Labor Market Polarization and The Great Divergence: Theory and Evidence*

Donald R. Davis Eric Mengus Tomasz Michalski

June 24, 2019

Incomplete, please do not circulate

Abstract

Two of the most important features of advanced labor markets in the past quarter century are labor market polarization and the great divergence. The first of these concerns the growth of jobs in high and low wage categories and the disappearance of middle wage jobs. The second is an explicitly spatial theory about the intensification of development particularly at the high end in large, already developed cities relative to smaller, less developed cities. This paper addresses how the two phenomena are interrelated. The great divergence is typically contemplated in a two factor setting with skill-biased technical change. Labor market polarization is instead considered in an explicitly three-factor setting, specifically rejecting the simpler framework. We develop a theory in which the driving forces of labor market polarization alone give rise to both phenomena, building on Autor and Dorn (2013), and Davis and Dingel (2014). Key to this is that the productivity advantages in large cities are biased toward high skilled tasks, so that a uniform shock to technology leads to labor market polarization with a biased impact on cities of different sizes, giving rise to the great divergence. We examine the model using detailed data for a sample of 117 French cities and find the patterns in the period 1994-2015 accord well with the theory.

Keywords: Labor Market Polarization, Great Divergence, System of Cities, Inequality

JEL Classification: J21, R12, R13

*Davis: Columbia and NBER (email: drdavis@columbia.edu); Mengus (mengus@hec.fr) and Michalski (michalski@hec.fr): HEC Paris and GREGHEC. We would like to thank Iain Bamford, Dominik Barthelme and Wolfgang Dauth as well as seminar participants at UAB, IGC, SMU, EIEF, UEA annual meeting 2018 (New York), Joint French Macro workshop 2018, NBER ITI Spring 2019, 2019 European meetings of the UEA (Amsterdam) for helpful comments and discussions. This work was supported by Investissements d'Avenir [grant number ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047]. All remaining errors are ours.

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

In this section, we build a model integrating the core framework of job polarization in Autor and Dorn (2013) with the system of cities model of Davis and Dingel (2014). This allows to obtain in equilibrium key features of data. First of all, it permits to have a log-supermodular distribution of skills across cities. Second, polarization of the job market occurs (both in the aggregate and in each city) when the price of capital or offshored tasks decreases. Furthermore, we provide conditions under which the labor market polarization is stronger (*magnitude effect*) and high-skill biased in the large city (*reallocation effect*) – with more destruction of middle-skilled jobs and more creation of high-skilled jobs – despite an initially lower share of exposure to middle-paying jobs. These conditions are that skills matter sufficiently more in high-paying occupations and that the comparative advantage of high-paying workers to live and work in the large city is sufficiently large.

2.1 The environment

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

Locations. The set of cities is $c \in \{1, 2\}$.[1]

In each city, there is a continuum of locations $\tau \in [0, \infty)$. τ denotes the distance from an ideal location inside a city. This can be interpreted in a variety of ways, including as commuting distance

[1] We extend our framework to N cities in Appendix A.2.

to a central business district, or as remoteness from the core of a productive cluster with positive spillovers. As this will become clear, having multiple locations within a city allows to introduce a trade-off between living in a better location in a smaller and less productive city or in a worse location ina larger and more productive city.[2]

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

Households. Households consume a single final good and 1 unit of housing.

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

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 $r(c, \tau) \geq 0$. As in Davis and Dingel (2014), locations are owned by absentee landlords who spend their rental income on the final good.s

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

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/or an offshorable input with the rest being technological parameters. As in Autor and Dorn (2013), we assume that capital/offshoring goods are relative substitutes with intermediate goods produced by the middle-skill sector but they are relative complements with the the intermediate goods produced by the low- and the high-skill sectors, that is $\gamma < \theta$.

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

$$\Pi = Q - p(h)q(h) - p(m)q(m) - p(l)q(l) - p(z)Z \tag{2}$$

[2]For further interpretations of the location τ and the connection with other models of cities in the literature, we refer the interested reader to Davis and Dingel (2014).

Intermediate goods. The intermediate goods $\{h, m, l\}$ 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 higpaying, m for middle-paying and and l for low-paying.

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

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

On this productivity, we make the following assumptions. First, $T(.)$ is a decreasing function identifying the cost in productivity of being remote from the most productive location in a city. Second, $H(\omega, \sigma, c)$ is log supermodular in (ω, σ) and $H(., \sigma, c)$ is increasing. As a result of log-supermodularity in (ω, σ), $H(\omega, m, c)/H(\omega, l, c)$ and $H(\omega, h, c)/H(\omega, m, c)$ are increasing functions of ω. On top of this assumption, we assume that H is such that skills matter much more for the high-paying occupation than for any other sectors. This amounts to assuming that $H(\omega, h, c)/H(\omega, m, c)$ is sufficiently more convex than $H(\omega, h, c)/H(\omega, m, c)$. To capture this idea, we assume H has the following functional form:

$$H(\omega, l, c) = A(l, c)\omega^{\phi} \text{ with } \phi \in (0, 1), \tag{4}$$

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

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

A similar assumption can be found in Autor and Dorn (2013) where skills matter relatively more for the middle-paying sector compared with the low-paying sector. More generally, this captures the idea that a marginal variation in skill is of very low importance for low-paying occupations; it starts to be more important in middle-paying occupation but it is key in high-paying occupations.[3]

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 focus of this paper. Davis and Dingel (2014, forthcoming) provide alternative approaches to endogenous productivity differences across cities even when fundamentals are symmetric.[4]

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

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

[3]See, for example, the literature on superstars that followed the seminal contribution by Rosen (1981).

[4]In appendix A.1, we provide a way to obtain endogenous productivity differences consistent with patterns of labor market polarization.

and a comparative advantage in higher-paying sectors:

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

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-skill 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, $p(l)$, $p(m)$ and $p(h)$.

The utility flow obtained by an agent with ability ω, location decisions (c,τ) and intermediate 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 ability 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 $\omega(h,c)$:

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

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

The threshold $\omega(m,c)$ is such that a household in a given city c is indifferent between working in the low-paying and the middle-paying sectors and the threshold $\omega(h,c)$ is such that the same household is indifferent between the middle- and the high-paying 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 ability ω works in sector l when $\omega \leq \omega(m,c)$, in sector m when $\omega \in (\omega(m,c),\omega(h,c))$ and in sector h when $\omega \geq \omega(h,c)$.*

Across cities, these thresholds satisfy:

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

Proof. See Appendix B.1 □

The differences in the thresholds across cities result from the comparative advantage of the larger cities in higher paying sectors associated with the increasing importance of individual abilities in higher paying sectors. For the same prices of intermediate goods, the same individual is relatively more productive in higher paying sectors in the larger city and, thus, has 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 $\omega(m,1) \leq \underline{\omega}$. In this case, there is no low-skill sector in city 1.

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

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

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

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

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 occupied 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 $\tau \in [0,\overline{\tau}(l,c)]$:*

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

Proof. See Appendix B.2 □

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

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

In addition, 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 B.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 specific to the city.

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 set of skills 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 $\omega \geq \omega(h,c)$ then $\tau \leq \overline{\tau}(h,c)$,*

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

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

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

Proof. See Appendix B.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 the skill ω as:

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

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 B.5 □

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.

In the end, Figure 1 summarizes the distribution of skills across sectors and cities.

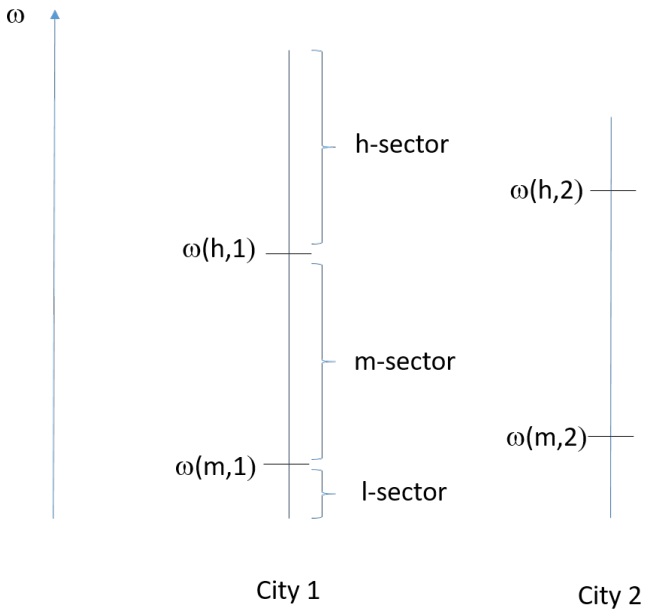


Figure 1 – Skills, sectors and cities

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 the relative prices $p(m)/p(h)$ and $p(l)/p(m)$:

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

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

Proof. See Appendix B.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 $\omega(h,1)$ in the larger city.[5]

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 $\omega(m,2)$ to increase by more than $\omega(m,1)$ for the same variation of intermediate good prices.

Figure 2 summarizes these findings.

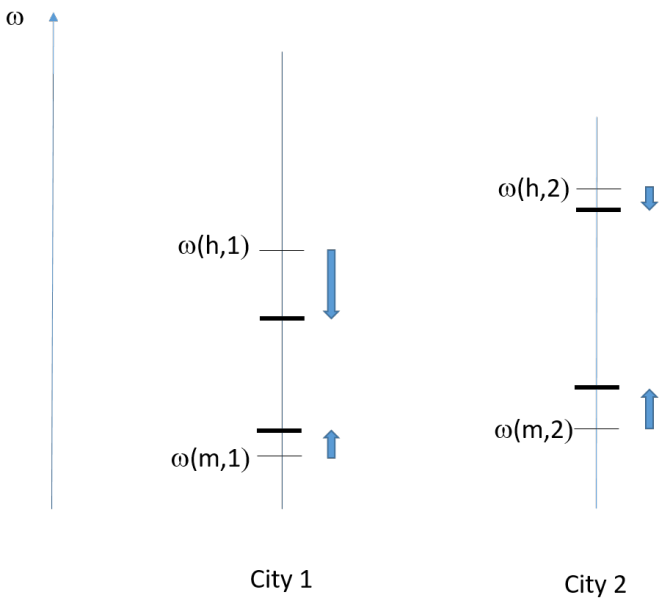


Figure 2 – The effects of a decline in the price of capital/offshored goods

2.3 Implications

We are now able to describe the main implications of our model. To start with, we describe the distribution of skills and sectors across cities. Then we investigate the effect of a labor market polarization shock (a decline in the price of the capital/offshoring good p_z) both in the aggregate and across cities. Finally, we provide further description of what happens to middle-paying jobs and to high-paying ones. In particular, we show that a labor market polarization shock can contribute to the Great Divergence across cities.

2.3.1 Allocation of skills and city exposure to the middle-skill sector

Let us first describe the allocation of skills and the exposures to different sectors across cities.

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

Initial exposure to middle-paying 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 1. *When* $\frac{A(h,1)}{A(h,2)}$ *is sufficiently large relative to* $\frac{A(m,1)}{A(m,2)}$*, the share of middle skill sector jobs is smaller in the larger city.*

Proof. See Appendix B.7. □

The lower exposure of the large city to middle-paying jobs reflects this city's comparative advantage in high-paying jobs. Indeed, 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 more exposed to high-paying jobs and less exposed to middle-paying jobs.

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

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

Proof. See Appendix B.8 □

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

Given our previous result in Proposition 1 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 1. *Under the conditions of Proposition 1, the elasticity for middle-paying workers with respect to the size of the city is lower than* 1*.*

One implication of this result associated with the fact that larger cities have a lower initial share of middle-paying workers as shown in Proposition 1 is that occupations do not need to be log-supermodular as this is the case for skills. More specifically, the total number of jobs in the middle-paying occupations may be lower in the larger city compared with the smaller one.

2.3.2 Labor market polarization in the aggregate and across cities

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). To start with, let us clarify the effect of a shock to the price of capital/offshoring intermediary goods affect the relative prices of the middle-skill sector with the high- and the low-skill sectors:

Lemma 7. *A decline in p_z leads to a decline of the relative prices of the middle-skill sector good: $p(m)/p(h)$ and $p(m)/p(l)$.*

Using this pattern of the relative price, let us investigate how the shock on p_z translates to the labor markets in the two cities.

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. From Lemma 6, we can infer that middle-paying jobs decline at both margins in both cities. This leads to the following proposition:

Proposition 3. *A decline in p_z reduces the share of middle-paying jobs in the aggregate and in each city.*

Proof. See Appendix B.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 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-paying 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 4. *When $\frac{A(h,1)}{A(h,2)}$ is sufficiently large relative to $\frac{A(m,1)}{A(m,2)}$, then a decline in p_z implies that in the large relative to the small city:*

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

(ii) *The increase in high-paying jobs is larger in percentage points.*

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

Proof. See Appendix B.10. □

Under the conditions of Proposition 4, labor market polarization is then not uniform across cities. More precisely, it differs both in terms of *magnitude* and in terms of *reallocation*. In terms of magnitude, the shock to middle-paying jobs is stronger in larger cities. In terms of reallocation, the reallocation following the shock is tilted towards high-paying jobs in the large city, while is the opposite in the small city.

Let us provide some intuition for these results. The decrease in the price of capital/offshored goods corresponds to a negative demand shock for middle-paying jobs. The effects of this shock are the strongest where the opportunity cost of being a middle-paying is the largest. For each level of skill ω, this opportunity cost is the one associated with not selecting a high- or a low-paying job.

When the comparative advantage of being a high-paying the large city is sufficiently large – that is when $A(h,1)/A(h,2)$ is sufficiently large compared with $A(m,1)/A(m,2)$ but also compared with $A(l,1)/A(l,2)$ given Assumption 1 – the opportunity cost to be a middle-paying is the highest for agents in the large city that are close to selecting a high-paying job. Thus, in that case, a decrease in the price of capital/offshored goods leads to a stronger decline in middle-paying jobs in the largest city and, for the same reason, in a larger increase in the share of high-paying jobs. In contrast, the smaller city has a comparative advantage in low-paying jobs, which leads to an increase in the share of these jobs in that city. Note that this result would not appear in Davis and Dingel (2014), given that there is no specific comparative advantage for high-paying jobs to be in the larger city.

To obtain these results, the mobility of workers across cities and sectors with play a key role: agents need to find optimal to live in a city and work in a given sector. In the absence of mobility of workers across sectors, the specialization of the large city in high-paying activities would mechanically lead the smaller city to host more middle-paying jobs – as a result, no labor market polarization would occur in that smaller city, which we will find to be at odds with what we observe in the data. The mobility of workers also ensures that the stronger destruction of middle-paying jobs in the large city does not lead to a stronger creation of low-paying jobs in that city. This contrasts with Autor and Dorn (2013), where for low levels of skill, there is mobility only between the low- and the middle-paying sectors and the corresponding workers cannot move from one city to another: there, the stronger destruction of middle-paying jobs leads to more creation of low-paying jobs in the large city.

Remark. We have obtained proposition 4 as an asymptotic result on the comparative advantage of the large city for the high-paying jobs in a context where productivities are relatively more elastic to skills for these high-paying jobs. In this way, our result does not depend on the skill distribution ($n(.)$). There exist alternative sets of conditions to obtain a polarization skewed as in Proposition 4. For example, we may obtain such an outcome in the case where there is only a constant absolute advantage to be in the large city but in which the distribution of skills is sufficiently skewed to

high-skill middle-paying jobs in the large city.

2.3.3 The effect of the labor market polarization on middle-paying jobs.

Let us investigate further the consequences of labor market polarization on middle-paying jobs. Our first important result is that the initial exposure to these jobs is not a predictor of the strength of their destruction. We then investigate which middle-paying jobs are destroyed and how this evolves across cities. Finally, we describe how the labor market polarization shock modifies the distribution of income among workers that were initially occupying middle-paying jobs.

Initial exposures to middle-paying jobs and labor market polarization Combining the result of Propositions 1 and 4 leads to the following corollary:

Corollary 2. *The destruction of middle-paying jobs is the largest in percentage points in the large city where there is, initially, the lower share of middle-paying jobs.*

Exposure is not the key driver that explains the destruction of middle-paying jobs. Our interpretation is that technology or offshoring are necessary ingredients for the destruction of middle-paying jobs but they are not sufficient and one only needs to think about incentives to destroy these jobs. In our model, the incentives to destroy middle-paying jobs depend on cities and the opportunity cost of keeping these jobs rather than creating new ones in other sectors.

A direct implication of Corollary 2 is that we cannot instrument future job destruction only by exposures or any other feasibility constraints for these destructions.

The heterogeneity among middle-paying jobs across cities Given that exposure is not the key driver of middle-paying job destructions, we further analyze in this paragraph the nature of middle-paying jobs that are destroyed in large and small cities.

Workers occupying middle-paying jobs are heterogenous with respect to their skills – they have different values for ω. It is then possible to further analyze how labor market polarization affects middle-paying jobs across cities depending on workers' skills. In this paragraph, we show that, in large cities, it is the most skilled (i.e. with the highest ω or, equivalently, with the highest wage) middle-paying jobs that are destroyed and replaced by high-paying jobs. In small cities, it is mainly the least skilled middle-paying jobs that are destroyed and replaced by low-paying jobs.

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

Proposition 5. *The share of higher wage middle-paying jobs decreases by more in percentage points in the large city.*

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

Proof. See Appendix B.11. □

What stands behind this proposition is the relative behaviors of the thresholds $\omega(h,c)$ and $\omega(m,c)$ across cities. These thresholds correspond to the indifference condition between, respectively, the high- and the middle-paying sectors and the middle- and the low-paying sectors. As shown by Lemma 6, the upper threshold $\omega(h,c)$ decreases by more in the large city (city 1) and the lower threshold $\omega(m,c)$ increases by more in the small city (city 2). As a result, more higher wage middle-paying jobs are destroyed in the large city following a labor market polarization shock and, similarly, more lower wage middle-paying jobs are destroyed in the small city.

The heterogeneity among middle-paying jobs and sectoral decisions Middle-paying sector workers also differ in terms of the income that they lose from the labor market polarization shock. In the current discussions (see Autor, 2019), most of the concerns on the consequences of technology and offshoring are concentrated on the outcomes of workers shifting from the middle- to the low-paying sector. In this paragraph, we revisit this result through the lens of our model.

Proposition 6. *In relative terms, workers that remain in the middle-paying sector face a larger income shock compared with middle-paying workers that shift to either the low-paying sector or to the high-paying sector.*

In particular, among agents initially in the middle sector, the least skilled agents shifting to low-paying sector face a smaller income decline in relative terms than the least skilled agents shifting to the high-paying sector.

The intuition of this result is straightforward: a worker in the middle-paying sector that stays in the middle-paying sector experience an income change proportional to the variation in the price of the middle-paying sector good: $d(p(m))/p(m)$. In contrast, a worker that change sector experience a smaller decline. Let us illustrate this in the case of a worker that shifts to the low-paying sector. To start with, this shift results from the fact that this agent receives a relatively larger nominal wage in the low-paying sector compared with the middle-paying sector, that is:

$$H(\omega,l,c)(p(l)+dp(l)) > H(\omega,m,c)(p(m)+dp(m)) \tag{17}$$

Using this inequality, we then obtain:

$$\frac{H(\omega,l,c)(p(l)+dp(l)) - H(\omega,m,c)p(m)}{H(\omega,m,c)p(m)} > \frac{dp(m)}{p(m)}. \tag{18}$$

Figure 3 allows to summarize this first result.

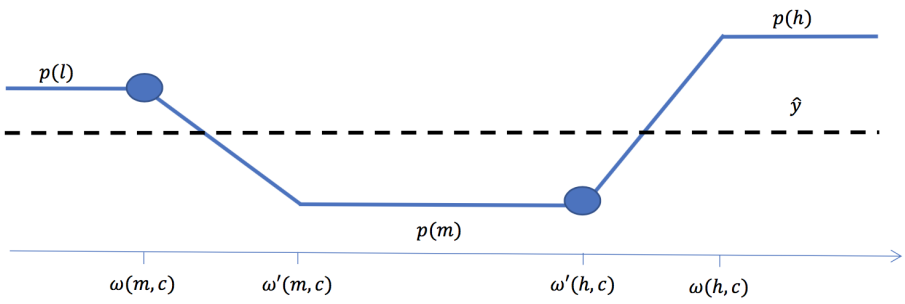


Figure 3 – Relative income variations across the skill distribution.

Note: $\hat{y}$ denotes the average wage growth following the shock on p_z. This graph is for arbitrary growth values for $p(l)$ and $p(h)$. In general, we only know that $p(m)$ declines relatively to $p(h)$ and $p(l)$. See Section 2.3.4 for a further discussion of this point.

An important implication is then that workers are actually partially protected by the fact that it is worth moving down. Workers that remain in the middle-paying sector do not have the same protection.

This protection is also different depending on how far an agent is from the initial indifference condition between two sectors. The further an agent is from this indifference condition, the larger should be the relative income decline to lead the agent to shift to the other sector. To me more concrete, Proposition 6 focuses on the situation of the least skilled agents shifting to the low-paying sector and the least skilled agents shifting to the high-paying sector. These two situations are illustrated by Figure 3 using blue circles.

In the former case, this agent immediately shifts to the low-paying sector thus mitigating the income loss. In the latter case, the least skilled agent shifting to the high-paying sector, because that agent is at the boundary between not shifting, faces almost the same income loss as agents staying in the middle-paying sector. By continuity, this result holds true for a continuum of agents shifting to the high- and to the low-paying sectors.

In the end, these conclusions then imply that some agents working in the high-paying sectors can be hit as severely as middle-paying workers by a labor market polarization shock, at least more than some agents that shift to the low-paying sector. This conclusion holds true, despite the fact that agents shifting to the high-paying sectors have on average a higher wage compared with agents in the low-paying sector: evaluating welfare effects cannot be summarized only by differences in sectoral wage means.

An important ingredient to make these comparisons is the possibility for middle-paying workers to shift to the high-paying sector. In particular, our conclusions that we obtained cannot be obtained in Autor and Dorn (2013) model where middle-paying workers can move only to the low-paying sector. From a more economic perspective, our results then contrast with the statements in Autor

(2019) that we should not be concerned by the ones who move up to the high-paying sector but mainly with the ones who move to the low-paying sector.

2.3.4 The effects on high-paying jobs and the great divergence

Let us now turn to what happens to high-paying jobs. Our main conclusion is the great divergence across cities.

The Great Divergence In the end, the decline in the price of capital/offshored goods leads to a reinforcement of the largest city into high-paying jobs: it is indeed in that city with the lowest exposure to middle-paying jobs (and the more specialized in high-paying jobs as noted by Proposition 1) that further specializes in high-paying jobs and reduces the number of middle-paying jobs. A direct implication from Proposition 4 is then the Great Divergence as dubbed by Moretti (2012). Indeed, the fact that larger cities are initial richer in high-paying jobs, we obtain that:

Corollary 3. *The share of high-paying jobs increases by more in cities with an initially larger share of high-skill jobs.*

By being skewed towards higher-paying jobs in larger cities, labor market polarization then contributes to make larger cities even richer in higher-paying jobs.

Supply or demand shock? What is the effect of labor market polarization on the wages of high-paying occupations? In Autor and Dorn (2013), the number of workers occupying high-paying jobs is assumed to be fixed. Thus, their wages are necessarily increasing: the labor market polarization shock is thus only a positive demand shock on high-paying labor given that this labor is a complement to capital. In our setting, this complementarity remains, but the possibility to shift to the high-paying sector also implies that more workers enter in the high-paying sector. As a result, the demand shock on high-paying labor is potentially mitigated by a supply shock for these jobs, so that the price of the high-paying good $p(h)$, and so the wage in the high-paying sector, can either increase or decrease.[6]

Let us now confront the predictions of our model with data from France on the period 1994-2015. In Section 3, we describe the different datasets that we are using and we report in Section 4 the different evidence on the distribution of skills and labor market polarization in France.

3 Data description

We focus on a few key questions: the characteristics and the evolution of labor market polarization in the aggregate and by metropolitan area and the log supermodularity of skills. These require

[6] A similar conclusion holds true for the low-paying sector that also faces both a demand and a supply shocks of workers. Nevertheless, the relative prices $p(m)/p(l)$ and $p(m)/p(h)$ both always decrease.

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 metropolitan area 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.[7] This data is provided by INSEE, the French national statistical institute, and is based on mandatory annual reports by all French companies.[8] 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.[9] 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.[10] 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.[11]

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

[7]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.

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

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

[10]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 metropolitan area-level data. Exclusion of these sectors from our analysis does not change our results considerably and does not impact our conclusions. We include Table 17 without these sectors as a replication of Table 8 in a robustness test.

[11]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.

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 (see Table 21 for representative occupations within each category).[12] 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 (based on the RTI measure of Autor et al. (2003)) occupation and offshorability (OFF-GMS) ranking from Goos et al. (2014). The exact index values for each category are given in the Appendix Table 22.

To classify high, medium and low-paying occupations and obtain the exposures to automation and offshoring we proceed as follows. We first merged into the 1994 French Labor Survey the exposure classifications of Goos et al. (2014) and mapped their 2-digit ISCO-based ones into the 2-digit CS used in French data.[13] 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 the low-paying occupations in the 2-digit CS classification, we retained those for which the share of low-paying occupations classified as such in Goos et al. (2014) in terms total hours worked in 1994 was over 50%. The resulting 4 occupations are indeed ranked as the least paid in 1994 in our data, and they are typically much less routine and offshorable than the least-paid middle-paying job category "unskilled industrial workers".

The low-paying occupations are then security workers, retail workers, personal service workers and unskilled manual laborers (CS 53, 55, 56, 68).[14] 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) based on actual offshoring patterns, 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)

[12]See Caliendo et al. (2015) for the use of CS 1-digit categories to analyse firms' hierarchies. Firms should report their data using much finer 4-digit codes, but many fail to do so especially before 2003. 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.

[13]The ISCO and CS categories are both available only in the French Labor Survey and not directly in the DADS data. We used hours worked in 1994 in the Survey as weights.

[14]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.

Table 1 – Basic statistics by 2 digit CS categories.

CS	Description	Employment Share percent		Average City Wage (in 2015 euros)		Routine rankings	Offshorable rankings
		1994	2015	1994	2015		
	high-paying occupations						
23	CEOs	1.0	0.9	42.82	59.20	16	17
37	**managers and professionals**	**6.2**	**10.2**	32.52	38.56	15	16
38	**engineers**	**5.1**	**9.0**	30.36	33.68	17	10
35	creative professionals	0.5	0.5	22.82	31.80	14	11
	medium-paying occupations						
48	**supervisors and foremen**	**4.1**	**2.7**	18.03	21.86	3	3
46	**mid-level professionals**	**12.3**	**7.6**	17.54	21.20	13	6
47	**technicians**	**5.7**	**6.3**	17.15	20.60	11	7
43	mid-level health professionals	0.8	1.5	15.05	18.05	10	13
62	**skilled industrial workers**	**14.1**	**9.3**	13.52	17.99	4	2
54	**office workers**	**11.8**	**11.2**	13.17	16.98	1	4
65	**transport and logistics personnel**	**2.9**	**3.0**	11.96	16.00	5	5
63	**skilled manual workers**	**8.0**	**8.3**	11.90	15.50	7	8
64	**drivers**	**5.0**	**5.5**	11.50	14.46	18	18
67	**unskilled industrial workers**	**10.9**	**5.7**	11.02	14.72	2	1
	low-paying occupations						
53	security workers	0.7	1.4	10.60	14.60	9	12
55	**sales-related occupations**	**5.4**	**8.3**	10.44	13.74	6	15
56	personal service workers	2.2	4.8	9.97	12.63	12	14
68	**unskilled manual workers**	**3.3**	**3.8**	9.11	13.27	8	9

Notes: CS refers to the PCS 2-digit codes. In-sample values. Employment share for metropolitan mainland France (excluding Corsica). "Routine" ranking based on the RTI measure of Autor et al. (2003) while "Offshorable" on the OFF-GMS measure from Goos et al. (2014), both mapped into PCS 2-digit employment categories. Occupations with employment shares above 2.5% in 1994 in bold. We borrow the translation of 2-digit CS categories from Harrigan et al. (2016).

as most offshorable.

This division into high-, middle-, and low-paying jobs, of course, is a heuristic that allows us to capture key features of the data. There is a wage gap in excess of 20 percent at the boundary between high- and middle-paying jobs in 1994, but only trivial wage variation at the boundary between middle- and low-paying jobs. Indeed if the boundary occupations at the lower margin were the same, but this were calculated using 2015 wages, then drivers, CS 64, would be in the low-paying segment, where it fits more naturally. In this, we follow the choices of Autor and Dorn, and Goos, Manning, Salomons, which we think highlight the correct forces at work.

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.

NOT SURE: 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 principally data on jobs that are performed in cities (metropolitan areas) above 50,000 inhabitants as of 2015 unless otherwise noted. We aggregate the commune-level data to the metropolitan area level ("unité urbaine") with city boundaries defined by INSEE as of 2010 unless otherwise indicated. There are 117 such cities in 2015 with the largest 55 above 100,000 inhabitants shown in Figure 4 with population data by category in Table 9. The characteristics of the final sample are given in Table 10.

WHAT ABOUT "Aires urbaines" ? ROBUSTNESS WITH CITIES ABOVE 20 000 INHABITANTS, ETC.

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. There were 396,540 (out of 596,430 for which we have data) and 633,845 (out of 998,455) firms active in these metropolitan areas in 1994 and 2015 respectively. After all above-mentioned exclusions we remain with 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.[15]

We group the cities into six major categories for our analysis. Paris, given its size (10.7m inhabitants in the metropolitan area and 37.5% of jobs in our final sample) is a category by itself. Then, we use 2 categories of cities above 0.5m: 0.5-0.75m and 0.75m and above (except Paris). Such a choice is warranted because there is a considerable size difference between the seventh largest metropolitan area – Bordeaux (904 thousand inhabitants) and the eighth – Nantes (634 thousand people). Moreover, cities with metropolitan areas of "0.75m and above" have also "urban areas" as defined by INSEE of over 1m inhabitants.[16] For other divisions we follow the ones of INSEE: 0.2-0.5m (size categories "71" and "72") , 0.1-0.2m (sizes "61" and "62") and 0.05-0.1m ("51" and "52"). We took the city size of 50,000 as a cutoff for our main discussion, although dropping this to 20,000 doesn't materially affect our results.

3.4 Other data

We also use other data sources to provide additional statistics. When we examine log-supermodularity, we use the detailed part of the Census of 1999 (5% of population), also provided by INSEE. It pro-

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

[16] This is the "unité urbaine" – all communes in the metropolitan area plus all communes where at least 40% of residents have employment in the same metropolitan area.

vides 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.

Proposition 2 shows, as in Davis and Dingel (2014), that the distribution of skills $f(\omega, c)$ is log-supermodular in city size. To obtain a measure of skills we turn to the 1999 Census, which has good data on both diplomas and commune of residence. We measure skills by the highest obtained diploma received by individuals. The results are shown in Table 2 with a more detailed classification given in the Appendix Table 11.[17] As expected, for low skilled categories we obtain population elasticity coefficient estimates that are significantly below 1. For high skill individuals, the coefficient is significantly above 1, signaling their great presence in large cities. In particular, the category of workers with a graduate diploma has a population elasticity of 1.18. 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. Importantly, France provides no evidence of extreme skill complementarity, a hypothesis advanced in Eeckhout et al. (2014).

Table 2 – Log-supermodularity, population elasticities by diploma (4 categories) in the 1999 Census data.

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
Below high school x ln pop	0.93**	0.91***	0.24	0.84
	(0.027)	(0.023)		
High school professional diploma (CAP, BEP) X ln pop	0.91***	0.91***	0.31	0.96
	(0.023)	(0.022)		
End of high school diploma (Bac) X ln pop	1.00	0.99	0.15	0.95
	(0.024)	(0.023)		
Higher education X ln pop	1.10***	1.09***	0.3	0.96
	(0.030)	(0.030)		

Notes: 112 cities > 50,000 inhabitants defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as in main sample. Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

[17] 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.

We reconfirm these results using our classification of high, middle and low paying jobs and the broad 1-digit CS categories in Appendix Tables 12-13. It is not a coincidence that the population elasticity coefficients for high-paying jobs and "cadres" (respectively 1.14 and 1.16) 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 1. Similar patterns in terms of population elasticities for the same 4 diploma categories can be obtained from the 1990 and 2013 Censuses (not shown), confirming the notion that log-supermodularity of skills holds for French cities over the entire studied period.

4.2 Employment shares of different occupation categories.

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

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.23	0.14	0.12	0.10	0.09	0.08
2015	0.37	0.25	0.21	0.16	0.14	0.12
change in ppct	0.13	0.11	0.09	0.06	0.05	0.04
growth in %	0.57	0.77	0.71	0.63	0.61	0.49

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 metropolitan areas (Tables 3-5).

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

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.65	0.74	0.75	0.77	0.79	0.79
2015	0.45	0.57	0.60	0.64	0.66	0.67
change in ppct	-0.20	-0.17	-0.15	-0.13	-0.14	-0.12
growth in %	-0.31	-0.23	-0.20	-0.17	-0.17	-0.15

Note: Percentage point changes and growth rates not directly calculable from the upper two lines of the table due to rounding.

The share of high-paying occupations in total employment increases monotonically with city size in both years, as required by log-supermodularity as shown by Proposition 2 and documented in the previous section. The differences are sizable, especially when comparing the extremes – the Parisian metropolitan area and cities between 50 and 100 thousand of population. In both 1994 and 2015, the fraction of high skill jobs in Paris was roughly three times as high as in cities of

50-100,000 population. Given the overall rise in shilled jobs, the gap rose from 15 percentage points to 25 percentage points.

Moreover, the share of middle-paying jobs monotonically declines with city size in line with Proposition 1 in both 1994 and 2015 (see Table 4). The share of lowest-paid occupations is highest in smallest cities in either of the years, although the cross-city variation is modest. The decline of low-paid occupation shares with city size is, however, very clear when one measures the share of hours worked for 3-lowest paying jobs (sales-related occupations, personal service workers and unskilled manual workers; see Table 14).

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 metropolitan area size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.12	0.13	0.12	0.13	0.12	0.13
2015	0.18	0.18	0.18	0.20	0.20	0.21
change in ppct	0.07	0.06	0.06	0.07	0.08	0.08
growth in %	0.59	0.48	0.48	0.55	0.68	0.64

Note: Percentage point changes and growth rates not directly calculable from the upper two lines of the table due to rounding.

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 in both 1994 and 2015 the employment shares in the four most routine and offshorable jobs are declining in city size. This is in line with the patterns for middle-paying occupations overall, as exhibited in Table 4, and with Proposition 1.

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

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.29	0.36	0.39	0.41	0.45	0.45
2015	0.19	0.25	0.27	0.29	0.31	0.32
change in ppct	-0.11	-0.12	-0.12	-0.12	-0.13	-0.12
growth in %	-0.36	-0.32	-0.31	-0.29	-0.29	-0.27

Note: Percentage point changes and growth rates not directly calculable from the upper two lines of the table due to rounding.

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; office workers; 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 mid-level associate professionals (CS 46), 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 6 and confirm the U-shaped relationship in France for the years 1994-2015 studied by Autor et al. (2006) and Autor and Dorn (2013) for the U.S. and documented by Goos and Manning (2007), Goos et al. (2009) and Goos et al. (2014) for Europe.[18] They are also consistent with observations made by Harrigan et al. (2016) for France for the time period 1994-2007.

Proposition 3 holds that labor market polarization is induced by increased automation or a lower offshoring cost. This 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 4) 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, and, as implied by Proposition 4, it should be such occupations exercised by the relatively higher-skilled that should be destroyed the most. Second, high-paying jobs should increase by more in larger metropolitan areas, while low-paying jobs increase more in smaller cities. The strength of these relationships need not depend on initial exposure to routine or offshorable occupations, and, indeed, such exposure should be lowest in large cities as indicated by Proposition 1 and exhibited in Section 4.2. All such patterns are upheld in our data.

[18] 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.

4.4.1 Labor market polarization in cities by 3 occupation groups.

We start by discussing the changes in shares in hours worked in different categories of occupations by metropolitan area sizes.

In Table 4 we see that the percentage point changes in employment shares od middle-paid occupations 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 metropolitan areas 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 also increase with city size.

The greater decrease in middle-paying jobs in larger cities can be accounted for by our model, as exhibited by Proposition 4. 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.[19]

Table 3 shows that the percentage point increase in high-paying jobs is also monotonic in metropolitan area size. In Paris and metropolitan areas above 0.75m 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 $\omega(m, c)$ 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.

4.4.2 Labor market polarization in cities among different middle-paying occupations.

Given that our initial shock for labor market polarization – automation and offshoring – affects most routine and offshorable occupations (RTI4 group) in the aggregate, we could expect that the differences across cities are also starkest in these employment categories. However, it is clear that this is not the case. Comparing these patterns with the changes experienced in the 4 most routine and offshorable jobs (Table 6) – we see that the fall of shares in this category of middle-paying jobs is similar across metropolitan areas without any clear relationship with size – between 10.5 and 13.1 percentage points. One of the reasons for this, as discussed in Appendix D, is that the decline in the most routine/offshorable jobs might have occurred more strongly in large cities before 1994.

[19]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 metropolitan area productivity gains $A(h, c)$ or a complementarity between high skill labor and capital (as in the Autor and Dorn (2013) model).

Given that the exposure of the largest metropolitan areas is lowest in 1994, the destruction rates of these occupations still monotonically rises with city population even if the differences across cities are not sizeable.[20]

The overall decline in middle-paying occupations in the largest metropolitan areas shown in Table 6 exceeds that of the destruction of the 4 most routine and offshorable jobs while it is smaller in cities between 50-100 thousand inhabitants (e.g. 12.1 vs. 12.2 percentage points respectively). This means that the non-routine middle-paying jobs decline the most in large cities. For Paris, the decline of such occupations overall (9.6 percentage points) is very close to that of RTI4 jobs (10.5 percentage points) while in metropolitan areas with more than 750,000 inhabitants the hours worked non-routine middle-paying jobs decline by 5.2 pp. This indicates that the automation and offshoring shocks should not be directly responsible for an important share of labor market adjustments in larger cities. Such tendencies are pictured in Figure 8 where – for the sake of clarity – we constrain ourself to the percentage point employment share evolutions among the largest (>0.5m people), medium-sized (0.1-0.5m) and smallest cities (between 50-100 thousand inhabitants). An important observation is that the middle-paying non-routine jobs on average are also better paid than routine ones in 1994 – indicating that they may require on average higher skills.

These observations require further scrutiny. Figures 9-11 point that it is the most skill-requiring middle-paying jobs that decline most in largest cities. In particular, it is true for the statistical category of "intermediate professionals", CS 43, 46, 47 and 48 that according to the PCS classification should by construction require higher skills than "employees" or "workers" (occupations with the first classification digit of "5" or "6") due to the associated task complexity but less than "cadres" (those starting with "3").[21] In contrast, in all cases the less-paid middle-skill occupations decline similarly across cities of different sizes. Other standard divisions – such as those with of white collar (CS 46 and 54) vs. other middle-paying jobs or purely blue collar workers (CS beginning with "6") vs. other middle-paying occupations deliver similar conclusions (see Appendix, Figures 12 and 13). This may seem surprising at first. Our theory, however, can directly account for such tendencies. As already mooted above, largest cities have a greater comparative advantage in performing high-skill tasks for agents. In equilibrium, there is thus a stronger "switch" in large cities due to the initial shock from the relatively more skill-intensive middle-task jobs to high-skill intensive occupations. The observable flip-side of such patterns is the higher growth of high-paying occupations there as well.[22][23] As a result, larger metropolitan areas become disproportionately

[20] In Appendix Table 15 we show the patterns for the 6 most offshorable jobs encompassing not only the 4 most routine occupations, but also transport and logistics personnel and mid-level professionals. The shares of such jobs in total employment are monotonically decreasing with city size whether in 1994 or 2015. Percentage point fall in the employment shares of these occupations between 1994 and 2015, however, is higher in larger cities.

[21] Indeed, such occupations were the four best-remunerated middle-paying jobs in 1994.

[22] Such tendencies would be reinforced with local agglomeration effects depending on the number of high-skilled that we do not model.

[23] Cortes (2016) studies the effects of labor market polarization in the aggregate in a model with occupational sorting driven by the comparative advantage of higher skilled in more complex tasks as in Gibbons et al. (2005).

richer in higher-skill jobs with time than small cities. In contrast, for low-skill tasks, larger cities have no such a productivity advantage. Consequently, the move of workers with skills that earlier performed middle-paying tasks into low-paid tasks is more similar in magnitude across different cities, resulting in a slightly higher growth of low-skilled jobs in smaller cities overall. Our theoretical implications are valid if the skill distribution is relatively stable through time across cities – but in data log-supermodularity of skills can be exhibited for French cities for 1990, 1999 and 2013. We conclude that although automation and offshoring shocks touching the most routine and offshorable occupations are an important trigger for labor market polarization everywhere, they cannot account for important transformations of labor markets across metropolitan areas.

4.4.3 Individual city-level evidence.

The changes in employment shares coming from overall summation of hours by metropolitan areas discussed above may not hold at the individual city level if they are driven by a few sizeable outliers. Therefore we refine our analysis by scrutinizing patterns at the individual city level as well.

First, we present the rank correlations (Spearman's ρ and Kendall's τ) between the city populations in 1990 and the percentage point changes of employment shares in different job categories (Table 7). We confirm the tendencies discussed at the more aggregate level. Rank correlations validate that respectively the percentage point change of high-paying jobs is positively while that of middle- and low-paying occupations negatively correlated with city size. In particular, it is the higher paid middle-paying jobs that disappear more in larger cities. The rank correlations for both the top-3 highest-remunerated middle-paying jobs (CS 48, 46 and 47, with average hourly wages above 17 EUR in 1994), the top-4 middle-paying "intermediate professions" and the middle-paying jobs with average wages in 1994 above the median (CS categories 48, 46, 47, 43 and 62) are all negative and statistically significant at the 1 % level. On the other hand, the rank correlations between the least paying middle-skill jobs changes and city sizes are small and not statistically significant. Routine jobs' decline appears not to be correlated with city size, and the same is true for purely white-collar (CS 46 and 54) or blue-collar occupations (CS starting with 6; statistics not shown).

In a further investigation, we compare the 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 (Table 8).[24] Calculations for cities between 100 thousand and 0.5

Three occupational groups ranked by ability (non-routine manual; routine and non-routine cognitive) are taken into account. As a result of increased automation those with higher-ability in routine occupations switch to non-routine cognitive jobs while those with low ability switch to non-routine manual jobs. These predictions are borne out in PSID data. Indeed those with highest abilities switch into non-routine cognitive occupations the most.

[24] In the Appendix Table 16 we show the means of employment shares in the presented categories across these cities while in the Appendix Table 17 the same changes and growth patterns as in in Table 8 in a sample without the FIRE industries.

Table 7 – Rank correlation statistics between city-level population in 1990 and percentage point changes in different occupation categories 1994-2015

Occupation category	Spearman's ρ	Kendall's τ
high-paying	0.49***	0.35***
middle-paying	-0.28***	-0.19***
low-paying	-0.30***	-0.21***
routine-4	-0.07	-0.04
middle-paying non-routine-4	-0.19**	-0.14**
top 3 middle-paying with highest wages (CS 46, 47, 48)	-0.25***	-0.18***
least-well paid middle-paying (CS 43, 54, 62, 63, 64, 65, 67)	-0.06	-0.04
intermediate professions	-0.26***	-0.19***
employees and blue-collar workers	-0.05	-0.04
middle paying with wages above median	-0.37***	-0.25***
middle paying with wages below median	0.10	0.07

Note: Sample: 117 cities >50,000 inhabitants in 2015. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

million inhabitants show intermediate patterns between the two groups.

As in the aggregate data by metropolitan areas, the hours worked in highest paying occupations increase most in large cities while the low-paying occupations increase more in smallest cities as well. 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 (11.6 percentage points).

When we scrutinize more detailed middle-paid categories, we arrive at the same findings as in the rank correlations. There is not a statistically significant difference in the percentage point change in the decline of the 4 most routine and offshorable jobs[25]. The difference in the evolution

[25]However, we observe a larger destruction in larger cities of jobs with the six most routine categories included (RTI6) or the six most offshorable ones (OFF6) – not shown. 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 the overall sample increases. However, it's employment share falls 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). We note that these two occupations are paid below the median wage (across considered occupations) in 1994. The CS 46 (mid-level professionals) is the second highest-paid middle-paying and the sixth most-offshorable occupation – though heterogeneous – and it contains several back-office jobs that may now be performed from afar (accountants, sales administration, banking and 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). 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. 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.

in the non-routine middle-paid jobs across different city sizes is striking. The decline in such non-routine professions employment shares in small cities is close to zero, while it is 7.3 pp on average in the largest cities. Once again, the relevant difference seems to be the evolution of the relatively better middle-paid occupations[26] which decline by 5.7 pp more in larger cities than in small ones over the period 1994-2015.

As lower panels of Table 8 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 metropolitan area sizes are confirmed: large and small cities task composition diverges considerably in the studied period.

Table 8 – Comparison of means of changes in different occupations, cities >0.5m vs. 0.05-0.1m.

Item	high-paid	middle-paid	low-paid	routine-4	non-routine-4	middle-paid above median	middle-paid below median
mean change ppct, cities >0.5m	0.116	-0.181	0.065	-0.108	-0.073	-0.130	-0.051
mean change ppct, cities 0.05-0.1m	0.037	-0.116	0.080	-0.111	-0.005	-0.073	-0.044
difference in pct points	0.079***	-0.065***	-0.014***	0.003	-0.068***	-0.057***	-0.008
mean growth, cities >0.5m	0.629	-0.265	0.543	-0.331	-0.201	-0.360	-0.159
mean growth, cities 0.05-0.1m	0.458	-0.149	0.621	-0.252	-0.005	-0.199	-0.101
difference in growth	0.172***	-0.116***	-0.078	-0.079***	-0.195***	-0.161***	-0.058***

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities. Differences remain statistically significant at least at the 1 % level without weighting or weighted by city population as of 2015 except for the difference in growth middle-paid below-median jobs for unweighted comparison. Individual mean changes or growth rates are statistically significant different from zero at the 1 % level.

4.4.4 Labor market polarization in cities at the 2-digit CS level.

We discuss now labor market polarization at a finer 2-digit level. We consider Paris and Lyon – the two largest cities; cities between 0.2-0.5m and 0.05-0.1m inhabitants in the aggregate and compare employment change patterns shown in Figure 14.

First of all, the figure reconfirms that labor market polarization occurred in different cities.[27]

[26]This is confirmed (not shown) whether one scrutinizes the behavior of top 3 middle-paying jobs with highest wages (CS 46, 47, 48) or the intermediate professionals (CS 43, 46, 47, 48, the four highest paid middle-skill occupations). The former decline over 1994-2015 by 5.2 pp while the latter by 5.6 pp more in largest metropolitan areas.

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

However, it is immediately clear that the patterns are markedly different across metropolitan areas. Figures 15 and 16, 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 between 0.05-0.1m inhabitants are larger (e.g. the average of absolute percentage point changes in the individual categories – 2.4 vs. 1.9 respectively – are higher), despite a lower exposure to jobs with the highest routine or offshorability indexes (Table 6).

Turning to individual categories, there is a much higher percentage point (pp) 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 the 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 CS 46 (mid-level associate professionals, ranked as the second best paid middle-paying and sixth most offshorable occupation) whose share shrinks most in Paris. The category that is most routine in sample – office workers (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) proportion of hours fell in Paris by 1.4 pp but increased 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) employment share by 4.5 pp in small cities. The patterns 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 data at our disposal.

First, the employment share of industrial workers (categories 62 and 67) 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 Census or 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 respectively larger cities or departments with large metropolitan areas. The evolution of data on individual occupations points that initially the process was especially related to deindustrialization, as the categories that decreased first and strongest among middle-paying jobs were the ones related with manufacturing (categories 48, 62 and 67). 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 and hence not related to automation and offshoring shocks. Further details are provided in Appendix D.

One explanation for the decline in the share of back-office or support jobs like office workers (CS 54) or technicians (CS 47) in Paris with their coincident expansion in small cities within the 1994-2015 period can be internal offshoring from large to small cities permitted by the Internet and communication technologies. Such tendencies are consistent with our model (all goods, including intermediates, are traded) though we do not model nor cannot fully verify empirically supply chain developments that are internal or external to firms.

4.4.5 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 17-24 with regressions reported in Appendix Tables 18 and 20 reveals instructive patterns. We compare the largest 11 cities (above 0.5m people) and 62 smallest cities between 0.05-0.1m inhabitants to highlight the observable differences.

First of all, it is clear that large cities above 0.5m people have lower initial exposures to routine and offshorable occupations (RTI4).[28] Although there is considerable variation, Figure 17 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. Regression analysis in Appendix Table 19 further confirms this point: the interaction of a dummy for large cities with exposure is robustly negatively different from zero. Routine and offshorable jobs in large cities are destroyed at a higher rate than in small cities in reaction to the same automation or trade shocks.

The picture turns different, however, when we scrutinize the relationships between exposure and the overall change in middle-paying jobs (that include the RTI4 category) in Figure 18. The population-weighted regression of changes in employment on initial exposure to 4 most routine or offshorable occupations reveals surprisingly a strongly positive relationship (though the non-population weighted relationship is close to zero). There is clearly a larger destruction of middle-paying jobs in the largest cities overall witnessed by the sign of the interaction of a dummy for large cities with RTI4 exposure (Appendix Table 18). 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.

This is confirmed by the patterns in Figure 19 which reveals a strong positive relation between the exposure to routine-4 group of jobs and a lower decline in non-routine-4 jobs. The largest decline in middle-paid non routine occupations occurs in cities least exposed to routine-4 jobs. The

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

subsequent Figures 20 and 21 show that the difference comes from the behavior of better paid (above the median in 1994) middle occupations that overall decline more than middle-paying jobs below the median wage, and that this decline is particularly stronger in larger cities that are also least exposed to RTI4 occupations (cf. Appendix Table 19).

These patterns of decline in different groups of middle-paying jobs seem thus at odds with the plain routine or offshoring-driven explanation of labor market polarization. Altogether, in the period 1994-2015, there may be some substitution of the 4 most routine or offshorable jobs by other middle-paying jobs in smaller cities while the destruction of middle-paying jobs – especially the more skill intensive ones, no matter whether routine/offshorable or not – is exacerbated in large cities.

Turning to the tails of the skill distribution, Figure 23 shows that the overall relationship between routine and offshorable job exposure is not correlated with the changes in the low paying jobs. Not only – as shown in Table 8 – increases in the low-paid job shares in large cities are lower on average than in small cities, but for the largest cities (assuming the same intercept) there is a negative correlation of exposure with the creation of these jobs. Considering each city category separately, however, these relationships are not statistically significant.

A striking pattern can be seen in Figure 22 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). Given that the largest cities are also least exposed, there is a negative population-weighted overall correlation between initial exposure to 4 most routine or offshorable jobs and the change in the employment of high-paying occupations. Larger cities create more high-skilled jobs than small cities for any level of exposure. Further regressions (not shown) reveal that the relationship (population-weighted) between initial most routine/offshorable occupation exposure and high-paid jobs creation within city categories is negative (though not statistically significant) for large cities while positive for small cities.

This overall 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 (especially those requiring higher skills) in larger cities (that are less exposed to technological shocks) as a result of labor market polarization is expected from our theory – Proposition 4. It is clear that what matters is the city size and the interaction of the technology changes on different middle-skill tasks.

Similar patterns to that for 4 most routine or offshorable jobs can be observed for the 6 most offshorable occupations. For example, in Figure 24 higher initial exposure to offshorable jobs leads to their greater decrease in the studied period. This time, however, large cities are relatively more exposed to offshorable jobs in comparison to RTI4 occupations only as – in particular – they have on average a higher share of the CS 46 category, mid-level professionals (cf. also Appendix Table 15). Again, conditional on exposure, offshorable jobs' employment shrinks by more in large cities

in the studied period.

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.

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, "Work of the Past, Work of the Future," NBER Working Papers 25588, National Bureau of Economic Research, Inc February 2019.

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.

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

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

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

Davis, Donald R. and Jonathan I. Dingel, "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.

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

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.

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

Rosen, Sherwin, "The economics of superstars," *The American economic review*, 1981, *71* (5), 845–858.

A Extensions of the model

A.1 Endogenous productivity

In the model, we treat productivity terms A as exogenous. In this subsection, we extend our results to the case where A is endogenous to the composition of labor in the city.

More specifically, let us consider the following form for productivity terms:

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

where $G(.,j)$ is an increasing function for all $j \in \{l,m,h\}$ and $G(x,j)$ is log-supermodular in $\{x,j\}$. Both assumptions ensure that, in equilibrium, Assumption 1 is satisfied.

Under these conditions, we obtain that:

Proposition 7. *A decline in p_z leads to*

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

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

By increasing the share of the population in the high-skill sector, the polarization shock increases productivity. Yet, as shown in Proposition 4, the increase in the share of the population in the high-skill sector is more important in the larger than in the smaller city. Thus, this increases productivity by more in the larger city, reinforcing the absolute advantage of this city. Given that we assume that productivity reacts by more for higher-skill occupations, this also leads to a reinforcement of the comparative advantage in the larger city for higher skill occupations.

Remark. A potential pitfall with endogenous productivity is that it can lead to multiple equilibria. For example, in our model, symmetric cities can also be an equilibrium outcome if productivity is endogenous. To extend our results to endogenous productivity would then require to maintain the assumption that we stay close to the selected equilibrium. We also refer the interested reader to Davis and Dingel (2014) for a discussion of the possibility of multiple equilibria in a related setting.

A.2 N cities

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

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

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

Sectoral decisions As a result of these assumptions and extending Lemma 1, the thresholds $\omega(m,c)$ and $\omega(h,c)$ are decreasing with city size. Furthermore, we can extend Lemma 6 and obtain that a change in p_z leads to a stronger decline in $\omega(h,c)$ in large cities and $\omega(m,c)$ increase by more in smaller cities.

Location decisions The description of location decision within a city as described in Lemmas 2, 3 and 4 given that these results apply for any city c. We then only need to describe how agents decide to choose locations between the different N cities.

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

$$H(\omega(\tau(c)), M(\omega(\tau(c)), c), c)T(\tau(c)) > H(\omega(\tau(c)), M(\omega(\tau(c)), i+1), i+1)T(0) \tag{19}$$

where $\omega(\tau(c)) = K(c, \tau(c))$ is the value of ω occupying location $\tau(c)$ in city c. This defines a maximum value for the skill in city $i+1$, $\bar{\omega}(i+1)$ for which inequality (15) holds with equality. As a result, any agent with a skill higher than $\bar{\omega}(i+1)$ will decide to live only in cities $c \leq i$.

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

$$H(\omega(\tau), M(\omega(\tau, c), c), c)T(\tau) = H(\omega(\tau), M(\omega(\tau), i+1), i+1)T(\tau'). \tag{20}$$

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

Results As for the 2-city case, labor market polarization will happen in the aggregate and across cities. This results from sectoral decisions. As for the 2-city case, the distribution of skills is going

to be log-supermodular.

Similarly, if, for any city $c \in \{1, ..., N-1\}$ the comparative advantage of c over city $c-1$ in high skill occupations is sufficiently large, i.e. $A(i,h)/A(i-1,h)$ is sufficiently large compared with $A(i,m)/A(i-1,m)$ for all $i \leq N-1$, we also obtain the results of Proposition 1 about initial exposures and of Proposition 4.

B Proofs of Propositions

B.1 Proof of Lemma 1

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

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

$$\frac{A(h,c)}{A(m,c)}\frac{p(h)}{p(m)} = \omega(h,c)e^{-\eta\omega(h,c)} \tag{21}$$

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 $\omega(h,1) < \omega(h,2)$.

B.2 Proof of Lemma 2

The proof of Lemma 1 in Davis and Dingel (2014) 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.

$$A(h,1)e^{\gamma\omega(h,1)}T(\tau_1)p(h) - r(1,\tau_1) = A(h,2)e^{\gamma\omega(h,1)}T(\tau_2)p(h) - r(1,\tau_2) \tag{22}$$

$$\geq A(m,2)\omega T(\tau_1)p(m) - r(1,\tau_1) \tag{23}$$

B.3 Proof of Lemma 3

Here, we follow Davis and Dingel (2014), lemma 2 and Lemma 1 in Costinot and Vogel (2010). Let us first define

$$f(\omega,c,\tau) = \int_\sigma f(\omega,c,\tau,\sigma)d\sigma$$

$$\Omega(\tau,c) = \{\omega \in \Omega, f(\omega,c,\tau) > 0\}$$

$$\mathcal{T}(\omega,c) = \{\tau \in [0,\bar{\tau}(c)], f(c,\omega,\tau) > 0\}$$

Using these objects, we obtain:

(i) $\Omega(\tau, c) \neq \emptyset$ for $0 \leq \tau \leq \bar{\tau}(c)$ and $\tau(\omega, c) \neq \emptyset$ 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:

$$T(\tau')G(\omega', c) - r(c, \tau') \geq T(\tau)G(\omega', c) - r(c, \tau) \tag{24}$$

$$T(\tau)G(\omega, c) - r(c, \tau) \geq T(\tau')G(\omega, c) - r(c, \tau') \tag{25}$$

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:

$$f(c, \omega, \tau) = f(\omega)\delta^{\text{Dirac}}\left(1 - 1_{\Omega(c,\tau)}\right) \text{ for almost all } \omega \in \Omega(c, \tau). \tag{26}$$

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 $K_c(\bar{\tau}(c)) = \underline{\omega}$.

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

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

Suppose then that $\omega < \omega^*$ is not in city 1. This implies that $r(1, \tau(1)^*) = 0$. In equilibrium, this also implies that $r(2, \tau(2)^*) = 0$ and, then, $r(2, \tau(2)) = 0$ for any $\tau(2) \geq \tau(2)^*$. Let us consider the least skill agent ($\underline{\omega}$). The location of this agent is then not optimal:

$$H(\underline{\omega}, M(\underline{\omega}, 2), 2)T(\bar{\tau}(2)) \leq H(\underline{\omega}, M(\underline{\omega}, 2), 2)T(\tau(2)^*)$$

thus, violating the definition of an equilibrium.

B.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 $K_c(\bar{\tau}(m, c)) = \omega(m, c)$.

B.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 $\omega(h, c)$ and $\omega(m, c)$. Given the definition of the thresholds, the function is continuous everywhere and, thus, given it is either constant or decreasing in each segment, it is globally weakly decreasing.

B.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:

$$\frac{H(\omega(h, c), h, c)}{H(\omega(h, c), m, c)} = \frac{p(m)}{p(h)} \tag{27}$$

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

$$\frac{d\left(\frac{H(\omega(h,c),h,c)}{H(\omega(h,c),m,c)}\right)}{\frac{H(\omega(h,c),h,c)}{H(\omega(h,c),m,c)}} = \frac{d\left(\frac{p(m)}{p(h)}\right)}{\frac{p(m)}{p(h)}} \tag{28}$$

Let us compute the different terms separately:

$$\frac{d\left(\frac{H(\omega(h,c),h)}{H(\omega(h,c),m)}\right)}{\frac{H(\omega(h,c),h)}{H(\omega(h,c),m)}} = \left(\frac{H_\omega(\omega(h,c),h)}{H(\omega(h,c),h)} - \frac{H_\omega(\omega(h,c),m)}{H(\omega(h,c),m)}\right) d\omega(h,c) \tag{29}$$

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

$$d\omega(h,c) = \frac{1}{\Gamma(\omega(h,c),c)} \frac{d\left(\frac{p(m)}{p(h)}\right)}{\frac{p(m)}{p(h)}} \tag{30}$$

Given that H is log-supermodular,

$$\Gamma(\omega(h,c),c) = \frac{H_\omega(\omega(h,c),h,c)}{H(\omega(h,c),h,c)} - \frac{H_\omega(\omega(h,c),m,c)}{H(\omega(h,c),m,c)} > 0. \tag{31}$$

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

$$d\omega(m,c) = \frac{1}{\Gamma(\omega(m,c),c)} \frac{d\left(\frac{p(l)}{p(m)}\right)}{\frac{p(l)}{p(m)}} \tag{32}$$

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

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

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

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

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

$$\frac{H_\omega(\omega(m,1),m,1)}{H(\omega(m,1),m,1)} - \frac{H_\omega(\omega(m,1),l,1)}{H(\omega(m,1),l,1)} - \frac{H_\omega(\omega(m,2),m,2)}{H(\omega(m,2),m,2)} + \frac{H_\omega(\omega(m,2),l,2)}{H(\omega(m,2),l,2)} \tag{34}$$

Let us make some further assumption on the H function. Using our assumption on H and given that $\phi < 1$, this simplifies the two expressions into:

$$\frac{1}{\omega(m,1)} - \frac{1}{\omega(m,2)} \geq 0 \tag{35}$$

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

$$\frac{H_\omega(\omega(h,1),h,1)}{H(\omega(h,1),h,1)} - \frac{1}{\omega(h,1)} - \frac{H_\omega(\omega(h,2),h,2)}{H(\omega(h,2),h,2)} + \frac{1}{\omega(h,2)} \tag{36}$$

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

$$\frac{H_\omega(\omega(h,1),h,1)}{H(\omega(h,1),h,1)} - \frac{H_\omega(\omega(h,2),h,2)}{H(\omega(h,2),h,2)} \leq \frac{1}{\omega(h,1)} - \frac{1}{\omega(h,2)} \tag{37}$$

which is satisfied.

B.7 Proof of Proposition 1

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

$$L\int_{\omega(m,c)}^{\omega(h,c)} f(x,c)dx = S\left(T^{-1}\left(h(\omega(m,c),c)\right)\right) - S\left(T^{-1}\left(h(\omega(h,c),c)\right)\right) \tag{38}$$

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

The share of middle skill agents in city c is:

$$s(m,c) = \frac{\int_{\omega(m,c)}^{\omega(h,c)} f(x,c)dx}{\int_{\underline{\omega}}^{\overline{\omega}(c)} f(x,c)dx} = \frac{S\left(T^{-1}\left(h(\omega(m,c),c)\right)\right) - S\left(T^{-1}\left(h(\omega(h,c),c)\right)\right)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{39}$$

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

B.8 Proof of Proposition 2

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

$$L\int_{\omega}^{\omega+d\omega} f(x,c)dx = S\left(T^{-1}\left(h(\omega,c)\right)\right) - S\left(T^{-1}\left(h(\omega+d\omega,c)\right)\right) \tag{40}$$

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

$$f(\omega,c) = -\frac{\partial}{\partial\omega} S\left(T^{-1}\left(h(\omega,c)\right)\right) = h'(\omega,c)V(h(\omega,c)) \tag{41}$$

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

Let us first note that $f(\omega, c)$ is log-supermodular if and only if, for all $\omega > \omega'$ and $c > c'$, we have:

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

When $f(\omega, c')$ and $f(\omega', c')$ are different than 0, this condition amounts to verifying than $f(\omega, c)/f(\omega, c')$ is strictly increasing or, equivalently that:

$$f'(\omega, c)f(\omega, c') > f'(\omega, c')f(\omega, c). \tag{42}$$

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

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

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

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

Replacing $f'(\omega, c)$ and $f'(\omega, c')$ by their values in (42), we then obtain the following condition:

$$\frac{h''(\omega, c)}{h'(\omega, c)} + \eta(V, h(\omega, c))\frac{h'(\omega, c)}{h(\omega, c)} > \frac{h''(\omega, c')}{h'(\omega, c')} + \eta(V, h(\omega, c'))\frac{h'(\omega, c')}{h(\omega, c')} \tag{43}$$

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

Lemma 8. **(i)** *If, for ω and ω' and for c and c', the occupation decisions are the same across cities, that is $M(\omega, c) = M(\omega, c')$ and $M(\omega', c) = M(\omega', c')$, then*

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

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

(ii) *If productivities are constant across occupations, $A(c, h) = A(c, m) = A(c, l)$ as in Davis and Dingel (2014), a necessary and sufficient condition for $f(\omega, c)$ to be log-supermodular is that $\eta(V, x)$ is decreasing in x.*

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

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

and thus $h(\omega, c) = h(\omega, c')$. By continuity, $M(\omega, c) = M(\omega, c')$ on a (right- or left-) neighborhood

of ω and thus $h(.,c) = h(.,c')$ on this neighborhood, thus ensuring that locally $h''(.,c) = h''(.,c')$ and $h'(.,c) = h'(.,c')$ and in particular that $h''(\omega,c) = h''(\omega,c')$ and $h'(\omega,c) = h'(\omega,c')$. In the end, (43) simplifies into:

$$\eta(V, h(\omega,c)) > \eta(V, h(\omega,c')) \tag{44}$$

which is satisfied as long as V features decreasing elasticity.

□

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

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

Lemma 9. *Suppose that $f(x,c)$ is log-supermodular in $\{x,c\}$ on $[\underline{x}, \overline{x}]$ and $[\overline{x}, \overline{\overline{x}}]$, then $f(x,c)$ is log-supermodular in $\{x,c\}$ on $[\underline{x}, \overline{\overline{x}}]$*

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

Using log-supermodularity on $[\overline{x}, \overline{\overline{x}}]$, we have:

$$\frac{f(x,c)}{f(x,c')} > \frac{f(\overline{x},c)}{f(\overline{x},c')}$$

Using log-supermodularity on $[\underline{x}, \overline{x}]$, we have:

$$\frac{f(\overline{x},c)}{f(\overline{x},c')} > \frac{f(x',c)}{f(x',c')}$$

Combining these two equations, we obtain:

$$\frac{f(x,c)}{f(x,c')} > \frac{f(x',c)}{f(x',c')}.$$

In the end, f is then log-supermodular on $[\underline{x}, \overline{\overline{x}}]$.

□

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

Let us start with some properties on the $h(\omega,c)$ function. The indifference condition between

location implies that:

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

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

$$h(\omega, 1)h(\omega', 2) \leq h(\omega', 1)h(\omega, 2)$$

and that

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

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

$$\eta(V, h(\omega, 1))\frac{h'(\omega, 1)}{h(\omega, 1)} > \eta(V, h(\omega, 2))\frac{h'(\omega, 2)}{h(\omega, 2)}. \tag{45}$$

Second, note that (43) is invariant to equilibrium prices.

In the end, when $\eta(V, h(\omega, 1))$ is sufficiently decreasing, condition (43) is satisfied.

B.9 Proof of Proposition 3

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

$$ds(m, c) = \frac{\left(S\left(T^{-1}\left(h(\omega(m, c), c)\right)\right)\right)' d\omega(m, c) - \left(S\left(T^{-1}\left(h(\omega(h, c), c)\right)\right)\right)' d\omega(h, c)}{S\left(T^{-1}\left(h(\underline{\omega}, c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c), c)\right)\right)} \tag{46}$$

Given that $d\omega(m, c) > 0$ and $d\omega(h, c) < 0$ 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.

B.10 Proof of Proposition 4

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

$$ds(m, c) = \frac{V\left(h(\omega(h, c), c)\right) d\omega(h, c) - V\left(h(\omega(m, c), c)\right) d\omega(m, c)}{S\left(T^{-1}\left(h(\underline{\omega}, c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c), c)\right)\right)} \tag{47}$$

We obtain:

$$ds(m,c) = \frac{\left(\frac{V(h(\omega(h,c),c))}{\Gamma(\omega(h,c),c,h)} + \frac{V(h(\omega(m,c),c))}{\Gamma(\omega(m,c),c,m)}\right)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)}\frac{dp}{p} \tag{48}$$

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

$$\frac{\left(V\left(h(\omega(h,c),c)\right)\frac{\omega(h,c)}{\gamma\omega(h,c)-1} + V\left(h(\omega(m,c),c)\right)\omega(m,c)(1-\phi)\right)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{49}$$

We have easily that:

$$V\left(h(\omega(h,1),1)\right)\frac{\omega(h,1)}{\gamma\omega(h,1)-1} > V\left(h(\omega(h,2),2)\right)\frac{\omega(h,2)}{\gamma\omega(h,2)-1}. \tag{50}$$

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

$$V\left(h(\omega(m,1),1)\right)\omega(m,1) - V\left(h(\omega(m,2),2)\right)\omega(m,2) \tag{51}$$

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

$$\frac{\left(V\left(h(\omega(h,1),1)\right)\frac{\omega(h,1)}{\gamma\omega(h,1)-1}\right)}{S\left(T^{-1}\left(h(\underline{\omega},1)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(1),1)\right)\right)} > \frac{\left(V\left(h(\omega(h,2),2)\right)\frac{\omega(h,2)}{\gamma\omega(h,2)-1}\right)}{S\left(T^{-1}\left(h(\underline{\omega},2)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(2),2)\right)\right)} \tag{52}$$

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

B.11 Proof of Proposition 5.

The variation in percentage points of higher wage middle skill jobs is:

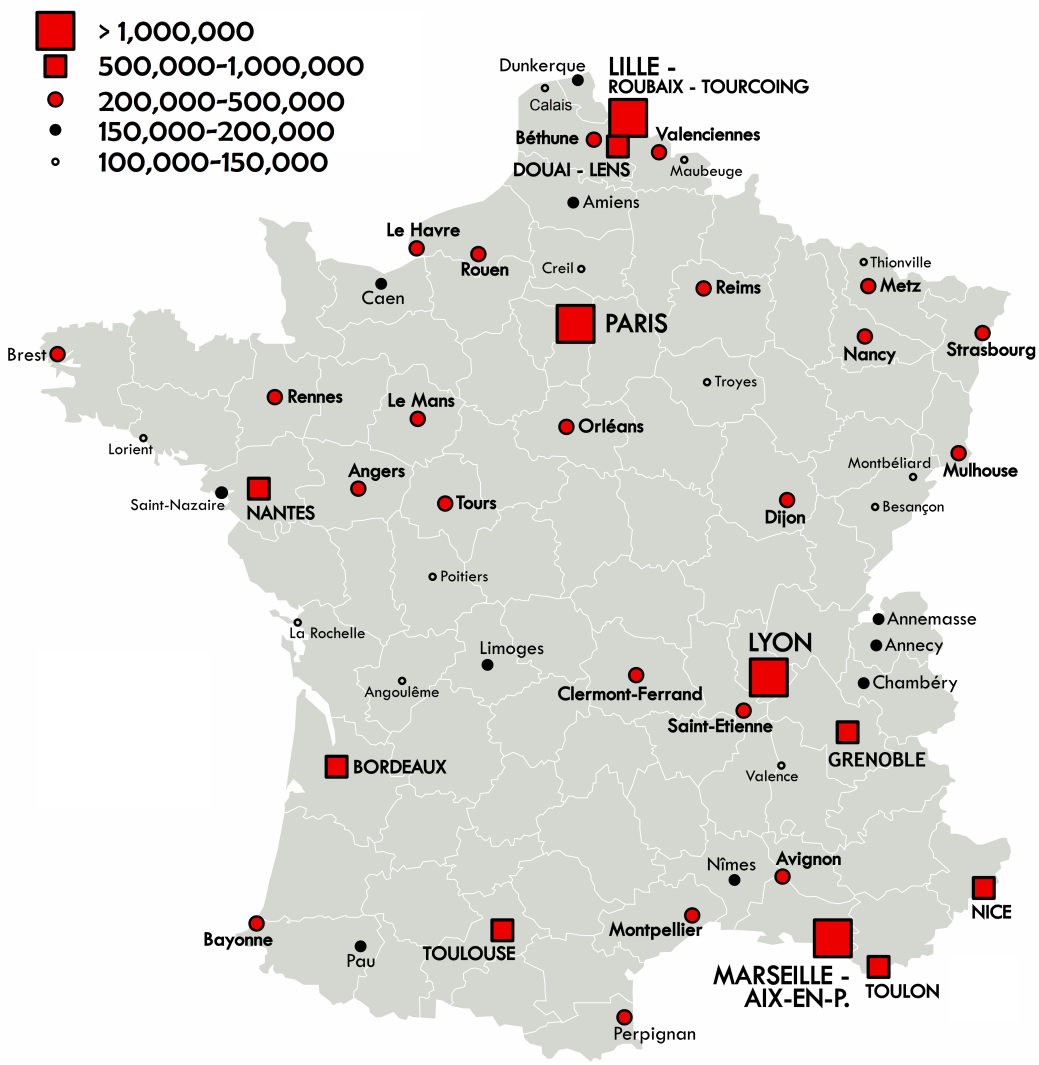
$$\frac{V\left(h(\omega(h,c),c)\right)d\omega(h,c)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{53}$$

and the variation in percentage points of lower wage middle skill jobs is:

$$\frac{-V\left(h(\omega(m,c),c)\right)d\omega(m,c)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{54}$$

Given that $d\omega(h,c)$ is larger in the large city

Figure 4 – Map of France with largest metropolitan areas in 2015.



C Figures and tables

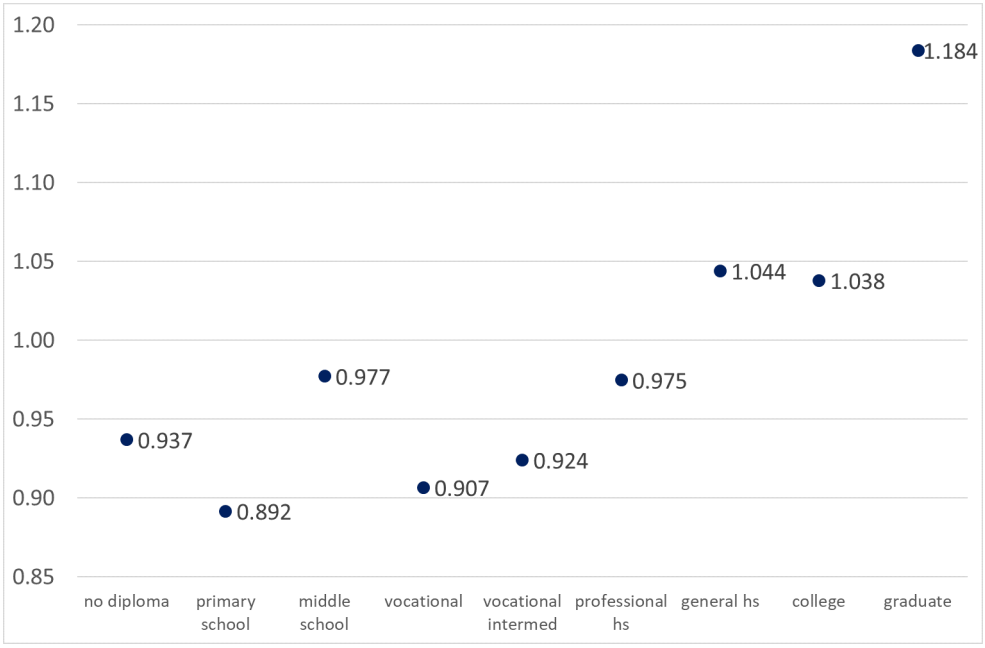
Table 9 – City categories: number of cities, population and the share of hours worked in 2015.

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 10 – 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

Figure 5 – Log-supermodularity, population elasticities by diploma (9 categories) in the 1999 Census data.



Notes: 112 cities > 50,000 inhabitants defined by INSEE as of 1999, population figures from 1999. Exclusions in terms of CS and age as in main sample.

Table 11 – Log-supermodularity, population elasticities by diploma (9 categories) in the 1999 Census data.

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

Notes: 112 cities > 50,000 inhabitants defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as in main sample. Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

Table 12 – Population elasticities by high, middle and low paying categories in the 1999 Census data.

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
High paying X ln pop	1.14***	1.14***	.17	.85
	(0.037)	(0.036)		
Middle-paying X ln pop	0.95*	0.95**	.65	.94
	(0.025)	(0.024)		
Low-paying X ln pop	0.94***	0.92***	.18	.96
	(0.018)	(0.015)		

Notes: 112 cities above 50,000 inhabitants as defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as in main sample. Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

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

Dependent variable: ln $f(\omega, c)$	All workers	French born	Population share	French born share
Cadres (CS 3) X ln pop	1.16***	1.15***	.17	.96
	(0.038)	(0.037)		
Intermediate professionals (CS 4) X ln pop	1.02	1.02	.30	.97
	(0.026)	(0.026)		
Low-skill employees (CS 5) X ln pop	0.97	0.96**	.27	.92
	(0.021)	(0.019)		
Blue-collar workers (CS 6) X ln pop	0.88***	0.86***	.26	.86
	(0.027)	(0.026)		

Notes: 112 cities above 50,000 inhabitants as defined by INSEE as of 1999. The variable "ln pop" is the natural logarithm of metropolitan area population from 1999. Exclusions in terms of CS and age as in main sample. CS 23 category – CEOs – not included in the category "cadres". Robust standard errors in parentheses. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the test of hypotheses whether a given coefficient is equal to one.

Table 14 – Share of 3 lowest-paying occupations per metropolitan area size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.11	0.12	0.12	0.12	0.11	0.12
2015	0.16	0.16	0.17	0.18	0.19	0.20
change in ppct	0.05	0.05	0.05	0.06	0.08	0.08
growth in %	0.51	0.42	0.43	0.54	0.69	0.64

Note: Percentage point changes and growth rates not directly calculable from the upper two lines of the table due to rounding.

Table 15 – Share of 6 most-offshorable occupations per metropolitan area size.

Agglo.size	Paris	> .75M	.5-.75M	.2-.5M	.1-.2M	.05-.1M
1994	0.49	0.54	0.54	0.56	0.58	0.59
2015	0.30	0.36	0.38	0.41	0.42	0.43
change in ppct	-0.19	-0.17	-0.16	-0.16	-0.16	-0.16
growth in %	-0.39	-0.33	-0.29	-0.28	-0.28	-0.27

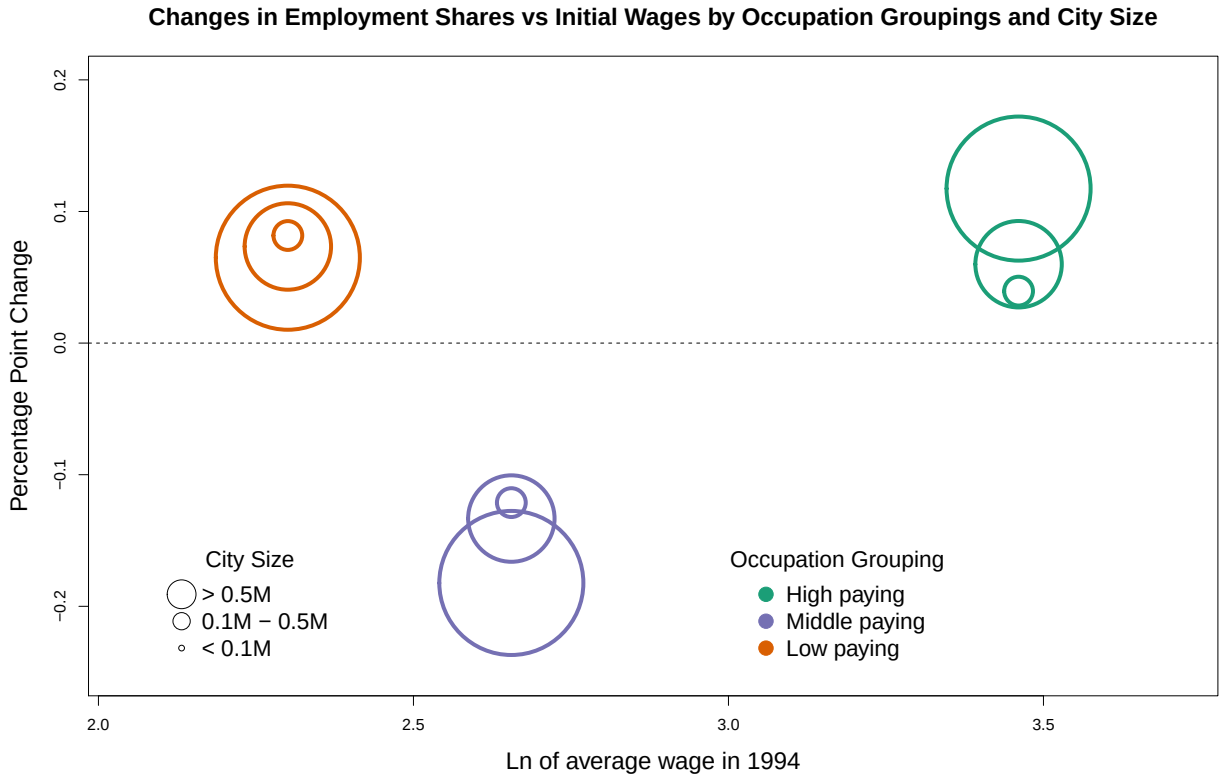
Note: Percentage point changes and growth rates not directly calculable from the upper two lines of the table due to rounding.

Figure 6 – Labor market polarization in France 1994-2015.



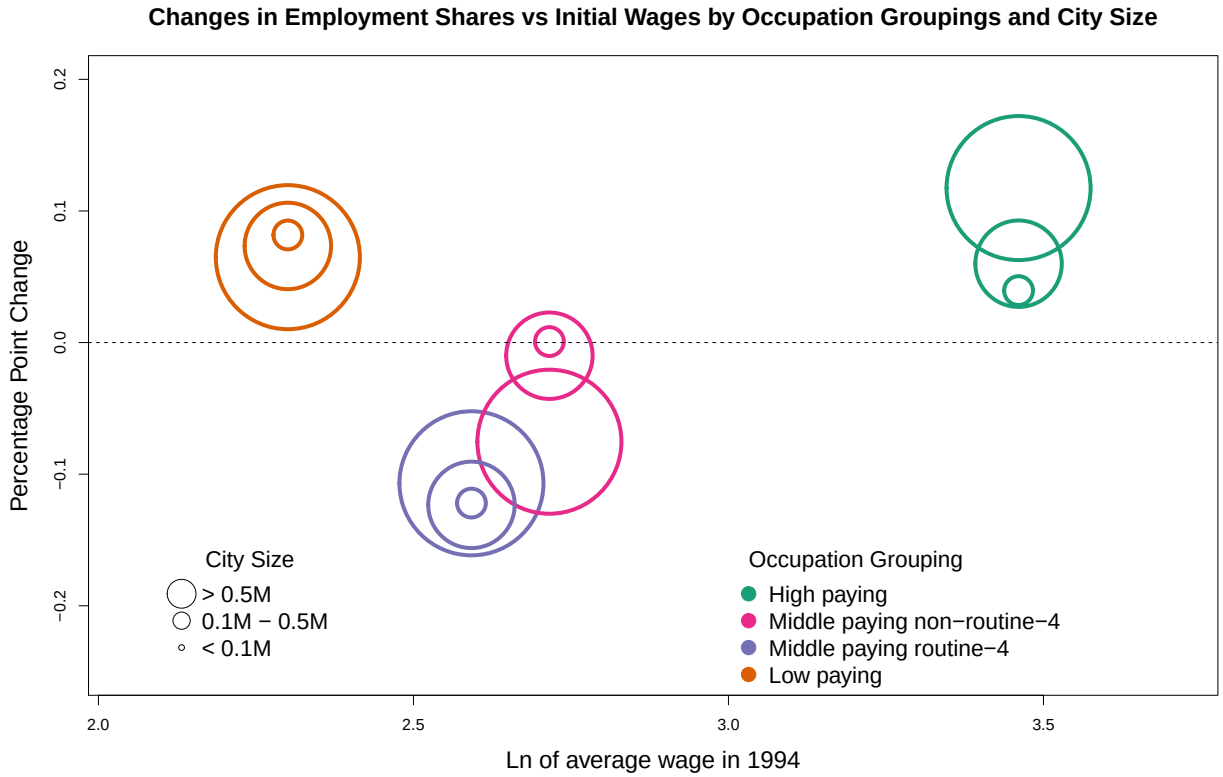
Notes: The figure shows the percentage point change in employment of the considered 2-digit CS categories plotted against their average wage in cities > .05m in 1994. Numbers pertain to 2-digit CS categories represented. Circle sizes correspond to the employment shares. The line shows a cubic relationship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

Figure 7 – Labor market polarization across three different city size groups, 1994-2015: 3 employment groups.



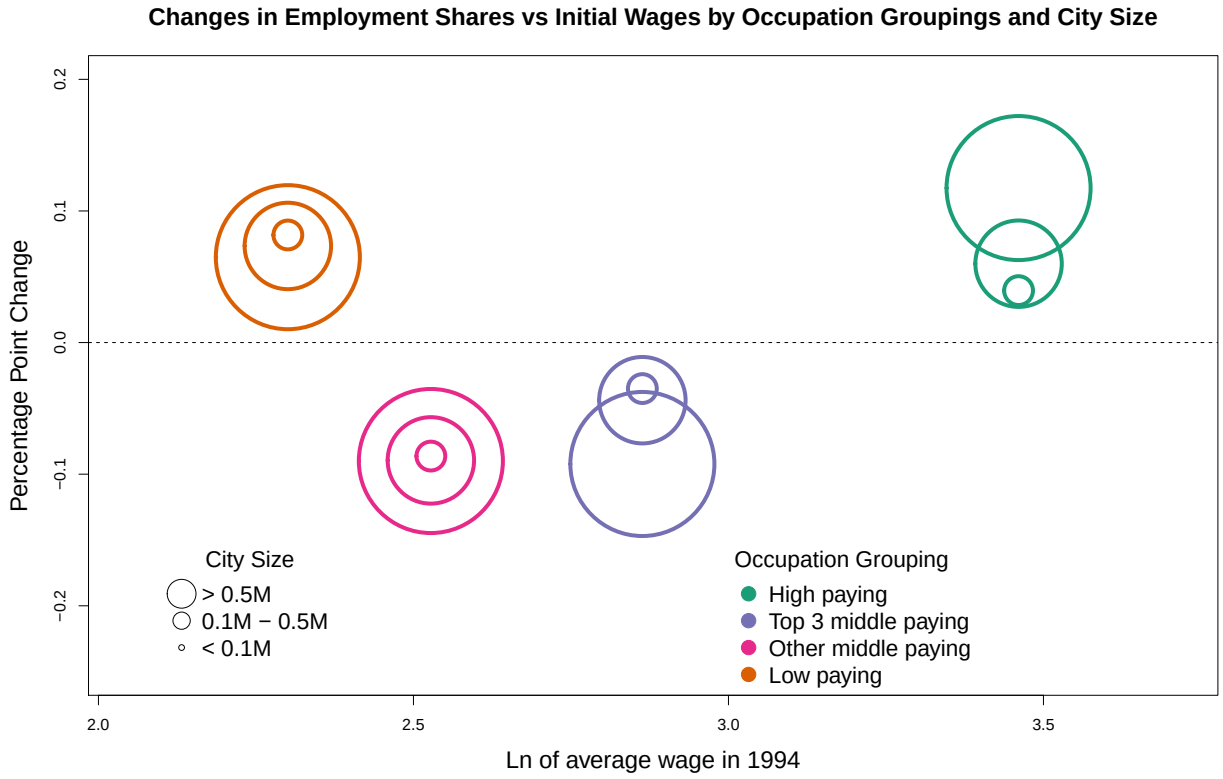
Notes: The figure shows the percentage point change in employment of the high- middle- and low- paying jobs plotted against the logarithm of average wage in cities > .05m in 1994 in different city groups (represented by circles)

Figure 8 – Labor market polarization across three different city size groups, 1994-2015: routine-4 vs. non-routine-4 jobs.



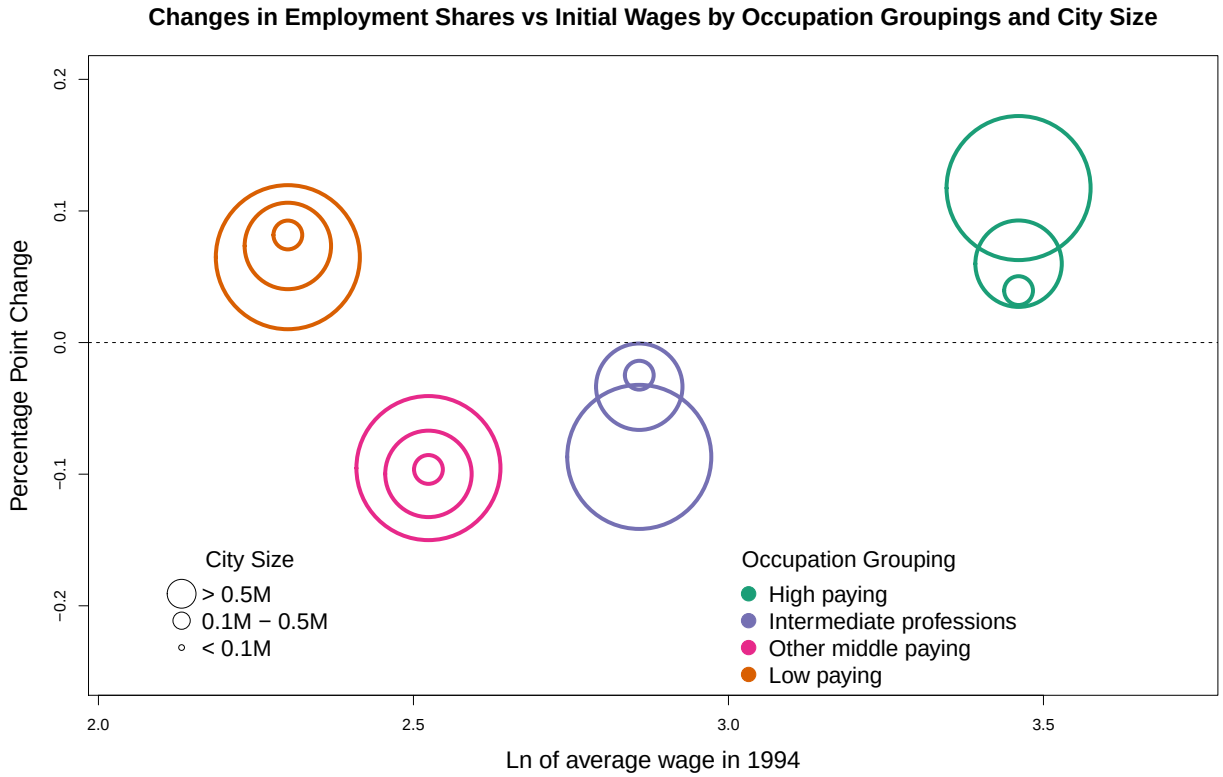
Notes: The figure shows the percentage point change in employment of the high-, low-, middle-paying routine-4 and non-routine-4 paying jobs plotted against the logarithm of average wage in cities > .05m in 1994 in different city groups.

Figure 9 – Labor market polarization across three different city size groups, 1994-2015: top three middle-paying vs. other middle-paying jobs.



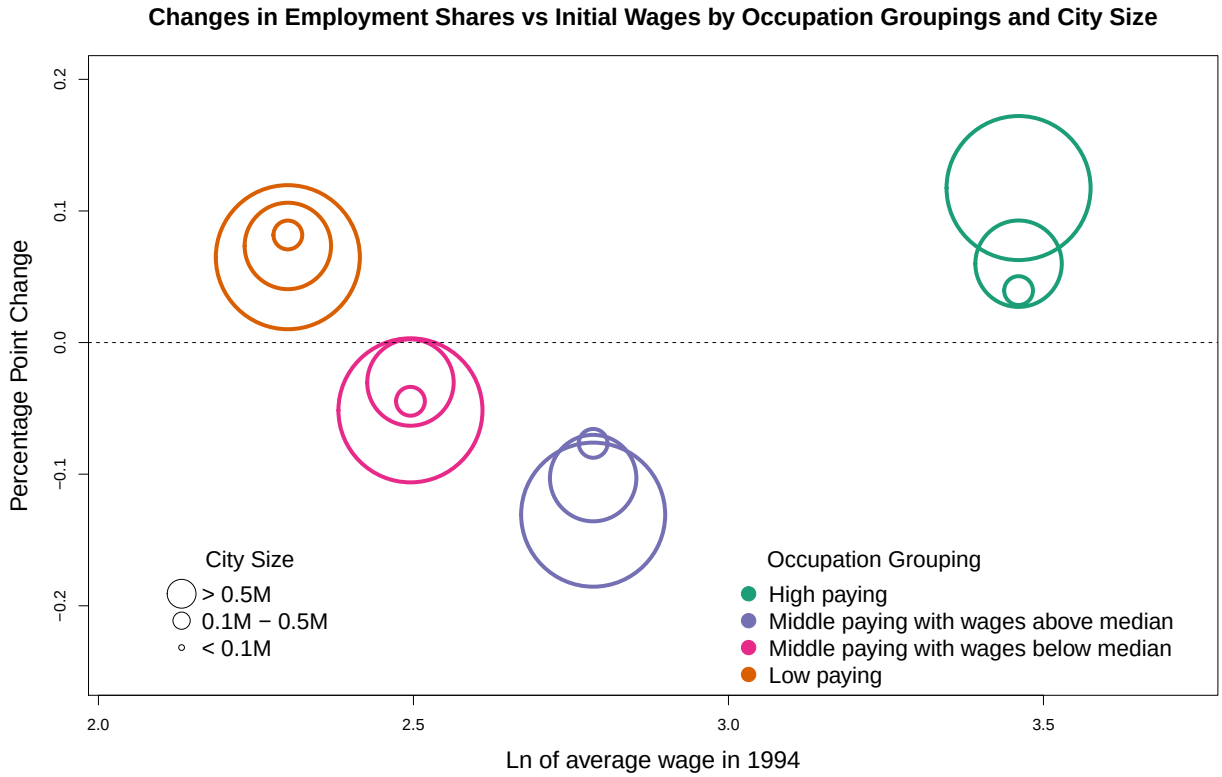
Notes: The figure shows the percentage point change in employment of the high-, low-, top three middle-paying and other middle-paying jobs plotted against the logarithm of average wage in cities > .05m in 1994 in different city groups.

Figure 10 – Labor market polarization across three different city size groups, 1994-2015: intermediate professions (CS 43, 46, 47 and 48) vs. other middle-paying jobs.



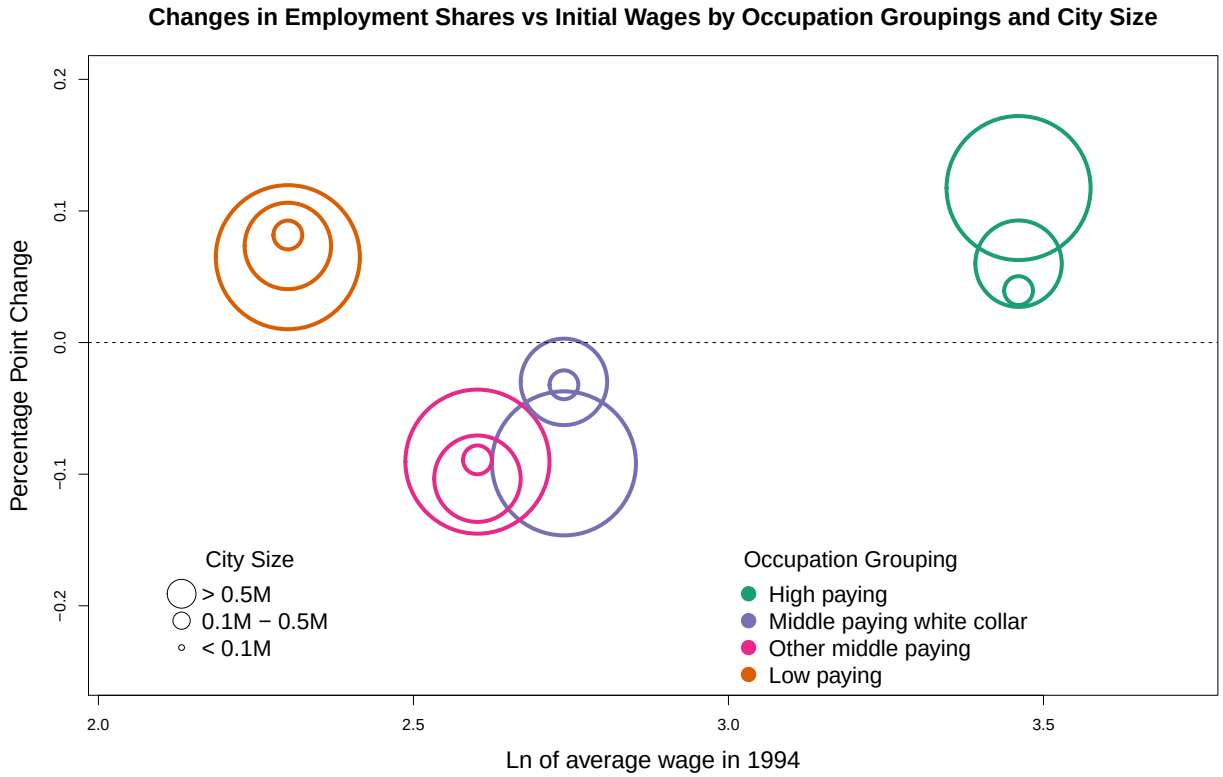
Notes: The figure shows the percentage point change in employment of the high-, low-, intermediate professions (CS 43, 46, 47 and 48) and other middle-paying jobs plotted against the logarithm of average wage in cities > .05m in 1994 in different city groups.

Figure 11 – Labor market polarization across three different city size groups, 1994-2015: middle-paying (CS 43, 46, 47, 48, 62) with wages above median in 1994 vs. other middle-paying jobs.



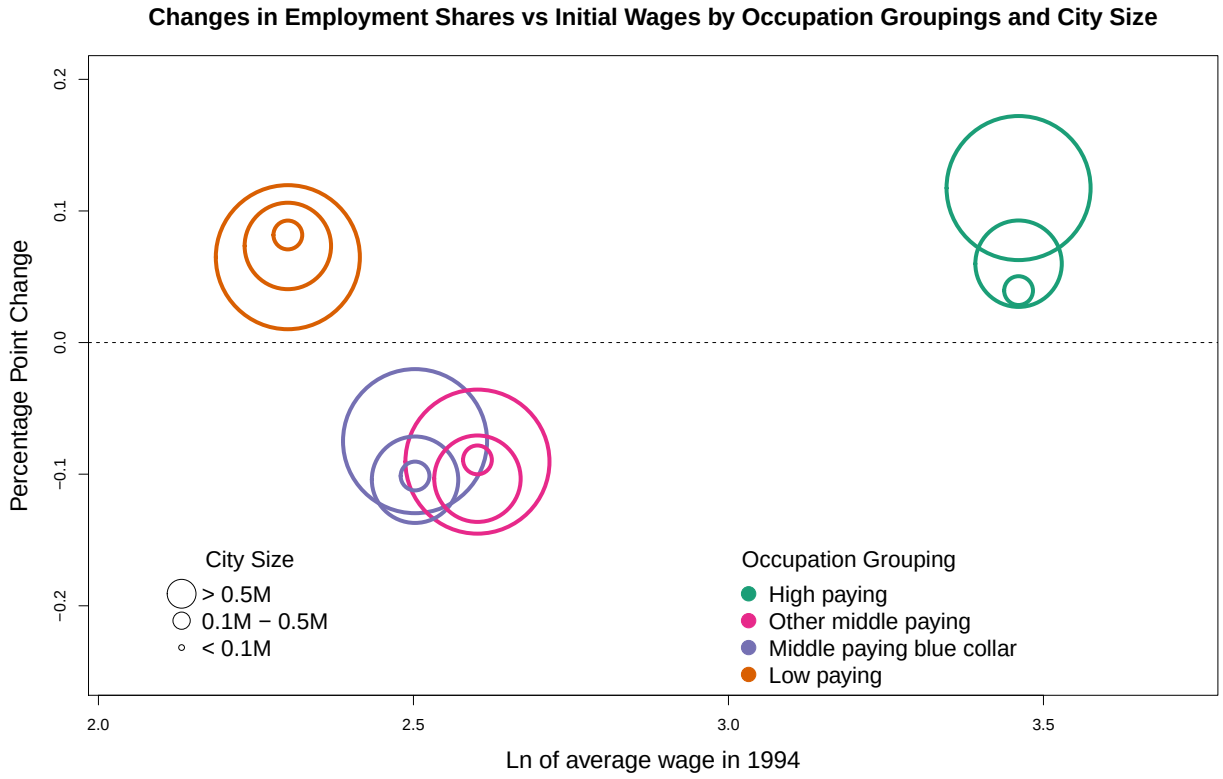
Notes: The figure shows the percentage point change in employment of the high-, low-, middle-paying (CS 43, 46, 47, 48, 62) with wages above median in 1994 and other middle-paying jobs plotted against the logarithm of average wage in cities > .05m in 1994 in different city groups.

Figure 12 – Labor market polarization across three different city size groups, 1994-2015: white-collar middle-paying (CS 46 and 54) vs. other middle-paying jobs.



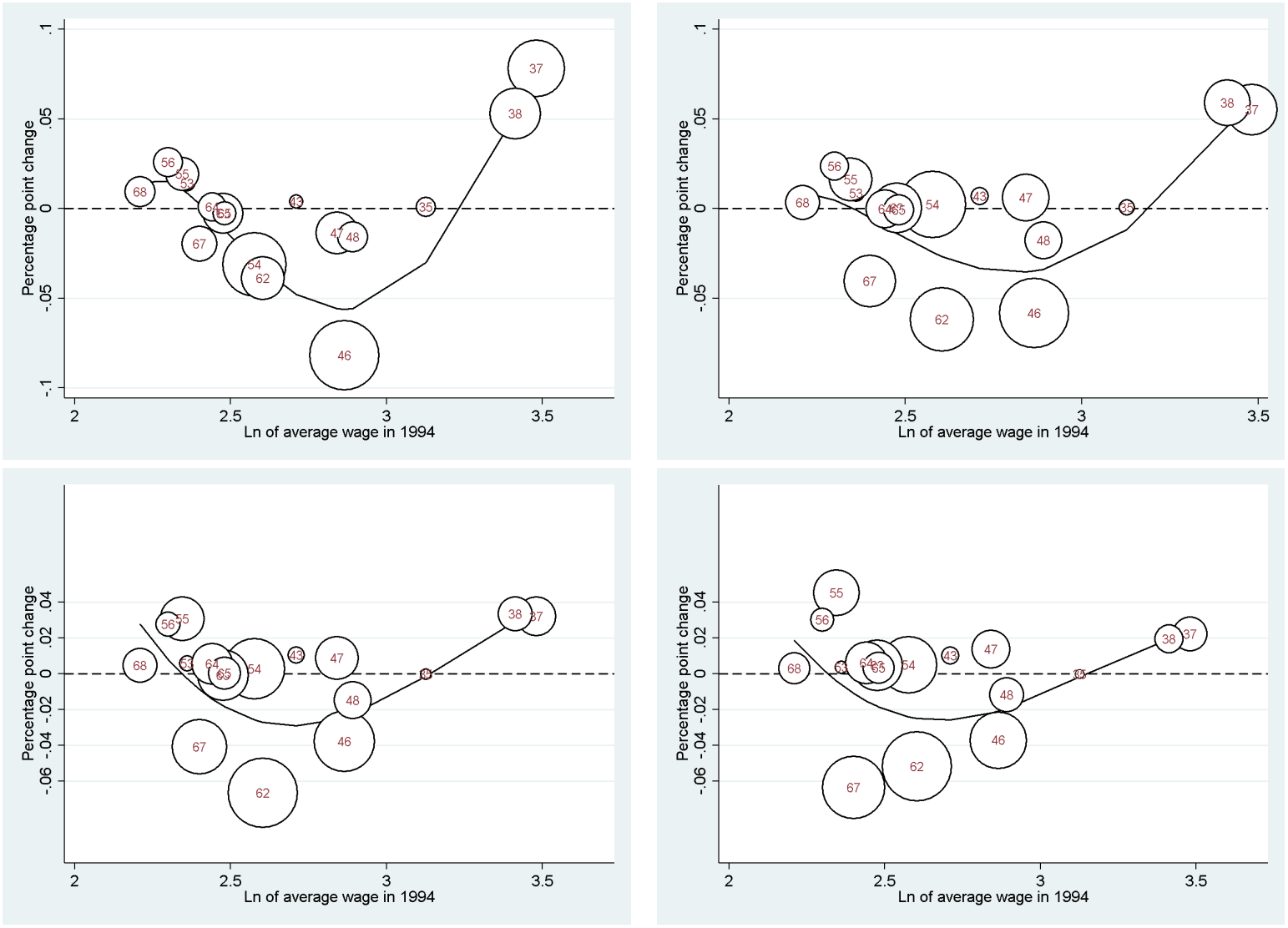
Notes: The figure shows the percentage point change in employment of the high-, low-, white-collar middle-paying (CS 46 and 54) and other middle-paying jobs plotted against the logarithm of average wage in cities > .05m in 1994 in different city groups.

Figure 13 – Labor market polarization across three different city size groups, 1994-2015: blue-collar middle-paying (CS 62, 63, 64, 65 and 67) vs. other middle-paying jobs.



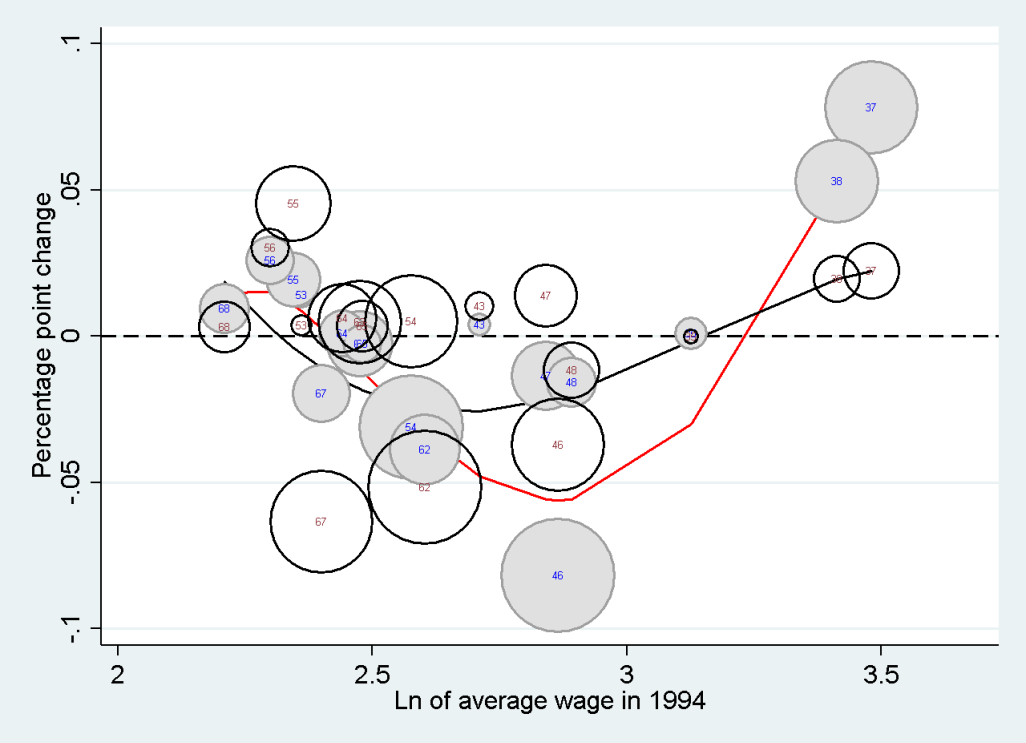
Notes: The figure shows the percentage point change in employment of the high-, low-, blue-collar middle-paying (CS 62, 63, 64, 65 and 67) and other middle-paying jobs plotted against the logarithm of average wage in cities > .05m in 1994 in different city groups.

Figure 14 – 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 15 – 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 16 – Comparison of means of employment shares in different occupations, cities >0.5m vs. 0.05-0.1m.

Item	high-paid	low-paid	middle-paid	routine-4	non-routine-4	middle-paid above median	middle-paid below median
1994							
mean, cities >0.5m	0.188	0.123	0.690	0.327	0.363	0.363	0.327
mean, cities 0.05-0.1m	0.080	0.139	0.780	0.427	0.354	0.360	0.421
difference in pct points	0.107***	-0.017*	-0.09***	-0.10***	0.009	0.004	-0.094***
2015							
mean, cities >0.5m	0.303	0.188	0.509	0.219	0.290	0.233	0.276
mean, cities 0.05-0.1m	0.117	0.219	0.664	0.316	0.348	0.287	0.377
difference in pct points	0.186***	-0.031***	-0.156***	-0.097***	-0.058***	-0.054***	-0.101***

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

Table 17 – 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	high-paid	middle-paid	low-paid	routine-4	non-routine-4	middle-paid above median	middle-paid below median
mean change ppct, cities >0.5m	0.104	0.072	-0.177	-0.109	-0.067	-0.128	-0.048
mean change ppct, cities 0.05-0.1m	0.032	0.084	-0.115	-0.119	0.004	-0.066	-0.049
difference in pct points	0.073***	-0.012**	-0.061***	0.010	-0.071***	-0.062***	0.001
mean growth, cities >0.5m	0.595	0.566	-0.259	-0.335	-0.185	-0.356	-0.150
mean growth, cities 0.05-0.1m	0.419	0.639	-0.148	-0.271	0.021	-0.182	-0.114
difference in growth	0.176***	-0.073	-0.111***	-0.064***	-0.206***	-0.174***	-0.035**

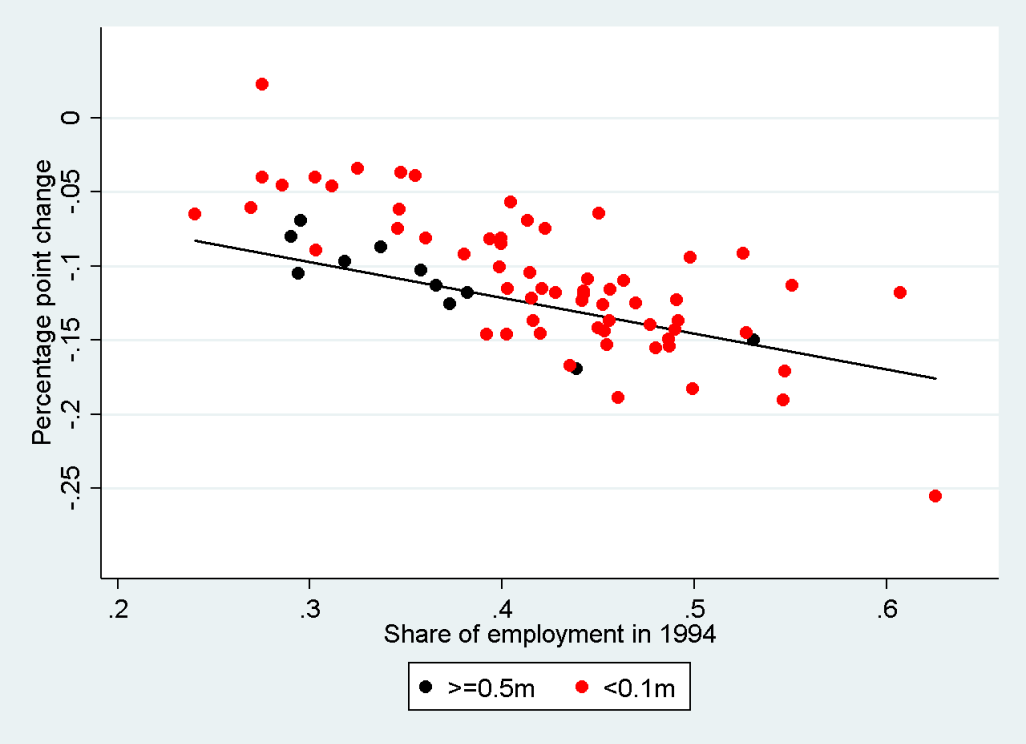
Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Figure 16 – Comparing labor market polarization 1994-2015 in Lyon and cities between .05-1m inhabitants.



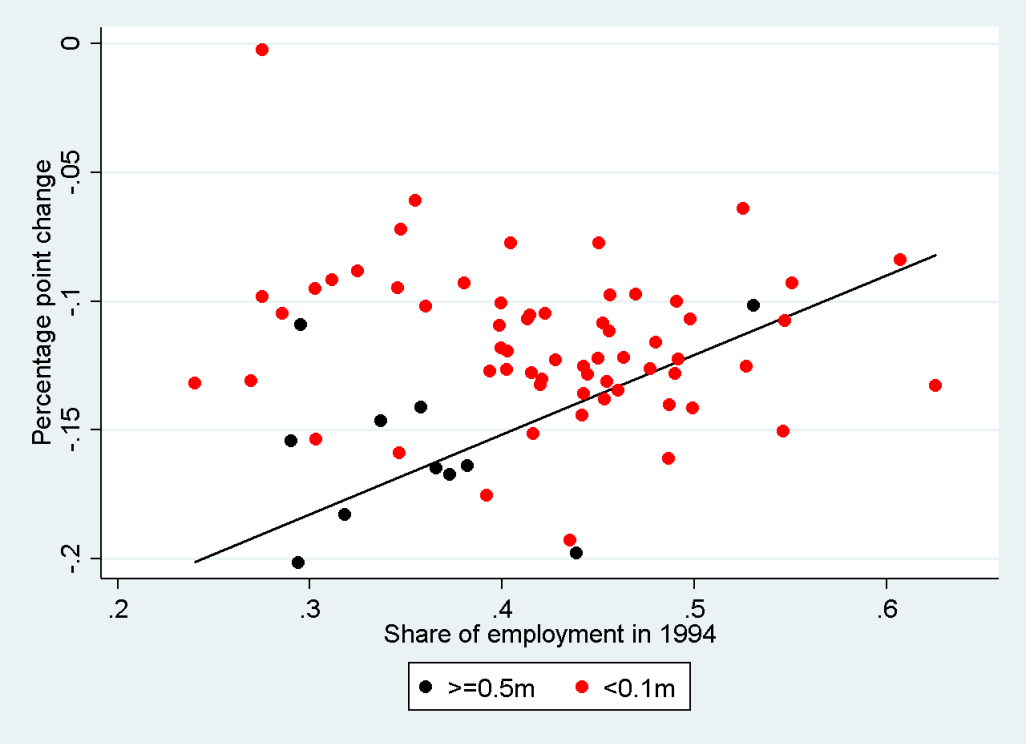
Notes: The figure shows the percentage point change in employment of the considered 2-digit CS categories plotted against their average wage in cities > 0.05m in 1994. Numbers pertain to 2-digit CS categories represented. Grey circles stand for Lyon while white for small city shares. Circle sizes correspond 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 17 – Exposure to RTI4 jobs and change in RTI4 jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



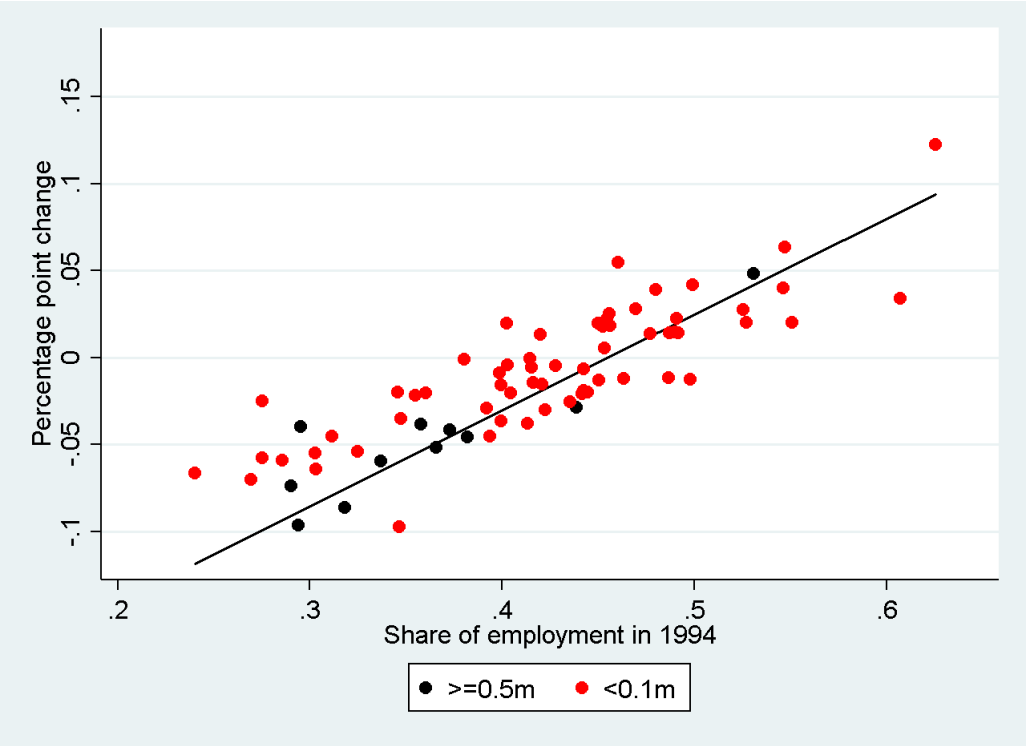
Notes: The figure shows the percentage point change in employment of RTI4 jobs (CS 48, 54, 62 and 67) between 1994-2015 plotted against the share of RTI4 jobs in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted (by 1990 population) fit of the relation between employment changes and the initial RTI4 exposure.

Figure 18 – Exposure to RTI4 jobs and change in middle-paying jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



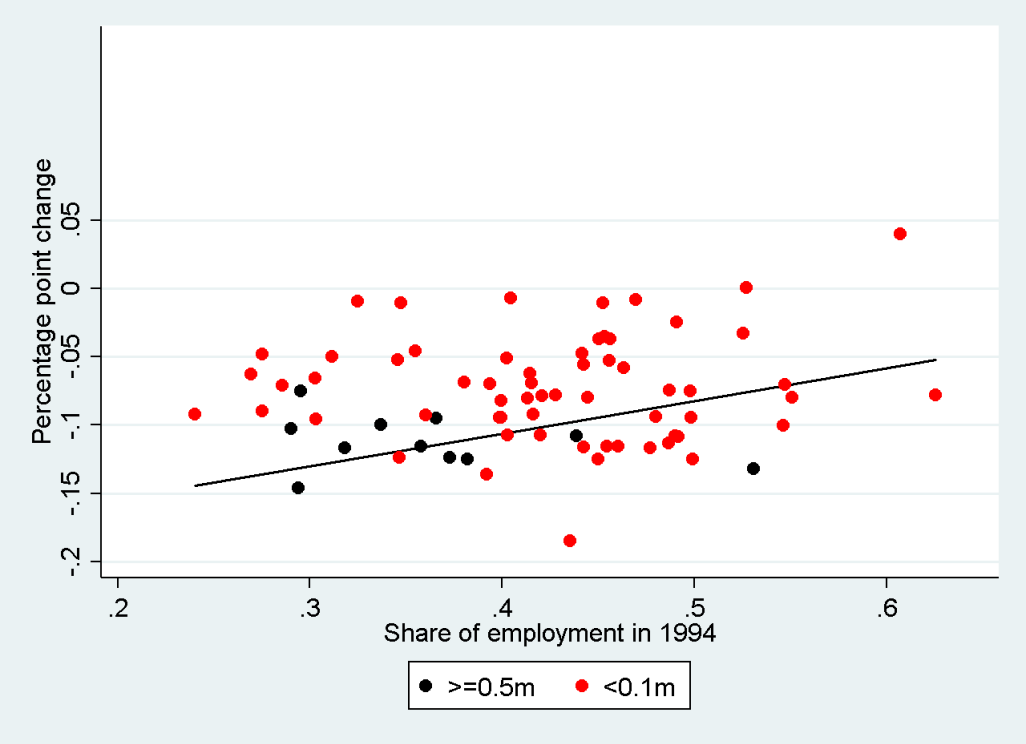
Notes: The figure shows the percentage point change in employment of middle-paying jobs between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above 0.5m (between 0.05-0.1m). The line shows a linear, population-weighted (by 1990 population) fit of the relation between employment changes and the initial RTI4 exposure.

Figure 19 – Exposure to RTI4 jobs and change in middle-paying non-routine-4 jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



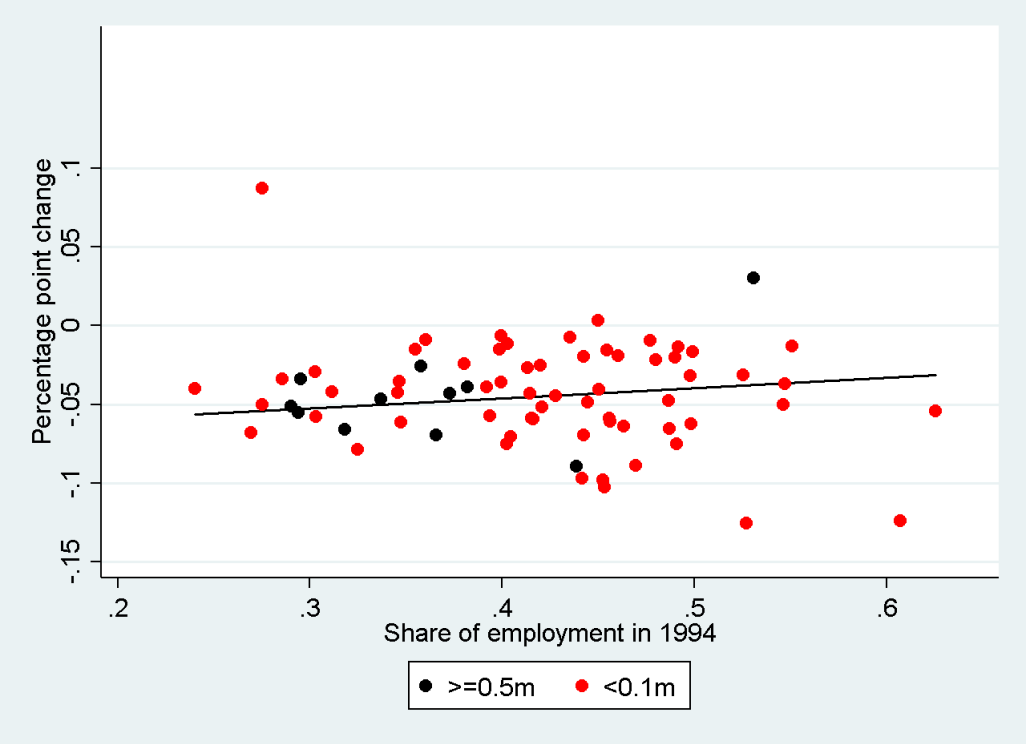
Notes: The figure shows the percentage point change in employment of middle-paying non-routine-4 jobs (CS 43, 46, 47, 63, 64 and 65) between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted (by 1990 population) fit of the relation between employment changes and the initial RTI4 exposure.

Figure 20 – Exposure to RTI4 jobs and change in middle-paying jobs (with wages above the median in 1994) in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



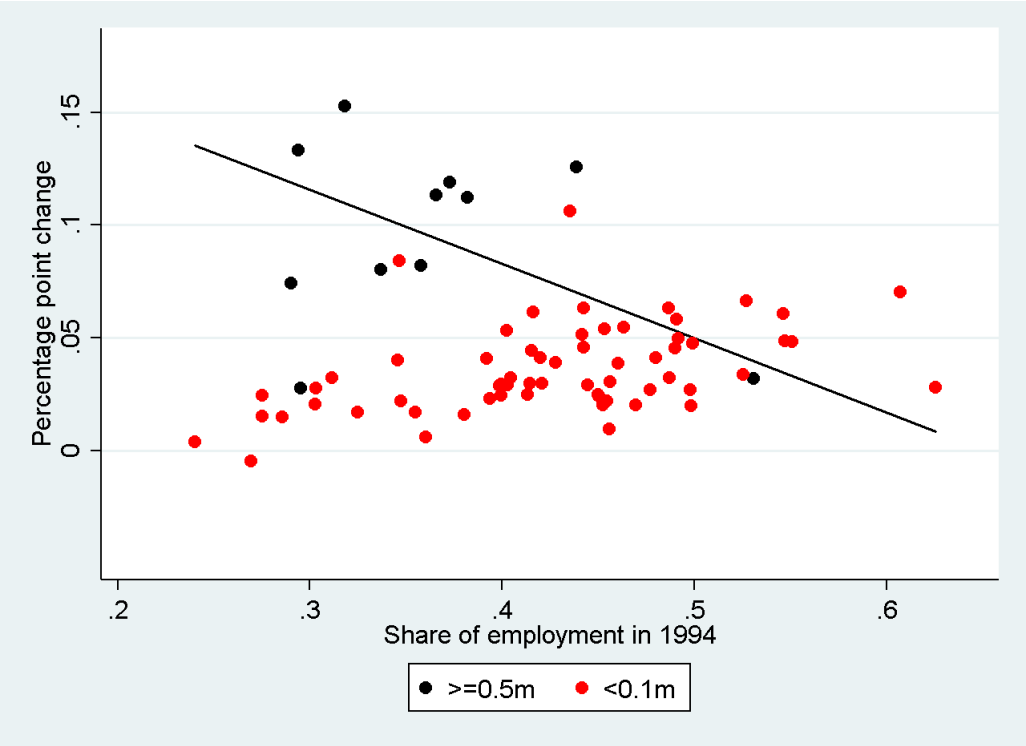
Notes: The figure shows the percentage point change in employment of middle-paying jobs with wages above the median in 1994 (CS 43, 46, 47, 48 and 62) between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted (by 1990 population) fit of the relation between employment changes and the initial RTI4 exposure.

Figure 21 – Exposure to RTI4 jobs and change in middle-paying jobs (with wages below the median in 1994) in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



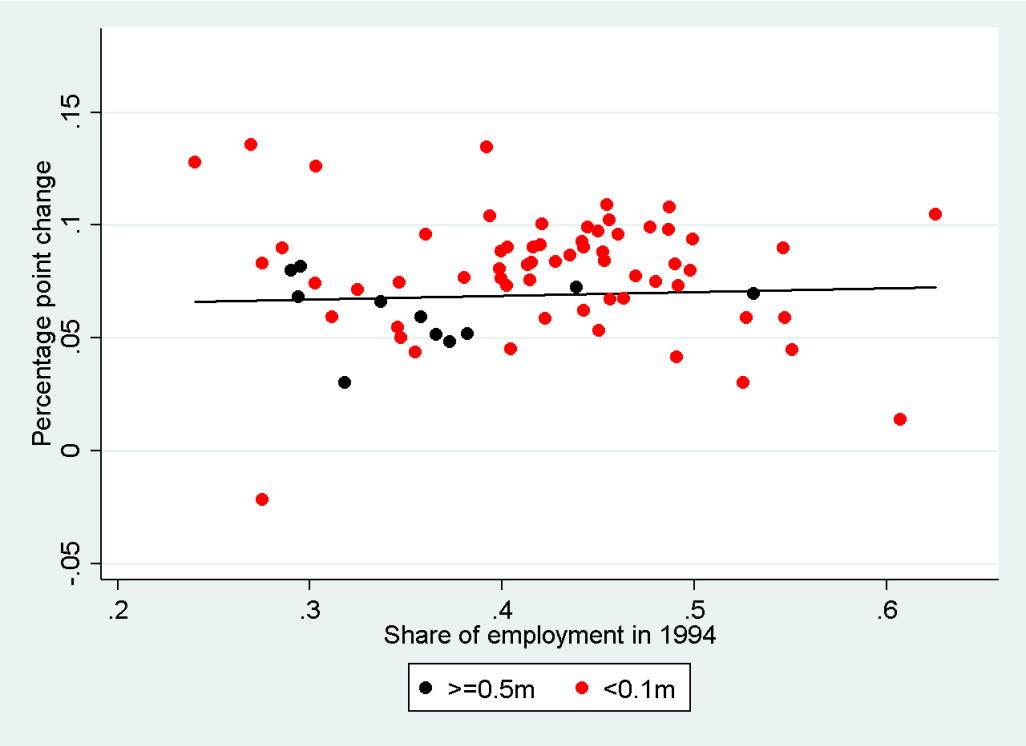
Notes: The figure shows the percentage point change in employment of middle-paying jobs with wages below the median in 1994 (CS 54, 63, 64, 65 and 67) between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted (by 1990 population) fit of the relation between employment changes and the initial RTI4 exposure.

Figure 22 – Exposure to RTI4 jobs and change in high-paying jobs in cities >0.5m and cities between 0.05-0.1m inhabitants, 1994-2015.



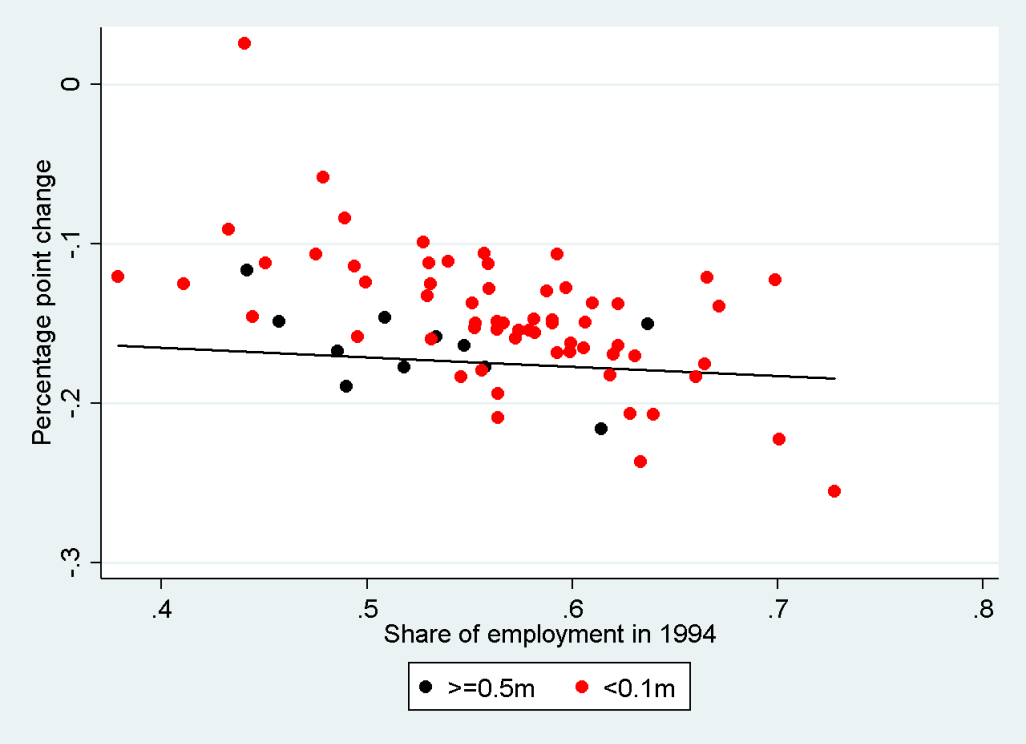
Notes: The figure shows the percentage point change in employment of high-paying jobs (CS 23, 35, 37 and 38) between 1994-2015 plotted against the share of RTI4 jobs (CS 48, 54, 62 and 67) in employment in 1994. Each black (red) dot represents a city above .5m (between 0.05-0.1m). The line shows a linear, population-weighted (by 1990 population) fit of the relation between employment changes and the initial RTI4 exposure.

Figure 23 – 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 (by 1990 population) fit of the relation between employment changes and the initial RTI4 exposure.

Figure 24 – 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.

Table 18 – Employment share changes 1994-2015 and exposure to 4 most routine/offshorable occupations (RTI4) in 1994, Part I

Employment share change of high-paid jobs								
employment share of RTI4 in 1994	-0.325***	0.007	-0.253**	0.060	-0.266**	0.060	-0.251**	0.059
	(0.093)	(0.048)	(0.108)	(0.048)	(0.109)	(0.056)	(0.112)	(0.045)
large X employment share of RTI4 in 1994			0.145***	0.160***	0.141***	0.158***	0.140***	0.142***
			(0.037)	(0.047)	(0.037)	(0.048)	(0.037)	(0.045)
constant	0.213***	0.043*	0.150***	0.012	0.156***	0.012	0.149***	0.012
	(0.038)	(0.022)	-0.048	-0.021	(0.048)	(0.024)	(0.050)	(0.020)
R^2	0.308	0.000	0.519	0.428	0.522	0.425	0.523	0.402
Employment share change of low-paid jobs								
employment share of RTI4 in 1994	0.016	-0.016	-0.006	-0.033	-0.006	-0.029	-0.009	-0.052
	(0.021)	(0.054)	(0.030)	(0.057)	(0.030)	(0.062)	(0.030)	(0.046)
large X employment share of RTI4 in 1994			-0.044***	-0.051***	-0.041***	-0.048**	-0.046***	-0.056***
			(0.015)	(0.019)	(0.015)	(0.019)	(0.015)	(0.018)
constant	0.062***	0.083***	0.082***	0.093***	0.080***	0.091***	0.083***	0.102***
	(0.007)	(0.023)	(0.013)	(0.025)	(0.013)	(0.028)	(0.013)	(0.020)
R^2	0.006	0.002	0.150	0.066	0.134	0.059	0.178	0.113
Employment share change of middle-paid jobs								
employment share of RTI4 in 1994	0.309***	0.009	0.259**	-0.027	0.272**	-0.031	0.044	-0.002
	(0.095)	(0.055)	(0.105)	(0.059)	(0.107)	(0.069)	(0.108)	(0.046)
large X employment share of RTI4 in 1994			-0.101***	-0.108***	-0.101***	-0.110***	-0.093**	-0.090**
			(0.036)	(0.039)	(0.036)	(0.041)	(0.04)	(0.036)
constant	-0.276***	-0.126***	-0.231***	-0.105***	-0.236***	-0.103***	-0.137***	-0.117***
	(0.039)	(0.025)	(0.046)	(0.027)	(0.047)	(0.031)	(0.048)	(0.020)
R^2	0.335	0.001	0.459	0.181	0.467	0.186	0.312	0.165
Observations	73	73	73	73	71	71	71	71
population weighted?	y	n	y	n	y	n	y	n
no outliers in RTI4 share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: RTI4 jobs are the 4 most routine or offshorable occupations with the highest RTI or OFF-GMS indexes (CS 48, 54, 62 and 67). Robust standard errors. 11 cities > 0.5m and 62 cities between 50,000-100,000 inhabitants. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 19 – Employment share changes 1994-2015 and exposure to 4 most routine/offshorable occupations (RTI4) in 1994, Part II

Employment share change of routine jobs								
employment share of RTI4 in 1994	-0.241***	-0.401***	-0.274***	-0.425***	-0.260***	-0.409***	-0.256***	-0.369***
	(0.058)	(0.057)	(0.063)	(0.058)	(0.062)	(0.058)	(0.060)	(0.049)
large X employment share of RTI4 in 1994			-0.064***	-0.071***	-0.066***	-0.073***	-0.062***	-0.064***
			(0.024)	(0.024)	(0.024)	(0.024)	(0.023)	(0.021)
constant	-0.025	0.057**	0.003	0.071***	-0.001	0.066***	-0.003	0.047**
	(0.023)	(0.023)	(0.027)	(0.024)	(0.027)	(0.024)	(0.026)	(0.020)
R^2	0.400	0.535	0.496	0.577	0.480	0.538	0.485	0.514
Employment share change of non-routine-4 jobs								
employment share of RTI4 in 1994	0.551***	0.410***	0.532***	0.398***	0.532***	0.378***	0.527***	0.363***
	(0.043)	(0.037)	(0.047)	(0.038)	(0.048)	(0.035)	(0.049)	(0.031)
large X employment share of RTI4 in 1994			-0.037**	-0.037**	-0.035**	-0.037*	-0.039**	-0.042**
			(0.015)	(0.019)	(0.015)	(0.020)	(0.015)	(0.020)
constant	-0.251***	-0.183***	-0.234***	-0.176***	-0.235***	-0.169***	-0.232***	-0.161***
	(0.018)	(0.015)	(0.020)	(0.016)	(0.021)	(0.015)	(0.021)	(0.013)
R^2	0.860	0.727	0.874	0.743	0.871	0.714	0.872	0.742
Employment share change of middle-paid jobs above median wage								
employment share of RTI4 in 1994	0.242***	0.065	0.188**	0.029	0.196**	0.028	0.180**	-0.013
	(0.091)	(0.063)	(0.076)	(0.059)	(0.079)	(0.070)	(0.079)	(0.043)
large X employment share of RTI4 in 1994			-0.109***	-0.107***	-0.108***	-0.108***	-0.112***	-0.113***
			(0.021)	(0.017)	(0.021)	(0.017)	(0.020)	(0.017)
constant	-0.203***	-0.105***	-0.155***	-0.084***	-0.158***	-0.084***	-0.152***	-0.067***
	(0.036)	(0.025)	(0.033)	(0.024)	(0.034)	(0.028)	(0.034)	(0.018)
R^2	0.262	0.018	0.443	0.141	0.445	0.143	0.456	0.177
Employment share change of middle-paid jobs below median wage								
employment share of RTI4 in 1994	0.068	-0.055	0.071	-0.056	0.077	-0.059	0.085	0.004
	(0.075)	(0.059)	(0.095)	(0.066)	(0.099)	(0.079)	(0.094)	(0.052)
large X employment share of RTI4 in 1994			0.007	-0.001	0.008	-0.002	0.010	0.009
			(0.044)	(0.041)	(0.045)	(0.043)	(0.044)	(0.037)
constant	-0.073***	-0.021	-0.076*	-0.021	-0.078*	-0.019	-0.082**	-0.047**
	(0.024)	(0.024)	(0.042)	(0.029)	(0.043)	(0.034)	(0.041)	(0.022)
R^2	0.048	0.020	0.050	0.020	0.055	0.019	0.080	0.002
Observations	73	73	73	73	71	71	71	71
population weighted?	y	n	y	n	y	n	y	n
no outliers in RTI4 share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: RTI4 jobs are the 4 most routine or offshorable occupations with the highest RTI or OFF-GMS indexes (CS 48, 54, 62 and 67). Robust standard errors. 11 cities > 0.5m and 62 cities between 50,000-100,000 inhabitants. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 20 – Employment of most offshorable jobs' share changes 1994-2015 and exposure to 6 most offshorable occupations (OFF6) in 1994

Employment share change of OFF6 jobs								
employment share of OFF6 in 1994	-0.058	-0.312***	-0.138	-0.340***	-0.120	-0.335***	-0.107	-0.259***
	(0.120)	(0.076)	(0.111)	(0.079)	(0.117)	(0.086)	(0.112)	(0.060)
large X employment share of OFF6 in 1994			-0.073***	-0.060***	-0.074***	-0.063***	-0.070***	-0.053***
			(0.020)	(0.019)	(0.020)	(0.020)	(0.019)	(0.017)
constant	-0.142**	0.027	-0.070	0.047	-0.079	0.046	-0.088	0.000
	(0.069)	(0.044)	(0.064)	(0.046)	(0.068)	(0.050)	(0.064)	(0.033)
R^2	0.018	0.291	0.255	0.366	0.256	0.338	0.253	0.305
Observations	73	73	73	73	71	71	71	71
population weighted?	y	n	y	n	y	n	y	n
no outliers in OFF6 share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: The category of 6 most offshorable occupations (OFF6) are those with highest OFF-GMS index (CS 46, 48, 54, 62, 65, 67). Robust standard errors. 11 cities > 0.5m and 62 cities between 50,000-100,000 inhabitants. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Table 21 – 2-digit PCS categories and representative 4-digit PCS categories (private sector)

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

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

Table 22 – Basic statistics by 2 digit CS categories: Full table.

CS 2-digit	description	employment share in %		average city wage (in 2015 euros)		Routine (index values)	Offshorable (index values)
		1994	2014	1994	2015		
	high-paying occupations						
23	CEOs	1.0	0.9	42.82	59.20	-0.75	-0.59
37	**managers and professionals**	**6.2**	**10.2**	32.52	38.56	-0.75	-0.59
38	**engineers**	**5.1**	**9.0**	30.36	33.68	-0.82	-0.39
35	creative professionals	0.5	0.5	22.82	31.80	-0.72	-0.49
	medium-paying occupations						
48	**supervisors and foremen**	**4.1**	**2.7**	18.03	21.86	0.42	1.23
46	**mid-level professionals**	**12.3**	**7.6**	17.54	21.20	-0.48	-0.16
47	**technicians**	**5.7**	**6.3**	17.15	20.60	-0.40	-0.29
43	mid-level health professionals	0.8	1.5	15.05	18.05	-0.35	-0.57
62	**skilled industrial workers**	**14.1**	**9.3**	13.52	17.99	0.38	1.24
54	**office workers**	**11.8**	**11.2**	13.17	16.98	2.03	0.87
65	**transport and logistics personnel**	**2.9**	**3.0**	11.96	16.00	0.33	0.27
63	**skilled manual workers**	**8.0**	**8.3**	11.90	15.50	0.17	-0.33
64	**drivers**	**5.0**	**5.5**	11.50	14.46	-1.50	-0.63
67	**unskilled industrial workers**	**10.9**	**5.7**	11.02	14.72	0.45	2.09
	low-paying occupations						
53	security workers	0.7	1.4	10.60	14.60	-0.28	-0.51
55	**sales-related occupations**	**5.4**	**8.3**	10.44	13.74	0.30	-0.57
56	personal service workers	2.2	4.8	9.97	12.63	-0.43	-0.57
68	**unskilled manual workers**	**3.3**	**3.8**	9.11	13.27	0.06	-0.36

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

D Pre-1994 Labor market developments

In this Appendix we document the relevant changes in the French labor market before the coverage of the detailed DADS data (starting in 1994).

Labor market polarization in France might have begun earlier than 1994: although the modern ICT were not widely used pre-1994 there was a significant advance in automation in manufacturing through CAD/CAM systems, early adoption of basic computer text editors or spreadsheets. There was also an increase in offshoring possibilities for French companies with such developments as the Spanish or Portuguese accession to the EEC in 1986 or the opening up of Eastern European countries in 1989. The strength of the automation or offshoring shocks is unclear, however, and the most offshorable occupations (CS 48, 62 and 67) related to manufacturing might have been affected the earliest. It is therefore instructive to detail some of the pre-1994 developments.

There are two data sources that allow to track occupations at the 2-digit PCS level back to 1982 when the PCS classification was introduced: the French Labor Market Survey (yearly data) and

the Census (1982, 1990 and 1999). The publicly available Labor Market Survey gives data at the department but not at the commune level, hence it is impossible to precisely characterize city-level labor markets. The Census, on the other hand, gives the commune location of respondents but does not give data about hours worked or wages. We use the Census as we are interested in the shares of employment in cities, but in contrast to data presented in main text the patterns will refer to shares of people employed and not actual hours worked. We use the publicly available individual data for the 1982, 1990 and 1999 censuses (covering 1/4th for 1982 and 1990, and 1/20th for 1999 of the entire population respectively). The publicly available census data for subsequent years is presented in such a way that misses important details (either in terms of communes or details of occupations) disallowing the same exercise.

The 1982-1999 counterparts to Tables 3-6 using Census data are as follows:

Table 23 – Share of 4 highest-paying occupations per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.18	0.12	0.10	0.10	0.09	0.09
1990	0.24	0.16	0.13	0.12	0.11	0.11
1999	0.26	0.17	0.14	0.13	0.12	0.11
change in ppct 1982-1990	0.06	0.04	0.03	0.03	0.02	0.02
change in ppct 1990-1999	0.02	0.01	0.01	0.00	0.00	0.00
change in ppct 1982-1999	0.08	0.05	0.04	0.03	0.03	0.02

Table 24 – Share of 10 middle-paying occupations per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.66	0.71	0.74	0.73	0.74	0.73
1990	0.61	0.68	0.69	0.70	0.71	0.71
1999	0.56	0.64	0.65	0.66	0.67	0.68
change in ppct 1982-1990	-0.05	-0.03	-0.04	-0.04	-0.04	-0.02
change in ppct 1990-1999	-0.05	-0.04	-0.04	-0.03	-0.04	-0.03
change in ppct 1982-1999	-0.10	-0.07	-0.08	-0.07	-0.07	-0.06

The conclusions from this exercise are as follows. First of all, exposure to the most routine and offshorable jobs is much higher in 1982 for large cities above 0.5m inhabitants than in 1994 in the DADS data, and the discrepancies in terms of shares of high- middle- and RTI4 jobs across city sizes are lower. Employment shares of the 4 most routine and offshorable occupations and, more generally, middle-paying jobs indeed decline faster in larger cities whether in 1982-1990 or in the entire 1982-1999 period. The labor market polarization across cities manifests itself as our theory predicts: high paying jobs' shares increase most in largest cities, as found in the exhaustive

Table 25 – Share of 4 lowest-paying occupations per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.15	0.17	0.16	0.17	0.17	0.18
1990	0.15	0.16	0.18	0.18	0.18	0.18
1999	0.18	0.19	0.21	0.21	0.21	0.22
change in ppct 1982-1990	0.00	0.00	0.01	0.01	0.01	0.01
change in ppct 1990-1999	0.02	0.02	0.03	0.03	0.03	0.03
change in ppct 1982-1999	0.02	0.02	0.04	0.04	0.05	0.04

Table 26 – Share of the 4 most routine and offshorable occupations (CS 48, 54, 62 and 67) per metropolitan area size, Census data 1982-1999.

Agglo.size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
1982	0.36	0.39	0.43	0.42	0.43	0.43
1990	0.28	0.33	0.34	0.35	0.37	0.38
1999	0.22	0.27	0.29	0.30	0.31	0.33
change in ppct 1982-1990	-0.07	-0.05	-0.08	-0.06	-0.06	-0.05
change in ppct 1990-1999	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06
change in ppct 1982-1999	-0.13	-0.12	-0.14	-0.12	-0.13	-0.11

DADS data for 1994-2015. Low-paying jobs do not increase at all in largest cities in 1982-1990 and increase less in terms of percentage points over 1990-1999 and the entire 1982-1999 period.

Similar patterns obtain for 1982-1994 using the Labor Market Survey data while classifying departments by largest city.

For individual occupational categories, the routine/offshorable job categories whose employment declines most in the studied years 1994-2015 in the DADS data in large cities are in particular PCS 46 and 54 (mid-level professionals and office workers respectively), whereas it is 62 and 67 (skilled and unskilled industrial workers respectively) for small cities (cf. the patterns in Figure 15).[29] A part of the answer of such a differential evolution may lay in the fact that large cities had different shares of these jobs at the beginning of the 1990s (see Table 27) than small cities, and such a discrepancy existed already in 1982. In particular, the share of mid-level professionals and office workers in employment was higher than that of industrial workers in 1982 in the largest cities above .75m inhabitants while the opposite is true for smaller cities. This feature of data may be explained by different deindustrialization across time and geographies as shown in Table 28 where the share of industry employment at Census years is given for the period 1968-2015. Already over the period 1968-1982 large cities experienced faster deindustrialization than small cities. Reports from research bodies as the INSEE or DATAR (Délégation interministérielle à l'aménagement du territoire et à

[29]The PCS 46 category contains heterogeneous professions that were differentially impacted by automation/offshoring. For example, occupations such as drafters, secretaries, photographers, sales in insurance, real estate, finance or advertising included in this category have RTI scores above 2; some of them are also very offshorable.

l'attractivité régionale) indicate the following. Internal offshoring of manufacturing tasks within France might have played a part due to the reduction in internal transport costs (both because of highway and railway construction), environmental regulations to keep polluting industries out of high density areas or a deliberate government policy to decentralize economic activity across France (e.g. moving public engineering schools outside Paris). For Ile-de-France, deindustrialization was largely due to the reorganization of the automobile (that moved out of large cities) and defense industries (idem, with aerospace moving to Toulouse in particular). To an unknown extent firm reorganization and shifting tasks outside the boundaries of firms (e.g. legal services, general and administrative or cleaning premises) that cannot be precisely measured was responsible for the fall in manufacturing value added overall. This, together with moving tasks within multi-establishment firms might have caused some of the tasks to be offshored within France from large to smaller cities.

Table 27 – Employment share of selected middle-paying occupations per metropolitan area size, Census data 1982-1999.

Agglo. Size		Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
mid-level professionals (CS 46) and office workers (CS 54)	1982	0.27	0.23	0.19	0.20	0.19	0.19
	1990	0.26	0.23	0.21	0.22	0.21	0.21
	1999	0.26	0.25	0.21	0.22	0.21	0.21
skilled (CS 62) and unskilled industrial workers (CS 67)	1982	0.14	0.19	0.25	0.24	0.25	0.25
	1990	0.10	0.16	0.18	0.19	0.20	0.22
	1999	0.07	0.11	0.14	0.15	0.16	0.18

Notes: Data from the publicly available individual data for the 1982, 1990 and 1999 censuses (covering 1/4th for 1982 and 1990, and 1/20th for 1999 of the entire population respectively). Aggregation to the city-level (based on unites urbaines as of 2015). Includes individuals between 25-64 of age.

Table 28 – Employment share in industry per metropolitan area size, Census data 1968-2015.

share of industry	Paris	>0.75m	0.5-0.75m	0.2-0.5m	0.1-0.2m	0.05-0.1m
1968	0.32	0.30	0.39	0.32	0.34	0.29
1975	0.29	0.28	0.34	0.30	0.34	0.30
1982	0.24	0.23	0.29	0.26	0.30	0.27
1990	0.19	0.20	0.22	0.22	0.24	0.24
1999	0.13	0.15	0.17	0.17	0.19	0.19
2015	0.07	0.11	0.11	0.11	0.13	0.13
ppct chng 68-82	-0.08	-0.07	-0.10	-0.06	-0.04	-0.02
ppct chng 68-90	-0.13	-0.11	-0.16	-0.10	-0.10	-0.05

Notes: Raw, exhaustive data from Censuses were prepared by the INSEE for each commune. Aggregation to the city-level (based on unites urbaines as of 2015) done by authors. Includes individuals between 25-54 of age.

E Toy model of middle skill sector

In this section, we further investigate the heterogeneity across middle-skill jobs. To this purpose, let us consider the following simplified version of our model that is also zoomed only on the middle skill workers.

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

$$\min \{q(m_l, c) + k_l, q(m_h, c) + \gamma_k k_h\} \tag{55}$$

Capital is still provided using an exogenous production function so that the price of capital is ξ.

The maximization of a firm in city c is then:

$$\max \left\{ \begin{array}{l} \min \{q(m_l, c) + k_l, q(m_h, c) + \gamma_k k_h\} \\ - w(l, c)q(m_l, c) - w(h, c)q(m_h, c) - \xi k_l - \xi k_h \end{array} \right\} \tag{56}$$

On productivities, we assume that $A(m_1, 1) > A(m_1, 2)$.

This leads to the following conditions:

$$w(l, c) \leq 1$$
$$w(h, c) \leq 1$$

Furthermore, only labor is used for task m_l when $1 - w(l, c) \geq 1 - \xi$. Only labor is used for task m_h when $1 - w(h, c) \geq \gamma_k - \xi$.

Location and sector decisions To also streamline the model in terms of location decisions, we assume that the cost to live in city c is $r(c)$. To be compensated against differences in costs of living, workers require a larger nominal wage. By denoting by $w(l, c)$ the wage to work in the lower-wage middle skill sector in city c, we need that:

$$w(l, 1) - r(1) = w(l, 2) - r(2)$$

for lower wage middle skill jobs to be in both cities and

$$w(h, 1) - r(1) = w(h, 2) - r(2).$$

We assume that if they are not in the middle-skill sector, households have a reservation utility of $\bar{U}_l$ if they are of the m_l type and of $\bar{U}_h > \bar{U}_l$ if they are of the m_l type. A necessary condition to have both type working in the middle-skill sector in equilibrium is that

$$\bar{U}_l = w(l,1) - r(1) = w(l,2) - r(2)$$
$$\bar{U}_h = w(h,1) - r(1) = w(h,2) - r(2)$$

In each city, the two middle-skill types hesitate to work in different sectors. This leads to the reservation wage $\bar{w}(c,\sigma)$.

We assume that m_h middle-skill workers have better options than m_l workers, which implies that $\bar{w}(c,l) < \bar{w}(c,h)$, for both cities $\{1,2\}$. We assume that reservation wages also satisfy

$$\bar{w}(1,l) - r(1) = \bar{w}(2,l) - r(2)$$
$$\bar{w}(1,h) - r(1) = \bar{w}(2,h) - r(2)$$

Given that the cost of live in the large city is larger, $r(1) > r(2)$, reservation wages also satisfy $\bar{w}(1,h) > \bar{w}(2,h)$ and $\bar{w}(1,l) > \bar{w}(2,l)$.

Equilibrium Let us investigate the presence of the two job categories in the two cities as a function of the price of capital.

In equilibrium, we have that: $\bar{w}(c,\sigma) = w(c,\sigma)$. Labor is predominantly used in the low wage middle-skilled sector in city c when $w(l,c) \leq \xi$ and labor is predominantly used in the high wage middle-skilled sector when $w(h,c) \leq \xi + 1 - \gamma_k$.

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

Lemma 10. *For both high- and low-wage middle skilled jobs, automation takes place first in the large city and then in the small city.*

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

$$w(l,c) \geq w(h,c) + \gamma_k - 1$$

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

Lemma 11. *When γ_k is sufficiently low, in both cities, automation of low-wage middle-skilled jobs takes place before automation of high-wage middle-skilled jobs.*

Lemma 10 and 11, we can then explain the timing of the destruction of jobs across and within cities. Lower-wage jobs first disappear in the large city and then in the small and, with some delay, we observe the same for higher-wage jobs.

What remains to be investigated is the relative timing of the disappearance of higher-wage middle-skilled jobs in the large city with the disappearance of jobs in the small city. The following Lemma clarifies this situation:

Lemma 12. *There exists a value γ_k satisfying the condition of lemma 11 such that the automation of high-wage middle skilled jobs in the large city takes place at the same time as automation of low-wage middle-skilled jobs in the small city:*

$$w(l,2) = w(h,1) + \gamma_k - 1$$

Proof. We can select γ_k such that $w(l,2) = w(h,1) + \gamma_k - 1$ consistently $w(l,2) \geq w(h,2) + \gamma_k - 1$ and $w(l,1) \geq w(h,1) + \gamma_k - 1$ when $w(h,1) \geq w(h,2)$, which we assumed. □

Extension Let us now suppose labor and capital are still marginally perfect substitutes but there is also a minimum of workers needed to run capital.

$$\min\{q(m_l,c) + \min\{q(m_l,c,k)/\epsilon, 1\}k_l, q(m_h,c) + \gamma_k \min\{q(m_h,c,k)/\epsilon, 1\}k_h\} \tag{57}$$

Our results go through with that production function except that, when there is automation, the amount of high- or low- wage middle-skilled labor is ϵ instead of 0. Indeed, given the linear structure of the model, either $q(m_l,c,k) = 0$ or $q(m_l,c,k) = \epsilon$. If it equals 0, then production takes place with labor only as the marginal productivity of capital is 0. If $q(m_l,c,k) = \epsilon$, we are back to the previous results as the marginal productivity of capital is 1.