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 August 2023

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 August 2023
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

A 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 analytic 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 the correct approach depends on the economic structure connecting shocks and jobs, and perhaps as well on the specific lens through which one examines the data. In our data, if we focus *only* on the middle-paid jobs deemed *most* exposed to the posited shocks, then we find 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. 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 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 (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 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.

The term "great divergence" was first applied to cities by Moretti (2012) and has been closely linked to models of skill-biased technical change.[1] 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

[1] 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.

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). 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 work is also related 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).

Our work here draws on various facets of these literatures. In order to link up to the labor market polarization literature, we go beyond a skilled-unskilled labor dichotomy. We focus on a setting with three key tasks and a continuum of labor, 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. Since we are focused on long run equilibria rather than transitional dynamics, we allow agents with all skills to choose a location and a sector without frictions. 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. Our framework seeks to unify the literatures on labor market polarization and the great urban divergence, so relies on the routinization and offshoring shocks in the former rather than the skill-biased technical change shocks common in the latter. In doing so, our paper focuses on how the polarization shocks translate into different evolutions in large and small cities. Relative to the prior literature on heterogeneous labor in spatial equilibrium, we simplify in some dimensions and enrich in others. Where the prior 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 not only of 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.

The paper is organized as follows. Section 2 describes our data. In Section 3 we discuss the empirical facts and results. Sections 3.1-3.2 consider labor market polarization in the aggregate and both within and across cities. Section 3.3 examines in greater detail the drivers of the decline of middle-paid jobs in large and small cities while Section 3.4 explores the resulting great urban divergence. In Section 3.5 we discuss whether existing theories can account for our facts. Then, in Section 4 we develop a model of labor market polarization in spatial equilibrium that is consistent with the presented empirical patterns. Section 4.4 exhibits further model results (e.g. on skewed polarization) through data calibration. Section 5 concludes. Rich additional material is provided in the Online Appendix.

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.[2] 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.[3] 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 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

[2] Detailed discussion and description of the constructed data set appears in the Online Appendix, F.1.

[3] 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.

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.[4] We exclude artisans (CS 21 and 22; many are not incorporated), 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) which in general are not available in data prior to the late 2000s.[5]

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.[6] 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 provided 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 used methodology). The exact RTI and OFF-GMS index values for each category are given in the Appendix Table F.2.

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

[4]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.

[5]Data for farmers (CS 10) or public sector workers are not available, while data for non-incorporated workers such as CS 21 and CS 22 are not well reported in the DADS data before the end of the 2000s.

[6]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 1994	2015	Average City Wage (in 2015 euros) 1994	2015	Routine (2015 vs. 1994)	Offshorable	Relative Wage change ranking ranking
	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).

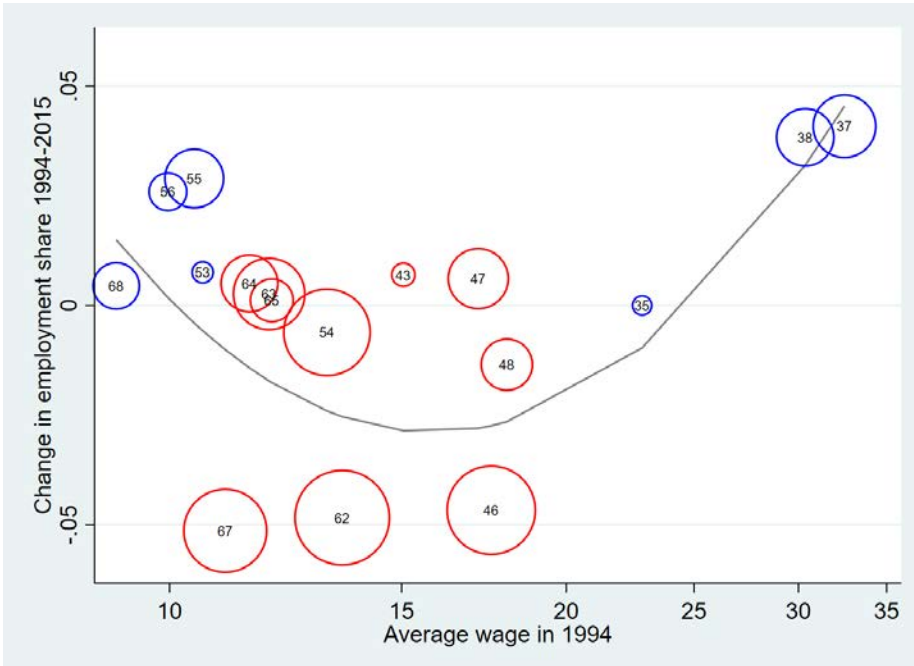
of high- and low-paid jobs. Following Autor and Dorn (2013) and Goos et al. (2014), we take the mean wage by occupation in the initial year of our data, here 1994, as a primitive ordering of our set of 18 occupations.[7] Forming the wage-occupation groups then requires criteria for the division into high-, medium-, and low-wage groups that respects this ordering.

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

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

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. 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.

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.[8] First, from above, all occupations with one percentage point or more decrease in jobs' share are

[8]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.

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.[9]

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 will divide middle-paid jobs is simply by one tier of higher skill and pay plus a second tier of lower skill and pay, as would be consistent

[9]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 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 4.4. In short, however we look at this, wage polarization emerges clearly. The estimation is discussed in more detail in Appendix F.2.1.

with continuum of skill approaches such as Cortes (2016) and our own multi-city approach. The simplest version of this divides the 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 is to connect with some of the prior work. While the original concern with labor market polarization was the 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 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 ranked in terms of offshorability. It's worth keeping in mind that 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 for a theoretical account).

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.[10]

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 68% of hours worked in the mainland in the non-farm private sector.[11] In 1994 and 2015 respectively, firms active in these cities for which we have data account for 396,637 (out of 596,368) and 633,851 (out of 998,467) firms. Noting exclusions as above (on worker age, types of jobs, cities), we retain a sample that accounts for 65% of total wages paid and 58% of hours

[10] They are shown in Online Appendix Figure F.1 and population data by city category in Table F.1. The characteristics of the final sample are given in Table F.5.

[11] Figures given our restrictions on data (e.g. age). In robustness checks 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.

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.[12]

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 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.

[12]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.

Table 1 provides detail on this evolution for all jobs in mainland France[13]. 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.[14] 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 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.[15]

In view of the above, we observe:

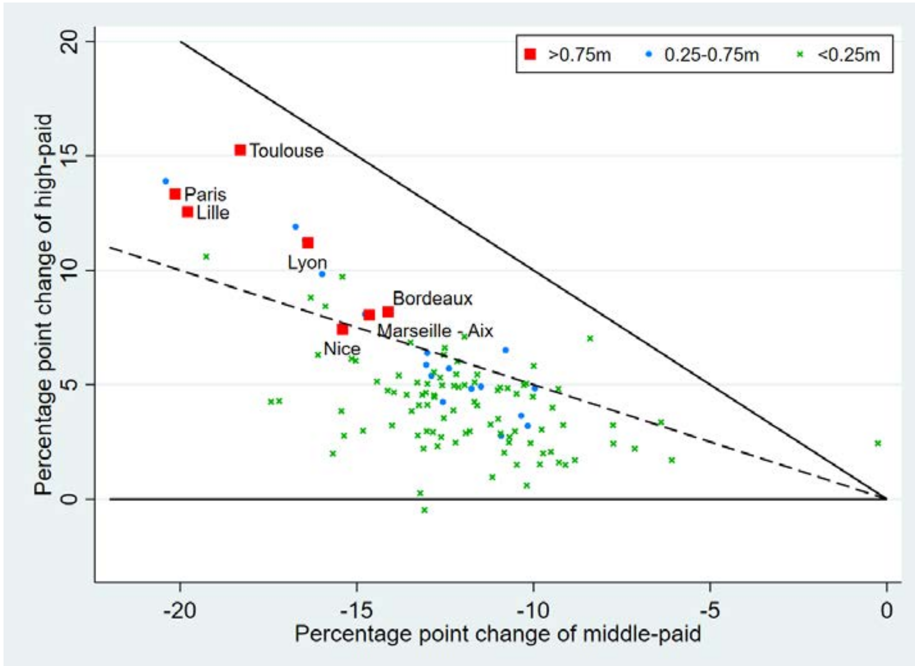
Fact 1 (Universal polarization)**.** *Over the period 1994-2015, French labor markets became more polarized in the aggregate and in nearly every individual city.*

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

[14] 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.

[15] 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.

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.

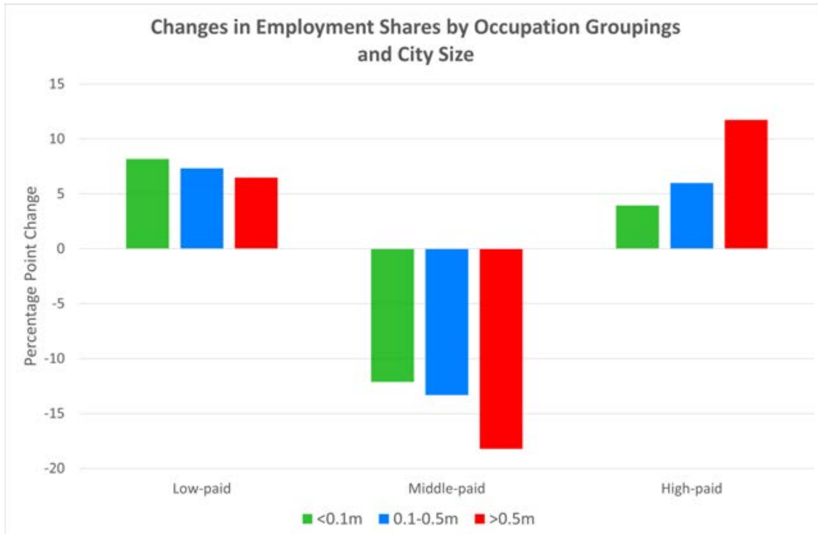
3.2 Middle-paid job loss and initial exposure

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 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.

Table 2 reveals a strong feature of the French data that may seem unexpected given prior

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.

literature. That table shows the evolution of the share of middle-paid jobs across six city sizes.[16] 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, 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.

These points are underscored jointly in Tables F.15 and 3.[17] 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%).[18] This leads to the following fact:

[16] One can observe a similar pattern using 2-digit CS categories as shown by Figure F.9 in the Online Appendix.

[17] 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.

[18] Appendix Table F.24 shows rank correlations between city size and middle-paid job loss are also statistically

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.

Fact 2 (Middle-paid job loss and initial exposure)**.** *Labor market polarization led to greater 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. Initial 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 Appendix F.5.1).

However we think it is important to return to the broader question of middle-paid job loss

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.

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.

that initially motivated the literature. 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 original, broader concern of middle-paid job loss.

3.3 Skewed middle-paid job loss

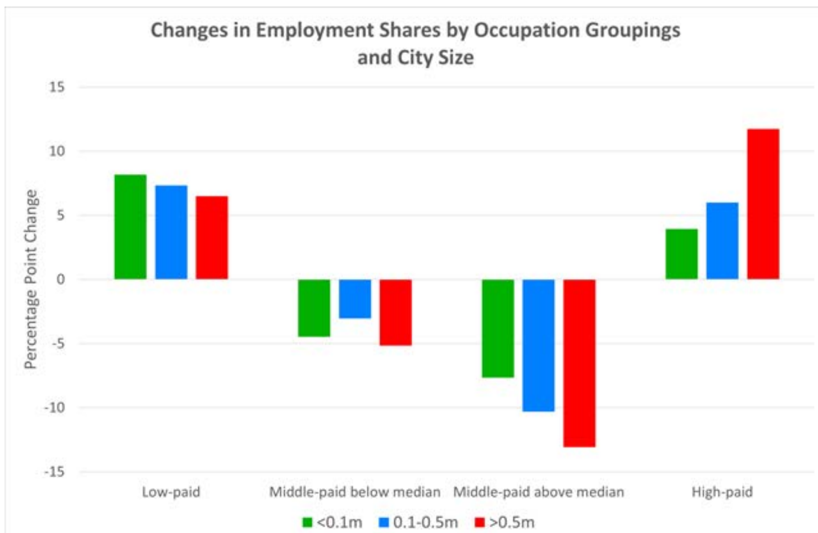
We investigate now the contrasting experience of large and small cities regarding middle-paid job loss. We split middle-paid occupations between those below versus above the median middle-paid wage. This will help us to understand if the contrasting experience of large and small cities arises due to job loss skewed toward one or the other margin in each.[19]

The results appear in Figure 4. This is identical to Figure 3 except for the new division of

[19] 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 of 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 middle-paid jobs (Table F.37). Detailed discussions are in the Online Appendices F.5.1-F.5.2; see also Online Appendix Figure F.11 and Tables F.38-F.39.

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 4 percentage point growth of high-paid jobs, even as there is an 8 percentage point decline in these upper-tier middle-paid jobs.[20] The contrasts by city size are much more modest for the lower-tier of middle-paid jobs.

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).

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.

[20]In Figure F.8, 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).

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.

3.4 The Great Urban Divergence

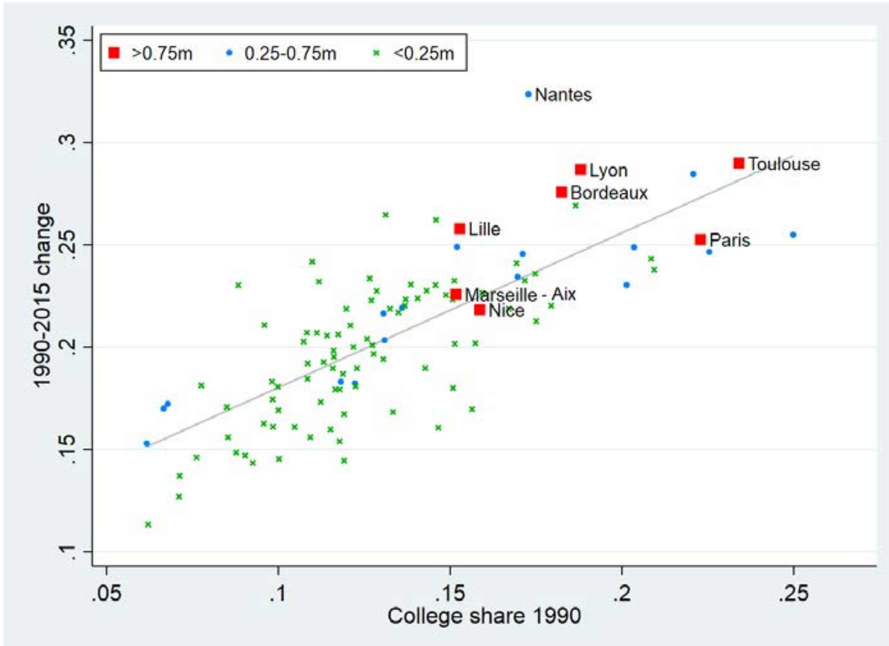
We begin our documentation of the Great Urban Divergence by examining skill as educational attainment. Following the original formulation of Moretti (2004), Figure 5 plots a city's change in the share of college graduates against the initial college share, here for 117 French cities. Our data cover a 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 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, 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

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%.

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.[21]

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.*

[21] 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.

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-skill jobs) rather than, as in the data, where these jobs were initially scarce per Fact 2.[22] 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.

[22] 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.

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.[23]

4 Theory

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).[24] 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.

[23]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. (2020) 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.

[24]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.

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 – that we report 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 share of 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 $c \in \{1, 2\}$.[25] 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 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.[26]

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 skills that we denote by ω, where ω is distributed on $[\underline{\omega}, \overline{\omega}]$ with a pdf $n(.)$.

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.

[25]We extend our framework to N cities in the Online Appendix B.2.

[26]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).

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 $\zeta < \theta$.[27] 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 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 $p(\sigma)$.

Each individual with skill ω, living in city c and in a location τ has a productivity:

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

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).

[27] 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.

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:*

$$\frac{A(h,1)}{A(h,2)} > \frac{A(m,1)}{A(m,2)} > \frac{A(l,1)}{A(l,2)}.$$

Assumptions 2 and 3 are consistent with the evidence that we gather in Appendix F.3.[28]

Capital good/offshoring intermediate good. The intermediate good z is produced by transforming final goods using the following technology:

$$Z = \frac{1}{\xi} q, \tag{3}$$

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

[28] 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 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 4.4. 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 appendix B.1, 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.

ω, location decisions (c, τ) and intermediate good sector σ is:

$$A(\sigma, c)H(\omega, \sigma)T(\tau)p(\sigma) - r(c, \tau) \tag{4}$$

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 (4) with respect to c, τ and σ.

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

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

$$A(h, c)H(\omega(h, c), h)p(h) = A(m, c)H(\omega(h, c), m)p(m) \tag{6}$$

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. 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:

$$\max_{\sigma,\tau} A(\sigma, 1)H(\omega, \sigma)T(\tau)p(\sigma) - r(1, \tau) \geq \max_{\sigma',\tau'} A(\sigma', 2)H(\omega, \sigma')T(\tau')p(\sigma') - r(2, \tau'). \tag{7}$$

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 $\tau \geq \bar{\tau}(c)$, so that $\bar{\tau}(c)$ defines the limits of the occupied area of city c.

Proof. See Appendix A.2 □

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 $r(c, \tau)$.[29]

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:

$$\begin{aligned} &A(M(\omega(\tau), 1), 1)H\left(\omega(\tau), M(\omega(\tau), 1)\right)T(\tau)p\left(M(\omega(\tau), 1)\right) = \cdots \\ &\cdots A(M(\omega(\tau), 2), 2)H\left(\omega(\tau), M(\omega(\tau), 2)\right)T(\tau')p\left(M(\omega(\tau), 2)\right). \end{aligned} \tag{8}$$

[29] See Appendix A.3 for more details on the intermediary steps on the results in this paragraph.

For $\omega \in [\omega, \bar{\omega}(2)]$, households are indifferent between a less desirable location in the more produ

City 2

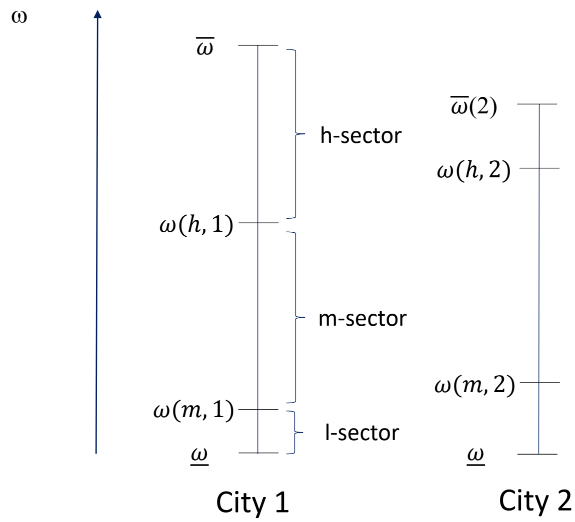


Figure 6: Skills, sectors and cities in equilibrium.

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 $\omega \in [\omega(\bar{2}), \bar{\omega}]$ 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.

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 more skilled than City 2.*

Proof. See Appendix A.4 and Appendix B.3 □

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.[31]

[30]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.

[31]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.

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:

Lemma 1. *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.*

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

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 4 (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.5. □

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.[32] Proposition 4 implies in the context of our model also wage polarization in the aggregate (that we exhibit in 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,

[32] 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.

polarization in any place immediately follows from polarization in the aggregate.[33] 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).[34]

4.4 Patterns of polarization across cities

Here we study whether our model leads to patterns of polarization consistent, at least qualitatively, with the one we observe in the data. In our model, patterns of polarization depend on the relative productivities of households across sectors and cities. As a start, we examine our within regression of individual wages on the DADS-Panel data for 1993-1995 (equation 32 in the Appendix), $ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_{co}} + \delta_t + v_i + e_{it}$, to 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) and use key properties from this to motivate our approach to simulations.[35]

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 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.[36][37]

[33] 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.

[34] Other examples of theory of spatial polarization where polarization does not arise everywhere or does not arise in aggregate include Cerina et al. (2022).

[35] 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. 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.

[36] We also obtain these results theoretically under some assumptions on productivity and when the comparative advantage of the large city is sufficiently strong in the high-paid sector. See Appendix D.

[37] Our results are developed as comparative statics in a fully frictionless model. This seems a valuable first approach given the long horizon, 1994-2015, we aim to understand. 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

Matching data on wages with productivities in the model. We first match productivities in the model with data on wages as 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:

$$\log w_i = \log A(\sigma, c)p(\sigma) + \log H(\omega, \sigma) \tag{9}$$

where (ω, σ, c) are, respectively, the skill, the sector and the city of individual i.[38] From this specification, we can directly obtain using regressions (equation 32) of log hourly wages on city/sector fixed effects with, 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. We take productivity parameters $A(\sigma, c)$ consistent with Table F.8. We then calibrate 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.

We assume that skills are distributed over Ω following a truncated normal distribution with a mean of 0 and standard deviation of 1 and that $H(\omega, \sigma) = \exp(\nu(\sigma)\omega + \mu(\sigma))$.[39] As a result, 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.

---

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.

[38] Implicitly, this means that we do not take into account the term $T(\tau)$. Assume 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 $w_i T(\tau_i)$.

[39] 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.

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 and see also Song et al. (2018) for a similar log-normal pattern of wages in the US). 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 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, Table F.6 provides information on the evolution of relative value marginal products. 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. Thus 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 Lemma 1.

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).

Implications. We now simulate the outcome of the model to a price change as we inferred from Table F.6. 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 4, 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

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

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.

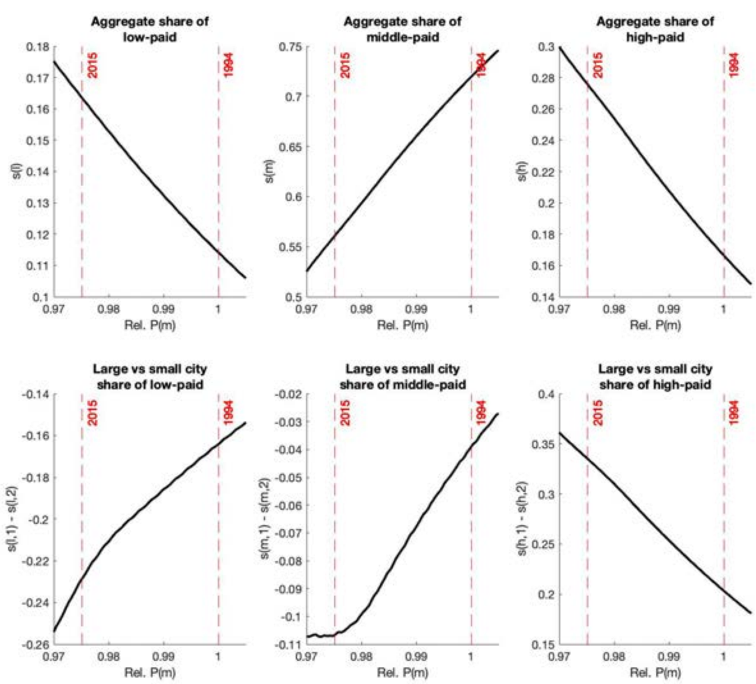


Figure 7: The effect of a decrease in the price of the middle-paid good

Middle-paid job loss and initial exposure. The first observation is that the share of middle-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.[40]

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

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

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. 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 that overall middle-paid jobs decline more in the large city. A way to understand this quantitative result is that the polarization shock leads to more reallocation to high-paid jobs overall, as in Table 4. 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.

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. We then 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.[41] We then obtain qualitatively similar results in terms of the patterns of polarization.[42]

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

[41]See Appendix D for a formal analysis of such a situation.

[42]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.

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.[43]

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 are mainly about the upper tier of these jobs, while it is about the lower tier in the small city.[44]

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.

[43] 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.

[44] 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.

5 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.

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.

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.

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, "Skilled Scalable Services: The New Urban Bias in Recent Economic Growth," 2020. mimeo.

Eeckhout, Jan, 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.

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.