Labor Market Polarization in a System of Cities: The Case of France*

Donald R. Davis Eric Mengus Tomasz Michalski

September 26, 2018

Incomplete, please do not circulate

Abstract

We study how labor market polarization impacts differentially large and small cities. Building on Autor and Dorn (2013), Davis and Dingel (2014) and Davis and Dingel (forthcoming), we add the relative advantage of large cities for the more skilled. Our theory implies that the large city will have smaller initial exposure to middle skill jobs, but that in the process of polarization these will decline even more sharply in the large than the small city. A sharper rise of high skill jobs in the large city and of low skill jobs in the small city should occur. Data for France from 1994-2015 strongly confirm key theoretical predictions.

Keywords: Labor Market Polarization, System of Cities, Inequality

JEL Classification: J21, R12, R13

*Davis: Columbia and NBER; Mengus and Michalski: HEC Paris and GREGHEC.

1 Introduction

Among the most important economic developments in the last three decades is labor market polarization – robust job growth at the bottom and top of the occupational skill distribution and a decline in middle skill jobs Acemoglu and Autor (2011). Researchers have documented labor market polarization in the United States (Autor et al., 2006; Autor and Dorn, 2013) and in many European countries (Goos and Manning, 2007; Goos et al., 2009).

Ongoing dramatic falls in the cost of automation, information and communications technologies loom large in the discussion of driving forces. On one hand, this is tied to the routinization hypothesis, whereby these technologies are particularly adept at substituting for routine tasks previously carried out by middle skill labor Autor et al. (2003). Likewise, advances in information and communications technologies also facilitated the growth in offshoring of tasks from high to low wage countries, notably China.

Researchers have begun to explore how exposure to affected occupations manifests itself in local labor markets (Autor et al., 2013; Goos et al., 2014; Acemoglu and Restrepo, 2017). Harrigan et al. (2016) explore how the effects of offshoring at the firm level are mediated by the export-import orientation of the firm and the presence of specific workers capable of translating offshoring opportunities into action.

The literature tying these impacts to local labor markets has made important contributions, but also has limitations. These have frequently assumed that some (e.g. Autor and Dorn, 2013) or all (Autor et al., 2013) types of skills are immobile across locations. Thus employment of some or all factors in a location, which should be an object of investigation, is instead an assumption. For a phenomenon such as labor market polarization, which develops over decades (see Barany and Siegel, 2018), this is an important limitation that needs to be relaxed.

Recently researchers have developed a literature exploring heterogeneity of workers or firms for cities in spatial equilibrium. These include Behrens et al. (2014), Gaubert (2018), and Davis and Dingel (2014, forthcoming). The latter of these provides the most suitable setting for our work.

Our theoretical work both replicates key prior theoretical results and goes beyond them. While our model departs in important ways from Davis and Dingel (2014) and Davis and Dingel (forthcoming), it likewise features log supermodularity of skills in city sizes and a skill premium that rises with city size. Similar to models of labor market polarization such as Autor and Dorn (2013) and Autor et al. (2013), our model features labor market polarization in the aggregate in response to technology and globalization shocks.

Our theory also goes beyond the prior work in a number of significant ways. First, with the distribution of all skills across space endogenous we are able to provide a condition for the larger city to have lower initial exposure to middle-skill jobs in equilibrium. Second we show that the forces that generate aggregate labor market polarization do so in both large and small cities considered separately. Third, we show that the response of the large and small city in our framework can differ importantly. Specifically we provide a condition under which in response to the shock the large city has a sharper percentage point decline in middle-skill jobs in spite of its low initial exposure; a greater percentage point rise in high-skill jobs; and a smaller percentage point increase in low-skill jobs.

We take these theoretical predictions to the data for France in the period 1994-2015, employing a rich variety of data sources covering wages, employment, educational attainment, and geographical location, among other variables. Our empirical work provides solid support for our theory. Large cities are more skilled and specialize in higher skill activities as per log supermodularity. They also have a lower exposure to middle-skill jobs than smaller cities. In the face of shocks that generate labor market polarization, these common forces work themselves out differently in large and small cities. There is a sharper percentage point decline in middle-skill jobs in large cities in spite of the lower initial exposure. The polarization is itself polarized in the sense that large cities have a strong bias toward job growth in the high-skill sectors and likewise for small cities in the low skill sectors. In sum, city size is a strong conditioning factor in the local effects of labor market polarization.

2 Theory

2.1 The environment

Let us first describe the model, which integrates the core model of job polarization in Autor and Dorn (2013) with the system of cities model of Davis and Dingel (2014). The set of cities is $c \in \{1, 2\}$. In each city, there is a continuum of locations $\tau \in [0, \infty)$.

The economy is populated by households that provide labor, consume and decide where to live and work.

Households. Each household inelastically provides 1 unit of labor but households have different skills that we denote by ω. ω is distributed on $[\underline{\omega}, \overline{\omega}]$ with a pdf $n(.)$ and $\underline{\omega} \geq 1$.

Households consume a single final good and 1 unit of housing. They 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 prize of unoccupied locations to 0 so that $r(c, \tau) \geq 0$.

Households can also decide in which sector σ they work.

Finally, we denote by $f(c, \tau, \omega, \sigma)$ the endogenous pdf of the distribution of households.

Production. Production in this economy involves different sectors: final goods are produced out of capital Z and intermediate goods $\{h, m, l\}$ that are also produced within the economy. All goods are traded with zero transportation costs except non-traded housing.

Final goods. The final good is produced by a continuum of identical competitive firms. To produce, they use intermediate goods $\{h, m, l\}$ as well as Z.

The production function of the representative firm is:

$$Q = \left(a_h q_h^{\gamma} + \left(a_m q_m^{\frac{1}{\theta}} + a_z Z^{\frac{1}{\theta}} \right)^{\gamma\theta} + a_l q_l^{\gamma} \right)^{1/\gamma} \tag{1}$$

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 the rest being technological parameters.

As we are using the final good of numeraire, the profits of the representative firms can be written as:

$$\Pi = Q - p_h q_h - p_m q_m - p_l q_l - p_z Z \tag{2}$$

Intermediate goods. The intermediate goods $\{h, m, l\}$ are freely tradable across cities and are produced with a constant return to scale technology only using labor.

There is one sector to produce each of the $\{h, m, l\}$ goods. Consistently, we label sectors by $\sigma \in \{h, m, l\}$ where h stands for high, m for middle and and l for low skilled.

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

$$H(\omega, \sigma, c)T(\tau) \tag{3}$$

Here 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 spillovers. Thus $T(.)$ is a decreasing function identifying the cost in productivity of being remote from the most productive location in a city. In each city c, we assume that the supply of locations at least as good as τ is $S(\tau)$ with $S(0) = 0$, $S(.)$ strictly increasing and twice continuously differentiable.

$H(\omega, \sigma, c)$ is log supermodular in (ω, σ) and $H(., \sigma, c)$ is increasing. For simplicity, we assume H has the following functional form:

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

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

$$H(\omega, h, c) = A(h, c)e^{\eta\omega} \text{ with } \eta > 1. \tag{6}$$

This assumption regarding the functional form simply means that individual skills matter for productivity even more for the high-skilled sector, while they don't matter for the low-skilled sector.[1]

[1] In the interest of simplicity we make the assumption of exogenous productivity differences between the two cities. Endogenous productivity differentiation across cities arises naturally in this type of framework, but is not the

Assumption 1. *City 1 has an absolute advantage in all sectors:*

$$A(1, j) > A(2, j) \text{ for } j \in \{l, m, h\}$$

and a comparative advantage in higher skilled sectors:

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

Assumption 1 directly implies that $H(\omega, \sigma, 1) \geq H(\omega, \sigma, 2)$ for all (ω, σ).

Finally, we assume that there is perfect competition in all the three sectors so that in each sector the wage per efficiency unit of labor equals the price of the intermediate good $p(\sigma)$.

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

$$Z = \frac{1}{\xi} q, \tag{7}$$

where q is the amount of final goods and ξ is a technology parameter. There, perfect competition implies:

$$p_z = \xi. \tag{8}$$

The intermediate good z has two interpretations. The first is that it is a capital good that substitutes for middle-paying 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 imported intermediates and ξ is the terms of trade. In this offshoring interpretation, the driving force is technical progress in, for example, the Chinese intermediate export sector (which it trades for imports of final output).

2.2 Household decisions

Let us first investigate location and sector decisions by agents and how these decisions depend on factor prices.

The utility flow obtained by an agent with skill ω, location decisions (c, τ) and intermediate

focus of this paper. Davis and Dingel (2014, forthcoming) provide alternative approaches to endogenous productivity differences across cities even when fundamentals are symmetric.

good sector σ is:

$$H(\omega, \sigma, c)T(\tau)p(\sigma) - r(c, \tau) \tag{9}$$

We are interested in understanding in which city and in which sector a household with intrinsic skill ω decides to work, that is, how the household maximizes (9) with respect to c, τ and σ.

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

$$H(\omega(m, c), m, c)p(m) = H(\omega(m, c), l, c)p(l) \tag{10}$$

$$H(\omega(c, h), h, c)p(h) = H(\omega(c, h), m, c)p(m) \tag{11}$$

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

The following lemma shows that these two thresholds are sufficient for characterizing sectoral decisions by households:

Lemma 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:

$$\omega(1, h) < \omega(2, h) \text{ and } \omega(1, m) < \omega(2, m) \tag{12}$$

Proof. See Appendix A.1 □

The differences in the thresholds across cities result from the comparative advantage of the larger cities in more skilled sectors associated with the increasing reliance on individual skills in higher skilled sectors. For the same prices of intermediate goods, the same individual is relatively more productive in higher skilled sectors in the larger city and, thus, have more incentive to work in these sectors.

Let us 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 should exist in at least one city. This happens, for example, when $\omega(1, m) \leq \underline{\omega}$. In this case, there is no low-skill sector in city 1.

This result contrasts with Davis and Dingel (2014) where the sectors and ωs in the less productive city were a strict subset of the sectors and ωs in the larger city. This comes from the assumption that the productivity gains in a given city are different depending on the sector ($A(c, \sigma)$ is a function of σ). In the case where these gains are constant across sectors ($A(c, \sigma) = A(c)$), the thresholds would be the same in the two cities and we would be back to the framework of Davis and Dingel (2014). In addition, note that different thresholds also imply the same ω need not produce in the same σ in both cities even when present.

In the end, Lemma 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. Let us first note that a household with skill ω will decide to work in city 1 and in a given location only if it is not better off working in the other city or in any other location τ, that is:

$$\max_{\sigma,\tau} H(\omega, \sigma, 1)T(\tau)p(\sigma) - r(1, \tau) \geq \max_{\sigma',\tau'} H(\omega, \sigma', 2)T(\tau')p(\sigma') - r(2, \tau'). \tag{13}$$

When this holds with equality the skill ω is present in the two cities.

Location within cities. Let us start by describing the decision location within each city. Here we closely follow Davis and Dingel (2014).

To start with, the set of locations occupied in city c is a bounded set. We denote by $\overline{\tau}(l, c)$ the maximum value of τ in city c. More desirable locations have higher rental prices:

Lemma 2. *Housing prices $r(c, \tau)$ are decreasing on $[0, \overline{\tau}(l, c)]$ and $r(c, \overline{\tau}(l, c)) = 0$. Finally, for all $\tau \in [0, \overline{\tau}(l, c)]$:*

$$S(\tau) = L \int_0^{\tau} \int_{\sigma} \int_{\omega} f(c, x, \omega, M(\omega, c)) d\omega d\sigma dx \tag{14}$$

Proof. See Appendix A.2 □

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

Lemma 3. *There exists a function K such that: $f(c, \tau, \omega, M(\omega, c)) > 0 \Leftrightarrow K(c, \tau) = \omega$. The function $K(c, .)$ is continuous and strictly decreasing.*

In addiction, there exist $\underline{\tau}(1)$ and $\underline{\tau}(2)$ such that $K(2, \underline{\tau}(2)) = K(1, \underline{\tau}(1)) = \underline{\omega}$ and $K(1, 0) = \overline{\omega}(1) = \overline{\omega}$ and $K(2, 0) = \overline{\omega}(2)$.

Proof. See Appendix A.3 □

The proof of this lemma closely follows Davis and Dingel (2014) but extended to the case where households' productivity in a given sector is city specific.

Previously, we noted that the different sectors need not be in both cities. In contrast, both cities have the least skilled person $\underline{\omega}$, so that the larger city's skill set is a strict superset of that in the smaller city.

Using the results of Lemma 1 and 3, we can connect location decisions with the sectoral decisions and show that more skill intensive sectors concentrate in the most attractive locations in the city.

Lemma 4 (Sorting within cities)**.** *In each city c, there exists $\overline{\tau}(h,c) \leq \overline{\tau}(m,c) \leq \overline{\tau}(l,c)$ such that:*

- *If $\omega \geq \omega(c,h)$ then $\tau \leq \overline{\tau}(h,c)$,*
- *If $\omega \in [\omega(m,c), \omega(c,h)]$ then $\tau \in [\overline{\tau}(h,c), \overline{\tau}(m,c)]$,*
- *If $\omega \leq \omega(m,c)$ then $\tau \in [\overline{\tau}(m,c), \overline{\tau}(l,c)]$.*

In particular, $f(\omega,\sigma,c,\tau) = 0$ for all ω, σ, c and $\tau \geq \overline{\tau}(l,c)$.

Proof. See Appendix A.4 □

Locations across cities. To start with, there are locations in city 1 where the productivity of worker is strictly higher than what it could be in city 2. This happens for locations τ where productivity in city 1 strictly exceeds what can be obtained in city 2, even in the best location. More formally:

$$H(\omega(\tau), M(\omega(\tau),1),1)T(\tau) > H(\omega(\tau), M(\omega(\tau),2),2)T(0) \tag{15}$$

where $\omega(\tau) = K(1,\tau)$ is the value of ω occupying location τ in city 1. This defines a maximum value for the skill in city 2, $\bar{\omega}(2)$ for which inequality (15) holds with equality.

Below the productivity $\bar{\omega}(2)$, 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:

$$H(\omega(\tau), M(\omega(\tau),1),1)T(\tau) = H(\omega(\tau), M(\omega(\tau),2),2)T(\tau'). \tag{16}$$

To give further intuition, when there are high-skill workers in both cities, (16) can be rewritten, depending on the value of ω as:

$$\begin{array}{rcl}
\forall \omega \in [\omega(2,h), \bar{\omega}(2)], & A(1,h)T(\tau) & = A(2,h)T(\tau') \\
\forall \omega \in [\omega(1,h), \omega(2,h)], & A(1,h)T(\tau)e^{\eta\omega}p(h) & = A(2,m)T(\tau')p(m) \\
\forall \omega \in [\omega(2,m), \omega(1,h)], & A(1,m)T(\tau) & = A(2,m)T(\tau') \\
\forall \omega \in [\omega(1,m), \omega(2,m)], & A(1,m)T(\tau)\omega p(m) & = A(2,l)T(\tau')p(l) \\
\forall \omega \leq \omega(1,m), & A(1,l)T(\tau) & = A(2,l)T(\tau')
\end{array}$$

In the end, for an $\omega \leq \bar{\omega}(2)$ and a τ, there exists a single τ' in city 2. This defines a function $\Gamma(\omega,\tau) = \tau'$. We then have that, for all ω:

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

In equilibrium, in the location τ, if the agent with skill ω is the marginal buyer, we then have that:

$$r(1,\tau) = r(2,\Gamma(\omega,\tau)).$$

In Davis and Dingel (2014), the function Γ would be constant with respect to ω, but, as the larger city has also a comparative advantage in higher-skill sector, we obtain the following result:

Lemma 5. *For all ω, $\Gamma(\omega, .)$ is continuously increasing in τ and, for any τ, $\Gamma(., \tau)$ is continuous and weakly decreasing in ω.*

Proof. See Appendix A.5 □

In the end, for $\omega \in [\underline{\omega}, \bar{\omega}(2)]$, households are indifferent between a less desirable location in the more productive and larger city 1 or a more desirable location in the less productive and smaller city 2.

Sectoral decisions and factor prices. As this can be observed from equations (10) and (11), the two thresholds are functions of intermediate good prices $p(l)$, $p(m)$ and $p(h)$. The following lemma clarifies how the thresholds moves as a function of $p(m)$.

Lemma 6. *A decline in $p(m)/p(h)$ implies a relatively larger decline for $\omega(1,h)$ than for $\omega(2,h)$. An increase in $p(l)/p(m)$ implies a relatively larger increase in $\omega(2,m)$ than for $\omega(1,m)$.*

Proof. See Appendix A.6 □

When the relative price of middle-paying work declines, the incentive for a middle-skill worker to become a high-skill worker increases for a larger set of people in the large than in the small city. The reason is that the difference in productivity in the middle-skill sector (that increases linearly with skill ω) and the high-skill sector (which increases exponentially) is lower in the large city for more workers given the lower threshold $\omega(1,h)$ in the larger city.[2]

Similarly, the incentive for a middle-skill worker to become a low skill worker also increases in both cities. Yet, as the relative productivity for middle-paying jobs is lower in the smaller city, this incentive for middle-skill workers to become low skill increases by more in the smaller city: this leads $\omega(2,m)$ to increase by more than $\omega(1,m)$ for the same variation of intermediate good prices.

2.3 Implications

We are now able to describe the main implications of our model. In particular, we show four empirical implications along the lines on the evidence that we report.

[2] More specifically, in the large city, the indifference condition between being the middle-skill and the high-skill sectors leads to a lower threshold $\omega(1,h)$ than the one in the smaller city. As a result, a similar variation in the gains of becoming a high-skill leads to a larger variation for $\omega(1,h)$, where productivity is relatively flatter with respect to ω than for $\omega(2,h)$ where productivity is steeper with respect to ω.

Allocation of skills and city exposure to the middle-skill sector Let us first describe the allocation of skills and the exposures to different skills across cities.

Initial exposure to middle-skill jobs Our first implication is about the exposure of a given city to middle-paying jobs, that is the share of total jobs in the middle-paying sector:

Proposition 7. *When the productivity of high skill workers is sufficiently large in the larger city ($A(1, h)$), the share of middle skill sector jobs is smaller in the larger city.*

Proof. See Appendix A.7. □

The relative importance of the three different sectors in the larger city depends on the relative productivity gains of concentrating in that city households working in each of these sectors. When the productivity gains of concentrating workers in the high-paying sectors in the large city are sufficiently large, the share of this sector becomes relatively large and crowds out the presence of the other sectors, thus leading the larger city to be less exposed to middle-paying jobs.

Log-supermodularity Our second implications concern the distribution of skills in the two cities:

Proposition 8. *Let us assume that the supply of locations in each city has decreasing elasticity (TBC). Then, the distribution of skills $f(c, \omega)$ is strictly log-supermodular.*

Proof. See Appendix A.8 □

Let us remind that a distribution is strictly log-supermodular when, for $c > c'$ and $\omega > \omega'$, $f(c, \omega)f(c', \omega') > f(c', \omega)f(c, \omega')$, which means that there are relatively more high skill workers in the larger city.

Given our previous result in Proposition 7 where we obtained conditions under which the share of middle-paying jobs is smaller in the larger city, we can also characterize the elasticity of the middle-paying jobs with respect to the size of the city:

Corollary 9. *Under the conditions of Proposition 7, the elasticity for middle-skill workers with respect to the size of the city is lower than 1.*

Labor market polarization Let us now 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).

Labor market polarization in the aggregate Let us first observe how this decline of the price of capital affects labor markets overall and in each city:

Proposition 10. *A decline in p_z reduces the share of middle-skill jobs in the aggregate and in each city.*

Proof. See Appendix A.9. □

As in Autor and Dorn (2013), a decline in the price of capital goods/offshoring intermediary goods leads firms to substitute middle-paying jobs by capital. This leads workers to reallocate, either to the high-paying or to the low-paying sectors, depending on workers' skills and, overall, the labor market becomes more polarized.

Importantly, this polarization does not only occur in the aggregate, as already is shown by Autor and Dorn (2013), but also in each city: in both cities, the share of middle-paying jobs declines.

Labor market polarization across cities Yet labor market polarization features some striking differences depending on the size of the city: how pronounced is the decline of middle-skill jobs and the rise in low- and high-paying jobs depends on the relative productivity gains of the different sectors in the different cities:

Proposition 11. *When $\frac{A(1,h)}{A(2,h)}$ is sufficiently large relative to $\frac{A(1,m)}{A(2,m)}$, then in the large city:*

(i) *The increase in high-skill jobs is larger in percentage points.*

(ii) *The decline in middle-skill jobs is larger in percentage points.*

(iii) *The increase in low-skill jobs is smaller in percentage points.*

Proof. See Appendix A.10. □

The decline in the price of capital goods/offshoring intermediary goods leads to a stronger decline of the threshold between the high-skilled and the medium-skilled sectors in the larger city. This thus leads to a stronger increase in the number of high-skill jobs there. Conversely, the threshold between medium- and low-skilled sectors increases by more in the smaller city, thus leading more low-skilled jobs to be created there. The evolution of the number of middle-skill jobs results from the evolution of the two thresholds: on the one hand, there are more middle-skilled jobs replaced at the top in the larger city but there are also more middle-jobs replaced at the bottom in the smaller city. In the case where $A(1,h)$ is sufficiently large and when $A(1,m)$ is sufficiently close to $A(2,m)$, the former effect dominates and the share of middle-skill jobs declines by more in the larger city.

3 Data description

We focus on a few key questions: the characteristics and the evolution of labor market polarization in the aggregate and by agglomeration; log supermodularity of skills; and the cross-city patterns of the skill premium. These require exhaustive data on job characteristics (e.g. their routine or offshorable nature), hours worked and wages by occupations. We also need a measure of skills such as educational attainment. The data should be geographically detailed at the agglomeration level and comparable through time.

3.1 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[3]. This data is provided by INSEE, the French national statistical institute, and is based on mandatory annual reports by all French companies[4]. It includes data about all legally held job positions (postes), detailed at the plant level. The initial year 1994 is chosen as this is the first year of data that has comprehensive coverage of hours worked while 2015 was used as the last vintage available. For each worker, for a particular job position, the main reported data are the hours worked, remuneration (total compensation before taxes), occupation type, age and gender[5]. Establishment location information is available at the *commune* level, the lowest administrative unit. There were 36,169 communes in metropolitan (mainland) France as of January 1, 2015.

We use data only for private companies, excluding privatized firms or those that changed status from public to private incorporation (which impacts for example the public or private law under which labor contracts are offered) in the period 1994-2015[6]. We use data for mainland France (without Corsica or the overseas departments). We limit the sample to workers 25-64 years of age. To minimize erroneous entries (for example employees recorded with few working hours with abnormally high income) one needs to filter the job positions. We retain all positions where there were at least 120 hours worked in a year[7].

[3]The French Labor Survey is available since 1982 but for early years has approximately 60,000 observations per year and has only data at the department level. This allows to document some general facts about labor market polarization starting from 1982, but the DADS-data is exhaustive and gives inter alia more geographical details.

[4]This includes public and private firms. Data on self-employed are not reported.

[5]We cannot observe education data or job tenure, and it is not possible to aggregate incomes by individual workers at each year level.

[6]Some firms in the finance, insurance and real estate sectors reported pre-2001 their employees from branches at few establishments for example at the department level which may introduce minor errors given the scale of the problem when we use agglomeration-level data. Exclusion of these sectors from our analysis does not change our results considerably and does not impact our conclusions. We include Table 14 without these sectors as a replication of Table 7 in a robustness test.

[7]We do not observe, however, a material difference in our results if no filtering is applied or filtering based on end-of-year presence with at least 30 days in the firm. INSEE provides filtering in the DADS data set, but it is not consistent between 1994 and 2015.

3.2 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-profesionnelles"). It has been developed by French statistical authorities to classify occupations according to their "socio-professional" status and does not have a clean and direct correspondence at this level to other internationally used classifications such as e.g. the International Standard Classification of Occupations (ISCO). The broad 1-digit codes represent CEOs or small-business owners (CS category "2"), "cadres" (high-skilled professionals, code "3"), medium-skilled professions (codes starting with "4"), low-skilled employees (codes with "5" as the first digit) and blue-collar workers (codes starting with "6"). The 2-digit categories provide more detail, allowing us to use 18 different CS 2-digit categories[8]. We exclude artisans, agriculture-related and public-sector occupations.

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 considered cities in 1994 and 2015, as well as their routine (RTI) occupation and offshorability (OFF-GMS) scores.

Table 1 – Basic statistics by 2 digit CS categories.

CS 2 code	description	employment share		average city wage		RTI	OFF-GMS
		1994	2014	1994	2015		
high-paying occupations							
23	CEOs	1.01%	0.86%	42.82	59.20	-0.75	-0.59
37	managers and professionals	6.15%	10.23%	32.52	38.56	-0.75	-0.59
38	engineers	5.14%	8.97%	30.36	33.68	-0.82	-0.39
35	creative professionals	0.54%	0.54%	22.82	31.80	-0.72	-0.49
medium-paying occupations							
48	supervisors and foremen	4.09%	2.74%	18.03	21.86	0.42	1.23
46	mid-level associate professionals	12.31%	7.64%	17.54	21.20	-0.48	-0.16
47	technicians	5.66%	6.27%	17.15	20.60	-0.40	-0.29
43	mid-level health professionals	0.81%	1.51%	15.05	18.05	-0.35	-0.57
62	skilled industrial workers	14.13%	9.29%	13.52	17.99	0.38	1.24
54	clerks	11.78%	11.17%	13.17	16.98	2.03	0.87
65	transport and logistics personnel	2.86%	2.97%	11.96	16.00	0.33	0.27
63	skilled manual workers	8.01%	8.27%	11.90	15.50	0.17	-0.33
64	drivers	5.03%	5.53%	11.50	14.46	-1.50	-0.63
67	unskilled industrial workers	10.86%	5.72%	11.02	14.72	0.45	2.09
low-paying occupations							
53	security workers	0.67%	1.42%	10.60	14.60	-0.28	-0.51
55	sales-related occupations	5.39%	8.28%	10.44	13.74	0.30	-0.57
56	personal service workers	2.20%	4.79%	9.97	12.63	-0.43	-0.57
68	unskilled manual workers	3.34%	3.78%	9.11	13.27	0.06	-0.36

Notes: In-sample values. Employment share for metropolitan (mainland) France. Average city wages in constant 2015 euros.

[8] Firms should report their data using much finer 4-digit codes, but many fail to do so, and as a result this data is not usable. After the 2003 revision the difference at the 2-digit level is a new category, 31 ("liberal" professionals such as lawyers etc.) that was previously included in 37. In all our data we merge the two together.

We classify as high paying occupations those of "cadres" and CEOs (CS codes 23, 35, 37, 38). Apart from a different legal (e.g. special retirement treatment) and social status, there is a clear gap in terms of wages between these and the remainder of the occupations ("non-cadres"). To determine low paying jobs and the exposures to automation and offshoring we used the data provided in Goos et al. (2014) and mapped their ISCO-based classification into the 2-digit CS ones using the French Labor Survey for 1994 where both are available.[9]

The low-paying occupations are then security workers, retail workers, personal service workers and unskilled manual laborers (CS 53, 55, 56, 68)[10]. As a measure of exposure to automation we retain the Routine Task Intensity index (RTI) used by Autor and Dorn (2013), where this is used to identify occupations for which computers may be able to substitute. For the measure of offshorability, we use the standardized index developed by Goos et al. (2014), where this indicates occupations that can be readily substituted by imports. We observe the CS category 54 (clerks) as being most routine and 67 (unskilled industrial workers) as most offshorable.

We note that the set of the four most routine and the four most offshorable occupations are the same (CS 48, 54, 62 and 67), comprising 40.9% of hours worked in 1994 in our private-sector employment sample and spanning the entire wage distribution of middle-paying jobs. We will refer to this group as RTI4 jobs. The focus on these 4 most routine and offshorable occupations corresponds well with our theory. In our setup medium skill intermediate goods (equation 7) can be produced either by capital goods or imported intermediates. The fall in the prices of both types of goods will have the same effects in our model, and will be captured similarly in our data.

3.3 Cities considered and final sample

For most of our empirical exercises, we are going to be concerned with across-city comparisons. We will therefore use data only on jobs that are performed in cities (agglomerations) above 50,000 inhabitants. We aggregate the commune-level data to the agglomeration level ("unité urbaine") with city boundaries defined by INSEE as of 2010 unless otherwise indicated. There are 117 such cities with the largest 55 above 100,000 inhabitants shown in Figure 1 with population data by category in Table 8. The characteristics of the final sample are given in Table 9.

The cities above 50,000 inhabitants have 54% of total population of metropolitan France. At the same time, the jobs therein are responsible in 2015 for 73% (73% in 1994) of wages paid and 68% (69% in 1994) of hours worked in the mainland in the non-farm private sector. 396,540 (out of 596,430 for which we have data) and 633,845 (out of 998,455) firms were active in these agglomerations in 1994 and 2015 respectively. After all above-mentioned exclusions we remain with

[9] We used hours worked as weights. In each of our 4 low-paying 2-digit CS categories, and only for those, the share of low-paying occupations classified as such in Goos et al. (2014) in total hours worked was over 50%.

[10] Our results are not materially affected when we use a narrower set of occupations – only CS 55, 56, 68. The reason for such a robustness check is that the CS 53 category is quite close in terms of wages in 1994 to the category CS 67 that we classify as a middle-paying job.

a sample that accounts for 65% of total wages paid and 58% of hours worked in metropolitan France in both 1994 and 2015 with data from 364,398 and 596,441 firms in 1994 and 2015 respectively[11].

We group the cities into six major categories for our analysis. Paris, given its size (10.7m inhabitants in the agglomeration and 37.5% of jobs in our final sample) is a category by itself. Then, we use 2 categories of cities above 0.5m cities: 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 agglomeration – Bordeaux (904 thousand inhabitants) and the eighth – Nantes (634 thousand people). All cities in our "0.75m and above" category have also urban areas as defined by INSEE[12] 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 - below which individual city labor markets are small.

3.4 Other data

We also use other data sources to provide additional statistics. For wage premia comparisons, we use the DADS-Panel data set also produced by INSEE out of the raw gathered data. This allows us to follow at least 4% yearly of the active worker population through time, and allows us to aggregate income at the worker level for each year. We proceed with the same exclusions as for the DADS-postes. When we examine log-supermodularity, we use the detailed part of the Census of 1999 (5% of population), also provided by INSEE. It provides data on education, nationality of respondents, and allows us to identify their location at the commune level. Population data was taken from the INSEE for the relevant years.

4 Core empirics

Our primary aim is to understand the consequences of technological and offshoring shocks on labor market polarization across cities. However these are best understood if we take as a preliminary understanding key cross-sectional patterns in the data.

4.1 Log-supermodularity of skills.

One implication – Proposition 8 – of the model is, as in Davis and Dingel (2014), that the distribution of skills $f(c, \omega)$ is log-supermodular in city size. To obtain a measure of skills we turn to the 1999 Census data that has both good data on diplomas and the commune of residence. The

[11] Establishment-level output data is unavailable in the French data.

[12] This is the "unité urbaine" plus all communes where at least 40% of residents have employment in the agglomeration.

Table 2 – Log-supermodularity, population elasticities by diploma (9 categories) in the 1999 Census data.

Dependent variable: ln $f(c,\omega)$	All workers	French born	Population share	French born share
No diploma X ln pop	0.938	0.906	.12	.75
	(0.032*)	(0.0254***)		
End of primary school X ln pop	0.892	0.877	.08	.89
	(0.032***)	(0.029***)		
End of middle school (collège) X ln pop	0.977	0.971	.07	.94
	(0.024)	(0.023)		
Vocational school diploma (CAP) X ln pop	0.907	0.902	.20	.96
	(0.024***)	(0.024)***		
Vocational high school intermediate diploma (BEP) X ln pop	0.924	0.92	.10	.96
	(0.022***)	(0.021***)		
High school vocational diploma (bac technologique or professionnel) X ln pop	0.975	0.972	.09	.97
	(0.026)	(0.025)		
General high school diploma (Bac) X ln pop	1.044	1.037	.06	.93
	(0.028)	(0.027)		
Undergraduate studies X ln pop	1.038	1.034	.15	.97
	(0.026)	(0.026)		
Graduate studies X ln pop	1.184	1.179	.14	.94
	(0.037***)	(0.036***)		

Notes: 112 cities > 50,000 inhabitants defined by INSEE as of 1999, population figures from 1999. Exclusions as in main sample. Standard errors, clustered by agglomeration, in parantheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

results are shown in Table 2.[13] We observe that, as expected, for low skilled categories we obtain coefficient estimates that are statistically significantly below 1 at conventional significance levels. The opposite is true for those with general high school degrees and higher: they are more prevalent in larger cities. In particular, the category of workers with a graduate diploma has a population elasticity of 1.184. We note that the elasticities for middle-skilled workers are not significantly different from 1. These observations carry over when we consider only French-born individuals: the presence of low-skilled immigrants does not change these patterns.

We reconfirm these patterns using our classification of high, middle and low paying jobs and the broad 1-digit CS categories in Tables 10-11. It is not a coincidence that the population elasticity coefficients for high-paying jobs and "cadres" (respectively 1.14 and 1.15) are similar: "cadres" perform the bulk of high-paying occupations. The coefficients on middle- and low- paying jobs that are below 1 (in a statistically significant manner) show that larger French cities have not only fewer low paying jobs, but also fewer middle-paying jobs. This conforms with Corollary 9.

[13]In this table, CAP or *Certificat d'aptitude professionnelle* is obtained at the age of 16, the BEP or *Brevet d'études professionnelles* is also obtained at the age of 16 but is a prerequisite for obtaining the more advanced bac professionnel at the age of 18.

Table 3 – Share of 4 highest-paying occupations per agglomeration size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.233	0.139	0.123	0.101	0.087	0.080
2015	0.367	0.247	0.210	0.164	0.139	0.119
change in ppct	0.133	0.108	0.087	0.063	0.053	0.039
growth in %	0.571	0.773	0.710	0.629	0.608	0.494

4.2 Employment shares of different occupation categories.

The model has implications on the cross-section employment shares across cities. We use the DADS-Postes data set for 1994 and 2015 and calculate hours worked in different occupations across agglomerations (Tables 3-5).

Table 4 – Share of 10 middle-paying occupations per agglomeration size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.651	0.736	0.753	0.773	0.793	0.793
2015	0.449	0.569	0.607	0.641	0.658	0.671
change in ppct	-0.201	-0.167	-0.146	-0.132	-0.135	-0.121
growth in %	-0.309	-0.227	-0.195	-0.171	-0.170	-0.153

We observe that the share of high-paying occupations in total employment increases monotonically with city size in both years, as required by log-supermodularity shown by Proposition 8. The differences are sizable, especially when comparing the extremes – Ile de France and cities between 50 and 100 thousand of population. In 1994, the share of high-paying occupations in Paris was 23.3 %, as opposed to 10.1 % for cities between 0.2 and 0.5m, and only 8.0 % in cities between 0.05 and 0.1m of population. This discrepancy grew to respectively 36.7 %, 16.4 %, and 11.9 % in 2015.

Moreover, the share of middle-paying jobs monotonically declines with city size in line with Proposition 7, both in 1994 and 2015 (see Table 4). The share of lowest-paid occupations is highest in smallest cities in either of the years, although the variation is modest. The decline of low-paid occupation shares with city size is very clear when one measures the share of hours worked for 3-lowest paying jobs (Table 12).

Overall, these patterns along with evidence on log-supermodularity support our theoretical model. They also contradict for France the extreme-skill complementarity hypothesis stated by Eeckhout et al. (2014).

Table 5 – Share of 4 lowest-paying occupations per agglomeration size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.116	0.125	0.124	0.126	0.121	0.127
2015	0.184	0.184	0.183	0.195	0.203	0.209
change in ppct	0.068	0.059	0.059	0.069	0.082	0.082
growth in %	0.589	0.476	0.477	0.547	0.678	0.642

As automation (Autor and Dorn (2013)) and offshoring (Goos et al. (2014)) are believed to be driving labor market polarization, it is important to observe the exposure of different cities to routine and offshorable tasks. In Table 6 we show that the employment shares across cities in 1994 and 2015 of most routine and offshorable jobs (the same 4 categories that belong to middle-paying group) are declining in city size – in line with the patterns for middle-paying occupations overall exhibited in Table 4 and our Proposition 7.

Table 6 – Share of the 4 most routine and offshorable occupations (CS 48, 54, 62 and 67) per agglomeration size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.294	0.363	0.392	0.413	0.445	0.445
2015	0.189	0.248	0.270	0.294	0.314	0.323
change in ppct	-0.105	-0.115	-0.122	-0.119	-0.131	-0.122
growth in %	-0.358	-0.316	-0.311	-0.289	-0.294	-0.274

4.3 Labor market polarization in France.

Turning to changes in employment patterns in time, we first state that labor market polarization – a fall in the employment share of medium-paying occupations and the rise in the share of high- and low-paying ones occurred in France in the period 1994-2015. The high-, medium, and low-paying occupations, respectively account for one-eighth, three-quarters, and one-eighth of employment in 1994 and roughly one-fifth, three-fifths, and one-fifth in 2015.

We can examine these changes in overall employment shares exhibited in Table 1. The share of middle-paying jobs declined from 75.5% to 61.1% between 1994-2015. The bulk of job losses in this category occurred in the RTI4 jobs – the 4 most automation and offshoring exposed occupations (supervisors and foremen; clerks; skilled and unskilled industrial workers), and their share in hours worked fell from 40.9% to 28.9%. The only other occupation experiencing a large employment share drop was CS 46 (mid-level associate professionals) the share of which fell from 12.3% to 7.6% in the labor force. It is ranked 6th highest in our classification of offshorable categories.

At the same time, the overall shares of high- and low-paying jobs increased respectively from 12.8% to 20.6% and from 11.6% to 18.3%.

The patterns detailed at the 2-digit CS-level are exhibited in Figure 2 and confirm the U-shaped relationship in France for the years 1994-2015 studied by (Autor et al., 2006; Autor and Dorn, 2013) for the U.S. and documented by (Goos and Manning, 2007; Goos et al., 2009, 2014) for Europe[14]. They are also consistent with observations made by Harrigan et al. (2016) for France for the time period 1994-2007.

[14] We exclude here the category of CEOs - CS category 23. It is an outlier with highest pay that has a rather constant population elasticity in sample – its employment share varies between 0.8% and 1.1% in 1994. The change in the share is less than 0.3 percentage points in absolute terms across cities in the given years.

Our theory – Proposition 10 – that labor market polarization is induced by increased automation and/or a lower offshoring cost – is compatible with the above patterns.

4.4 Labor market polarization in large and small cities.

One can observe labor market polarization in France in the aggregate data. Our model (Proposition 11) implies that we should observe polarization at the city level as well. However, labor market polarization should have different consequences for large and small cities. First, middle-paying jobs should decline the most in large cities. Second, high-paying jobs should increase by more in larger agglomerations, while low-paying jobs increase more in smaller cities. Notice that the strength of these relationships does not depend on initial exposure to routine or offshorable occupations, and, indeed, such exposure should be lowest in large cities as indicated by Proposition 7 and exhibited in Section 4.2. All such patterns are upheld in our data.

Labor market polarization in cities at the 2-digit CS level. We start by discussing labor market polarization in cities of different sizes in France (Paris and Lyon – the two largest cities; cities between 0.2-0.5m and 0.05-0.1m inhabitants) comparing employment change patterns shown in Figure 3.

First of all, the figure confirms that labor market polarization occurred in different cities[15]. However, it is immediately clear that the patterns are markedly different across agglomerations. Figures 4 and 5, where we superimpose the employment share changes in Paris and Lyon vs. those in cities with population between 0.05-0.1m, further attest to this. The changes in the Parisian labor market in comparison to smallest cities are larger, despite a lower exposure to jobs with high RTI or OFF-GMS indexes (see Table 6). Both the average of absolute percentage point changes in the individual categories (2.4 vs. 1.9) and their standard deviation (3.6 vs. 2.8) are higher for Paris than for cities between 0.05-0.1m inhabitants.

Turning to individual categories, there is a much higher percentage point increase in the employment share of categories CS 37 (managers and professionals) and 38 (engineers) in Paris and Lyon in comparison to other cities. Moreover, different categories of middle-paying occupations are declining in large and small cities. While in smallest cities highest employment declines are observed for industrial workers (whether skilled or unskilled - categories 62 and 67) that happen to be the two most offshorable categories in our data as well, it is category 46 (mid-level associate professionals) that declines most in Paris. The category that is most routine in sample – clerks (CS 54) – declines also most strongly in Paris (by 3.1 pp) while in small cities employment in this category grows (by 0.5 pp). The CS category 47 (technicians) decline in Paris by 1.4 pp but increase by 1.3 pp in the smallest cities. Among the low-paying occupations, there is a relatively strong increase in the CS category 55 (sales related occupations) by 4.5 pp in small cities. The patterns

[15] Similar patterns can be observed for all size categories considered in Tables 3-6.

in Lyon and cities between 0.2-0.5m inhabitants show intermediate developments between the two polar cases, suggesting common forces at play that may be related to city size.

Some of these differences in labor market developments for individual categories across cities may be due to several factors that should be mentioned but cannot be fully addressed empirically given the limitations of our data that start in 1994.

First, the share of industrial workers (categories 62 and 67) in employment in large cities is lower in 1994 at the outset of our sample than in small cities. Therefore, the additional adjustment in terms of percentage points we observe in these categories over the period 1994-2015 may be less pronounced as well. Data from the French Labor market survey (available since 1982) allow us to document that labor market polarization was ongoing for several years beforehand and it was strongest (in terms of middle-paying jobs' employment share losses) in departments with large agglomerations. One of the reasons manufacturing was moving out of large cities in the earlier period could be the development of road and train networks in France.

One explanation for the decline in the share of back-office or support jobs like clerks (CS 54) or technicians (CS 47) in Paris with their coincident expansion in small cities can be internal offshoring permitted by the Internet and communication technologies. Such tendencies are consistent with our model (all goods, including intermediates, are traded) though we do not model nor cannot fully verify empirically supply chain developments that are internal or external to firms.

Labor market polarization in cities by 3 occupation groups. We discuss now the changes in shares in hours worked in different categories of occupations by agglomeration sizes.

In Table 4 we see that the percentage point changes in employment shares are highest in the largest cities. In Paris, over the period 1994-2015 middle-paying jobs share declined by 20.1 percentage points. In contrast, this decline was lower in smaller cities - only 12.1 percentage points in agglomerations between 50 and 100 thousand inhabitants, and this is despite the lower initial share of middle-paying jobs in larger cities. We note that the middle-paying job destruction rates monotonically increase with city size. Comparing these patterns with the changes experienced in the most routine and offshorable jobs – the RTI4 group (Table 6) – we see that the fall of shares in this category of middle-paying jobs is similar across agglomerations – between 10.5 and 13.1 percentage points. The overall decline in middle-paying occupations in the largest agglomerations exceeds that of the destruction of the most routine and offshorable jobs while it is smaller in the cities below 200 thousand inhabitants (e.g. 12.1 vs. 12.2 percentage points for cities between 0.05-0.1m range). Given that the exposure of the largest agglomerations is lowest in 1994, the destruction rates of these occupations still monotonically rises with city population.

The greater decrease in middle-paying jobs in larger cities can be accounted for by our model, as exhibited by Proposition 11. There is a direct effect of the substitution of labor by capital or imported intermediates in the middle-skill task that may push – given the differences in the production functions across sectors – many more workers from the middle-skill task into the high-skill task in the large city in comparison to the smaller one[16].

Table 3 shows that the percentage point increase in high-paying jobs is monotonic in agglomeration size. In Paris and agglomerations above 1m inhabitants the increase in such occupations is above 10 percentage points over the period 1994-2015. The smallest cities – those between 50 and 100 thousand inhabitants – have the lowest gain of 3.9 pp. Finally, as depicted in Table 5 the percentage point increase in low-paying jobs is highest for smallest cities. Indeed, our theory predicts that the change in the indifference thresholds $\omega(c, m)$ between working in the l and m-skill sectors will be higher in small cities by Lemma 6. We note that there is a much a sharper tradeoff of high tasks for medium skills (especially in the large cities than in smaller cities) while the changes in the low-paying jobs, although displaying behavior according with the model, are less important in terms of size.

Individual city-level evidence. The changes in employment shares coming from overall summation of hours by agglomerations discussed above may not hold at the individual city level if they are driven by a few outliers. Therefore we present in Table 7 the comparison of mean percentage point changes and growth rates between the 11 largest (above 0.5m of inhabitants) and 62 smallest (between 50-100 thousand people) cities in our sample[17]. Calculations for cities between 100 thousand and 0.5 million inhabitants show intermediate patterns between the two.

There is not a statistically significant difference in the percentage point change in the decline of the RTI4 group – the most routine and offshorable jobs (but see the discussion of patterns in Figure 6 below), though we observe a larger destruction of such jobs in larger cities with the six most routine categories included (RTI6) or the six most offshorable ones (OFF6). Average destruction of any such jobs is always stronger in larger cities. The added occupations when one moves from the narrower RTI4 category to the RTI6 and OFF6 deserve thus closer attention. Category CS 65 (transport and logistics workers) is the fifth most routine and offshorable occupation in our sample, but its employment share in overall sample increases. However, it declines in cities >0.5m by 0.2 pp while it increases in the smallest cities by 0.3 pp. The CS 55 (sales-related occupations) ranks as the sixth most automatizable job - it contains inter alia such jobs as cashiers or telemarketers. It is a low-paying occupation, and even though some positions might have been supplanted by machines, its share in employment increases in all cities – and in accordance with our theory more strongly in smaller cities (4.5 pp) than in large ones (2 pp). The CS 46 (mid-level associate professionals) occupation is the sixth most-offshorable – though heterogenous – and it contains several back-office jobs that may now be performed from afar (accountants, sales administration, banking and

[16]This tendency could be reinforced by several mechanisms that we do not model. For example, it could be due to productivity gains for high-skill workers either because of an increase in agglomeration productivity gains $A(h, c)$ or a complementarity between high skill labor and capital (as in the Autor and Dorn (2013) model).

[17]In Table 13 we show the means of employment shares in the presented categories across these cities while in Table 14 the same changes and growth patterns as in in Table 7 in a sample without FIRE industries.

insurance staff, interpreters, graphic designers etc.). It was the one that declined considerably more in larger cities than in smaller ones (7.1 pp vs. 3.7 pp respectively)[18]. The differences in the behavior of these CS 65, 55 and 46 categories between large and small cities may be further driven by ongoing internal offshoring of tasks within France.

As in the aggregate data by agglomerations, we observe a higher destruction overall of middle-paying jobs (by 18.1 percentage points on average) in cities of over 0.5m in comparison with smallest cities (11.5 percentage points). Low-paying occupations increase more in smallest cities while the highest paying in large cities as well. As lower panels of Table 7 indicate, there is a statistically significant difference in the growth rates of the discussed categories as well except in the low-paying occupations. The patterns exhibited by the aggregate data by agglomeration sizes are confirmed.

Table 7 – Comparison of means of changes in different occupations, cities >0.5m vs. 0.05-0.1m, sample with finance, insurance and real-estate observations.

Item	RTI4	RTI6	OFF6	middle-paid	low-paid	high-paid
mean change ppct, cities >0.5m	-0.107	-0.091	-0.178	-0.181	0.065	0.116
mean change ppct, cities 0.05-0.1m	-0.109	-0.063	-0.145	-0.115	0.079	0.036
difference in pct points	0.002	-0.027***	-0.033***	-0.065***	-0.015***	0.080***
mean growth, cities >0.5m	-0.33	-0.228	-0.351	-0.265	0.539	0.633
mean growth, cities 0.05-0.1m	-0.249	-0.114	-0.255	-0.149	0.61	0.453
difference in growth	-0.081***	-0.114***	-0.096***	-0.116***	-0.071	0.18***

Notes: population weighted, allowing for unequal variances. N=73 (11 cities > 0.5m). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Initial exposure to routine and offshorable jobs. We disregarded so far the role of initial exposure to routine and offshorable jobs. Further scrutiny of employment changes at the city level shown in Figures 6-10 reveals instructive patterns.

First of all, it is clear that large cities above 0.5m people have lower initial exposures to routine and offshorable occupations (RTI4)[19]. Although there is considerable variation, Figure 6 confirms the observation of Autor and Dorn (2013) that the initial exposure to the most routine (and, in our context, offshorable) jobs is strongly negatively correlated with their decline as the technological shocks occur. The observations for large cities lie in the lower envelope of observations: conditional on initial exposure, the changes in the employment shares in these cities are larger than in cities with a population between 0.05-0.1m inhabitants.

[18] Another hypothesis for the decline in this category might have to do with the internal organization of the firm. Some of the professions in this category – such as interpreters, photographers, graphic designers, journalists embedded in companies – might have been internally offshored to freelancers. If that were the case, our DADS data would not capture those as we do not have data on self-employed.

[19] The large cities with the highest initial exposure are the Douai-Lens and Lille agglomerations, both located in the old industrial region in the North of France.

The picture turns different, however, when we scrutinize the relationships between exposure and the change in middle-paying jobs (that include the RTI4 category) in Figure 7. The population-weighted regression of changes in employment on initial RTI4 exposure reveals a strongly positive relationship (indeed, the non-population weighted relationship is close to zero). There is clearly a larger destruction of middle-paying jobs in the largest cities that happen to be also the least-exposed ones. For many small but highly-exposed cities, the drop in RTI4 jobs is larger than the decline in middle-paying jobs while the opposite is true for largest cities. There may be hence some substitution of RTI4 jobs by other middle-paying jobs in smaller cities while the destruction of middle-paying jobs is exacerbated in large cities.

Figure 8 shows that the relationship between routine and offshorable job exposure is not correlated with the changes the low paying jobs. Again, as in Table 5 increases in large cities are lower on average than in small cities.

A striking pattern can be seen in Figure 9 where the two groups of cities form disparate sets of observations. The increase in high-paying occupations is distinctively higher for most large cities above 0.5m inhabitants (the outliers are Douai-Lens and Toulon). As a result, there is a negative population-weighted correlation between initial RTI4 exposure and the change in the employment of high-paying occupations. For smaller cities, the increase in high-paying jobs is positively correlated with initial RTI4 exposure.

This lack of relationship between changes in low-paid occupations, and exposure to RTI4 jobs, a negative correlation between high-paying occupations and initial RTI4 share plus a larger destruction of middle-paying jobs in larger cities (that are less exposed) as a result of labor market polarization is expected from our theory – Proposition 11. It is clear that what matters is the city size and the interaction of the technology changes on middle-skill tasks.

Similar patterns to that for RTI4 jobs can be observed for the 6 most offshorable occupations. For example, in Figure 10 higher initial exposure to offshorable jobs leads to their greater decrease in the studied period. This time, however, large cities are more exposed to offshorable jobs as – in particular – they have on average a higher share of the CS 46 category, mid-level associate professionals. Again, conditional on exposure, large cities lose more percentage points of offshorable jobs.

5 Conclusions

In a country subject to the forces of labor market polarization, how will this play out differently in large and small cities? As a framework for understanding this, we build on elements of Autor and Dorn (2013), Davis and Dingel (2014) and Davis and Dingel (forthcoming) with amendments that strengthen the relative advantage of large cities for the more skilled. Our theoretical results predict that the large city will have smaller initial exposure to middle skill jobs, but that in the process of polarization these will decline even more sharply in the large than the small city. Under these same conditions, our theory predicts a sharper rise of high skill jobs in the large city and of low skill jobs in the small city. We take this to the data for France from 1994-2015. The results strongly confirm the key implications of the theory. In large cities there is much stronger growth of high skill jobs and decline in middle skill jobs than in small cities. Small cities do have stronger growth in low skill jobs, although the differences are modest.

Acknowledgements

This work was supported by Investissements d'Avenir [grant number ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047].

References

Acemoglu, Daron 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 Pascual Restrepo**, "Robots and Jobs: Evidence from US Labor Markets," NBER Working Papers 23285, National Bureau of Economic Research, Inc March 2017.

Autor, David H. and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, August 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," *The 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.

Barany, Zsofia L. and Christian Siegel, "Job Polarization and Structural Change," *American Economic Journal: Macroeconomics*, January 2018, *10* (1), 57–89.

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

Davis, Donald R. and Jonathan I. Dingel, "The Comparative Advantage of Cities," NBER Working Papers 20602, National Bureau of Economic Research, Inc October 2014.

_ **and** _ , "A Spatial Knowledge Economy," *American Economic Review*, forthcoming.

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

Gaubert, Cecile, "Firm sorting and agglomeration," Technical Report, National Bureau of Economic Research 2018.

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," *The review of economics and statistics*, 2007, *89* (1), 118–133.

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, National Bureau of Economic Research, Inc March 2016.

A Proofs of Propositions

A.1 Proof of Lemma 1

First, $\omega(m,c) = A(l,c)/A(m,c)p(l)/p(m)$. This latter ratio is larger in city 2 than in city 1 given Assumption 1 and thus $\omega(1,m) < \omega(2,m)$.

Second, $\omega(h,c)$ solves the following equation:

$$\frac{A(h,c)}{A(m,c)}\frac{p(h)}{p(m)} = \omega(h,c)e^{-\eta\omega(h,c)} \tag{17}$$

On $[\underline{\omega}, \overline{\omega}]$, the rhs term is a decreasing function of ω. Given Assumption 1, the ratio $A(h,c)/A(m,c)$ is larger in city 1 than in city 2, thus implying $\omega(1,h) < \omega(2,h)$.

A.2 Proof of Lemma 2

The proof of Lemma 1 in Davis and Dingel (2017) still holds: otherwise, there exists $\tau' < \tau''$ such that $r(c,\tau') \leq r(c,\tau'')$. Thus, $U(c,\tau',\sigma,\omega) > U(c,\tau'',\sigma,\omega)$ for all σ and all ω. This contradicts the fact that τ'' has to maximize utility for some individual with some skill ω and sectoral decision σ.

$$A(h,1)e^{\gamma\omega(1,h)}T(\tau_1)p(h) - r(1,\tau_1) = A(h,2)e^{\gamma\omega(1,h)}T(\tau_2)p(h) - r(1,\tau_2) \tag{18}$$

$$\geq A(m,2)\omega T(\tau_1)p(m) - r(1,\tau_1) \tag{19}$$

A.3 Proof of Lemma 3

We follow here Davis and Dingel (2017), lemma 2 and Lemma 1 in Costinot and Vogel (2010).

We define $f(\omega,c,\tau) = \int_\sigma f(\omega,c,\tau,\sigma)d\sigma$, $\Omega(\tau,c) = \{\omega \in \Omega, f(\omega,c,\tau) > 0\}$ and $\mathcal{T}(\omega,c) = \{\tau \in [0,\bar{\tau}(c)], f(c$

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

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

$$T(\tau')G(\omega',c) - r(c,\tau') \geq T(\tau)G(\omega',c) - r(c,\tau) \tag{20}$$

$$T(\tau)G(\omega,c) - r(c,\tau) \geq T(\tau')G(\omega,c) - r(c,\tau') \tag{21}$$

This jointly implies that $(T(\tau') - T(\tau))(G(\omega',c) - G(\omega,c)) \geq 0$, but this cannot be with $\tau' > \tau$ and $\omega' > \omega$. The same reasoning can be applied when $\tau' < \tau$. We can also conclude that for any $\tau < \tau'$, if $\omega \in \Omega(\tau)$ and $\omega' \in \Omega(\tau')$, then $\omega \geq \omega'$.

(iii) $\Omega(\tau, c)$ is a singleton for all but a countable subset of $[0, \bar{\tau}(c)]$. For any $\tau \in [0, \bar{\tau}(c)]$, $\Omega(\tau, c)$ is measurable as it a non-empty interval. Let $\mathcal{T}_0(c)$ denote the subset of locations τ such that $\mu(\Omega(\tau, c)) > 0$, μ being the Lebesgue measure over $\mathcal{R}$. Let us show that $\mathcal{T}_0(c)$ is a countable sets – any other $\Omega(\tau, c)$ where $\tau \notin \mathcal{T}_0(c)$ is a singleton as it is a interval with measure 0. For any $\tau \in \mathcal{T}_0(c)$, let us define $\underline{\omega}(\tau) \equiv \inf \Omega(\tau, c)$ and $\overline{\omega}(\tau) \equiv \sup \Omega(\tau, c)$. As $\mu(\Omega(\tau, c)) > 0$, $\underline{\omega}(\tau) < \overline{\omega}(\tau)$. Thus there exists a integer j such that $j\,(\overline{\omega}(\tau) - \underline{\omega}(\tau)) > (\bar{\omega}(c) - \underline{\omega})$. Given that $\mu(\Omega(\tau, c) \cap \Omega(\tau', c)) = 0$ for $\tau \neq \tau'$, for any j, we can then have at most j elements $\{\tau_1, ...\tau_j\} \equiv \mathcal{T}_j^0$ verifying $j\,(\overline{\omega}(\tau_i) - \underline{\omega}(\tau_i)) > (\bar{\omega}(c) - \underline{\omega})$. Thus $\mathcal{T}_j^0$ is countable. Given that $\mathcal{T}^0 = \bigcup_{j=1}^{\infty} \mathcal{T}_j^0$ and that the countable union of countable sets is also countable, we conclude that $\mathcal{T}^0$ is countable.

(iv) $\mathcal{T}(\omega, c)$ is a singleton for all but a countable subset of Ω. As in Davis and Dingel (2017), we use the arguments as in steps 2 and 3.

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

$$f(c, \omega, \tau) = f(\omega)\delta^{\text{Dirac}}\left(1 - 1_{\Omega(c,\tau)}\right) \text{ for almost all } \omega \in \Omega(c, \tau). \tag{22}$$

This contradicts assumptions on $S(\tau)$ as this implies that $S'(\tau) = \infty$. TBC.

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

A.4 Proof of Lemma 4

By using the function $K_c(\tau)$ that is continuous and weakly decreasing from Lemma 3, there exist unique $\bar{\tau}(h, c)$ such that $K_c(\bar{\tau}(h, c)) = \omega(h, c)$ and $\bar{\tau}(m, c)$ such that $K_c(\bar{\tau}(m, c)) = \omega(m, c)$.

A.5 Proof of Lemma 5

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

A.6 Proof of Lemma 6

Let us now compute how a change in price of intermediate goods modifies the thresholds. By rewriting the indifference condition as:

$$\frac{H(\omega(c,h),h,c)}{H(\omega(c,h),m,c)} = \frac{p(m)}{p(h)} \tag{23}$$

we obtain, by differentiating both the right and the left hand terms:

$$\frac{d\left(\frac{H(\omega(c,h),h,c)}{H(\omega(c,h),m,c)}\right)}{\frac{H(\omega(c,h),h,c)}{H(\omega(c,h),m,c)}} = \frac{d\left(\frac{p(m)}{p(h)}\right)}{\frac{p(m)}{p(h)}} \tag{24}$$

Let us compute the different terms separately:

$$\frac{d\left(\frac{H(\omega(c,h),h)}{H(\omega(c,h),m)}\right)}{\frac{H(\omega(c,h),h)}{H(\omega(c,h),m)}} = \left(\frac{H_\omega(\omega(c,h),h)}{H(\omega(c,h),h)} - \frac{H_\omega(\omega(c,h),m)}{H(\omega(c,h),m)}\right) d\omega(c,h) \tag{25}$$

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

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

Given that H is log-supermodular,

$$\Gamma(\omega(c,h),c) = \frac{H_\omega(\omega(c,h),h,c)}{H(\omega(c,h),h,c)} - \frac{H_\omega(\omega(c,h),m,c)}{H(\omega(c,h),m,c)} > 0. \tag{27}$$

As a result, a decline in $p(m)/p(h)$ then leads to a decline in $\omega(c,h)$. Similarly, we obtain:

$$d\omega(c,m) = \frac{1}{\Gamma(\omega(c,m),c)} \frac{d\left(\frac{p(l)}{p(m)}\right)}{\frac{p(l)}{p(m)}} \tag{28}$$

As a result, an increase in $p(l)/p(m)$ then leads to an increase in $\omega(c,m)$.

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

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

$$\frac{H_\omega(\omega(1,h),h,1)}{H(\omega(1,h),h,1)} - \frac{H_\omega(\omega(1,h),m,1)}{H(\omega(1,h),m,1)} - \frac{H_\omega(\omega(2,h),h,2)}{H(\omega(2,h),h,2)} + \frac{H_\omega(\omega(2,h),m,2)}{H(\omega(2,h),m,2)} \tag{29}$$

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

$$\frac{H_\omega(\omega(1,m),m,1)}{H(\omega(1,m),m,1)} - \frac{H_\omega(\omega(1,m),l,1)}{H(\omega(1,m),l,1)} - \frac{H_\omega(\omega(2,m),m,2)}{H(\omega(2,m),m,2)} + \frac{H_\omega(\omega(2,m),l,2)}{H(\omega(2,m),l,2)} \tag{30}$$

Let us make some further assumption on the H function. As Autor-Dorn, we assume that $H(\omega,l,c) = 1$ and $H(\omega,m,c) = \lambda(c)\omega$. This simplifies the two expressions into:

$$\frac{1}{\omega(1,m)} - \frac{1}{\omega(2,m)} \geq 0 \tag{31}$$

which is positive as $\omega(1,m) \leq \omega(2,m)$ and:

$$\frac{H_\omega(\omega(1,h),h,1)}{H(\omega(1,h),h,1)} - \frac{1}{\omega(1,h)} - \frac{H_\omega(\omega(2,h),h,2)}{H(\omega(2,h),h,2)} + \frac{1}{\omega(2,h)} \tag{32}$$

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

$$\frac{H_\omega(\omega(1,h),h,1)}{H(\omega(1,h),h,1)} - \frac{H_\omega(\omega(2,h),h,2)}{H(\omega(2,h),h,2)} \leq \frac{1}{\omega(1,h)} - \frac{1}{\omega(2,h)} \tag{33}$$

which is satisfied.

A.7 Proof of Proposition 7

In city c, the population of individuals in the middle-skill task, i.e. with skill ω is between $\omega(c,h)$ and $\omega(c,m)$, is:

$$L\int_{\omega(c,m)}^{\omega(c,h)} f(x,c)dx = S\left(T^{-1}\left(h(\omega(c,m),c)\right)\right) - S\left(T^{-1}\left(h(\omega(c,h),c)\right)\right) \tag{34}$$

where $K(T^{-1}(h(\omega,c)),c) = \omega$.

The share of middle skill agents in city c is:

$$s(m,c) = \frac{\int_{\omega(c,m)}^{\omega(c,h)} f(x,c)dx}{\int_{\underline{\omega}}^{\overline{\omega}(c)} f(x,c)dx} = \frac{S\left(T^{-1}\left(h(\omega(c,m),c)\right)\right) - S\left(T^{-1}\left(h(\omega(c,h),c)\right)\right)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{35}$$

Using the continuity of the different function and given that $\omega(c,h)$ is decreasing in $A(c)$ and $s(m,c) = 0$ when $A(c) \to \infty$, we thus obtain that, taking all other parameters constant including $A(2)$, there exists a sufficient large $A(1)$ such that shares satisfy $s(m,1) \leq s(m,2)$.

A.8 Proof of Proposition 8

The population of individuals with skills between ω and $\omega + d\omega$ is:

$$L \int_{\omega}^{\omega+d\omega)} f(x,c)dx = S\left(T^{-1}\left(h(\omega,c)\right)\right) - S\left(T^{-1}\left(h(\omega+d\omega,c)\right)\right) \tag{36}$$

Taking the derivative with respect to $d\omega$ and taking $d\omega \to 0$ yield:

$$f(\omega,c) = -\frac{\partial}{\partial\omega} S\left(T^{-1}\left(h(\omega,c)\right)\right) = \frac{h'(\omega,c)V(h(\omega,c))}{L} \tag{37}$$

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

To show that $f(c,\omega)$ is log-supermodular in (c,ω), it is then sufficient to show that $V(h(\omega,c))$ is log-supermodular in (c,ω) which happens if and only if:

$$\frac{V'(h(\omega,1))h'(\omega,1)}{V(h(\omega,1)} - \frac{V'(h(\omega,2))h'(\omega,2)}{V(h(\omega,2)} > 0 \tag{38}$$

Let us assume that V features decreasing elasticity, that is, given that $h(\omega,1) < h(\omega,2)$:

$$\frac{V'(h(\omega,1))h(\omega,1)}{V(h(\omega,1))} > \frac{V'(h(\omega,2))h(\omega,2)}{V(h(\omega,2))} \tag{39}$$

Thus, we obtain that:

$$\frac{V'(h(\omega,1))h'(\omega,1)}{V(h(\omega,1))} > \frac{V'(h(\omega,2))h'(\omega,2)}{V(h(\omega,2))} \left(\frac{\frac{h'(\omega,1)}{h(\omega,1)}}{\frac{h'(\omega,2)}{h(\omega,2)}}\right) \tag{40}$$

Lemma 12. *For all ω, we have:*

$$\frac{h'(\omega,1)}{h(\omega,1)} \geq \frac{h'(\omega,2)}{h(\omega,2)} \tag{41}$$

Proof. TBA. □

As a result of Lemma 12, we obtain that $V(h(\omega,c))$ is log-supermodular in (c,ω).

A.9 Proof of Proposition 10

The change in percentage points of the share of middle skill is:

$$ds(m,c) = \frac{\left(S\left(T^{-1}\left(h(\omega(c,m),c)\right)\right)\right)' d\omega(c,m) - \left(S\left(T^{-1}\left(h(\omega(c,h),c)\right)\right)\right)' d\omega(c,h)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{42}$$

Given that $d\omega(c,m) > 0$ and $d\omega(c,h) < 0$ when $dp(k) < 0$, we obtain that $s(m,c)$ declines in both cities c. As these shares decline in both cities, it also declines overall.

A.10 Proof of Proposition 11

We want to compute the variation in the share of middle-skill workers in a given city,

$$ds(m,c) = \frac{V\left(h(\omega(c,h),c)\right) d\omega(c,h) - V\left(h(\omega(c,m),c)\right) d\omega(c,m)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{43}$$

We obtain:

$$ds(m,c) = \frac{\left(\frac{V(h(\omega(c,h),c))}{\Gamma(\omega(c,h),c,h)} + \frac{V(h(\omega(c,m),c))}{\Gamma(\omega(c,m),c,m)}\right)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \frac{dp}{p} \tag{44}$$

Using the expressions for $\Gamma(\omega(c,m),c,m)$ and $\Gamma(\omega(c,h),c,h)$, the coefficient can be rewritten as:

$$\frac{\left(V\left(h(\omega(c,h),c)\right) \frac{\omega(c,h)}{\gamma\omega(c,h)-1} + V\left(h(\omega(c,m),c)\right)\omega(c,m)\right)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{45}$$

We have easily that:

$$V\left(h(\omega(1,h),1)\right) \frac{\omega(1,h)}{\gamma\omega(1,h)-1} > V\left(h(\omega(2,h),2)\right) \frac{\omega(2,h)}{\gamma\omega(2,h)-1}. \tag{46}$$

On the other hand, the sign of the following term is ambiguous:

$$V\left(h(\omega(1,m),1)\right)\omega(1,m) - V\left(h(\omega(2,m),2)\right)\omega(2,m) \tag{47}$$

When $A(1,m)$ is close to $A(2,m)$, we then find that a sufficient condition for the decline to be larger in the smaller city is:

$$\frac{\left(V\left(h(\omega(1,h),1)\right) \frac{\omega(1,h)}{\gamma\omega(1,h)-1}\right)}{S\left(T^{-1}\left(h(\underline{\omega},1)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(1),1)\right)\right)} > \frac{\left(V\left(h(\omega(2,h),2)\right) \frac{\omega(2,h)}{\gamma\omega(2,h)-1}\right)}{S\left(T^{-1}\left(h(\underline{\omega},2)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(2),2)\right)\right)} \tag{48}$$

this condition is satisfied when $A(1,h)$ is sufficient large. Indeed, it can be observed that the left hand term diverge to ∞ when $A(1,h) \to \infty$.

B Figures and tables

Figure 1 – Map of France with largest agglomerations in 2015.

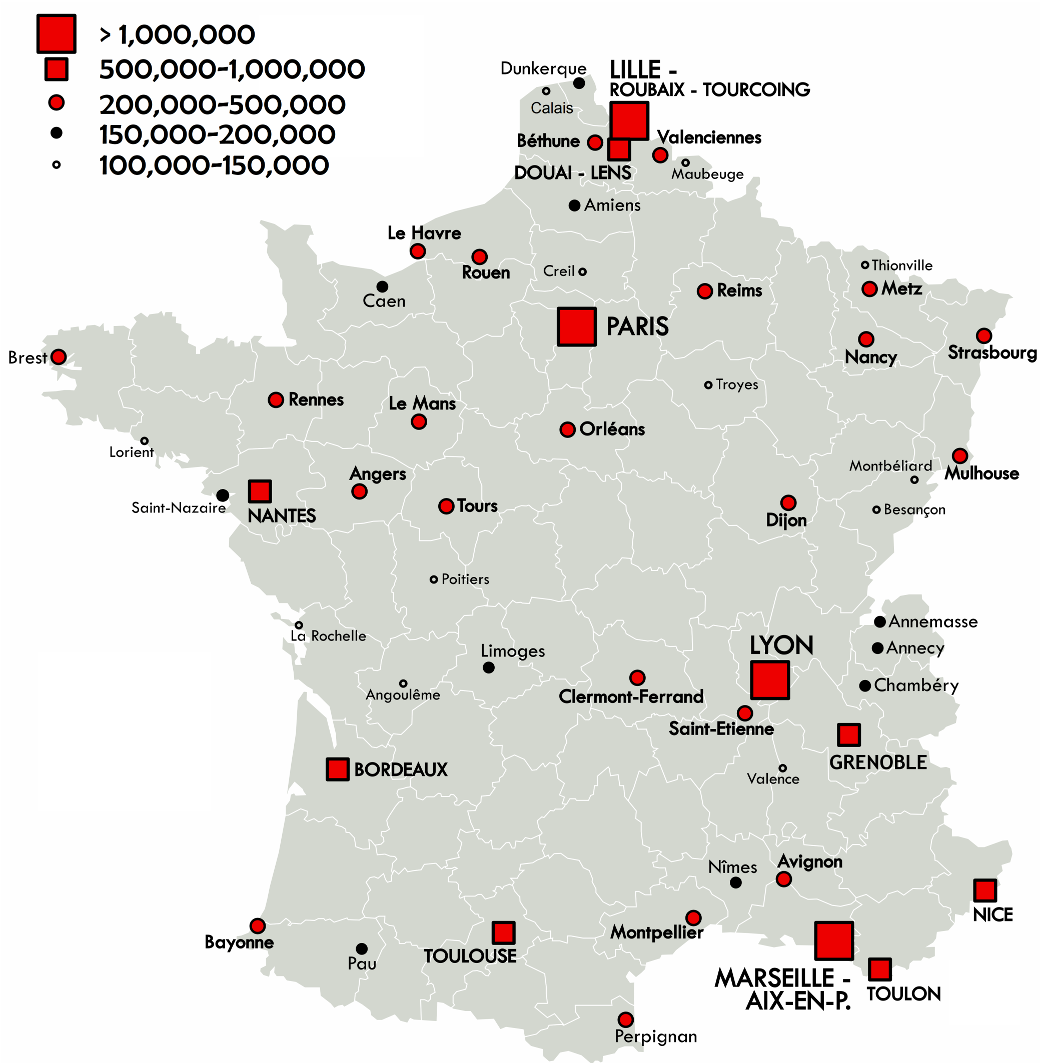


Table 8 – City categories: number of cities, population and the share of hours worked in 2015.

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

Table 9 – Summary statistics at the city level.

Item		year	mean	stdev	min	max
population		2015	293,485	1,007,302	50,571	10,706,072
number of firms with jobs in the city		1994	3,522	141,922	529	13,283
		2015	5,728	222,251	881	20,903
employment share	high paying jobs	1994	0.090	0.027	0.052	0.233
		2015	0.139	0.049	0.080	0.367
	middle-paying jobs	1994	0.775	0.050	0.601	0.872
		2015	0.651	0.059	0.449	0.799
	low paying jobs	1994	0.135	0.041	0.063	0.312
		2015	0.210	0.051	0.084	0.440
	RTI4 jobs	1994	0.419	0.078	0.240	0.625
		2015	0.307	0.055	0.175	0.489
	RTI6 jobs	1994	0.523	0.067	0.355	0.705
		2015	0.451	0.059	0.266	0.581
	OFF6 jobs	1994	0.564	0.065	0.379	0.728
		2015	0.415	0.055	0.258	0.576
employment share percentage change 1994-2015	high paying jobs		0.048	0.029	- 0.005	0.153
	middle-paying jobs		- 0.124	0.030	- 0.204	- 0.003
	low paying jobs		0.076	0.024	- 0.022	0.137
	RTI4 jobs		- 0.112	0.041	- 0.255	0.022
	RTI6 jobs		- 0.072	0.036	- 0.172	0.044
	OFF6 jobs		- 0.150	0.036	- 0.255	0.026
employment share growth 1994-2015	high paying jobs		0.532	0.251	- 0.049	1.508
	middle-paying jobs		- 0.161	0.042	- 0.309	- 0.004
	low paying jobs		0.601	0.231	- 0.097	1.420
	RTI4 jobs		- 0.263	0.074	- 0.411	0.081
	RTI6 jobs		- 0.136	0.065	- 0.298	0.095
	OFF6 jobs		- 0.265	0.057	- 0.387	0.059

Table 10 – Population elasticities by high, middle and low paying categories in the 1999 Census data.

Dependent variable: ln $f(c,\omega)$	All workers	French born	Population share	French born share
High paying X ln pop	1.141	1.138	.17	.85
	(0.036***)	(0.036***)		
Middle-paying X ln pop	0.954	0.949	.64	.94
	(0.024*)	(0.023**)		
Low-paying X ln pop	0.943	0.921	.19	.96
	(0.017***)	(0.014***)		

Notes: 112 cities above 50,000 inhabitants as defined by INSEE as of 1999, population figures from 1999. Exclusions as in main sample. Standard errors, clustered by agglomeration, in parantheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

Table 11 – Population elasticities by 1-digit CS categories in the 1999 Census data.

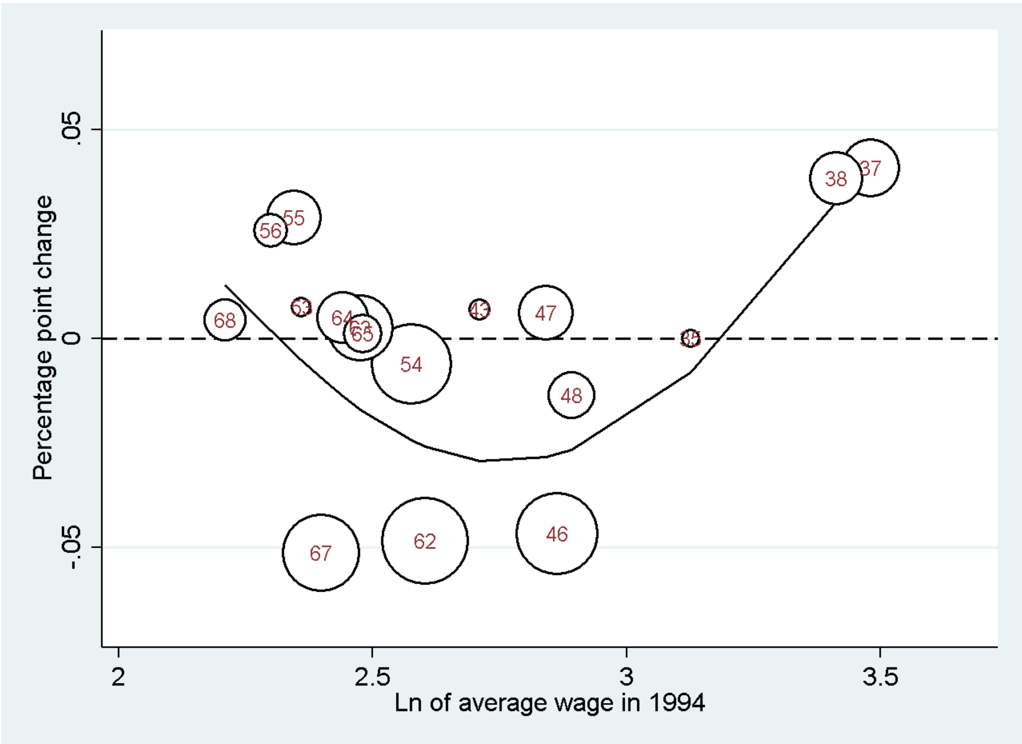
Dependent variable: ln $f(c,\omega)$	All workers	French born	Population share	French born share
Cadres (CS 3) X ln pop	1.157	1.154	.17	.96
	(0.037***)	(0.037***)		
Intermediate professionals (CS 4) X ln pop	1.02	1.018	.28	.97
	(0.026)	(0.025)		
Low-skill employees (CS 5) X ln pop	0.97	0.96	.27	.92
	(0.020)	(0.018**)		
Blue-collar workers (CS 6) X ln pop	0.878	0.856	.26	.86
	(0.026***)	(0.026***)		

Notes: 112 cities above 50,000 inhabitants as defined by INSEE as of 1999, population figures from 1999. Exclusions as in main sample. CS 23 category – CEOs – not included in the category "cadres". Standard errors, clustered by agglomeration, in parantheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

Table 12 – Share of 3 lowest-paying occupations per agglomeration size.

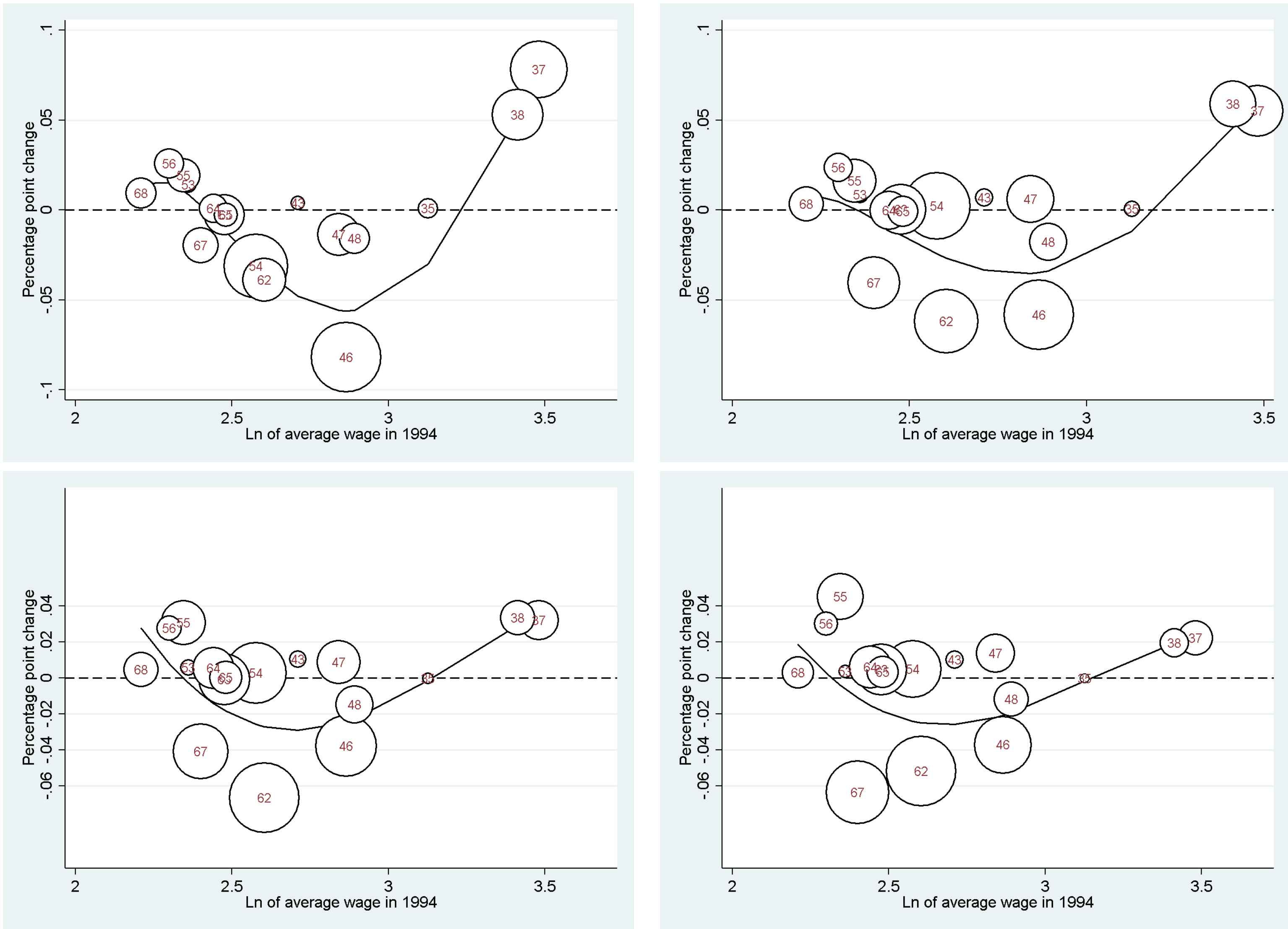
Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.107	0.116	0.118	0.118	0.114	0.122
2015	0.161	0.164	0.168	0.182	0.192	0.200
change in ppct	0.054	0.049	0.050	0.063	0.079	0.078
growth in %	0.508	0.423	0.425	0.535	0.692	0.643

Figure 2 – Labor market polarization in France 1994-2015.



Notes: The figure shows the percentage point change in employment of the considered 2-digit CS categories plotted against their average wage in cities > .05m in 1994. Numbers pertain to 2-digit CS categories represented. Circle sizes correspond to the employment shares. The line shows a cubic relationship between the average wage and the percentage point change. The CS category "23" - CEOs excluded.

Figure 3 – Labor market polarization across cities 1994-2015.



Left-upper panel: Paris. Right-upper panel: Lyon. Left-lower panel: cities between 0.2-0.5m. Right-lower panel: cities between .05-.1m inhabitants.
Notes: Figures show the percentage point change in employment of the considered 2-digit CS categories plotted against their average wage in cities > 0.05m in 1994. Numbers pertain to 2-digit CS categories represented. Circle sizes correspond to the employment shares. The line shows a cubic relationship between the average wage and the percentage point change. The CS category “23” - CEOs excluded.

Figure 4 – Comparing labor market polarization 1994-2015 in Paris and cities between .05-.1m inhabitants.

Notes: The figure shows the percentage point change in employment of the considered 2-digit CS categories plotted against their average wage in cities > .05m in 1994. Numbers pertain to 2-digit CS categories represented. Grey circles stand for Parisian while white for small city shares. Circle sizes correspond to the employment shares (same scale for the two compared groups). The two lines shows a cubic relationship between the average wage and the percentage point changes in employment for Paris (red) and cities between .05-.1m inhabitants (black) respectively. The CS category "23" - CEOs excluded.

Table 13 – Comparison of means of employment shares in different occupations, cities >0.5m vs. 0.05-0.1m. Sample with finance, insurance and real estate sectors.

Item	RTI4	RTI6	OFF6	middle-paid	low-paid	high-paid
1994						
mean, cities >0.5m	0.326	0.401	0.509	0.690	0.123	0.187
mean, cities 0.05-0.1m	0.423	0.534	0.566	0.778	0.141	0.081
difference in pct points	-0.097***	-0.133***	-0.057***	-0.088***	-0.019**	0.107***
2015						
mean, cities >0.5m	0.219	0.310	0.331	0.509	0.187	0.303
mean, cities 0.05-0.1m	0.314	0.471	0.421	0.663	0.220	0.117
difference in pct points	-0.095***	-0.161***	-0.090***	-0.153***	-0.033***	0.187***

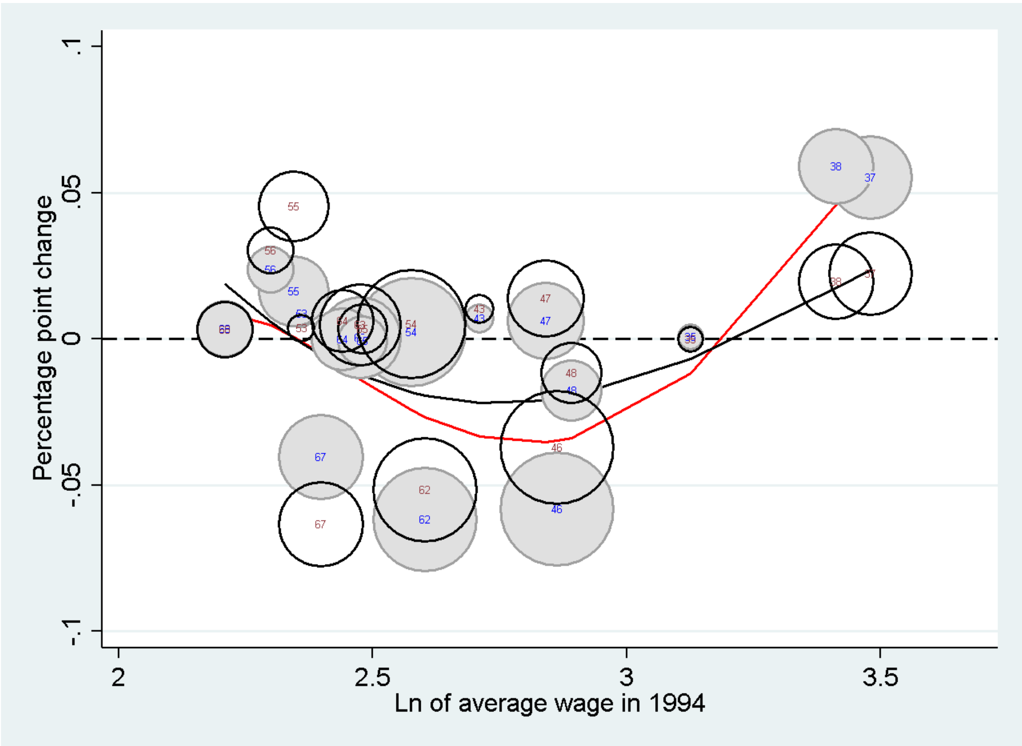
Notes: population weighted, allowing for unequal variances. N=73 (11 cities > 0.5m). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table 14 – Comparison of means of changes in different occupations, cities >0.5m vs. 0.05-0.1m. Sample without finance, insurance and real estate sectors.

Item	RTI4	RTI6	OFF6	middle-paid	low-paid	high-paid
mean change ppct, >0.5m	-.109	-.09	-.173	-.177	.072	.105
mean change ppct, 0.05-0.1m	-.117	-.069	-.147	-.115	.084	.031
difference in pct points	.008	-.021***	-.026***	-.062***	-.012**	.074***
mean growth, >0.5m	-.335	-.224	-.354	-.259	.562	.60
mean growth, 0.05-0.1m	-.268	-.124	-.263	-.148	.628	.415
difference in growth	-.067***	-.100***	-.091***	-0.111***	-.066	.185***

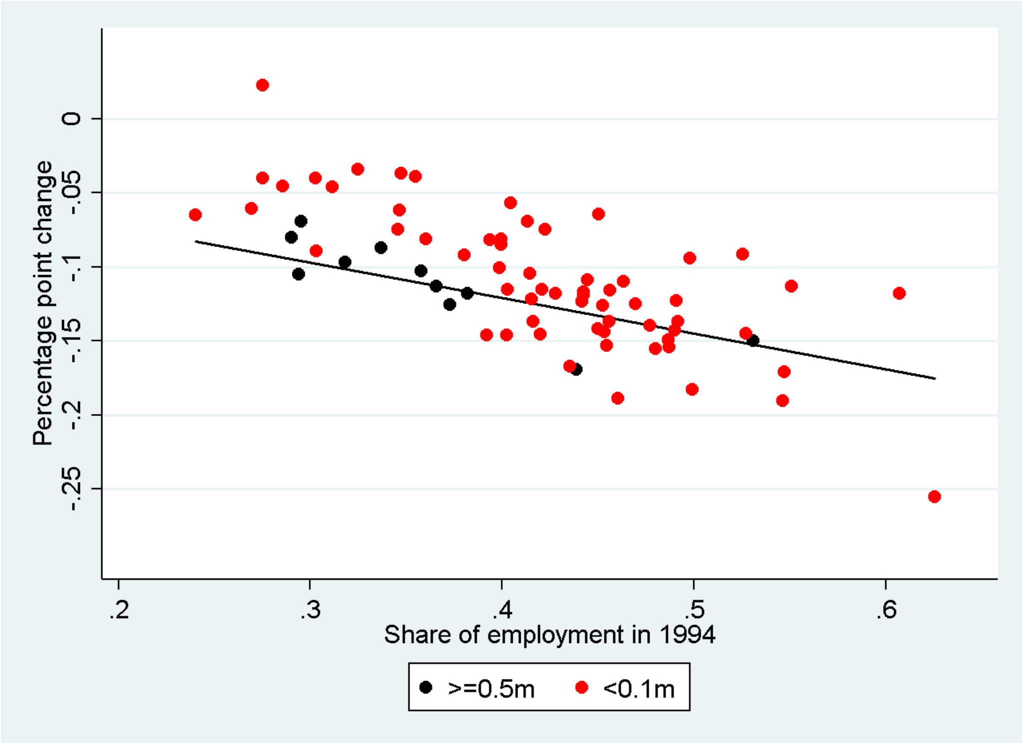
Notes: population weighted, allowing for unequal variances. N=73 (11 cities > 0.5m). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Figure 5 – Comparing labor market polarization 1994-2015 in Lyon and cities between .05-1m inhabitants.



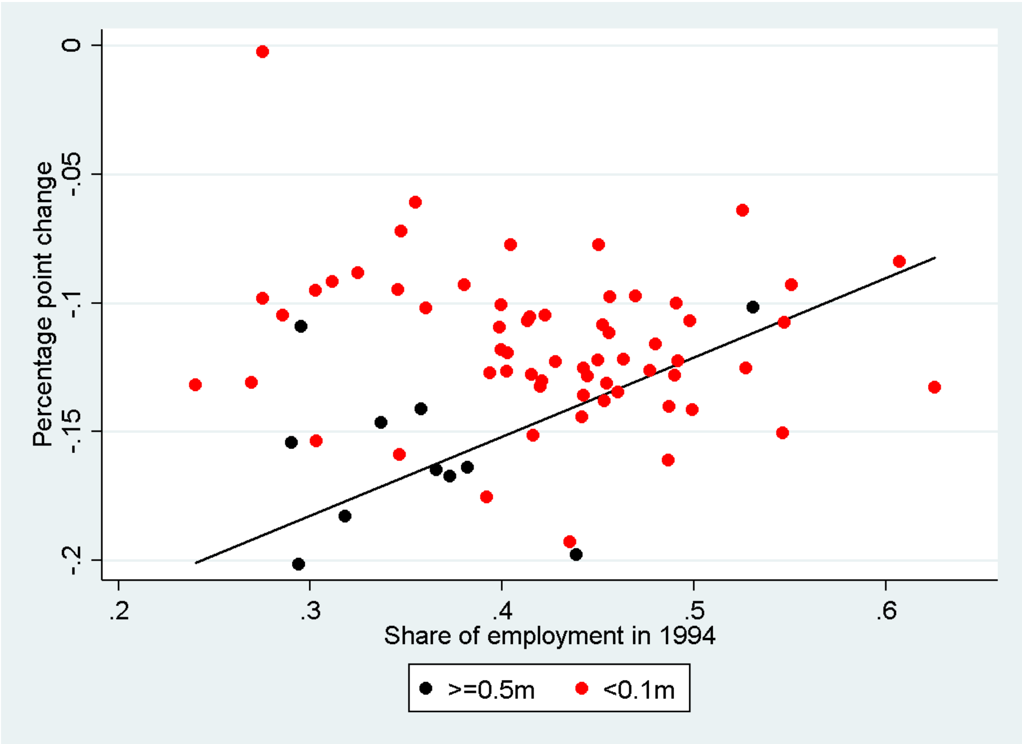
Notes: The figure shows the percentage point change in employment of the considered 2-digit CS categories plotted against their average wage in cities > 0.05m in 1994. Numbers pertain to 2-digit CS categories represented. Grey circles stand for Lyon while white for small city shares. Circle sizes correspond to the employment shares (same scale for the two compared groups). The two lines shows a cubic relationship between the average wage and the percentage point changes in employment for Lyon (red) and cities between 0.05-0.1m inhabitants (black) respectively. The CS category "23" - CEOs excluded.

Figure 6 – Exposure to RTI4 jobs and change in RTI4 jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



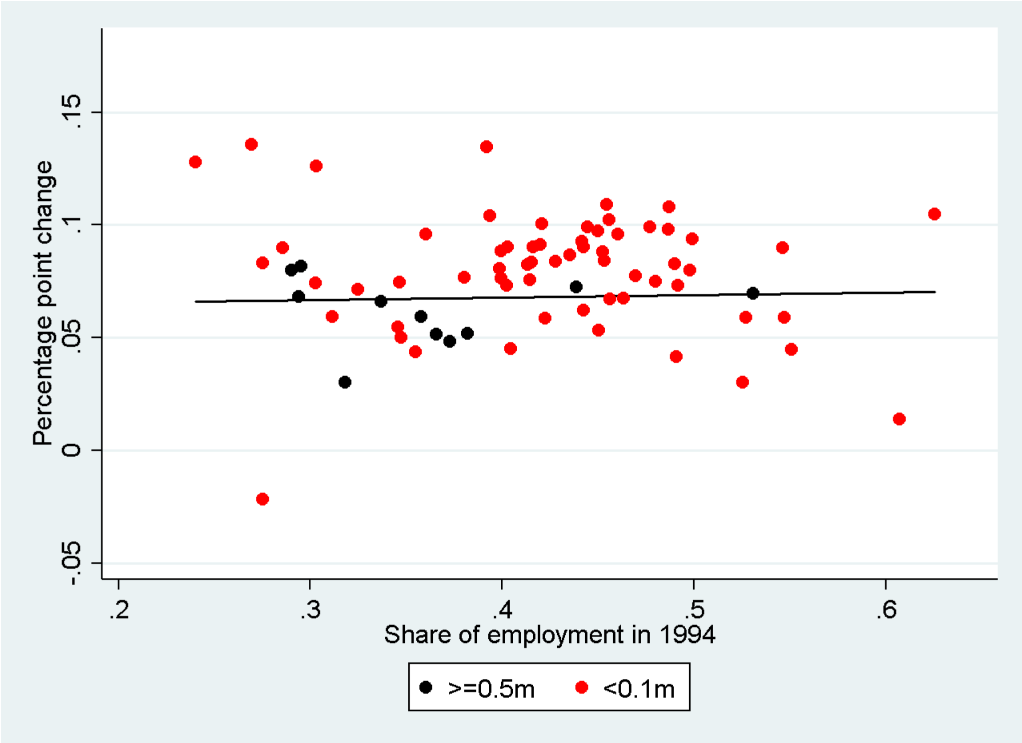
Notes: The figure shows the percentage point change in employment of RTI4 jobs (CS 48, 54, 62 and 67) between 1994-2015 plotted against the share of RTI4 jobs in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted fit of the relation between employment changes and the initial RTI4 exposure.

Figure 7 – Exposure to RTI4 jobs and change in middle-paying jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



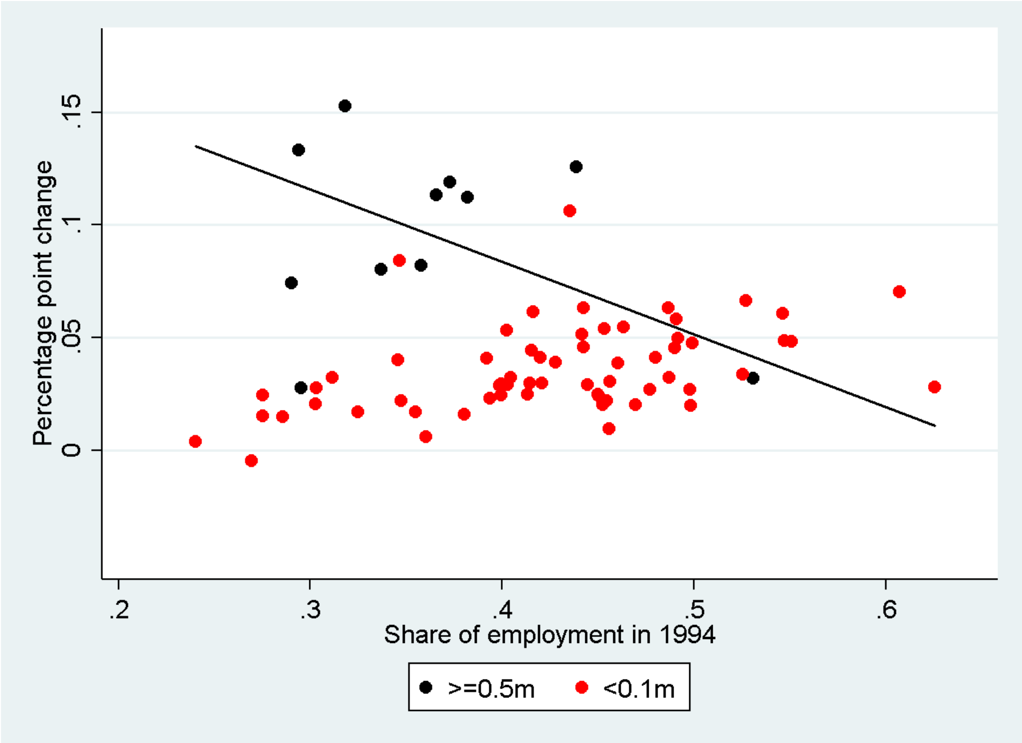
Notes: The figure shows the percentage point change in employment of middle-paying jobs between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above 0.5m (between 0.05-0.1m). The line shows a linear, population-weighted fit of the relation between employment changes and the initial RTI4 exposure.

Figure 8 – Exposure to RTI4 jobs and change in low-paying jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



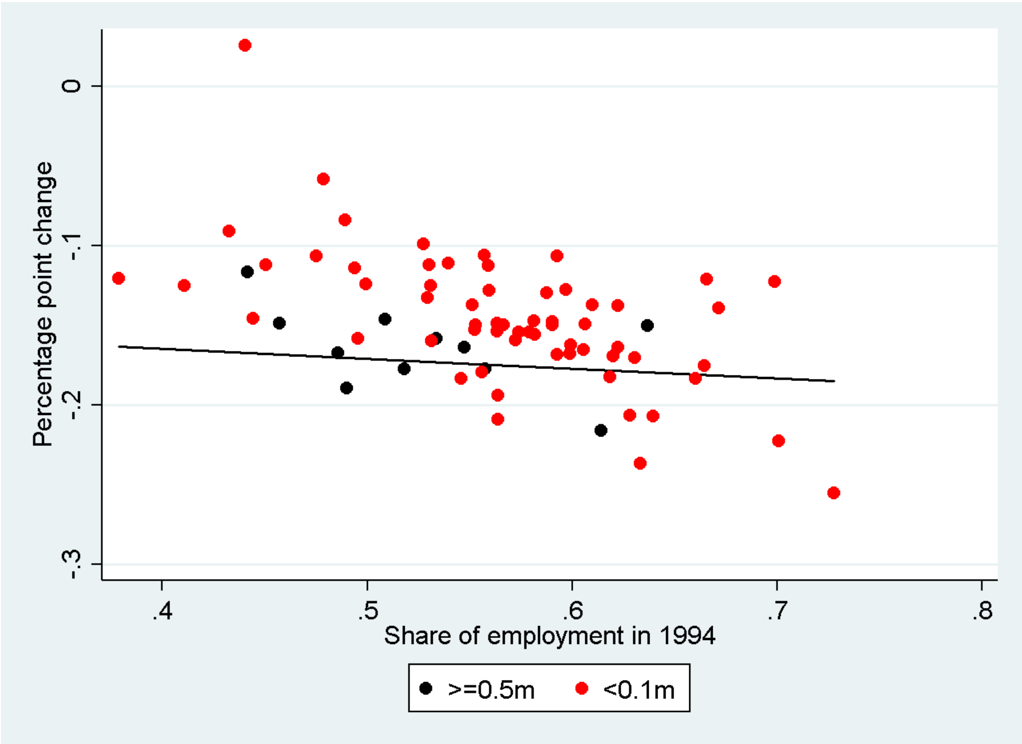
Notes: The figure shows the percentage point change in employment of low-paying jobs (CS 53, 55, 56 and 68) between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted fit of the relation between employment changes and the initial RTI4 exposure.

Figure 9 – Exposure to RTI4 jobs and change in high-paying jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



Notes: The figure shows the percentage point change in employment of high-paying jobs (CS 23, 35, 37 and 38) between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted fit of the relation between employment changes and the initial RTI4 exposure.

Figure 10 – Exposure to 6 most offshorable jobs and their employment share change in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



Notes: The figure shows the percentage point change in employment of jobs with highest OFF-GMS index (CS 46, 48, 54, 62, 65, 67) between 1994-2015 plotted against the share of such jobs in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted fit of the relation between employment changes and the initial exposure.