11.4***	-10.2***	-2.8

Notes: Robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.32: Comparison of means of changes in employment shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample without finance, insurance and real estate sectors.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.104	-0.177	0.072	-0.109	-0.067	-0.128	-0.048
mean change, cities 0.05-0.1m	0.032	-0.116	0.084	-0.119	0.004	-0.066	-0.049
difference	0.073***	-0.061***	-0.012**	0.010	-0.071***	-0.062***	0.001
Growth in percent							
mean growth, cities >0.5m	59.6	-25.9	56.6	-33.5	-18.4	-35.6	-15.0
mean growth, cities 0.05-0.1m	41.9	-14.9	64.0	-27.1	2.1	-18.2	-11.4
difference in growth	17.7***	-11.1***	-7.4	-6.4***	-20.5***	-17.4***	-3.5**

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.33: Comparison of means of changes in employment shares of different occupations across young and old cohorts, cities >0.5m vs. 0.05-0.1m.

Percentage changes	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above the median	middle-paid below the median
workers in the 25-34 age cohort							
mean change, cities >0.5m	0.111	-0.181	0.070	-0.091	-0.090	-0.116	-0.064
mean change, cities 0.05-0.1m	0.025	-0.126	0.101	-0.104	-0.022	-0.062	-0.064
difference	0.086***	-0.055***	-0.031***	0.013*	-0.068***	-0.055***	-0.000
workers in the 55-64 age cohort							
mean change, cities >0.5m	0.087	-0.136	0.049	-0.105	-0.031	-0.100	-0.036
mean change, cities 0.05-0.1m	0.018	-0.064	0.046	-0.068	0.004	-0.061	-0.003
difference	0.070***	-0.072***	0.003	-0.037***	-0.035***	-0.039***	-0.033***

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.34: Additional statistics for the main sample for middle-paid jobs.

Item	top three middle-paid with highest wages	Bottom seven least-well-paid middle-paid	intermediate professions	employees and blue-collar workers	white-collar workers	blue-collar workers
Changes						
mean change, cities >0.5m	-0.089	-0.092	-0.083	-0.097	-0.086	-0.079
mean change, cities 0.05-0.1m	-0.037	-0.079	-0.027	-0.089	-0.036	-0.093
difference	-0.052***	-0.013**	-0.056***	-0.008	-0.050***	0.014
Growth in percent						
mean growth, cities >0.5m	-33.1	-21.8	-30.0	-23.6	-28.2	-28.3
mean growth, cities 0.05-0.1m	-18.4	-13.4	-12.4	-15.4	-15.1	-19.7
difference in growth	-14.7***	-8.5***	-17.5***	-8.2***	-13.0***	-8.6***

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities. Top 3 middle-paid with highest wages: CS 46, 47, 48. Bottom 7 least-well-paid middle-paid: CS 43, 54, 62, 63, 64, 65, 67. Intermediate professions: CS 43, 46, 47, 48. Employees and blue-collar workers: CS 54, 62, 63, 64, 65, 67. White-collar workers: CS 46 & 54. Blue-collar workers: CS 62, 63, 64, 65, 67.

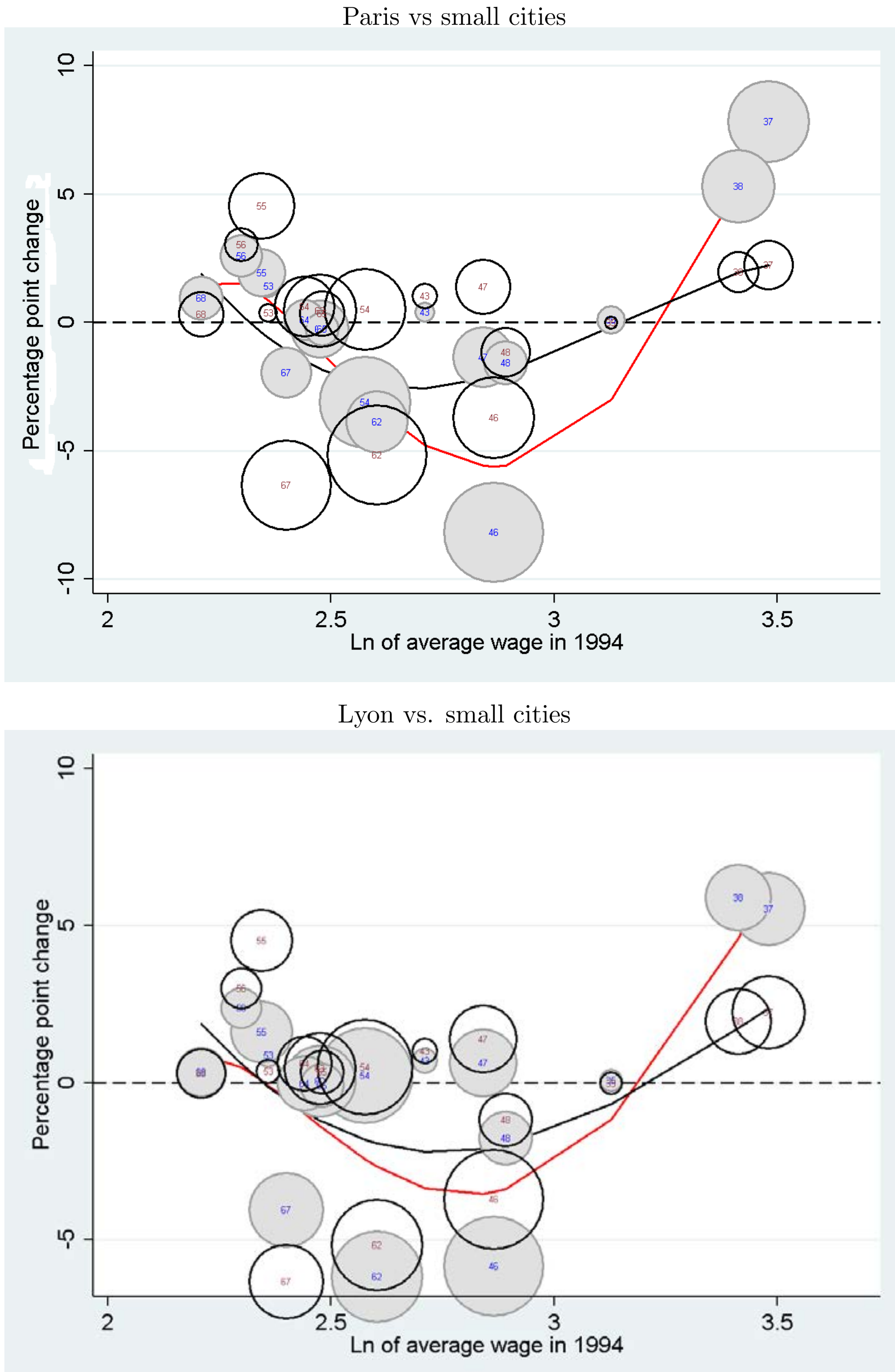
F.4.2 Other graphs and figures

Table F.35: Educational divergence across cities 1990-2013.

1990						
Agglo. size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
none or at most middle school	32.44	32.28	30.27	31.69	32.66	34.54
CAP, BEP	22.69	26.13	29.44	29.82	30.16	30.29
High school (general, vocational)	20.01	20.64	21.54	20.2	20.18	20.34
At least college	24.86	20.95	18.75	18.28	17	14.83
2013						
Agglo. size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
none or at most middle school	19.21	16.87	16.21	17.66	18.44	19.85
CAP, BEP	14.85	19.12	22.57	23.65	24.49	26.9
High school (general, vocational)	16.67	17.88	19.18	18.37	18.6	19.15
At least college	49.28	46.13	42.04	40.31	38.47	34.1

Note: Data from 1990 and 2013 Censuses for workers aged 25-64 years in non-agricultural sectors in cities >0.05m.

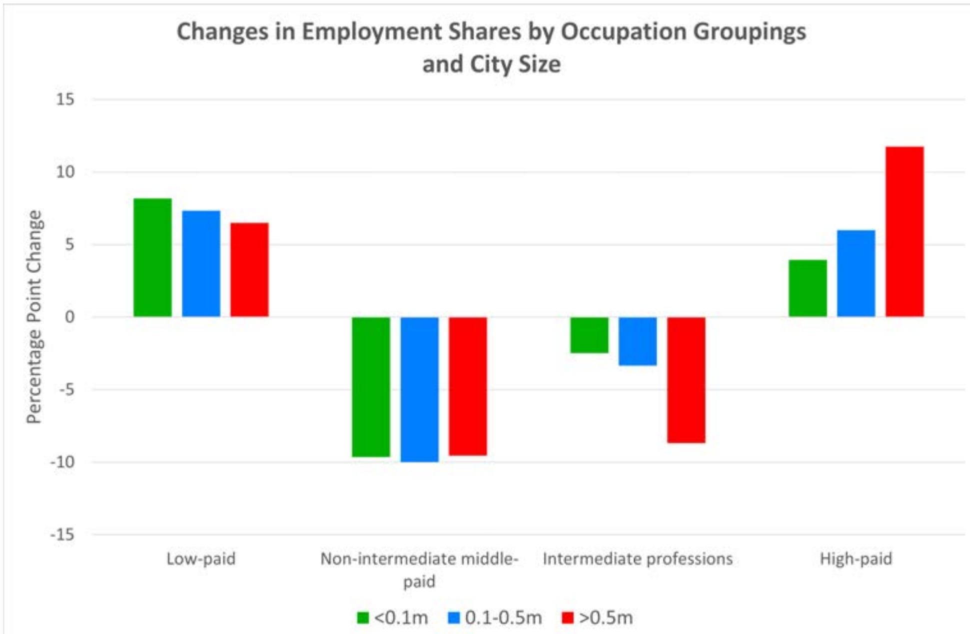
Figure F.8: Comparison of labor market polarization 1994-2015 in Paris/Lyon and cities between .05-.1m inhabitants.

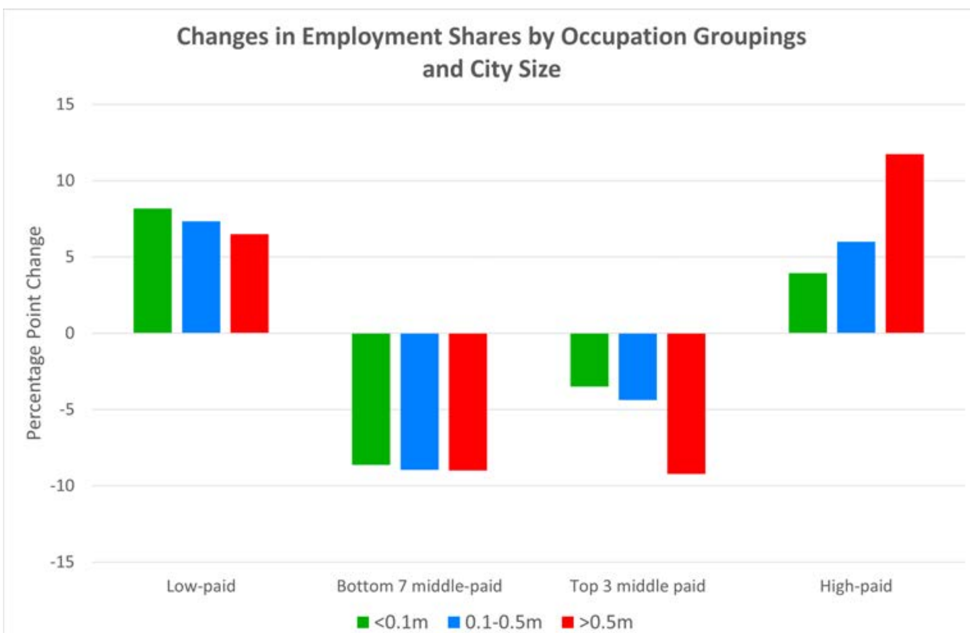


The figure shows the percentage point change in employment shares of the considered 2-digit CS categories plotted against their average wage in cities > .05m in 1994. Numbers pertain to 2-digit CS categories represented. Grey circles stand for Paris (upper panel) or Lyon (lower panel) while white for small city shares. Circle sizes correspond to the employment shares (same scale for the two compared groups) in 1994. The two lines shows a cubic relationship between the average wage and the percentage point changes in employment for Paris (red) and cities between .05-.1m inhabitants (black) respectively. The CS category "23" - CEOs excluded.

This figure documents a stronger decline in middle-paid jobs in larger cities and differential reallocation effects (higher creation of high-paid jobs in larger agglomerations) of labor market polarization across cities of different sizes using the contrast between Paris / Lyon and cities of 0.05-0.1m as a group.

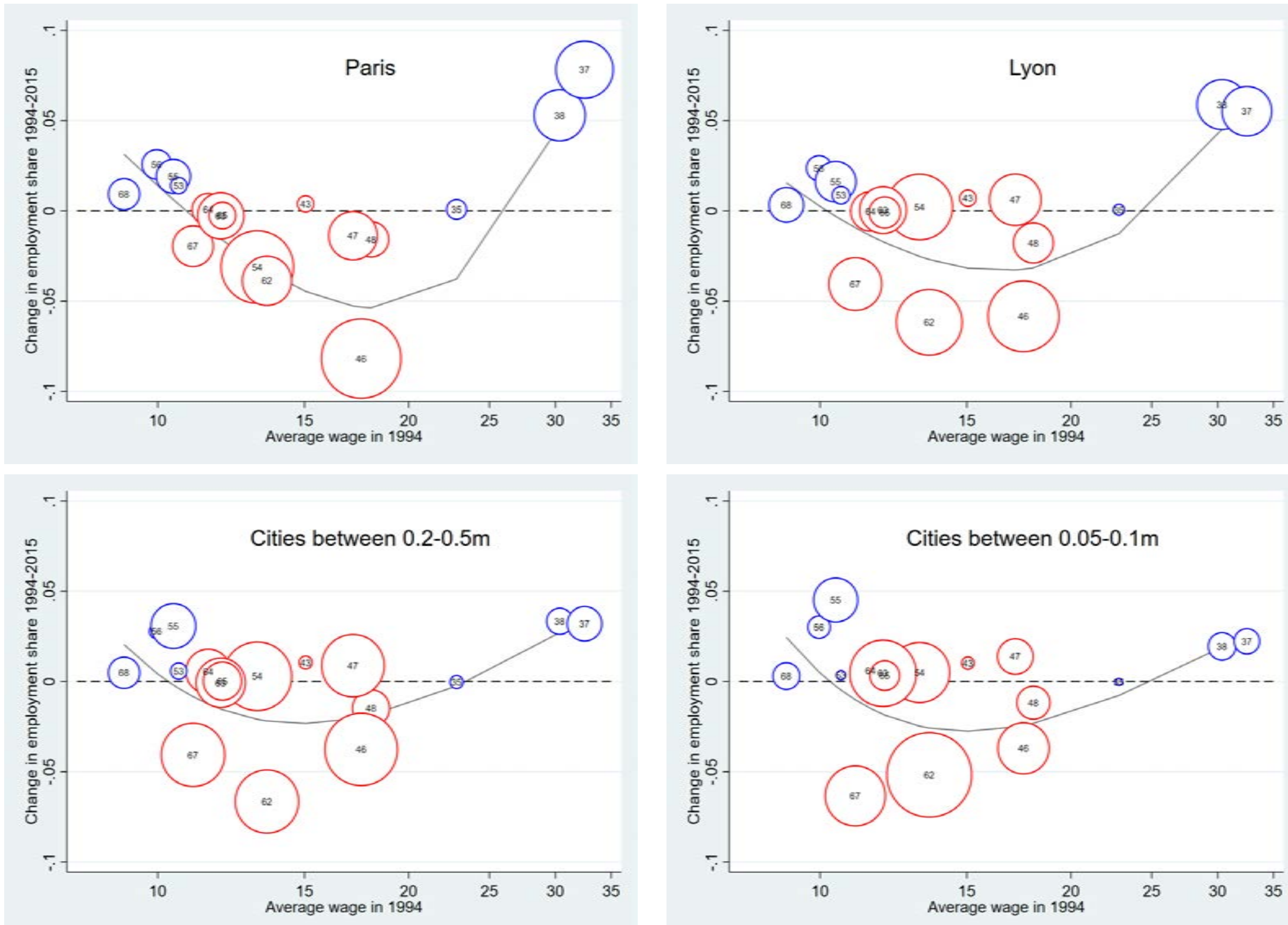
Figure F.9: Labor market polarization across three different city size groups, 1994-2015: 4 employment groups





This figure shows percentage point changes in employment shares of high-, low- and different types of middle-paid jobs with hours worked summed by job types and 3 city sizes: large (above >0.5m inhabitants), medium-sized (0.1-0.5m) and small (0.05-0.1m) in the period 1994-2015. The various partitions of middle-paid jobs in each panel order these jobs by median wage. In the upper panel the middle-paid jobs are divided into intermediate professions (CS 43, 46, 47, 48) and other middle-paid jobs. In the lower panel these are divided into top-3 paying middle-paid occupations (CS 48, 46, 47) and the remainder of middle-paid jobs. Top-3 paid middle-paid jobs are a subset of intermediate professions that are also in the upper tier of middle-paid occupations.
All of the panels show that, for all these partitions of the middle skill jobs, the destruction of the lower-paid jobs was similar across all city sizes. At the same time, the panels show clearly that the destruction of the highest-paid middle-skill jobs rises monotonically with city size.

Figure F.10: Labor market polarization across cities 1994-2015.



Left-upper panel: Paris. Right-upper panel: Lyon. Left-lower panel: cities between 0.2-0.5m. Right-lower panel: cities between .05-.1m inhabitants.
Figures show the percentage point change in employment shares of the considered 2-digit CS categories plotted against their average wage in cities > 0.05m in 1994. Numbers pertain to 2-digit CS categories represented. Circle sizes correspond to the employment shares in 1994. The line shows a cubic relationship between the average wage and the percentage point change. The CS category "23" - CEOs excluded.
This figure documents that labor market polarization occurred in all types of cities. The magnification and differential reallocation effects of labor market polarization for larger cities in contrast to smallest cities are clearly visible: there was a weaker destruction of middle-paid jobs in smaller cities, and low-paid jobs were created there more strongly than high-paid occupations.

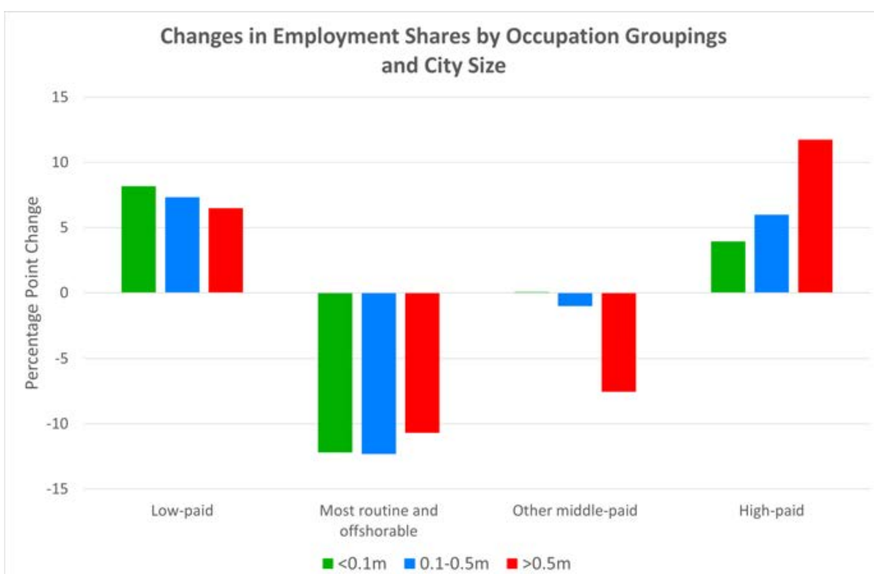
F.5 Initial exposure and the evolution of different job categories

In this section, we reconfirm prior work on polarization and show that locations' exposure to MRO jobs is a good predictor of loss of these jobs. However, exposure to MRO jobs, to the contrary, is *negatively* correlated with the loss of OMP jobs and in general *not* correlated with the destruction of middle-paid jobs taken as a whole.

We start by contrasting the MRO and OMP middle-paid jobs evolution in Figure F.11. Noting that MRO jobs are on average lower-paid than OMP jobs, we can order these on a wage axis. We have grouped cities into three sizes, with those above 0.5m at the top and those below 0.1m at the bottom. There indeed we see that large cities (that have a lower initial exposure to these occupations) have a smaller percentage point loss of MRO jobs, although the differences are modest. We also see that the strong contrast in experience comes in the OMP jobs. There are striking declines in OMP jobs in the largest, small losses in middle-sized, and essentially zero change in the smallest cities. That is, the contrasts in experience across city size in middle-paid job loss across cities of different sizes is precisely in the segment of jobs that in previous work was dropped from the discussion.

An important point is that OMP jobs may also be sensitive to trade or automation shocks. It may be indirectly, such as middle-managers (CS 46) who would lose MRO jobs to manage, or it occurs later in comparison to MRO occupations. Furthermore, what may matter in a large city labor market is the *relative* degree of routinizability and offshorability in comparison to high-paid occupations.

Figure F.11: Labor market polarization and the great urban divergence across three different city size groups, 1994-2015: MRO and OMP split of middle-paid jobs



This figure shows percentage point changes in employment shares of high-, low- and different types of middle-paid jobs with hours worked summed by job types and 3 city sizes: large (above >0.5m inhabitants), medium-sized (0.1-0.5m) and small (0.05-0.1m) in the period 1994-2015. The bars for high- and low-paid jobs are exactly as in Figure 3. The division of middle-paid occupations is between the most routine and offshorable (MRO) and other middle-paid occupations (OMP).
The figure shows that the destruction of the MRO jobs was similar across all city sizes. At the same time, the destruction of the OMP jobs rises monotonically with city size. Indeed, OMP occupations actually *grow* very modestly in the smallest cities.

F.5.1 Exposure and the loss of the most routinizable and offshorable (MRO) jobs

We investigate whether cities with a higher exposure to the most routinizable and offshorable jobs (our MRO group) see the largest decline in the share of these jobs. In the lowest panel of Table F.36 we report both the 1994 and 2015 employment shares in the four CS 2-digit MRO occupations in six city groups. The employment share in these four MRO jobs are declining in city size both in 1994 and in 2015 in line with the patterns for middle-paid occupations overall. This observation is confirmed using rank correlations in Table F.16.

In the lowest panel of Table F.36, the evolution of these shares over this period is relatively constant in percentage points: we see that the fall of shares in this category of middle-paid jobs is similar across metropolitan areas without any clear relationship with size – between 10.5 and 13.1 percentage points. Table 3 shows no statistically significant difference between large and small cities in the change in these MRO occupations. The same conclusion arises in rank-correlation tests (Table F.24).

Table F.36: Share of high-, middle- and low-paid occupations in hours worked per metropolitan area size in 1994 and 2015.

MRO							
Agglo.size	Paris	> .75m	.5-.75m	.2-.5m	.1-.2m	.05-.1m	All cities
1994	0.29	0.36	0.39	0.41	0.45	0.45	0.36
2015	0.19	0.25	0.27	0.29	0.31	0.32	0.25
change	-0.11	-0.11	-0.12	-0.12	-0.13	-0.12	-0.12
growth in %	-36	-32	-31	-29	-29	-27	-32

This Table shows the means of shares of hours in total employment of different occupational groups in 1994 and 2015 for all 117 cities in our sample allocated in 6 bins according to city size (with Paris being a separate category), showing the percentage point changes and growth rates between 1994-2015. One observation per bin of the hours totals.
The share of MRO jobs (CS 48, 54, 62 and 67) in total employment is decreasing with city size whether in 1994 or 2015. Percentage point destruction of these MRO jobs is similar across city sizes despite their lower initial share in employment for larger cities.

However, these statements do not control for actual exposure to these specific jobs at the city level. It is clear from Figure F.12 that large cities above 0.5m people have lower initial exposures to the MRO occupations.[60] Although there is considerable variation, Figure F.12 confirms the observation of Autor and Dorn (2013) that the initial exposure to the most routine (and, in our context, also offshorable) jobs is strongly negatively correlated with their change as technological shocks occur. The observations for large cities lie in the lower envelope of observations. Thus conditional on initial exposure, the changes in the employment shares in these cities are larger than in other cities. Regression analysis in Table F.39, top panel, confirms these points: initial exposure is negatively correlated with change in the MRO employment shares and the interaction of a dummy for large cities with exposure is robustly negatively different from zero. MRO jobs in large cities are destroyed at a higher rate than in small cities with the same initial exposure in reaction to the same automation or trade shocks.

In the end, consistent with Autor and Dorn (2013), we obtain that the initial exposure to the most

[60]The large cities with the highest initial exposure to MRO jobs are the Douai-Lens and Lille metropolitan areas, both located in the old industrial region in the North of France.

routinizable and offshorable jobs is a good predictor of the most routinizable and offshorable jobs loss themselves. However, such jobs, conditional on exposure, are decimated more in larger cities.

F.5.2 Exposure to MRO jobs and the broad loss of middle-paid jobs

We see that initial exposure to MRO jobs is strongly associated with subsequent loss of these jobs. But this raises the question of whether exposure to MRO jobs is also associated with the loss of other middle-paid (OMP) jobs, or indeed with middle-paid jobs taken as a whole.

We can look at the relation between initial exposure to MRO jobs and the subsequent change in OMP jobs in Figure F.12. There is a strikingly strong negative relation between exposure to MRO jobs and subsequent loss of OMP jobs, confirmed in the second panel of Table F.39. The big losses of OMP jobs are in large cities, which have initially small exposure to MRO jobs.

We also know, though, from Fact 2 that large cities experienced a larger decline in middle-paid jobs overall (that include the MRO category). This leads us to suspect that the exposure to MRO jobs by itself may not be a good predictor of the overall change in middle-paid jobs.

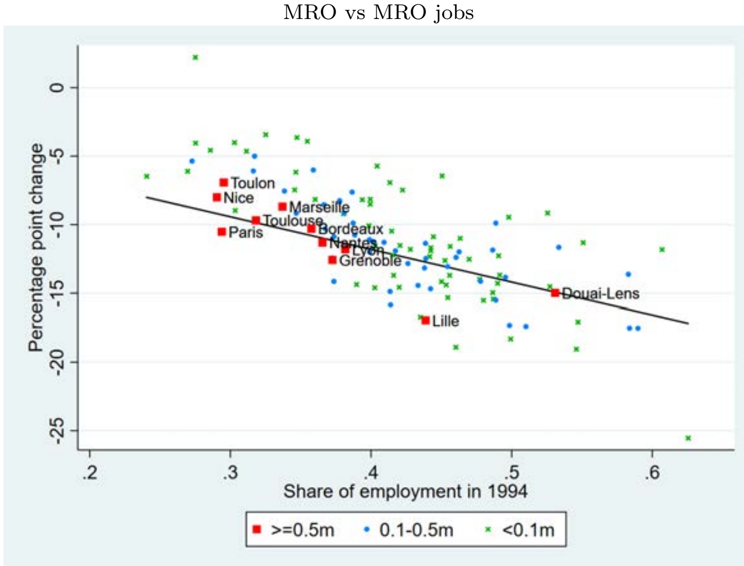
Indeed, the population-weighted regression of changes in employment for the entire middle-paid category on initial exposure to the four MRO occupations (in Table F.38, bottom panel) reveals a strong positive relationship (though the non-population weighted relationship is zero). There is clearly a larger destruction of middle-paid jobs in the largest cities conditioning on exposure, witnessed by the sign of the interaction of a dummy for large cities with MRO job exposure. For many small but highly-exposed cities, the drop in MRO jobs is larger than the decline in middle-paid jobs while the opposite is true for the largest cities. The initial exposure to the most routinizable and offshorable (MRO) jobs is not a good predictor of the evolution of the entire class of middle-paid jobs across cities. Thus, initial exposures to the MRO jobs are not a key driver of a broad measure of labor market polarization in local labor markets.

Even if automation and/or offshoring (Autor and Dorn, 2013; Goos et al., 2014) are driving labor market polarization, the extent to which these affect the broad category of middle-paid jobs does not depend only on a city's initial exposure to the most routinizable and offshorable jobs. These results are robust to considering a larger set of occupations for the most routinizable and offshorable jobs[61] and also a longer time period (See Section F.6 of this Appendix).

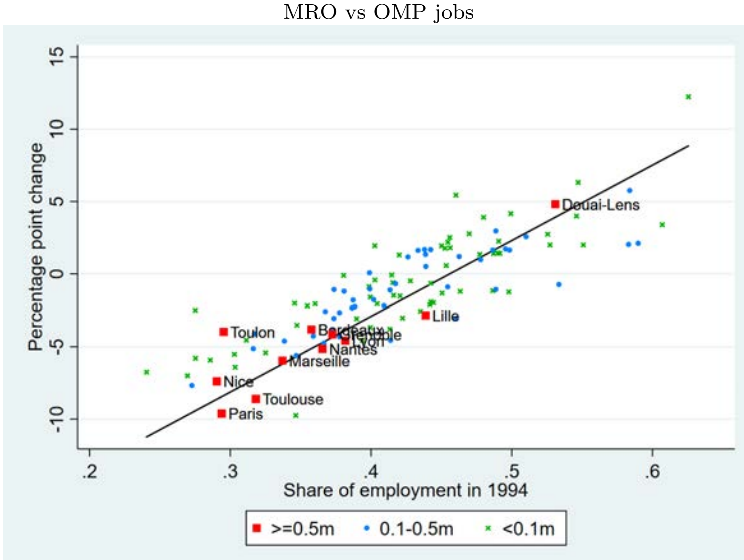
[61] In Table F.20 we show the patterns for the 6 most offshorable jobs encompassing not only the four MRO occupations, but also categories such as transport and logistics personnel (CS 65) and mid-level professionals (CS 46). The shares of such jobs in total employment are monotonically decreasing with city size whether in 1994 or 2015. Percentage point fall in the employment shares of these occupations between 1994 and 2015 is higher in larger cities, thus confirming our results. Moreover, it can be seen in Figure F.15 that a higher initial exposure to offshorable jobs leads to their greater decrease in the studied period. This time, however, large cities are relatively more exposed to offshorable jobs in comparison to MRO occupations only as – in particular – they have on average a higher share of the CS 46 category, mid-level professionals (cf. also Table F.20). In Table F.40 we demonstrate that the initial exposure to this wider set of occupations is also a good predictor of their employment share change. Again, conditional on exposure, offshorable jobs' employment shrinks by more in large cities in the studied period.

Figure F.12: Exposure to MRO jobs and change in the employment shares of MRO and other middle-paid OMP jobs, 1994-2015.

MRO vs MRO jobs



MRO vs OMP jobs



The figures show the percentage point change in employment shares of MRO jobs (CS 48, 54, 62 and 67) between 1994-2015 plotted against their share in employment (upper panel) or other middle-paid OMP jobs (lower panel) in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The lines show a linear, population-weighted (by 1990 population) fit of the relationship between employment changes and the initial exposure to MRO jobs. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.

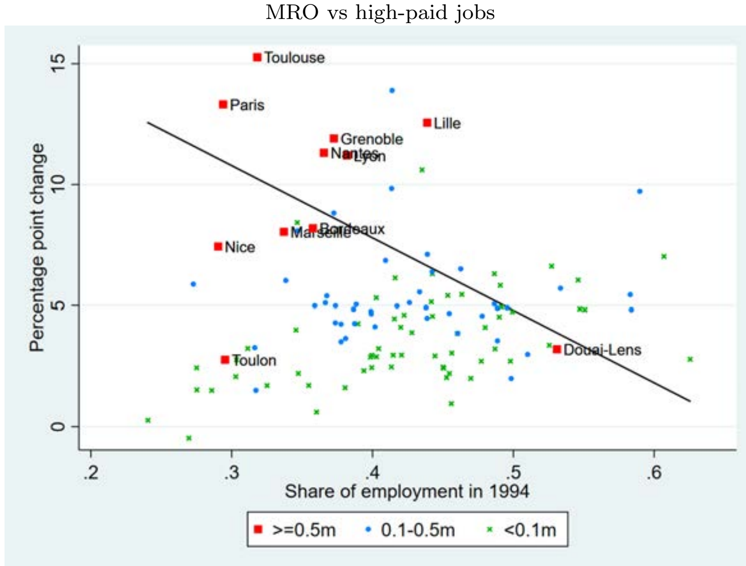
The initial exposure of largest cities to the most routine and offshorable occupations (MRO) in 1994 is on average lowest in largest cities. The exceptions are Douai-Lens and Lille in the industrial North.

The relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of these jobs over the period 1994-2015 (upper panel) is negative as predicted by Autor and Dorn (2013) (cf. Table F.39 on the robustness of the slope of the fitted line). Conditional on the initial exposure, however, the decline in the MRO occupations is highest in largest cities. Moreover, the average decline in the MRO jobs is not significantly different (cf. Table 3, column 4) between the largest and smallest cities in the sample (which are on average more exposed to those occupations – see Table F.15).

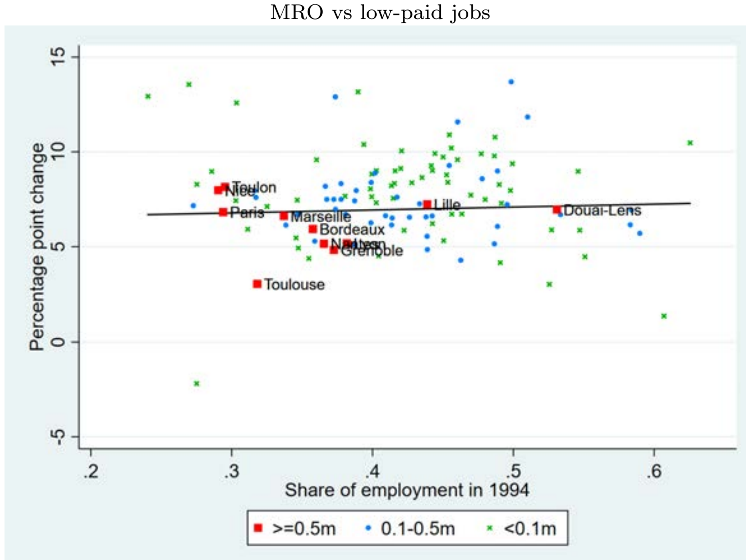
The lower-panel figure depicts a strongly positive population-weighted relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of OMP jobs (cf. Table F.39 on robustness of the slope). In *all but one* cities with initial exposure to MRO jobs above 0.5 the share of OMP jobs in total employment increased over the period 1994-2015. The decline of the OMP jobs in percentage points is on average stronger in large than in small cities (cf. Table 3, column 5), significant at 1% level.

Figure F.13: Exposure to MRO jobs and change in the employment share of high-paid and low-paid jobs, 1994-2015.

MRO vs high-paid jobs



MRO vs low-paid jobs



These figures show the percentage point change in employment shares of high-paid jobs (upper panel) and low paid jobs (lower panel) between 1994-2015 plotted against the share of MRO jobs (CS 48, 54, 62 and 67) in employment in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The line shows a linear, population-weighted (by 1990 population) fit of the relationship between employment changes and the initial MRO exposure. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.

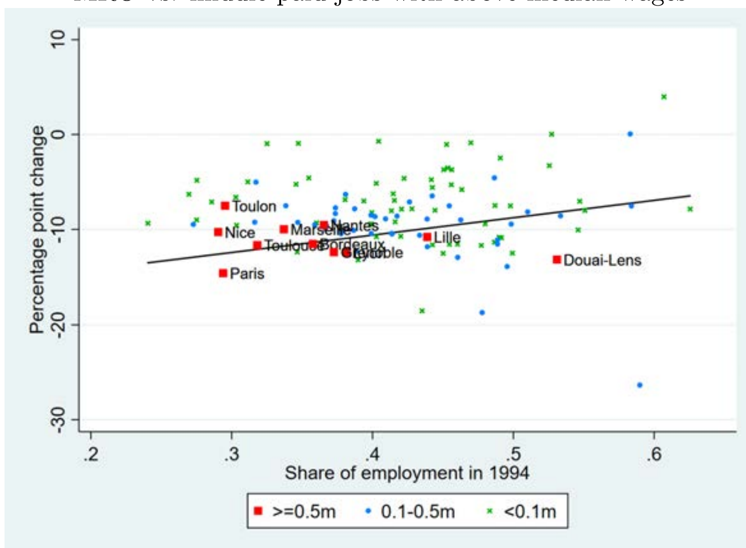
The upper panel figure documents a strongly negative population-weighted relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of high-paid jobs over the period 1994-2015 (cf. Table F.38 on the robustness of the slope of the fitted line). The creation of the high-paid occupations is on average much stronger in large than in small cities (cf. Table 3, first column) and this difference is significant at the 1% level.

The lower panel figure depicts no relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of low-paid jobs over the period 1994-2015. However, the increase in the employment shares of the low-paid occupations is on average significantly higher (at the 1% level) in small cities (cf. Table 3, column 3).

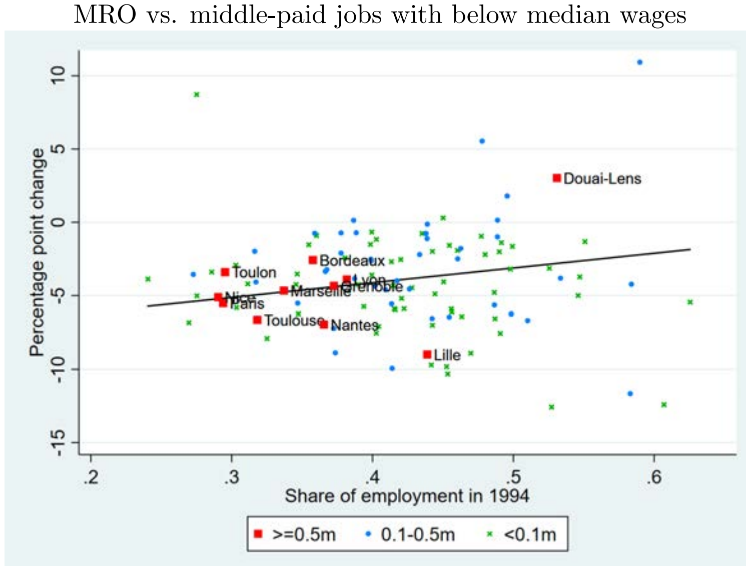
Such patterns are incompatible with the Autor and Dorn (2013) model that predicts that local labor markets with the highest local exposure to routine jobs would experience the strongest creation of high-paid and low-paid jobs.

Figure F.14: Exposure to MRO jobs categories and 1994-2015 changes in employment shares of middle-paid jobs above- and below-median in terms of wages in 1994.

MRO vs. middle-paid jobs with above median wages



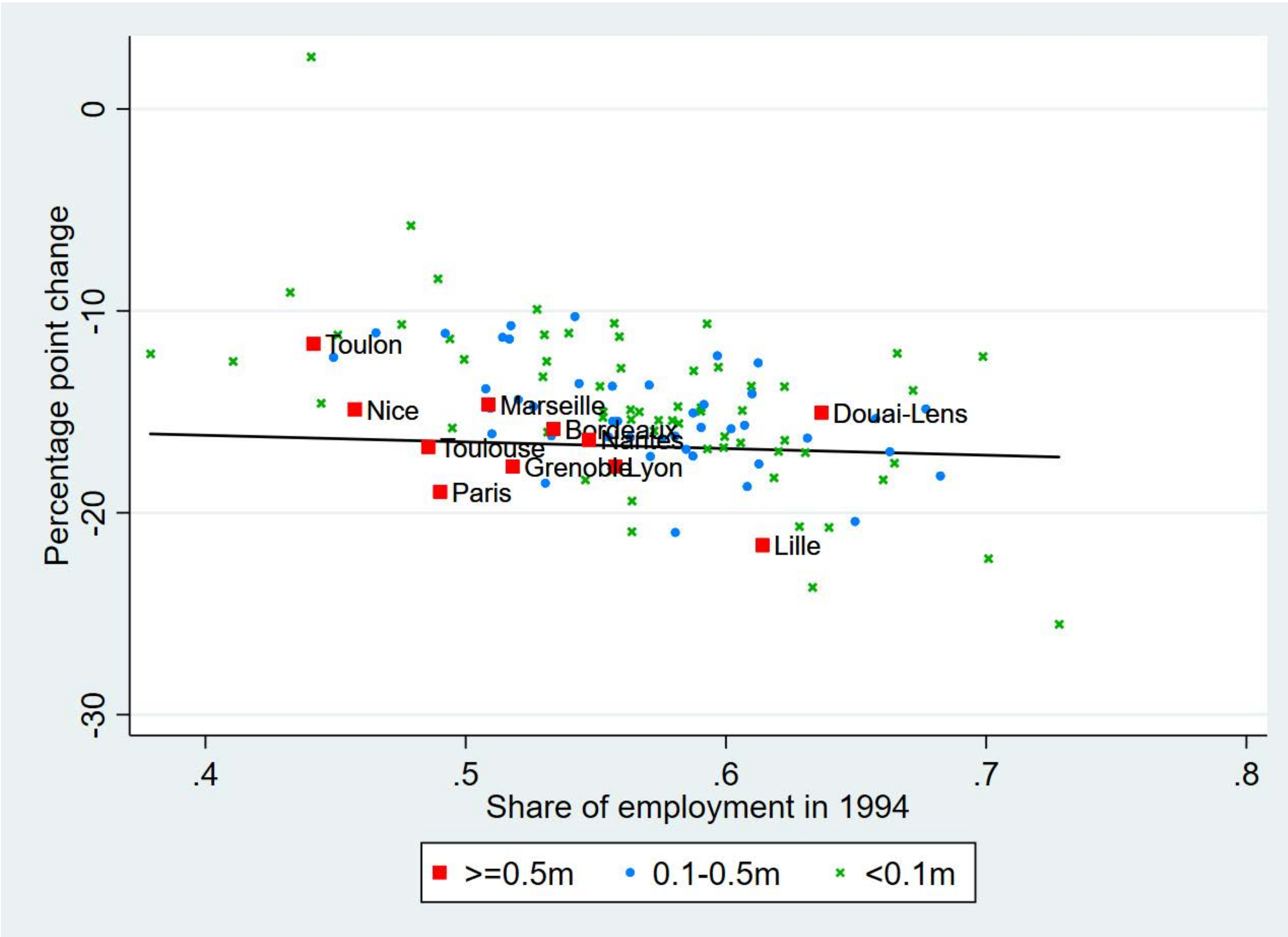
MRO vs. middle-paid jobs with below median wages



The figure shows the percentage point change in employment shares of middle-paid jobs with wages above (upper panel) and below (lower panel) the median in 1994 between 1994-2015 plotted against the share of MRO jobs (CS 48, 54, 62 and 67) in employment in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The lines show a linear, population-weighted (by 1990 population) fit of the relationship between employment changes and the initial MRO exposure. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.

There is a positive population-weighted relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share (over the period 1994-2015) of middle-paid jobs with the 1994 average wages both above and below the median (cf. Table F.39 on the robustness of the slope of the fitted line). The decline of the middle-paid occupations with average wages above the median is on average stronger in large than in small cities (cf. Table 3, next-to-last column) and this discrepancy is significant at 1% level. The decline of the middle-paid occupations with average wages below the median in percentage points is on average not statistically significantly different between large and small cities (cf. Table 3, last column).

Figure F.15: Exposure to 6 most offshorable jobs and their employment share change in cities, 1994-2015.



The figure shows the percentage point change in employment shares of 6 occupations with the highest OFF-GMS index (CS 46, 48, 54, 62, 65, 67) between 1994-2015 plotted against their share in employment in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The line shows a linear, population-weighted fit of the relationship between employment changes and the initial exposure to these most offshorable jobs. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.

The initial exposure of largest cities to the 6 most offshorable occupations in 1994 is on average lowest in largest cities. The relationship between the employment share of the 6 most offshorable occupations at the city level in 1994 and the change in the employment share of these jobs over the period 1994-2015 is weakly negative (cf. Table F.40 on the robustness of the slope of the fitted line). The average decline in these 6 most offshorable jobs, however, is statistically significantly stronger (at the 1% level) in the largest than in smallest cities in the sample (which tend to be initially more exposed to them).

Table F.37: Changes in the employment shares of middle-paid jobs between 1994-2015 and exposure to middle-paid occupations in 1994.

Employment share change of middle-paid jobs								
employment share of middle-paid jobs in 1994	0.41***	0.03	0.27***	-0.03	0.28***	-0.04	0.27***	0.01
	(0.08)	(0.07)	(0.10)	(0.07)	(0.10)	(0.08)	(0.10)	(0.05)
middle × employment share of middle-paid in 1994			-0.01*	-0.02**	-0.01*	-0.02**	-0.01	-0.01*
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of middle-paid in 1994			-0.06***	-0.06***	-0.06***	-0.06***	-0.05***	-0.05***
			(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
constant	-0.45***	-0.14***	-0.33***	-0.09*	-0.33***	-0.09	-0.33***	-0.13***
	(0.07)	(0.05)	(0.08)	(0.05)	(0.08)	(0.06)	(0.08)	(0.04)
R^2	0.47	0.00	0.59	0.16	0.59	0.17	0.63	0.18
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in middle-paid share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of the middle-paid jobs in total hours worked over the period 1994-2015 on their initial employment share at the individual city level in 1994. The first two columns report regression coefficients respectively with and without population (as of 1990) weighting. In the regressions reported in columns 3 and 4 an additional slope for medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.
The relationship between the exposure to the most offshorable jobs at the city level and their share change over the period 1994-2015 is positive (but not significant in the non-population-weighted regressions). The interpretation of these results together with the intercept is that, on average, cities that were *less* initially exposed to middle-skill jobs experienced their stronger destruction. However, conditioning on exposure, the destruction of middle-paid jobs was stronger in larger cities for all specifications.
The evidence in this Table points that there is no simple relationship between a higher initial exposure to middle-paid jobs at the city level and their subsequent destruction as a result of automation or offshoring shocks.

Table F.38: Employment share changes 1994-2015 and exposure to 4 most routine/offshorable occupations (MRO) in 1994, Part I.

Employment share change of high-paid jobs								
employment share of MRO jobs in 1994	-0.30***	0.01	-0.19**	0.04	-0.20**	0.03	-0.19**	0.03
	(0.10)	(0.04)	(0.09)	(0.04)	(0.09)	(0.04)	(0.10)	(0.03)
middle × employment share of MRO jobs in 1994			0.04***	0.04***	0.04***	0.03***	0.04***	0.04***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of MRO jobs in 1994			0.16***	0.16***	0.16***	0.15***	0.16***	0.14***
			(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
constant	0.20***	0.04**	0.12***	0.02	0.13***	0.02	0.12***	0.02
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.01)
R^2	0.31	0.00	0.56	0.34	0.57	0.33	0.56	0.31
Employment share change of low-paid jobs								
employment share of MRO jobs in 1994	0.02	-0.01	0.00	-0.02	0.00	-0.01	-0.01	-0.05
	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.03)
middle × employment share of MRO jobs in 1994			-0.02**	-0.01	-0.02*	-0.01	-0.02**	-0.02*
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of MRO jobs in 1994			-0.04***	-0.05***	-0.04***	-0.04***	-0.05***	-0.06***
			(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
constant	0.06***	0.08***	0.08***	0.09***	0.08***	0.08***	0.08***	0.10***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
R^2	0.00	0.00	0.11	0.05	0.10	0.04	0.13	0.10
Employment share change of middle-paid jobs								
employment share of MRO jobs in 1994	0.28***	0.00	0.19**	-0.02	0.20**	-0.02	0.19**	0.00
	(0.10)	(0.04)	(0.09)	(0.04)	(0.09)	(0.05)	(0.09)	(0.03)
middle × employment share of MRO jobs in 1994			-0.02	-0.03**	-0.02	-0.03**	-0.01	-0.02*
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of MRO jobs in 1994			-0.12***	-0.11***	-0.12***	-0.11***	-0.12***	-0.10***
			(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
constant	-0.26***	-0.12***	-0.20***	-0.11***	-0.20***	-0.11***	-0.20***	-0.12***
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)
R^2	0.33	0.00	0.51	0.15	0.51	0.16	0.54	0.16
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in MRO share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: MRO jobs are the 4 most routine or offshorable occupations with the highest RTI or OFF-GMS indexes (CS 48, 54, 62 and 67). Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of high-, low- and middle-paid occupational categories in total hours worked over the period 1994-2015 on the initial employment share of MRO jobs in 1994 at the individual city level. The first two columns report regression coefficients respectively with and without population weighting. In the regressions reported in columns 3 and 4 an additional slope for the medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.
Regressions show that the relationship between initial exposure to the most routine and offshorable jobs at the city level in 1994 is not a good predictor of increases of the shares of high- or low-paid or declines of middle-paid jobs. In population-weighted regressions, larger initial exposure to the most routine and offshorable jobs leads to a *lower* increase of high-paid occupations' share, no tendency for low-paid jobs, and a *lower* destruction of middle-paid jobs. At the same time, non-population weighted regressions (column 2) do not show any relationship between this exposure and the changes in the studied occupational shares. Allowing for a separate slope for large cities reveals that labor market responses in these metropolitan areas are always different from those of small ones.

Table F.39: Employment share changes 1994-2015 and exposure to 4 most routine/offshorable occupations (MRO) in 1994, Part II.

Δ *Emp. sh., MRO jobs*								
emp. sh., MRO jobs in 1994	-0.24*** (0.05)	-0.39*** (0.04)	-0.29*** (0.05)	-0.40*** (0.04)	-0.28*** (0.05)	-0.38*** (0.04)	-0.28*** (0.05)	-0.36*** (0.04)
middle × emp. sh., MRO jobs in 1994			-0.01 (0.01)	-0.01 (0.01)	-0.02 (0.01)	-0.02 (0.01)	-0.01 (0.01)	-0.02 (0.01)
large × emp. sh., MRO jobs in 1994			-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.06*** (0.02)
constant	-0.02 (0.02)	0.05*** (0.02)	0.01 (0.02)	0.06*** (0.02)	0.01 (0.02)	0.05*** (0.02)	0.01 (0.02)	0.04*** (0.01)
R^2	0.43	0.55	0.53	0.58	0.52	0.55	0.53	0.54
Δ *Emp. sh., OMP jobs*								
emp. sh., MRO jobs in 1994	0.52*** (0.05)	0.38*** (0.03)	0.48*** (0.05)	0.38*** (0.03)	0.48*** (0.05)	0.36*** (0.03)	0.48*** (0.05)	0.35*** (0.02)
middle × emp. sh., MRO jobs in 1994			-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
large × emp. sh., MRO jobs in 1994			-0.05*** (0.02)	-0.04** (0.02)	-0.05*** (0.02)	-0.04* (0.02)	-0.05*** (0.02)	-0.04** (0.02)
constant	-0.24*** (0.02)	-0.17*** (0.01)	-0.21*** (0.02)	-0.17*** (0.01)	-0.21*** (0.02)	-0.16*** (0.01)	-0.21*** (0.02)	-0.16*** (0.01)
R^2	0.84	0.71	0.87	0.73	0.86	0.71	0.87	0.73
Δ *Emp. sh., middle-paid > median wage*								
emp. sh., MRO jobs in 1994	0.18** (0.09)	0.01 (0.07)	0.13* (0.07)	0 (0.06)	0.13*(0.07)	0 (0.07)	0.14** (0.06)	0.01 (0.04)
middle × emp. sh., MRO jobs in 1994			-0.05*** (0.02)	-0.06*** (0.02)	-0.05*** (0.02)	-0.06*** (0.02)	-0.04** (0.02)	-0.04** (0.02)
large × emp. sh., MRO jobs in 1994			-0.13*** (0.02)	-0.11*** (0.02)	-0.13*** (0.02)	-0.11*** (0.02)	-0.12*** (0.02)	-0.10*** (0.02)
constant	-0.18*** (0.04)	-0.09*** (0.03)	-0.13*** (0.03)	-0.07*** (0.03)	-0.13*** (0.03)	-0.07** (0.03)	-0.14*** (0.03)	-0.08*** (0.02)
R^2	0.17	0.00	0.36	0.14	0.36	0.15	0.41	0.13
Δ *Emp. sh., middle-paid < median wage*								
emp. sh., MRO jobs in 1994	0.10** (0.05)	-0.01 (0.06)	0.06 (0.07)	-0.02 (0.06)	0.06 (0.08)	-0.02 (0.07)	0.05 (0.07)	-0.04 (0.05)
middle × emp. sh., MRO jobs in 1994			0.03** (0.02)	0.03* (0.02)	0.03* (0.02)	0.03* (0.02)	0.02 (0.02)	0.02 (0.02)
large × emp. sh., MRO jobs in 1994			0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)
constant	-0.08*** (0.02)	-0.04 (0.02)	-0.07** (0.03)	-0.04 (0.03)	-0.07** (0.03)	-0.04 (0.03)	-0.07** (0.03)	-0.03 (0.02)
R^2	0.09	0.00	0.12	0.04	0.12	0.04	0.10	0.02
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in MRO share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: MRO jobs are the 4 most routine or offshorable occupations with the highest RTI or OFF-GMS indexes (CS 48, 54, 62 and 67). Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of different middle-paid occupational categories in total hours worked over the period 1994-2015 on the initial employment share of MRO jobs at the individual city level in 1994. The first two columns report regression coefficients respectively with and without population weighting. In the regressions reported in columns 3 and 4 an additional slope for medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.
It can be observed in the top panel that initial exposure to MRO jobs is negatively correlated with the change in their employment shares. The interaction of a dummy for large cities with exposure is robustly negatively different from zero. Routine and offshorable jobs in large cities are destroyed at a higher rate than in small cities.
There is a strong and significant positive relation between the exposure to MRO group of jobs and the employment share of OMP jobs. We see that the largest decline in OMP occupations occurs in cities *least* exposed to MRO jobs. The simple linear fit indicates that in cities with largest initial MRO job exposure (above 0.5), the employment share of the OMP jobs increases over the period 1994-2015 (cf. Figure F.12).
We cannot observe any robustly significant relationship between exposure to MRO jobs and the changes in the employment share of middle-paid jobs with wages below the median average wage in 1994, regardless of the city size.

Table F.40: Employment share changes of the most offshorable jobs over the period 1994-2015 and exposure to 6 most offshorable occupations (OFF6) in 1994.

Employment share change of OFF6 jobs								
employment share of OFF6 jobs in 1994	-0.03	-0.29***	-0.17**	-0.31***	-0.15*	-0.30***	-0.14*	-0.25***
	(0.12)	(0.06)	(0.08)	(0.06)	(0.08)	(0.06)	(0.08)	(0.05)
middle × employment share of OFF6 jobs in 1994			-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of OFF6 jobs in 1994			-0.08***	-0.06***	-0.08***	-0.06***	-0.07***	-0.05***
			(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
constant	-0.15**	0.02	-0.05	0.03	-0.06	0.03	-0.07	0
	(0.07)	(0.03)	(0.05)	(0.03)	(0.05)	(0.04)	(0.05)	(0.03)
R^2	0.00	0.28	0.33	0.34	0.33	0.32	0.34	0.30
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in OFF6 share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: The category of 6 most offshorable occupations (OFF6) are those with highest OFF-GMS index (CS 46, 48, 54, 62, 65, 67). Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015. Weighting by population as of 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of 6 most offshorable occupations in total hours worked over the period 1994-2015 on their initial employment share at the individual city level in 1994. The first two columns report regression coefficients respectively with and without population weighting. In the regressions reported in columns 3 and 4 an additional slope for medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.
The relationship between the exposure to the most offshorable jobs at the city level and their share change over the period 1994-2015 is negative (but not significant in the population weighted regressions). However, this relationship is revealed to be significantly (at the 1% level) negative for largest cities once we allow for a separate slope. The 6 most offshorable jobs in large cities are destroyed at a higher rate than in small cities conditioning on the same level of exposure.

F.6 Pre-1994 Labor market developments

In this Appendix we document the relevant changes in the French labor markets before the coverage of the detailed DADS data (starting in 1994).

Some of the differences in labor market developments for individual categories across cities may be due to several factors that should be mentioned but cannot be fully addressed empirically given the limitations of data at our disposal.

Labor market polarization in France might have begun earlier than 1994: although the modern ICT were not widely used pre-1994 there was a significant advance in automation in manufacturing through CAD/CAM systems, early adoption of basic computer text editors or spreadsheets. There was also an increase in offshoring possibilities for French companies with such developments as the Spanish or Portuguese accession to the EEC in 1986 or the opening up of Eastern European countries in 1989. The strength of the automation or offshoring shocks is unclear, however, and the most offshorable occupations (CS 48, 62 and 67) related to manufacturing might have been affected the earliest. It is therefore instructive to detail some of the pre-1994 developments.

There are two data sources that allow to track occupations at the 2-digit PCS level back to 1982 when the PCS classification was introduced: the French Labor Market Survey (yearly data) and the Census (1982, 1990 and 1999). The publicly available Labor Market Survey gives data at the department but not at the commune level, hence it is impossible to precisely characterize city-level labor markets. The Census, on the other hand, gives the commune location of respondents but does not give data about hours worked or wages. We use the Census as we are interested in the shares of employment in cities, but in contrast to data presented in main text the patterns will refer to shares of people employed and not actual hours worked. We use the publicly available individual data for the 1982, 1990 and 1999 censuses (covering 1/4th for 1982 and 1990, and 1/20th for 1999 of the entire population respectively).

The 1982-1999 counterparts to Table 2 using Census data are in Table F.41.

Table F.41: Share of 4 highest-paid occupations per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.18	0.12	0.10	0.10	0.09	0.09
1990	0.24	0.16	0.13	0.12	0.11	0.11
1999	0.26	0.17	0.14	0.13	0.12	0.11
change 1982-1990	0.06	0.04	0.03	0.03	0.02	0.02
change 1990-1999	0.02	0.01	0.01	0.00	0.00	0.00
change 1982-1999	0.08	0.05	0.04	0.03	0.03	0.02

The conclusions from this exercise are as follows. First of all, exposure to the most routine and offshorable jobs is much higher in 1982 for large cities above 0.5m inhabitants than in 1994 in the DADS data, and the discrepancies in terms of shares of high- middle- and MRO jobs across city sizes are lower. Employment shares of the MRO and, more generally, middle-paid jobs indeed decline faster

Table F.42: Share of 10 middle-paid occupations per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.66	0.71	0.74	0.73	0.74	0.73
1990	0.61	0.68	0.69	0.70	0.71	0.71
1999	0.56	0.64	0.65	0.66	0.67	0.68
change 1982-1990	-0.05	-0.03	-0.04	-0.04	-0.04	-0.02
change 1990-1999	-0.05	-0.04	-0.04	-0.03	-0.04	-0.03
change 1982-1999	-0.10	-0.07	-0.08	-0.07	-0.07	-0.06

Table F.43: Share of 4 lowest-paid occupations per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.15	0.17	0.16	0.17	0.17	0.18
1990	0.15	0.16	0.18	0.18	0.18	0.18
1999	0.18	0.19	0.21	0.21	0.21	0.22
change 1982-1990	0.00	0.00	0.01	0.01	0.01	0.01
change 1990-1999	0.02	0.02	0.03	0.03	0.03	0.03
change 1982-1999	0.02	0.02	0.04	0.04	0.05	0.04

in larger cities whether in 1982-1990 or in the entire 1982-1999 period. The labor market polarization across cities manifests itself as our theory predicts: high-paid jobs' shares increase most in largest cities, as found in the exhaustive DADS data for 1994-2015. Low-paid jobs do not increase at all in largest cities in 1982-1990 and increase less in terms of percentage points over 1990-1999 and the entire 1982-1999 period.

Similar patterns obtain for 1982-1994 using the Labor Market Survey data while classifying departments by largest city.

For individual occupational categories, the routine/offshorable job categories whose employment declines most in the studied years 1994-2015 in the DADS data in large cities are in particular PCS 46 and 54 (mid-level professionals and office workers respectively), whereas it is 62 and 67 (skilled and unskilled industrial workers respectively) for small cities (cf. the patterns in Figure F.8).[62] A part of the answer of such a differential evolution may lay in the fact that large cities had different shares of these jobs at the beginning of the 1990s (see Table F.45) than small cities, and such a discrepancy existed already in 1982. In particular, the share of mid-level professionals and office workers in employment was higher than that of industrial workers in 1982 in the largest cities above .75m inhabitants while the opposite is true for smaller cities. Therefore, the additional adjustment in terms of percentage points we observe in these blue-collar categories over the period 1994-2015 may be less pronounced as well. This feature of data may be explained by different deindustrialization across time and geographies as

[62]The PCS 46 category contains heterogeneous professions that were differentially impacted by automation/offshoring. For example, occupations such as drafters, secretaries, photographers, sales in insurance, real estate, finance or advertising included in this category have RTI scores above 2; some of them are also very offshorable.

Table F.44: Share of the 4 most routine and offshorable occupations (CS 48, 54, 62 and 67) per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.36	0.39	0.43	0.42	0.43	0.43
1990	0.28	0.33	0.34	0.35	0.37	0.38
1999	0.22	0.27	0.29	0.30	0.31	0.33
change 1982-1990	-0.07	-0.05	-0.08	-0.06	-0.06	-0.05
change 1990-1999	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06
change 1982-1999	-0.13	-0.12	-0.14	-0.12	-0.13	-0.11

shown in Table F.46 where the share of industry employment at Census years is given for the period 1968-2015. Already over the period 1968-1982 large cities experienced faster deindustrialization than small cities. Reports from research bodies as the INSEE or DATAR (Délégation interministérielle à l'aménagement du territoire et à l'attractivité régionale) indicate the following. Internal offshoring of manufacturing tasks within France might have played a part due to the reduction in internal transport costs (both because of highway and railway construction), environmental regulations to keep polluting industries out of high density areas or a deliberate government policy to decentralize economic activity across France (e.g. moving public engineering schools outside Paris), and hence not related to automation and offshoring shocks. For Île-de-France, deindustrialization was largely due to the reorganization of the automobile (that moved out of large cities) and defense industries (idem, with aerospace moving to Toulouse in particular).

To an unknown extent firm reorganization and shifting tasks outside the boundaries of firms (e.g. legal services, general and administrative or cleaning premises) that cannot be precisely measured was responsible for the fall in manufacturing value added overall. This, together with moving tasks within multi-establishment firms might have caused some of the tasks to be offshored within France from large to smaller cities.

One explanation for the decline in the share of back-office or support jobs like office workers (CS 54) or technicians (CS 47) in Paris with their coincident expansion in small cities within the later 1994-2015 period can be internal offshoring from large to small cities permitted by the Internet and communication technologies. Such tendencies are consistent with our model (all goods, including intermediates, are traded) though we do not model nor cannot verify empirically supply chain developments that are internal or external to firms.

Table F.45: Employment share of selected middle-paid occupations per metropolitan area size, Census data 1982-1999.

Agglo. Size		Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
mid-level professionals (CS 46) and office workers (CS 54)	1982	0.27	0.23	0.19	0.20	0.19	0.19
	1990	0.26	0.23	0.21	0.22	0.21	0.21
	1999	0.26	0.25	0.21	0.22	0.21	0.21
skilled (CS 62) and unskilled industrial workers (CS 67)	1982	0.14	0.19	0.25	0.24	0.25	0.25
	1990	0.10	0.16	0.18	0.19	0.20	0.22
	1999	0.07	0.11	0.14	0.15	0.16	0.18

Notes: Data from the publicly available individual data for the 1982, 1990 and 1999 censuses (covering 1/4th for 1982 and 1990, and 1/20th for 1999 of the entire population respectively). Aggregation to the city-level (based on unites urbaines as of 2015). Includes individuals between 25-64 of age.

Table F.46: Employment share in industry per metropolitan area size, Census data 1968-2015.

share of industry	Paris	>0.75m	0.5-0.75m	0.2-0.5m	0.1-0.2m	0.05-0.1m
1968	0.32	0.30	0.39	0.32	0.34	0.29
1975	0.29	0.28	0.34	0.30	0.34	0.30
1982	0.24	0.23	0.29	0.26	0.30	0.27
1990	0.19	0.20	0.22	0.22	0.24	0.24
1999	0.13	0.15	0.17	0.17	0.19	0.19
2015	0.07	0.11	0.11	0.11	0.13	0.13
pp change 68-82	-0.08	-0.07	-0.10	-0.06	-0.04	-0.02
pp change 68-90	-0.13	-0.11	-0.16	-0.10	-0.10	-0.05

Notes: Raw, exhaustive data from Censuses were prepared by the INSEE for each commune. Aggregation to the city-level (based on unites urbaines as of 2015) done by authors. Includes individuals between 25-54 of age.

F.7 Public sector employment 1990-2015.

DADS-Postes data does not contain public sector employee data prior to the end of the 2000s. We can, however, assess the evolution of public sector jobs (number of positions but not hours) using the harmonized Census SAPHIR. In this exercise, we cannot restrict the private sector only to incorporated firms (catégorie juridique "5" in INSEE's classification) as in the main sample.

We restrict attention to all 2-digit CS categories between 23 and 68 apart from CS 44 (clergy) among employed in cities considered in our sample within the age range 25-64 for 1990 and 2015. We classify CS 33 (category A: public administration managers) and 34 (higher education professors, scientists in public employment) as high-paid, CS 42 (teachers) and CS 45 (category B: intermediate public administration workers) as middle-paid, CS 52 (category C and D employees of public administration) and the public employment portion of CS 53 (policemen, military) as low-paid.

Public sector employment grew nationally by 0.5 pp in total from 23.3% to 23.8% of total employment between 1990 and 2015. In 1990, small cities with populations below 0.1m were relatively more abundant in public sector jobs with a 24.8% of total versus 21.3% in Paris. The share of high-paid public sector jobs increased in the period 1990-2015 by 1.2 pp (to 6.8% of total jobs).

The breakdown of the change within cities is given in Table F.47. There was growth of public sector employment in cities with population below 0.5m (between 1.9 pp in cities between 0.2-0.5m and 2.7 pp in cities between 0.1-0.2m). High-paid public sector jobs increased in cities of all sizes (from 0.7 pp in cities with a population between 0.5m-0.75m to 1.7 pp in cities between 0.1-0.2m). Low-paid jobs increased in smallest cities by 1.2 pp and decreased by 2.1 pp in Paris, with intermediate values for other cities. Changes in the shares of middle-paid public sector jobs are negligible.

Based on this data, we see that there was no comparable shock to middle-paid jobs in the public sector as in the private sector discussed in the paper. The public sector does not seem to play an important role in the evolution of jobs nationally, within (or across, not shown) cities, and does not exhibit any spatially interesting patterns. In particular, the decrease in middle-paid jobs observed in the private sector was not absorbed by middle-paid or low-paid public sector jobs, either nationally or at the city level. Moreover, the changes in low-paid public sector jobs (decreases in larger cities and increases in smaller cities) do not "compensate" the lower growth of low-paid jobs in larger cities and lead to a higher growth of low-paid occupations there in the aggregate but actually exacerbate the differential growth in these types of jobs across cities overall.

Table F.47: Change in the share of private and public sector jobs within city groups and in the aggregate 1990-2015 from Census data.

job type	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M	All cities
private sector	0.007	-0.001	0.002	-0.019	-0.027	-0.025	-0.005
public sector high-paid	0.011	0.011	0.007	0.016	0.017	0.009	0.012
public sector middle-paid	0.003	0.000	-0.003	0.002	0.005	0.004	0.002
public sector low-paid	-0.021	-0.010	-0.006	0.001	0.005	0.012	-0.009

Notes: Aggregation to the city-level based on metropolitan areas as of 2015. Includes employed individuals between 25-64 of age for all 2-digit CS categories 23-68 with the exception of CS category 44 (clergy).