A Proofs of Propositions - for online publication only

A.1 Proof of Proposition 1

First, note that log-supermodularity of H implies that $\forall \omega > \omega'$ and $\forall \sigma > \sigma'$, $H(\omega,\sigma)H(\omega',\sigma') \geq H(\omega',\sigma)H(\omega,\sigma')$. Manipulating this expression, we obtain that $H(\omega,\sigma)/H(\omega,\sigma')$ with $\sigma > \sigma'$ is increasing in ω as:

$$\frac{H(\omega,\sigma)}{H(\omega,\sigma')} > \frac{H(\omega',\sigma)}{H(\omega',\sigma')}.$$

Note then that $H(\omega(m,c),m)/H(\omega(m,c),l) = A(l,c)/A(m,c)p(l)/p(m)$. This latter ratio is larger in City 2 than in City 1 given Assumption 3 and thus $\omega(m,1) < \omega(m,2)$.

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

$$\frac{A(m,c)}{A(h,c)}\frac{p(m)}{p(h)} = H(\omega(h,c),h)/H(\omega(h,c),m) \tag{10}$$

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

A.2 Proof of Proposition 2

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

A.3 Intermediary results on location and sectoral decisions

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

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

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

$$S(\tau) = L\int_0^{\tau}\int_{\sigma}\int_{\omega} f(\omega, M(\omega,c), c, x)d\omega d\sigma dx \tag{11}$$

Proof. We closely follow here the proof of Lemma 1 in Davis and Dingel (2020).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$f(c, \omega, \tau) = f(\omega)\delta^{\text{Dirac}}\left(1 - 1_{\Omega(c,\tau)}\right) \text{ for almost all } \omega \in \Omega(c, \tau). \tag{12}$$

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

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

Indeed, the least skill agent, $\underline{\omega}$ is in both cities when the low-paid sector is in both cities.

Suppose it is not the case. Let us denote by $\omega^* > \underline{\omega}$ the agent with the lowest skill that live in both cities.

This agent is indifferent to live in both cities, that is:

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

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

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

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

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

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

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

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

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

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

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

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

$$A(M(\omega, 1), 1)H(\omega, M(\omega, 1))T(\tau) = A(M(\omega, 2), 2)H(\omega, M(\omega, 2))T(\Gamma(\omega, \tau), 2) \tag{13}$$

In equilibrium, in the location τ, if the agent with skill ω is the marginal buyer, we then have that $r(1, \tau) = r(2, \Gamma(\omega, \tau))$.

In Davis and Dingel (2020), the function Γ would be constant with respect to ω, but, as the larger city has also a comparative advantage in higher-paid sector, we obtain the following result:

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

Proof. Let us consider the function Γ as defined by equation (13), that is:

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

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

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

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

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

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

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

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

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

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

However, this contradicts the optimality of location $(2, \tau_2)$ as this location is strictly dominated by location $(1, \tau_1')$:

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

□

A.4 Proof of Proposition 3

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

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

A.5 Proof of Proposition 4

The change in percentage points of the share of agents in the middle-paid sector in city c is:

$$ds(m,c) = \frac{\left(S\left(T^{-1}\left(h(\omega(m,c),c)\right)\right)\right)' d\omega(m,c) - \left(S\left(T^{-1}\left(h(\omega(h,c),c)\right)\right)\right)' d\omega(h,c)}{S\left(T^{-1}\left(h(\underline{\omega},c)\right)\right) - S\left(T^{-1}\left(h(\overline{\omega}(c),c)\right)\right)} \tag{14}$$

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 Extensions of the model - for online publication only

B.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 3 is satisfied.

Under these conditions, we obtain that:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

B.3 Log-supermodularity of skills across cities

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

and thus $h(\omega, c) = h(\omega, c')$. By continuity, $M(\omega, c) = M(\omega, c')$ on a (right- or left-) neighborhood of ω and thus $h(., c) = h(., c')$ on this neighborhood, thus ensuring that locally $h''(., c) = h''(., c')$

and $h'(.,c) = h'(.,c')$ and in particular that $h''(\omega,c) = h''(\omega,c')$ and $h'(\omega,c) = h'(\omega,c')$. In the end, (20) simplifies into $\xi(V, h(\omega,c)) > \xi(V, h(\omega,c'))$, which is satisfied as long as V features decreasing elasticity.

□

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

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

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

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

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

□

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

C Simulating the model – for online publication only

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

C.1 Algorithm

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

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

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

$$A(1,h)H(\bar{\omega}(2),M(\bar{\omega}(2),1))T(\bar{\tau}(2)) = A(2,h)H(\bar{\omega}(2),M(\bar{\omega}(2),2))T(0)$$
$$\int_{\bar{\omega}(2)}^{\overline{\omega}} f(\omega)d\omega = S(\bar{\tau}(2))$$

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

- Determine which sectors prevail in City 1 and City 2 by comparing ω and $\omega(m,c)$ and $\omega(h,c)$, in $c \in \{1,2\}$.
- Solve in $x \geq 0$ the following equation $L \times f(\omega) = S(x+\tau_1) - S(\tau_1) + S(\tau_2 + g(x)) - S(\tau_2)$ where $g(x)$ is defined as:

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

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

C.2 Connecting data to the model

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

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

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

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

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

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

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

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

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

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

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

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

Let us now turn to function H that maps skills and sectors into productivity. Assuming a distribution $f(.)$ for ω and denoting by g the distribution of v, we can then infer:

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

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

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

C.3 Upper and lower tier middle paid job loss

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

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

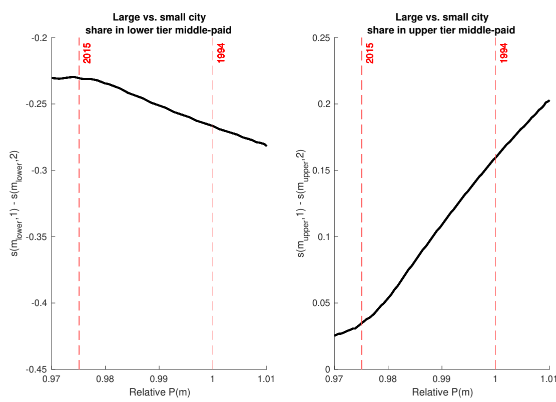


Figure C.1: Evolution across cities of upper-tier and lower-tier middle-paid jobs

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

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

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

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

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

C.4 Results – uniform distribution

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

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

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

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

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

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

C.5 Normal distribution – robustness – shock

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Table C.8: Model-based productivity estimates

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

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

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

Table C.10: Model-based productivity estimates

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

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

D Additional theoretical results – for online publication only

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

D.1 Middle-paid job loss and initial exposure

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

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

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

The evolution of this share depends on two factors. On the one hand, this share increases in each city due to the reallocation of middle-paid workers that already were in city c to the high-paid sector. The strength of this first channel depends on the density of middle-paid jobs close to the threshold, $f(\omega(h,c),c)/L(c)$, and the variation of the threshold due to the relative price change. This latter variation is $d\omega(h,c) = 1/\Theta(\omega(h,c),c,h)dp/p$ and results from the indifference condition (Equation (6)) between the high- and the middle-paid sectors. On the other hand, the relative price decline makes city 1 more attractive for households with productivity between $\omega(h,1)$ and $\omega(h,2)$: these households are initially indifferent between the high-paid sector in City 1 and the middle-paid in City 2. For this range of skills, a decline in the relative price of the good produced by the middle-paid sector then reduces the attractiveness of City 2 compared with City 1, thus leading to a reallocation of workers to City 1 – the term of across-city reallocation is then positive. In the end, when the comparative advantage of the large city in the high-paid sector is sufficiently large, the share of high-paid workers increases only in the large city, proving the result.

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

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

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

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

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

This form of productivity is the one we use in the case of a uniform distribution of skills in Appendix C.4. It helps to capture the increasing convexity of individual fixed effects that we obtain in the data.

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

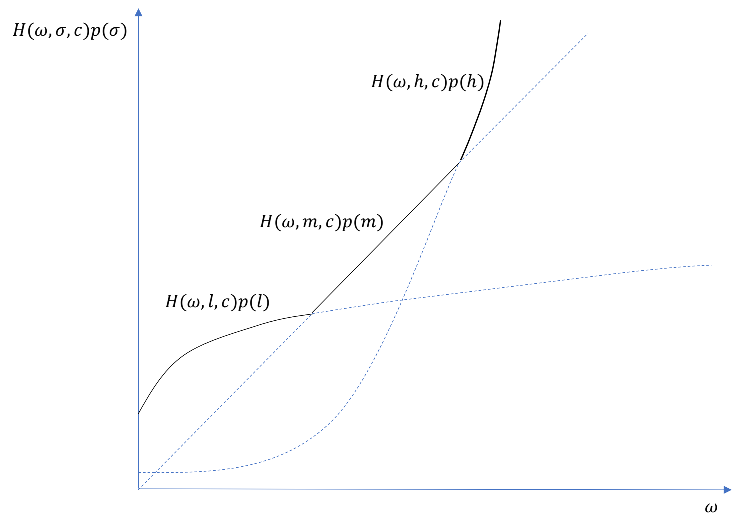


Figure D.1: Value marginal products in the different sectors.

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

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

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

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

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

Proof. See Appendix D.4.2. □

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

The decrease in the price of capital/offshored goods corresponds to a negative demand shock for middle-paid jobs, but its effects across location depends on local comparative advantages. Under Assumption D.1 such a decline leads, on the one hand, to a stronger decrease in the threshold $\omega(h,1)$ than in $\omega(h,2)$ ($\Theta(\omega(h,c),c,h)$ is smaller in the large city): households have a stronger incentive to shift to the high-paid sector in the large city than in the small city due to the comparative advantage of the large city in the high-paid sector. On the other hand, the threshold $\omega(m,1)$ increases by less than $\omega(m,2)$: the incentive to shift to the low-skill sector increases by more in the small city. We illustrate this point in Figure D.2 in the special case where $A(m,1) = A(m,2)$. In the case where the comparative advantage of the large city in the high-paid sector is large enough, i.e., $A(h,1)/A(h,2)$ is sufficiently large relative to $A(m,1)/A(m,2)$ and, by transitivity, to $A(l,1)/A(l,2)$, the first effect dominates. Not only are more high-paid jobs created in the large city, but also more middle-paid jobs are destroyed there. There is a rise in low-paid jobs in the large city, but of a lesser magnitude than of high-paid jobs. Figure D.3 summarizes all these results.[45]

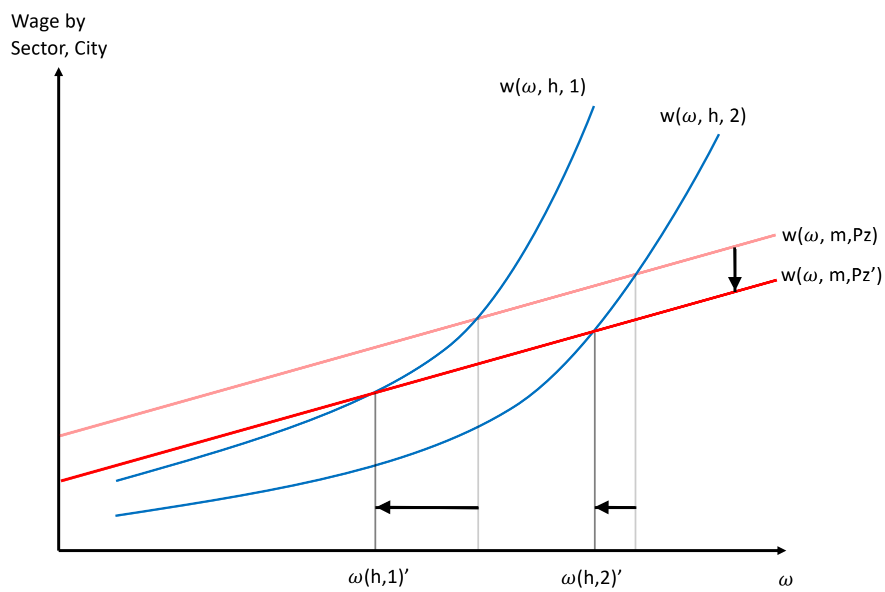


Figure D.2: Effect of a decline of the price of the capital/offshoring good on sector decisions – special case of equal productivities of cities in the m sector

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

Result (ii) in Proposition D.1 implies that the increase in low-paid jobs is smaller in larger cities. The logic is as discussed earlier. The comparative advantage of the large relative to small city in the

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

middle- relative to low-paid occupations implies that the cutoff for middle- relative to low-paid occupations is lower in the large city. However at this margin, this implies that the large city adjusts less elastically. The contrast between the cities will be stronger when either the concavity of productivity with respect to skill in the low-paid sector is strong or the comparative advantage of the large city in the middle- to low-paid sector is strong.

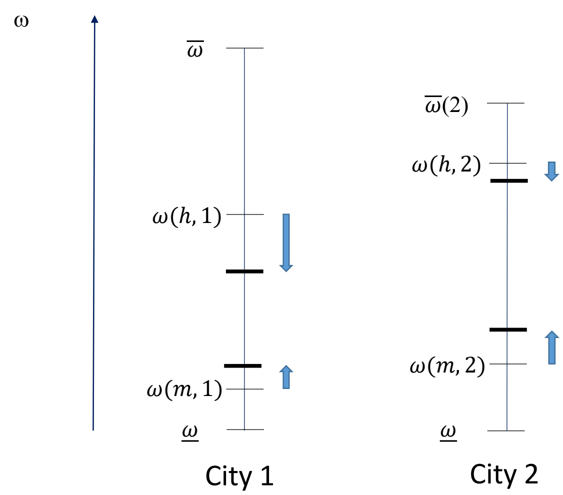


Figure D.3: The effects of a decline in the price of capital/offshored goods.

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

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

Proposition D.2 (Middle-paid job loss and initial exposure)**.** *When* $\frac{A(h,1)}{A(h,2)}$ *is sufficiently large relative to* $\frac{A(m,1)}{A(m,2)}$*, the share of middle skill sector jobs is smaller in the larger city.*

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

Proof. See Appendix D.4.3. □

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

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

We now look more closely at heterogeneity in adjustments within the middle-paid sector in each city. As in the low- and high-paid sectors, workers occupying middle-paid jobs are heterogeneous with respect to their skills – they have different values for ω. It is then possible to further analyze how labor market polarization affects middle-paid jobs across cities depending on workers' skills. Here, we show that, in large cities, it is the most skilled (i.e. with the highest ω or, equivalently, with the highest wage) middle-paid jobs that are destroyed and replaced by high-paid jobs. In small cities, it is mainly the least-skilled middle-paid jobs that are destroyed and replaced by low-paid jobs.[46]

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

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

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

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

In terms of the model's notation, what stands behind this proposition is the relative behavior of the thresholds $\omega(h, c)$ and $\omega(m, c)$ across cities. These thresholds correspond to the indifference condition between, respectively, the high- and the middle-paid sectors and the middle- and the low-paid sectors. Under Assumption D.1, the upper threshold $\omega(h, c)$ decreases by more in the large city (City 1) and the lower threshold $\omega(m, c)$ increases by more in the small city (City 2). As a result, more higher wage middle-paid jobs are destroyed in the large city following a labor market polarization shock while more lower wage middle-paid jobs are destroyed in the small city.

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

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

[46] Our partition of occupations into low-, middle-, and high-paid sectors in Table 1 emphasized vulnerabilities to our posited shocks. Heterogeneity within groups notwithstanding, the middle-paid sector really is different. All eight of the most offshorable occupations and six of eight of the most routinizable occupations are in the middle-paid sector. The lowest-paid of these, CS 67 unskilled industrial workers is the most offshorable occupation and the second most routinizable one. At the other end, the highest-paid middle-wage occupation, CS 48 supervisors and foremen is third both in offshorability and routinizability. These underscore the value in our framework of examining margins with both low- and high-paid jobs

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

Proof. This proposition first results from Proposition D.1, which shows that the large city experiences a larger increase in high-paid jobs. We obtain that the share of high-paid workers is larger in the large city by combining Corollary 2 in Appendix B.3 and Proposition D.2, which show that the share of low-paid and middle-paid are both smaller in the large city. □

D.4 Proofs

D.4.1 Sectoral decisions and factor prices.

As this can be observed from equations (5) and (6), the two thresholds are functions of intermediate good prices $p(l)$, $p(m)$ and $p(h)$. The following lemma clarifies how these thresholds move as a function of the relative prices $p(m)/p(h)$ and $p(l)/p(m)$ when Assumption D.1 holds.

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

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

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

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

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

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

Let us compute the different terms separately:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

which is satisfied. □

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

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

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

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

Figure D.3 summarizes these findings.

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

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

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

D.4.2 Proof of Proposition D.1

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

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

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

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

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

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

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

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

D.4.3 Proof of Proposition D.2

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

F.1 More detailed data description

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

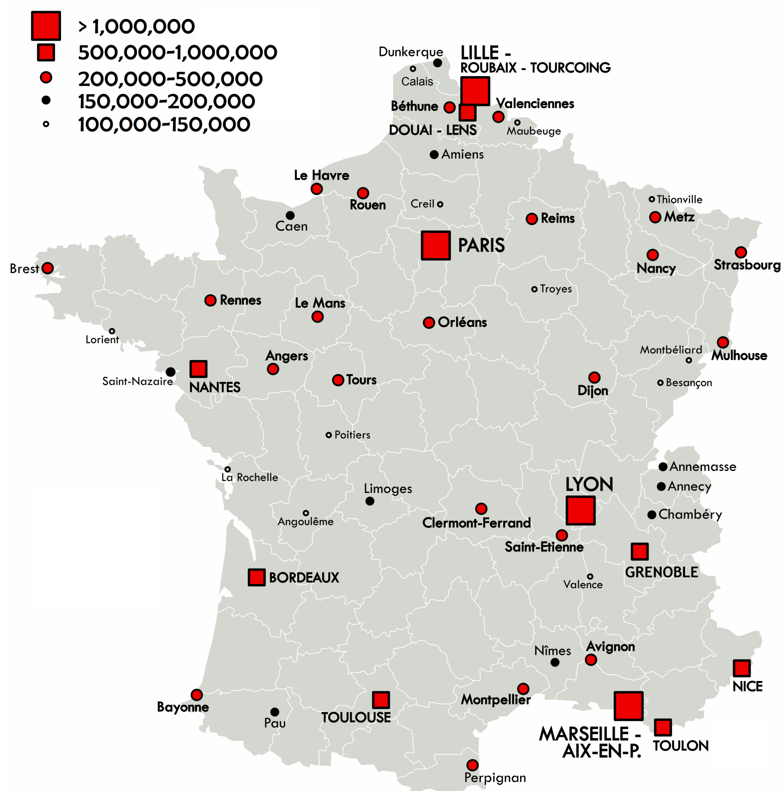


Table F.1: City size categories, number of cities, population and the share of hours worked in 2015.

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

Notes on dataset construction. There are two alternative sources of data to DADS-Postes on hours and wages by occupation in the studied period. The DADS-panel data set is constructed by the

INSEE on the basis of DADS-Postes, retaining a fraction (1/25 of total pre-2001) of individuals in the main data set. It shares the advantages and limitation of DADS-Postes; therefore for our purposes of analyzing city labor markets at two distant dates it offers no advantages at the cost of lower precision.

Another data set is the French Labor Survey is available since 1982. For early years it has approximately 60,000 observations per year and has only data at the department level. This allows to document some general facts about labor market polarization starting from 1982, but disallows constructing precise changes in labor market evolutions at the city level for 2-digit CS occupations. The DADS-data is exhaustive and gives inter alia more geographical details.

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

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

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

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

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

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

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

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

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

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

Trying to use the remaining 1-digit CS codes 4, 5, and 6 to define a boundary between middle- and low-paid groups immediately runs into problems. The 2-digit CS 5 and CS 6 occupations have no clear ordering by initial mean wage, so cannot be sensibly separated. This is also by construction according to the PCS classification as CS 5 ("employees", typically service workers) and 6 ("workers", typically blue-collar workers working in industry and various artisans) need not differ much in skills.[47] If they are combined as the low-paid sector, then this would include two of the sectors with the largest absolute and relative job declines, CS 62 and 67, in the low-paid sector. Again, this is against the spirit of the labor market polarization approach.

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

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

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

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

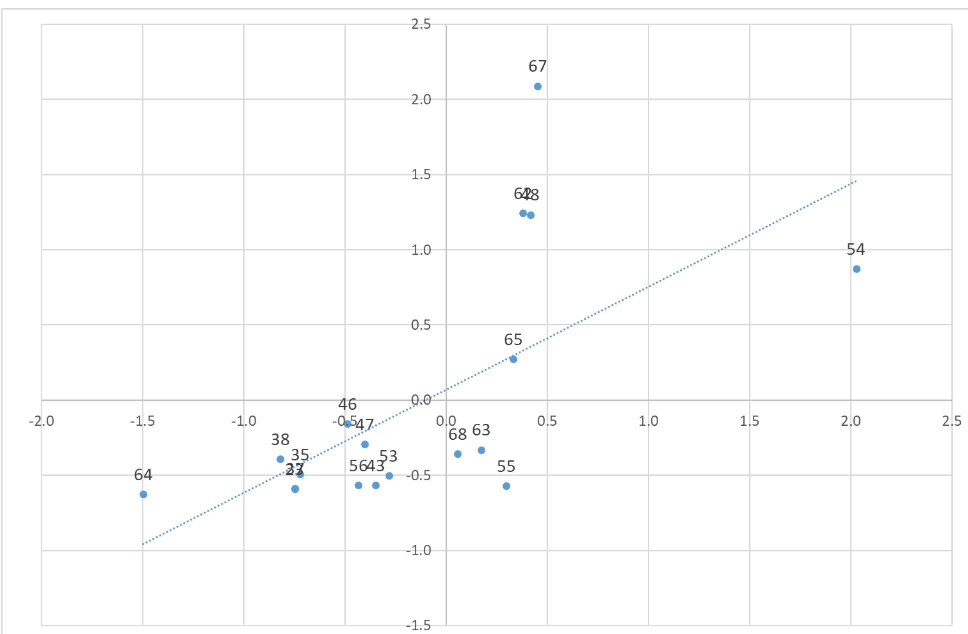
Notes: In-sample values. Employment share for mainland France. Average city wages in constant 2015 euros. Categories in bold are those with employment shares above 2.5% in 1994 in sample. Translation from French of category names other than PCS 23, 35, 43 and 53 taken from Table 2 of Harrigan et al.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

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

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

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

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



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

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

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

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

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

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

F.2 Evidence on SBTC and LMP

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

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

where X_{it} are time-varying worker characteristics — age, tenure at current firm (both up to their quartic terms) or an indicator of obtaining a new job; $\gamma_{p_o A_o}$ are occupation (high-, middle- and low- paid) fixed effects composed of sectoral prices p_o and productivities A_o that we cannot identify separately; δ_t are year-fixed effects; and v_i are worker-fixed effects.[48] We use DADS-panel data for 1993-1995 to exploit the panel data dimension around 1994 and (in a separate regression) 2014-2016 around 2015. The low-paid sector is treated as the base sector. Sectoral price × productivity ratios inferred from occupation fixed effects are exhibited in Table F.6.[49]

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

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

Table F.6 shows that sectoral wages polarize. Taking the sector-component of the wage in the low-paid sector as a base, middle-paid wages decline by 0.31 log points, while the high-paid sector component rises in an equal magnitude (a symmetry feature we employ in Section 4.4).

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

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

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

[48]Without further data, we cannot separate at the sectoral level the changes in prices p_o from A_o.

[49]The differences between parameter values (within year/across time) are statistically significant.

However, automation and offshoring shocks affecting middle-paid jobs and inducing labor market polarization would reveal themselves through the following relative sectoral price evolution: $\left(\frac{p'_m}{p'_h}\right) < \left(\frac{p_m}{p_h}\right)$ and $\left(\frac{p'_m}{p'_l}\right) < \left(\frac{p_m}{p_l}\right)$. With such changes in relative sectoral prices alone we are able to rationalize the observed patterns absent of sectoral productivity changes[50].

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

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

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

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

F.2.1 Evolution of wages 1994-2015

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

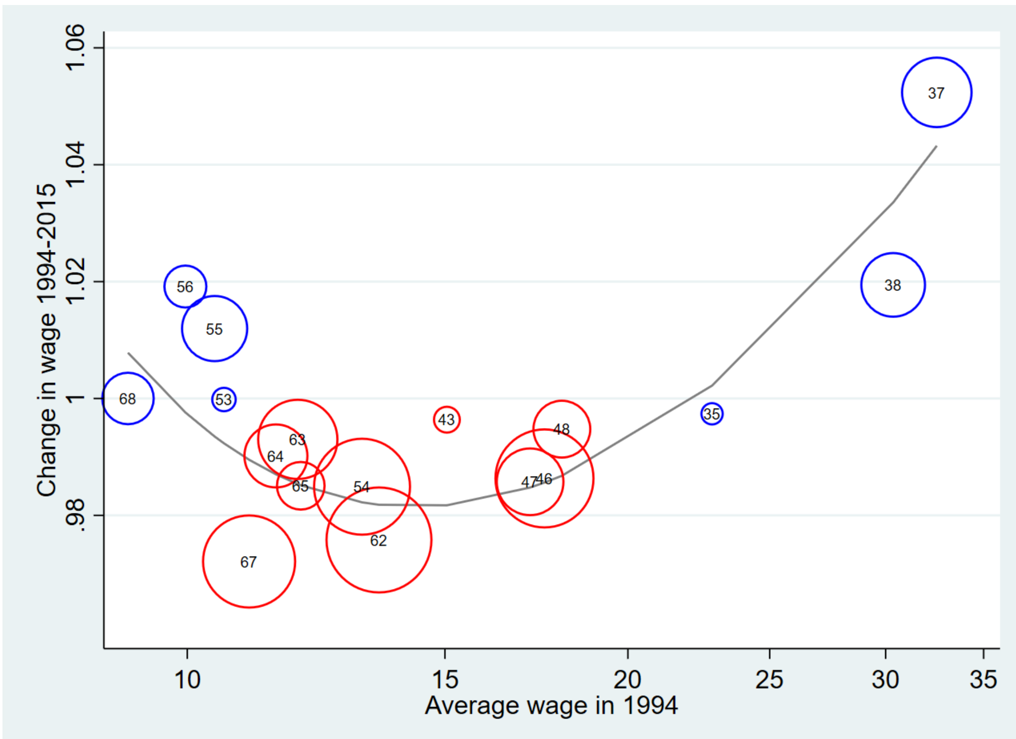
We estimate equation (30) using all 18 2-digit CS job categories instead of 3 sectors.

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

[50] We can further explore the extent of the labor market polarization implied by data constrained by different assumptions on the evolution of technology. Suppose that $\left(\frac{p'_hA'_h}{p_hA_h}/\frac{p'_lA'_l}{p_lA_l}\right) = \left(\frac{A'_h}{A_h}/\frac{A'_l}{A_l}\right)$ gives the extent of SBTC (meaning the relative p_h/p_l remain constant). Whenever $\left(\frac{A'_m}{A_m}/\frac{A'_L}{A_l}\right) \geq 0$ this implies given data $\ln\left(\frac{p'_m}{p'_h}\right) - \ln\left(\frac{p_m}{p_h}\right) = \ln\left(\frac{p'_m}{p'_l}\right) - \ln\left(\frac{p_m}{p_l}\right) < 0$ or labor market polarization. In other words, with our estimates, this implies that if there is SBTC in data, it *exacerbates* the needed relative decline in middle-paid sector prices to match the obtained estimates. In particular, when $\frac{A'_h}{A_h} \geq \frac{A'_m}{A_m}$ that the fall in $\frac{p_m}{p_h}$ and $\frac{p_m}{p_l}$ is restricted to the range $[-0.031, -0.063]$ required by our model simulation based on the Normal skill distribution (see Section 4.4) to obtain skewed polarization.

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

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



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

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

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

F.3 Model validation with data

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

F.3.1 Wages and productivity across cities and occupations

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

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

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

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

Whatever the grouping of large vs. small cities that is used, larger cities exhibit larger absolute productivities in all sectors and comparative advantage in high- relative to middle-paid sectors, as well as middle- relative to low-paid sectors. That is, large cities have a comparative advantage in more skilled sectors.[51] This data further justifies Assumption 3 made in Section 4.

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

Table F.8: Relative productivity across sectors and cities

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

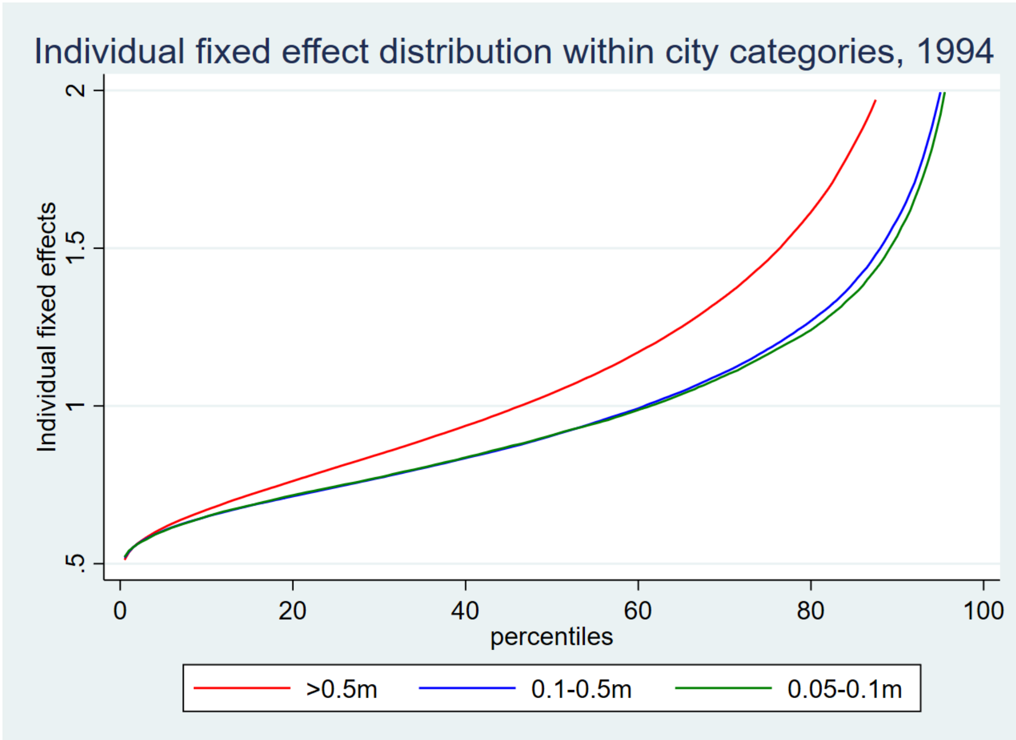
This Table presents the relative productivities across different groups of largest vs. smallest cities inferred from the terms $\gamma_{p_o A_{co}}$ in within regressions $ln(wage_{it}) = \alpha + \beta X_{it} + \gamma_{p_o A_{co}} + \delta_t + v_i + e_{it}$ on DADS-Panel data for 1993-1995 using different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

F.3.2 Individual fixed effects from wage data

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

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

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



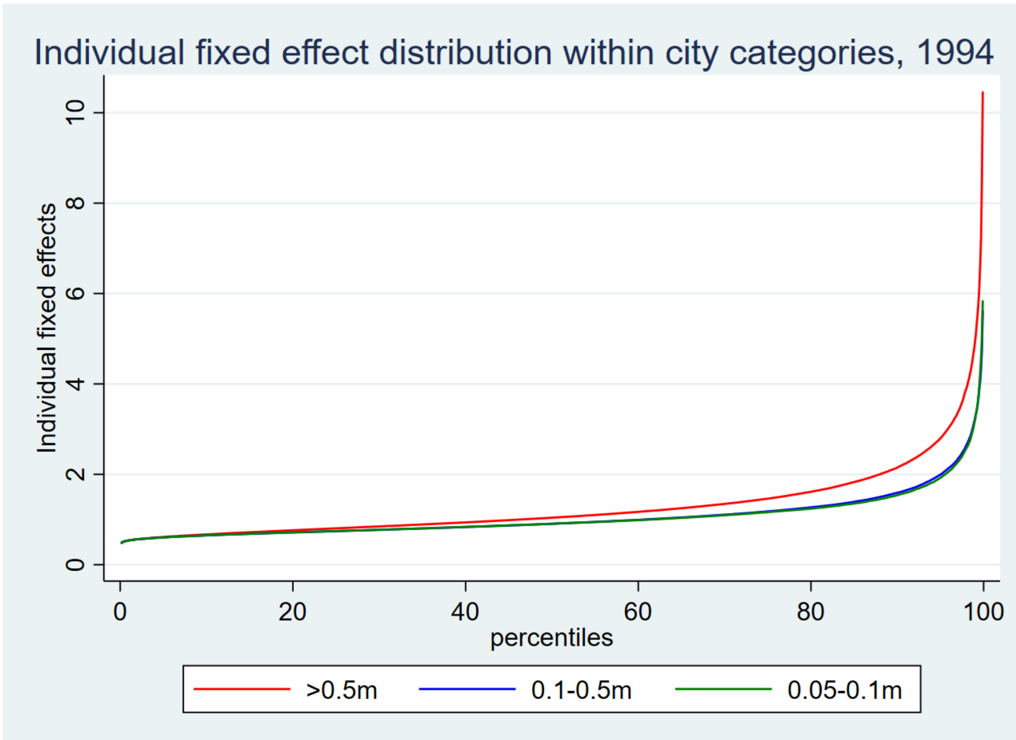
The figure shows the distribution of recovered individual fixed effects from a within regression (32) on a yearly worker DADS-panel data for men in 1993-1995 for cities above 0.5m, between 0.1-0.5m and 0.05-0.1m inhabitants:

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

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

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

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



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

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

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

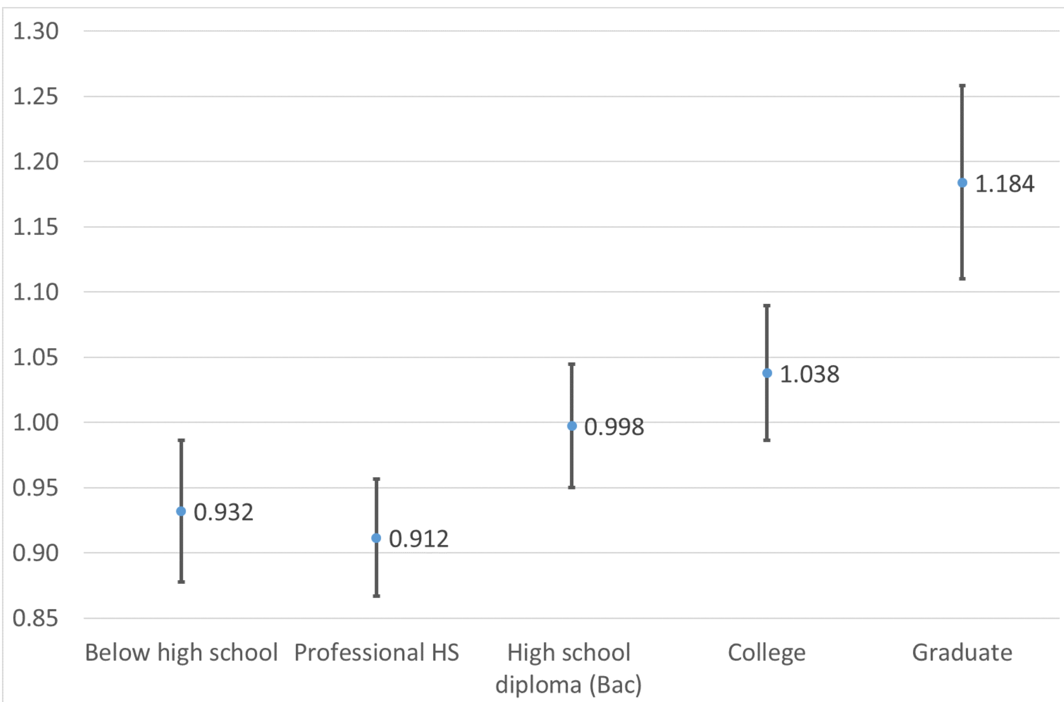
F.3.3 Log-supermodularity in data

Proposition B.2 provides conditions, as in Davis and Dingel (2020), under which the distribution of skills $f(\omega, c)$ is log-supermodular in city size. To obtain a measure of skills we turn to the 1999 Census, which has the best data on both diplomas and commune of residence among the Censuses spanning our time period.[52] We measure skills by the highest diploma received by individuals. The results are illustrated in Figure F.6 with further results in Tables F.9-F.14). As expected, Figure F.6 shows there is an ordering of the population elasticicities of skills, with the two lowest skill groups having an elasticity statistically significantly below 1; two middle skill groups (with high school diplomas and some college) having an elasticity insignificantly different from 1; and a high skill category of workers with a graduate diploma that has a significant population elasticity of 1.18. These observations carry

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

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

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



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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

F.3.4 Occupation shares across cities

We now turn to occupation share patterns that can be obtained from the detailed DADS data on hours worked. As indicated in the theory section, people with the same skills may perform different occupations depending on the city where they choose to reside, and this has implications on the observable employment shares across cities.[53]

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

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

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

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

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

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

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

The share of the middle-paid jobs monotonically declines with city size – accounted for by Proposition D.2 in both 1994 and 2015. The share of the lowest-paid occupations is highest in the smallest cities in either of the years, although the cross-city variation is modest.[54]

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

F.3.5 Job and city transitions

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

We use 1993-1995 DADS-Panel data on working age men using 6 city groups as in Table 2.[55] We first show in Table F.21 cross-tabulations of categorical variables carrying information on occupation and city transitions. We code the occupation change variable as -1, 0, +1 for a change to a worse-paid job category, no change (of a job or category) or an upgrade to a better-paid job respectively within the 3 year window around 1994. Similarly, we code the variable capturing city change as -1 or +1 for moving into smaller/larger cities and 0 for no change in city size or no move). In this data we keep workers that did not change their place of work nor a job in the studied time period.

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

Table F.21: Job type and city size transitions

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

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

To further explore these relations, we run an ordered logit regression of occupation change on city change, age up to a quartic term (to account for the likelihood of moving connected with age) and recovered individual fixed effects (to account for possible ongoing sorting) from a regression as in equation (32) using 6 city groups[56]:

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

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

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

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

Table F.22: Ordered Logit Results

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

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

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

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

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

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

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

F.4 Labor Market Polarization: Additional Results

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

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

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

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

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

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

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

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

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

In Table F.28 we assign PCS 53 "Security workers" as middle-paid jobs (belonging to those paid below the median) instead of low-paid jobs while in Table F.29 we assign PCS 67 "Low skill industrial workers" as low-paid jobs instead of middle-paid ones. The latter assignment is counter to the spirit of our base classification as they are the second-most routine occupation and the most offshorable in our data.

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

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

Table F.25: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample without Paris.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.095	-0.156	0.062	-0.111	-0.045	-0.110	-0.046
mean change, cities 0.05-0.1m	0.037	-0.116	0.080	-0.111	-0.006	-0.073	-0.044
difference	0.058***	-0.04***	-0.018***	0.000	-0.040***	-0.037***	-0.003
Growth in percent							
mean growth, cities >0.5m	70.0	-21.3	49.0	-29.9	-11.8	-29.1	-13.0
mean growth, cities 0.05-0.1m	45.7	-14.9	62.2	-25.2	-0.6	-19.9	-10.2
difference in growth	24.3***	-6.3***	-13.2**	-4.7**	-11.3***	-9.2***	-2.9

Notes: 1990 population weighted, robust standard errors. N=72 (10 cities > 0.5m as of 2015, without Paris). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.26: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample using "aires urbaines" as the definition of the city.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.104	-0.168	0.064	-0.109	-0.059	-0.120	-0.048
mean change, cities 0.05-0.1m	0.030	-0.109	0.080	-0.110	0.001	-0.067	-0.042
difference	0.074***	-0.059***	-0.015***	0.002	-0.060***	-0.052***	-0.007
Growth in percent							
mean growth, cities >0.5m	62.0	-24.1	54.3	-31.7	-16.2	-33.0	-14.4
mean growth, cities 0.05-0.1m	41.1	-14.0	61.5	-24.6	1.4	-19.2	-9.5
difference in growth	20.8***	-10.1***	-7.2	-7.1***	-17.7***	-13.8***	-4.9***

Notes: 1990 population weighted, robust standard errors. N=76 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.27: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.02-0.05m.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.116	-0.181	0.065	-0.108	-0.073	-0.130	-0.051
mean change, cities 0.05-0.1m	0.033	-0.116	0.083	-0.113	-0.003	-0.073	-0.043
difference	0.082***	-0.065***	-0.017***	0.005	-0.07***	-0.057***	-0.008
Growth in percent							
mean growth, cities >0.5m	63.0	-26.5	54.4	-33.1	-20.1	-35.9	-16.0
mean growth, cities 0.05-0.1m	44.4	-14.8	68.0	-24.3	0.5	-19.8	-9.6
difference in growth	18.6***	-11.8***	-13.7***	-8.8***	-20.6***	-16.1***	-6.3***

Notes: 1990 population weighted, robust standard errors. N=144 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities.

Table F.28: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample with PCS 53 "Security workers" counted towards middle-paid jobs (below the median).

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.116	-0.169	0.053	-0.108	-0.061	-0.130	-0.039
mean change, cities 0.05-0.1m	0.037	-0.113	0.076	-0.111	-0.002	-0.073	-0.040
difference	0.079***	-0.056***	-0.023***	0.003	-0.059***	-0.057***	0.001
Growth in percent							
mean growth, cities >0.5m	63.0	-24.4	47.7	-33.1	-16.3	-36.0	-11.8
mean growth, cities 0.05-0.1m	45.7	-14.3	62.3	-25.2	0.4	-19.9	-9.1
difference in growth	17.2***	-10.1***	-14.7***	-7.9***	-16.7***	-16.0***	-2.7*

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

Table F.29: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample with PCS 67 "Low skill industrial workers" counted towards low-paid jobs.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.116	-0.153	0.037	-0.108	-0.073	-0.130	-0.023
mean change, cities 0.05-0.1m	0.037	-0.059	0.023	-0.111	-0.006	-0.073	0.013
difference	0.079***	-0.094***	0.015	0.003	-0.068***	-0.057***	-0.036***
Growth in percent							
mean growth, cities >0.5m	63.0	-24.5	21.2	-33.1	-20.1	-36.0	-8.9
mean growth, cities 0.05-0.1m	45.7	-8.9	10.2	-25.2	-0.6	-19.9	5.3
difference in growth	17.2***	-15.7***	11.0*	-7.9***	-19.5***	-16.0***	-14.2***

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

Table F.30: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample keeping hours worked <120 hours / year (no "filtering" of observations by hours).

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.115	-0.181	0.066	-0.108	-0.073	-0.130	-0.051
mean change, cities 0.05-0.1m	0.036	-0.117	0.080	-0.111	-0.006	-0.073	-0.044
difference	0.079***	-0.064***	-0.014***	0.003	-0.067***	-0.057***	-0.008
Growth in percent							
mean growth, cities >0.5m	62.7	-26.6	54.7	-33.2	-20.0	-36.0	-15.9
mean growth, cities 0.05-0.1m	45.4	-15.0	62.4	-25.3	-0.6	-20.0	-10.1
difference in growth	17.3***	-11.6***	-7.6	-7.9***	-19.5***	-16.0***	-5.8***

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

Table F.31: Comparison of means of changes in the shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample without weighting the results by city population in 1990.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.096	-0.157	0.062	-0.111	-0.047	-0.113	-0.045
mean change, cities 0.05-0.1m	0.036	-0.116	0.079	-0.109	-0.006	-0.072	-0.044
difference	0.059***	-0.042***	-0.018***	-0.001	-0.040***	-0.041***	-0.001
Growth in percent							
mean growth, cities >0.5m	67.5	-21.8	48.9	-30.4	-12.2	-30.2	-12.9
mean growth, cities 0.05-0.1m	45.9	-14.9	61.5	-24.9	-0.8	-20.0	-10.1
difference in growth	21.6***	-6.9***	-12.6**	-5.5***	-11.4***	-10.2***	-2.8

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

Table F.32: Comparison of means of changes in employment shares of different occupations, cities >0.5m vs. 0.05-0.1m. Sample without finance, insurance and real estate sectors.

Item	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above median	middle-paid below median
Changes							
mean change, cities >0.5m	0.104	-0.177	0.072	-0.109	-0.067	-0.128	-0.048
mean change, cities 0.05-0.1m	0.032	-0.116	0.084	-0.119	0.004	-0.066	-0.049
difference	0.073***	-0.061***	-0.012**	0.010	-0.071***	-0.062***	0.001
Growth in percent							
mean growth, cities >0.5m	59.6	-25.9	56.6	-33.5	-18.4	-35.6	-15.0
mean growth, cities 0.05-0.1m	41.9	-14.9	64.0	-27.1	2.1	-18.2	-11.4
difference in growth	17.7***	-11.1***	-7.4	-6.4***	-20.5***	-17.4***	-3.5**

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

Table F.33: Comparison of means of changes in employment shares of different occupations across young and old cohorts, cities >0.5m vs. 0.05-0.1m.

Percentage changes	high-paid	middle-paid	low-paid	MRO	OMP	middle-paid above the median	middle-paid below the median
workers in the 25-34 age cohort							
mean change, cities >0.5m	0.111	-0.181	0.070	-0.091	-0.090	-0.116	-0.064
mean change, cities 0.05-0.1m	0.025	-0.126	0.101	-0.104	-0.022	-0.062	-0.064
difference	0.086***	-0.055***	-0.031***	0.013*	-0.068***	-0.055***	-0.000
workers in the 55-64 age cohort							
mean change, cities >0.5m	0.087	-0.136	0.049	-0.105	-0.031	-0.100	-0.036
mean change, cities 0.05-0.1m	0.018	-0.064	0.046	-0.068	0.004	-0.061	-0.003
difference	0.070***	-0.072***	0.003	-0.037***	-0.035***	-0.039***	-0.033***

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

Table F.34: Additional statistics for the main sample for middle-paid jobs.

Item	top three middle-paid with highest wages	Bottom seven least-well-paid middle-paid	intermediate professions	employees and blue-collar workers	white-collar workers	blue-collar workers
Changes						
mean change, cities >0.5m	-0.089	-0.092	-0.083	-0.097	-0.086	-0.079
mean change, cities 0.05-0.1m	-0.037	-0.079	-0.027	-0.089	-0.036	-0.093
difference	-0.052***	-0.013**	-0.056***	-0.008	-0.050***	0.014
Growth in percent						
mean growth, cities >0.5m	-33.1	-21.8	-30.0	-23.6	-28.2	-28.3
mean growth, cities 0.05-0.1m	-18.4	-13.4	-12.4	-15.4	-15.1	-19.7
difference in growth	-14.7***	-8.5***	-17.5***	-8.2***	-13.0***	-8.6***

Notes: 1990 population weighted, robust standard errors. N=73 (11 cities > 0.5m as of 2015). ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels for the tests of equality of means between the groups of small and large cities. Top 3 middle-paid with highest wages: CS 46, 47, 48. Bottom 7 least-well-paid middle-paid: CS 43, 54, 62, 63, 64, 65, 67. Intermediate professions: CS 43, 46, 47, 48. Employees and blue-collar workers: CS 54, 62, 63, 64, 65, 67. White-collar workers: CS 46 & 54. Blue-collar workers: CS 62, 63, 64, 65, 67.

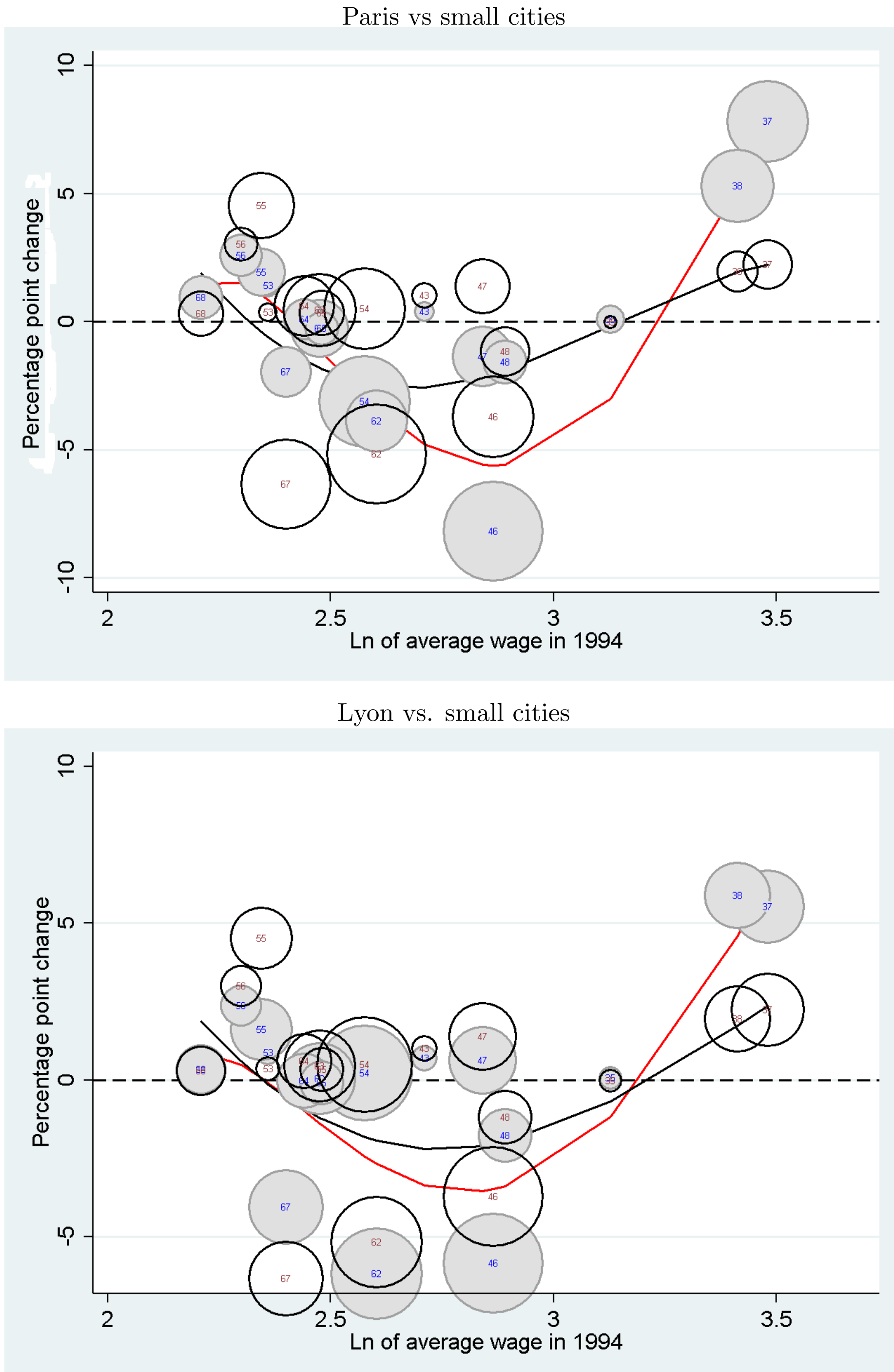
F.4.2 Other graphs and figures

Table F.35: Educational divergence across cities 1990-2013.

1990						
Agglo. size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
none or at most middle school	32.44	32.28	30.27	31.69	32.66	34.54
CAP, BEP	22.69	26.13	29.44	29.82	30.16	30.29
High school (general, vocational)	20.01	20.64	21.54	20.2	20.18	20.34
At least college	24.86	20.95	18.75	18.28	17	14.83
2013						
Agglo. size	Paris	> 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M
none or at most middle school	19.21	16.87	16.21	17.66	18.44	19.85
CAP, BEP	14.85	19.12	22.57	23.65	24.49	26.9
High school (general, vocational)	16.67	17.88	19.18	18.37	18.6	19.15
At least college	49.28	46.13	42.04	40.31	38.47	34.1

Note: Data from 1990 and 2013 Censuses for workers aged 25-64 years in non-agricultural sectors in cities >0.05m.

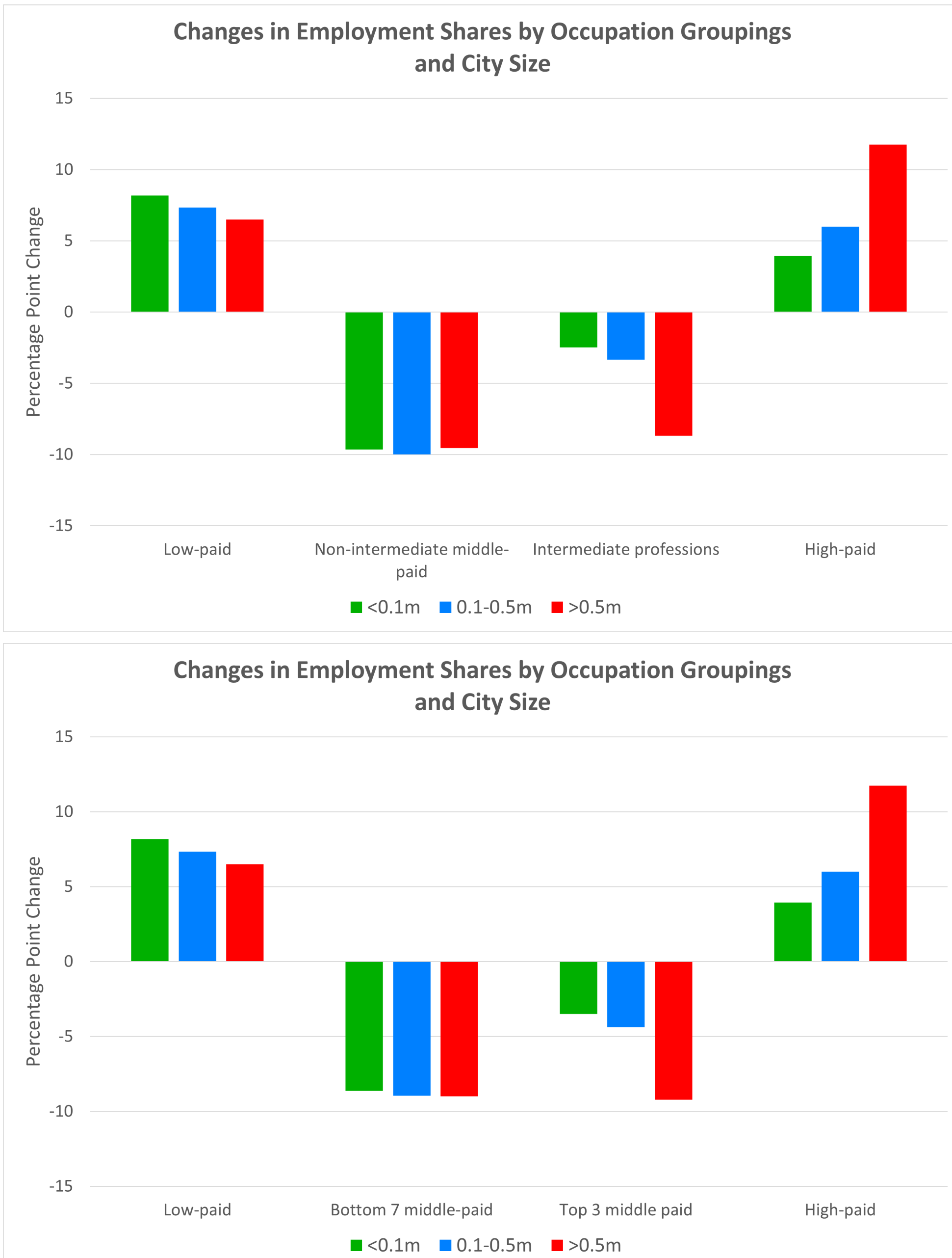
Figure F.7: Comparison of labor market polarization 1994-2015 in Paris/Lyon and cities between .05-.1m inhabitants.



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

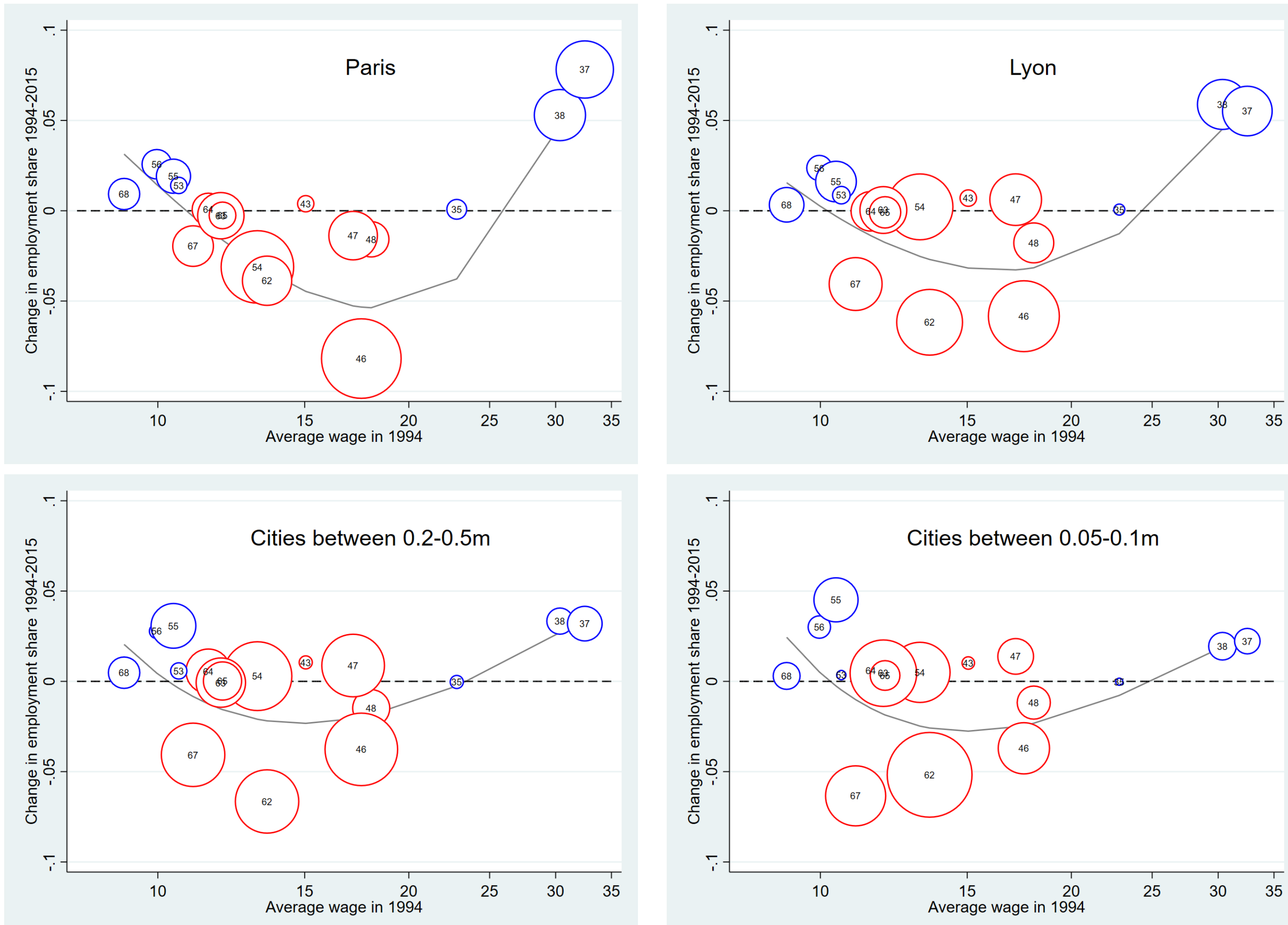
This figure documents a stronger decline in middle-paid jobs in larger cities and differential reallocation effects (higher creation of high-paid jobs in larger agglomerations) of labor market polarization across cities of different sizes using the contrast between Paris / Lyon and cities of 0.05-0.1m as a group.

Figure F.8: Labor market polarization across three different city size groups, 1994-2015: 4 employment groups



This figure shows percentage point changes in employment shares of high-, low- and different types of middle-paid jobs with hours worked summed by job types and 3 city sizes: large (above >0.5m inhabitants), medium-sized (0.1-0.5m) and small (0.05-0.1m) in the period 1994-2015. The various partitions of middle-paid jobs in each panel order these jobs by median wage. In the upper panel the middle-paid jobs are divided into intermediate professions (CS 43, 46, 47, 48) and other middle-paid jobs. In the lower panel these are divided into top-3 paying middle-paid occupations (CS 48, 46, 47) and the remainder of middle-paid jobs. Top-3 paid middle-paid jobs are a subset of intermediate professions that are also in the upper tier of middle-paid occupations.
All of the panels show that, for all these partitions of the middle skill jobs, the destruction of the lower-paid jobs was similar across all city sizes. At the same time, the panels show clearly that the destruction of the highest-paid middle-skill jobs rises monotonically with city size.

Figure F.9: Labor market polarization across cities 1994-2015.



Left-upper panel: Paris. Right-upper panel: Lyon. Left-lower panel: cities between 0.2-0.5m. Right-lower panel: cities between .05-.1m inhabitants.
Figures show the percentage point change in employment shares of the considered 2-digit CS categories plotted against their average wage in cities > 0.05m in 1994. Numbers pertain to 2-digit CS categories represented. Circle sizes correspond to the employment shares in 1994. The line shows a cubic relationship between the average wage and the percentage point change. The CS category “23” - CEOs excluded.
This figure documents that labor market polarization occurred in all types of cities. The magnification and differential reallocation effects of labor market polarization for larger cities in contrast to smallest cities are clearly visible: there was a weaker destruction of middle-paid jobs in smaller cities, and low-paid jobs were created there more strongly than high-paid occupations.

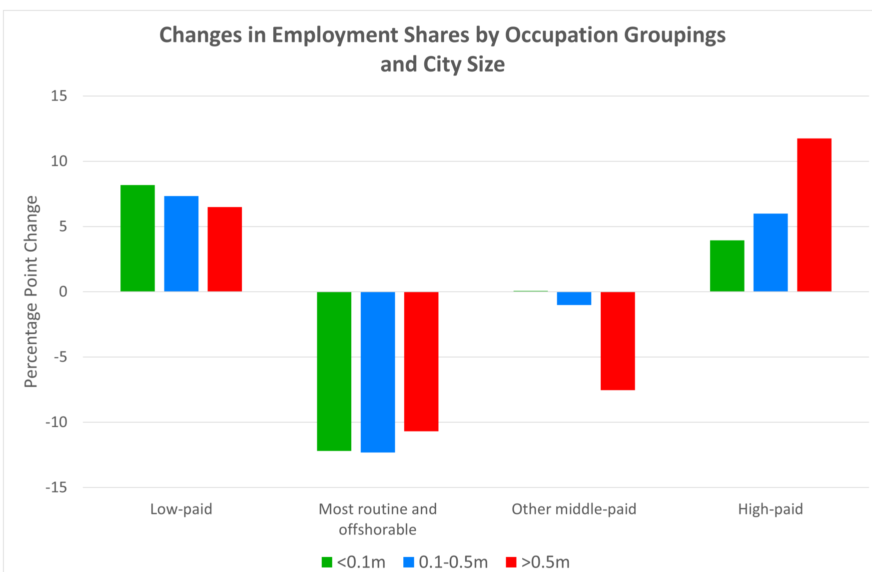
F.5 Initial exposure and the evolution of different job categories

In this section, we reconfirm prior work on polarization and show that locations' exposure to MRO jobs is a good predictor of loss of these jobs. However, exposure to MRO jobs, to the contrary, is *negatively* correlated with the loss of OMP jobs and in general *not* correlated with the destruction of middle-paid jobs taken as a whole.

We start by contrasting the MRO and OMP middle-paid jobs evolution in Figure F.10. Noting that MRO jobs are on average lower-paid than OMP jobs, we can order these on a wage axis. We have grouped cities into three sizes, with those above 0.5m at the top and those below 0.1m at the bottom. There indeed we see that large cities (that have a lower initial exposure to these occupations) have a smaller percentage point loss of MRO jobs, although the differences are modest. We also see that the strong contrast in experience comes in the OMP jobs. There are striking declines in OMP jobs in the largest, small losses in middle-sized, and essentially zero change in the smallest cities. That is, the contrasts in experience across city size in middle-paid job loss across cities of different sizes is precisely in the segment of jobs that in previous work was dropped from the discussion.

An important point is that OMP jobs may also be sensitive to trade or automation shocks. It may be indirectly, such as middle-managers (CS 46) who would lose MRO jobs to manage, or it occurs later in comparison to MRO occupations. Furthermore, what may matter in a large city labor market is the *relative* degree of routinizability and offshorability in comparison to high-paid occupations.

Figure F.10: Labor market polarization and the great urban divergence across three different city size groups, 1994-2015: MRO and OMP split of middle-paid jobs



This figure shows percentage point changes in employment shares of high-, low- and different types of middle-paid jobs with hours worked summed by job types and 3 city sizes: large (above >0.5m inhabitants), medium-sized (0.1-0.5m) and small (0.05-0.1m) in the period 1994-2015. The bars for high- and low-paid jobs are exactly as in Figure 3. The division of middle-paid occupations is between the most routine and offshorable (MRO) and other middle-paid occupations (OMP).
The figure shows that the destruction of the MRO jobs was similar across all city sizes. At the same time, the destruction of the OMP jobs rises monotonically with city size. Indeed, OMP occupations actually *grow* very modestly in the smallest cities.

F.5.1 Exposure and the loss of the most routinizable and offshorable (MRO) jobs

We investigate whether cities with a higher exposure to the most routinizable and offshorable jobs (our MRO group) see the largest decline in the share of these jobs. In the lowest panel of Table F.36 we report both the 1994 and 2015 employment shares in the four CS 2-digit MRO occupations in six city groups. The employment share in these four MRO jobs are declining in city size both in 1994 and in 2015 in line with the patterns for middle-paid occupations overall. This observation is confirmed using rank correlations in Table F.16.

In the lowest panel of Table F.36, the evolution of these shares over this period is relatively constant in percentage points: we see that the fall of shares in this category of middle-paid jobs is similar across metropolitan areas without any clear relationship with size – between 10.5 and 13.1 percentage points. Table 3 shows no statistically significant difference between large and small cities in the change in these MRO occupations. The same conclusion arises in rank-correlation tests (Table F.24).

Table F.36: Share of high-, middle- and low-paid occupations in hours worked per metropolitan area size in 1994 and 2015.

MRO							
Agglo.size	Paris	> .75m	.5-.75m	.2-.5m	.1-.2m	.05-.1m	All cities
1994	0.29	0.36	0.39	0.41	0.45	0.45	0.36
2015	0.19	0.25	0.27	0.29	0.31	0.32	0.25
change	-0.11	-0.11	-0.12	-0.12	-0.13	-0.12	-0.12
growth in %	-36	-32	-31	-29	-29	-27	-32

This Table shows the means of shares of hours in total employment of different occupational groups in 1994 and 2015 for all 117 cities in our sample allocated in 6 bins according to city size (with Paris being a separate category), showing the percentage point changes and growth rates between 1994-2015. One observation per bin of the hours totals.
The share of MRO jobs (CS 48, 54, 62 and 67) in total employment is decreasing with city size whether in 1994 or 2015. Percentage point destruction of these MRO jobs is similar across city sizes despite their lower initial share in employment for larger cities.

However, these statements do not control for actual exposure to these specific jobs at the city level. It is clear from Figure F.11 that large cities above 0.5m people have lower initial exposures to the MRO occupations.[57] Although there is considerable variation, Figure F.11 confirms the observation of Autor and Dorn (2013) that the initial exposure to the most routine (and, in our context, also offshorable) jobs is strongly negatively correlated with their change as technological shocks occur. The observations for large cities lie in the lower envelope of observations. Thus conditional on initial exposure, the changes in the employment shares in these cities are larger than in other cities. Regression analysis in Table F.39, top panel, confirms these points: initial exposure is negatively correlated with change in the MRO employment shares and the interaction of a dummy for large cities with exposure is robustly negatively different from zero. MRO jobs in large cities are destroyed at a higher rate than in small cities with the same initial exposure in reaction to the same automation or trade shocks.

In the end, consistent with Autor and Dorn (2013), we obtain that the initial exposure to the most

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

routinizable and offshorable jobs is a good predictor of the most routinizable and offshorable jobs loss themselves. However, such jobs, conditional on exposure, are decimated more in larger cities.

F.5.2 Exposure to MRO jobs and the broad loss of middle-paid jobs

We see that initial exposure to MRO jobs is strongly associated with subsequent loss of these jobs. But this raises the question of whether exposure to MRO jobs is also associated with the loss of other middle-paid (OMP) jobs, or indeed with middle-paid jobs taken as a whole.

We can look at the relation between initial exposure to MRO jobs and the subsequent change in OMP jobs in Figure F.11. There is a strikingly strong negative relation between exposure to MRO jobs and subsequent loss of OMP jobs, confirmed in the second panel of Table F.39. The big losses of OMP jobs are in large cities, which have initially small exposure to MRO jobs.

We also know, though, from Fact 2 that large cities experienced a larger decline in middle-paid jobs overall (that include the MRO category). This leads us to suspect that the exposure to MRO jobs by itself may not be a good predictor of the overall change in middle-paid jobs.

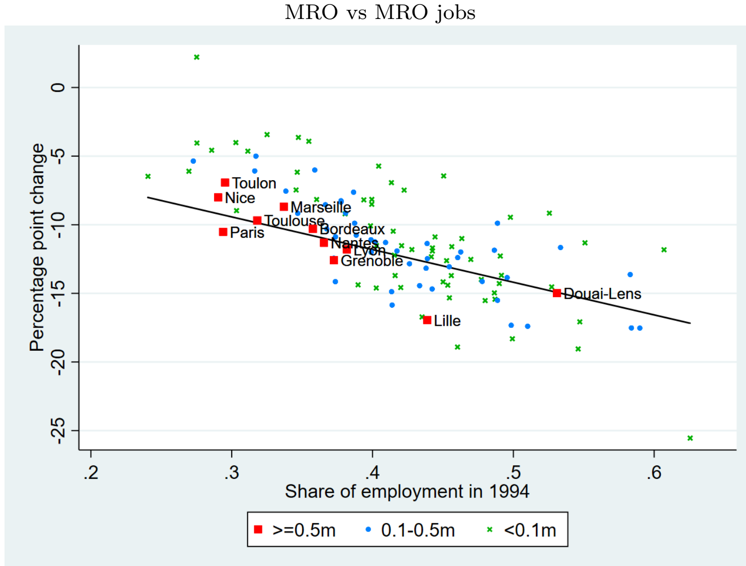
Indeed, the population-weighted regression of changes in employment for the entire middle-paid category on initial exposure to the four MRO occupations (in Table F.38, bottom panel) reveals a strong positive relationship (though the non-population weighted relationship is zero). There is clearly a larger destruction of middle-paid jobs in the largest cities conditioning on exposure, witnessed by the sign of the interaction of a dummy for large cities with MRO job exposure. For many small but highly-exposed cities, the drop in MRO jobs is larger than the decline in middle-paid jobs while the opposite is true for the largest cities. The initial exposure to the most routinizable and offshorable (MRO) jobs is not a good predictor of the evolution of the entire class of middle-paid jobs across cities. Thus, initial exposures to the MRO jobs are not a key driver of a broad measure of labor market polarization in local labor markets.

Even if automation and/or offshoring (Autor and Dorn, 2013; Goos et al., 2014) are driving labor market polarization, the extent to which these affect the broad category of middle-paid jobs does not depend only on a city's initial exposure to the most routinizable and offshorable jobs. These results are robust to considering a larger set of occupations for the most routinizable and offshorable jobs[58] and also a longer time period (See Section F.6 of this Appendix).

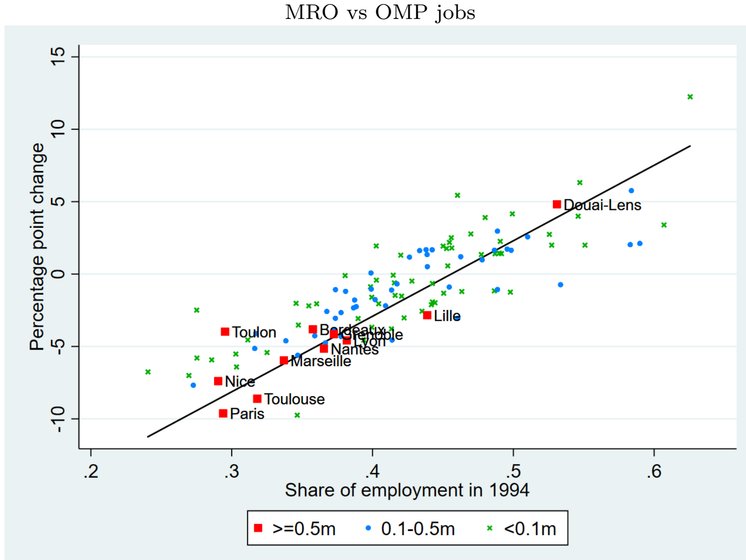
[58]In Table F.20 we show the patterns for the 6 most offshorable jobs encompassing not only the four MRO occupations, but also categories such as transport and logistics personnel (CS 65) and mid-level professionals (CS 46). The shares of such jobs in total employment are monotonically decreasing with city size whether in 1994 or 2015. Percentage point fall in the employment shares of these occupations between 1994 and 2015 is higher in larger cities, thus confirming our results. Moreover, it can be seen in Figure F.14 that a higher initial exposure to offshorable jobs leads to their greater decrease in the studied period. This time, however, large cities are relatively more exposed to offshorable jobs in comparison to MRO occupations only as – in particular – they have on average a higher share of the CS 46 category, mid-level professionals (cf. also Table F.20). In Table F.40 we demonstrate that the initial exposure to this wider set of occupations is also a good predictor of their employment share change. Again, conditional on exposure, offshorable jobs' employment shrinks by more in large cities in the studied period.

Figure F.11: Exposure to MRO jobs and change in the employment shares of MRO and other middle-paid OMP jobs, 1994-2015.

MRO vs MRO jobs



MRO vs OMP jobs



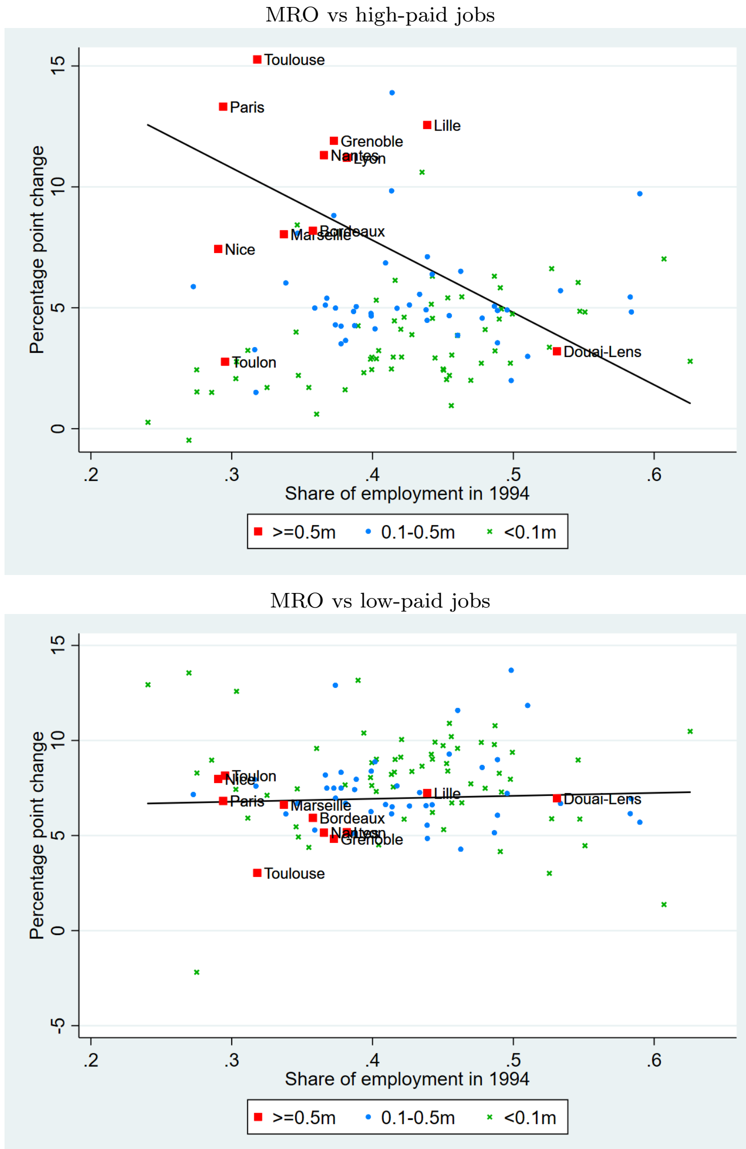
The figures show the percentage point change in employment shares of MRO jobs (CS 48, 54, 62 and 67) between 1994-2015 plotted against their share in employment (upper panel) or other middle-paid OMP jobs (lower panel) in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The lines show a linear, population-weighted (by 1990 population) fit of the relationship between employment changes and the initial exposure to MRO jobs. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.

The initial exposure of largest cities to the most routine and offshorable occupations (MRO) in 1994 is on average lowest in largest cities. The exceptions are Douai-Lens and Lille in the industrial North.

The relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of these jobs over the period 1994-2015 (upper panel) is negative as predicted by Autor and Dorn (2013) (cf. Table F.39 on the robustness of the slope of the fitted line). Conditional on the initial exposure, however, the decline in the MRO occupations is highest in largest cities. Moreover, the average decline in the MRO jobs is not significantly different (cf. Table 3, column 4) between the largest and smallest cities in the sample (which are on average more exposed to those occupations – see Table F.15).

The lower-panel figure depicts a strongly positive population-weighted relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of OMP jobs (cf. Table F.39 on robustness of the slope). In *all but one* cities with initial exposure to MRO jobs above 0.5 the share of OMP jobs in total employment increased over the period 1994-2015. The decline of the OMP jobs in percentage points is on average stronger in large than in small cities (cf. Table 3, column 5), significant at 1% level.

Figure F.12: Exposure to MRO jobs and change in the employment share of high-paid and low-paid jobs, 1994-2015.



These figures show the percentage point change in employment shares of high-paid jobs (upper panel) and low paid jobs (lower panel) between 1994-2015 plotted against the share of MRO jobs (CS 48, 54, 62 and 67) in employment in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The line shows a linear, population-weighted (by 1990 population) fit of the relationship between employment changes and the initial MRO exposure. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.

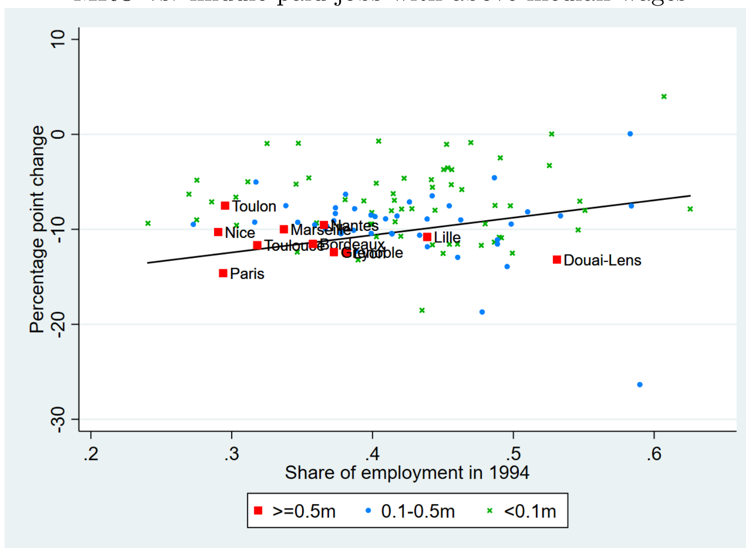
The upper panel figure documents a strongly negative population-weighted relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of high-paid jobs over the period 1994-2015 (cf. Table F.38 on the robustness of the slope of the fitted line). The creation of the high-paid occupations is on average much stronger in large than in small cities (cf. Table 3, first column) and this difference is significant at the 1% level.

The lower panel figure depicts no relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share of low-paid jobs over the period 1994-2015. However, the increase in the employment shares of the low-paid occupations is on average significantly higher (at the 1% level) in small cities (cf. Table 3, column 3).

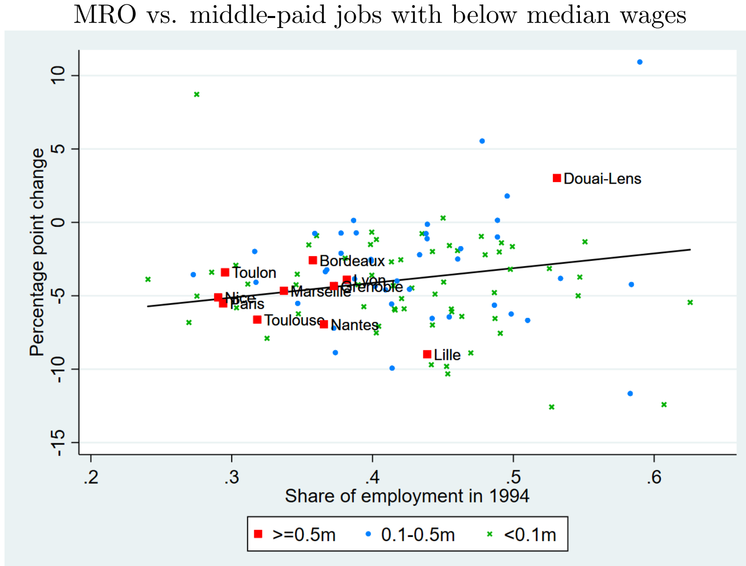
Such patterns are incompatible with the Autor and Dorn (2013) model that predicts that local labor markets with the highest local exposure to routine jobs would experience the strongest creation of high-paid and low-paid jobs.

Figure F.13: Exposure to MRO jobs categories and 1994-2015 changes in employment shares of middle-paid jobs above- and below-median in terms of wages in 1994.

MRO vs. middle-paid jobs with above median wages

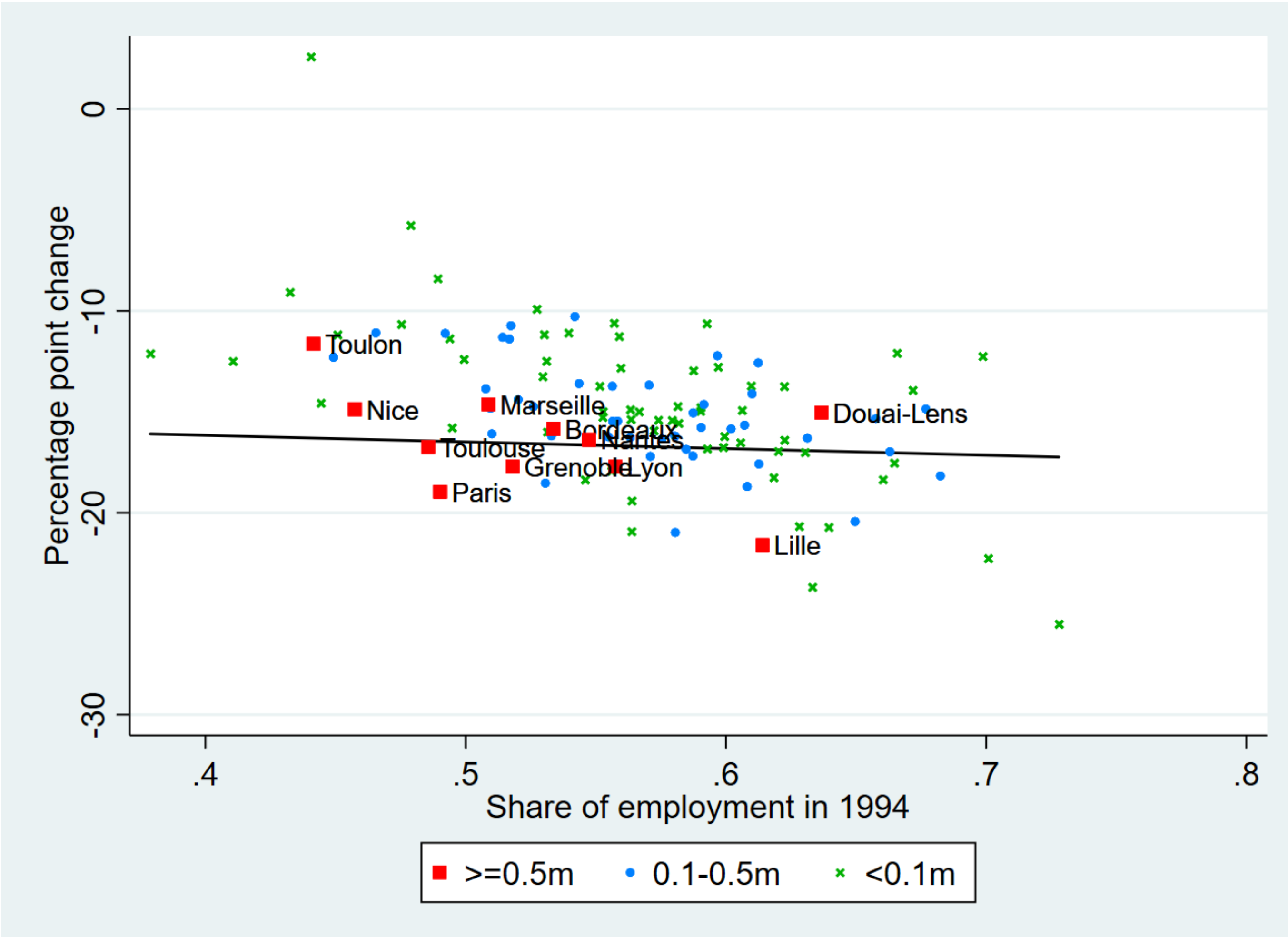


MRO vs. middle-paid jobs with below median wages



The figure shows the percentage point change in employment shares of middle-paid jobs with wages above (upper panel) and below (lower panel) the median in 1994 between 1994-2015 plotted against the share of MRO jobs (CS 48, 54, 62 and 67) in employment in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The lines show a linear, population-weighted (by 1990 population) fit of the relationship between employment changes and the initial MRO exposure. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.
There is a positive population-weighted relationship between the employment share of the MRO occupations at the city level in 1994 and the change in the employment share (over the period 1994-2015) of middle-paid jobs with the 1994 average wages both above and below the median (cf. Table F.39 on the robustness of the slope of the fitted line). The decline of the middle-paid occupations with average wages above the median is on average stronger in large than in small cities (cf. Table 3, next-to-last column) and this discrepancy is significant at 1% level. The decline of the middle-paid occupations with average wages below the median in percentage points is on average not statistically significantly different between large and small cities (cf. Table 3, last column).

Figure F.14: Exposure to 6 most offshorable jobs and their employment share change in cities, 1994-2015.



The figure shows the percentage point change in employment shares of 6 occupations with the highest OFF-GMS index (CS 46, 48, 54, 62, 65, 67) between 1994-2015 plotted against their share in employment in 1994 at the city level. Each red square, blue dot or green check symbolizes, respectively, a large (above >0.5m inhabitants), medium-sized (0.1-0.5m) or small (0.05-0.1m) city. The line shows a linear, population-weighted fit of the relationship between employment changes and the initial exposure to these most offshorable jobs. Names of cities with more than 0.5m inhabitants are shown. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015.

The initial exposure of largest cities to the 6 most offshorable occupations in 1994 is on average lowest in largest cities. The relationship between the employment share of the 6 most offshorable occupations at the city level in 1994 and the change in the employment share of these jobs over the period 1994-2015 is weakly negative (cf. Table F.40 on the robustness of the slope of the fitted line). The average decline in these 6 most offshorable jobs, however, is statistically significantly stronger (at the 1% level) in the largest than in smallest cities in the sample (which tend to be initially more exposed to them).

Table F.37: Changes in the employment shares of middle-paid jobs between 1994-2015 and exposure to middle-paid occupations in 1994.

Employment share change of middle-paid jobs								
employment share of middle-paid jobs in 1994	0.41***	0.03	0.27***	-0.03	0.28***	-0.04	0.27***	0.01
	(0.08)	(0.07)	(0.10)	(0.07)	(0.10)	(0.08)	(0.10)	(0.05)
middle × employment share of middle-paid in 1994			-0.01*	-0.02**	-0.01*	-0.02**	-0.01	-0.01*
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of middle-paid in 1994			-0.06***	-0.06***	-0.06***	-0.06***	-0.05***	-0.05***
			(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
constant	-0.45***	-0.14***	-0.33***	-0.09*	-0.33***	-0.09	-0.33***	-0.13***
	(0.07)	(0.05)	(0.08)	(0.05)	(0.08)	(0.06)	(0.08)	(0.04)
R^2	0.47	0.00	0.59	0.16	0.59	0.17	0.63	0.18
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in middle-paid share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of the middle-paid jobs in total hours worked over the period 1994-2015 on their initial employment share at the individual city level in 1994. The first two columns report regression coefficients respectively with and without population (as of 1990) weighting. In the regressions reported in columns 3 and 4 an additional slope for medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.
The relationship between the exposure to the most offshorable jobs at the city level and their share change over the period 1994-2015 is positive (but not significant in the non-population-weighted regressions). The interpretation of these results together with the intercept is that, on average, cities that were *less* initially exposed to middle-skill jobs experienced their stronger destruction. However, conditioning on exposure, the destruction of middle-paid jobs was stronger in larger cities for all specifications.
The evidence in this Table points that there is no simple relationship between a higher initial exposure to middle-paid jobs at the city level and their subsequent destruction as a result of automation or offshoring shocks.

Table F.38: Employment share changes 1994-2015 and exposure to 4 most routine/offshorable occupations (MRO) in 1994, Part I.

Employment share change of high-paid jobs								
employment share of MRO jobs in 1994	-0.30***	0.01	-0.19**	0.04	-0.20**	0.03	-0.19**	0.03
	(0.10)	(0.04)	(0.09)	(0.04)	(0.09)	(0.04)	(0.10)	(0.03)
middle × employment share of MRO jobs in 1994			0.04***	0.04***	0.04***	0.03***	0.04***	0.04***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of MRO jobs in 1994			0.16***	0.16***	0.16***	0.15***	0.16***	0.14***
			(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
constant	0.20***	0.04**	0.12***	0.02	0.13***	0.02	0.12***	0.02
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.01)
R^2	0.31	0.00	0.56	0.34	0.57	0.33	0.56	0.31
Employment share change of low-paid jobs								
employment share of MRO jobs in 1994	0.02	-0.01	0.00	-0.02	0.00	-0.01	-0.01	-0.05
	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.03)
middle × employment share of MRO jobs in 1994			-0.02**	-0.01	-0.02*	-0.01	-0.02**	-0.02*
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of MRO jobs in 1994			-0.04***	-0.05***	-0.04***	-0.04***	-0.05***	-0.06***
			(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
constant	0.06***	0.08***	0.08***	0.09***	0.08***	0.08***	0.08***	0.10***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
R^2	0.00	0.00	0.11	0.05	0.10	0.04	0.13	0.10
Employment share change of middle-paid jobs								
employment share of MRO jobs in 1994	0.28***	0.00	0.19**	-0.02	0.20**	-0.02	0.19**	0.00
	(0.10)	(0.04)	(0.09)	(0.04)	(0.09)	(0.05)	(0.09)	(0.03)
middle × employment share of MRO jobs in 1994			-0.02	-0.03**	-0.02	-0.03**	-0.01	-0.02*
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of MRO jobs in 1994			-0.12***	-0.11***	-0.12***	-0.11***	-0.12***	-0.10***
			(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
constant	-0.26***	-0.12***	-0.20***	-0.11***	-0.20***	-0.11***	-0.20***	-0.12***
	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)
R^2	0.33	0.00	0.51	0.15	0.51	0.16	0.54	0.16
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in MRO share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: MRO jobs are the 4 most routine or offshorable occupations with the highest RTI or OFF-GMS indexes (CS 48, 54, 62 and 67). Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of high-, low- and middle-paid occupational categories in total hours worked over the period 1994-2015 on the initial employment share of MRO jobs in 1994 at the individual city level. The first two columns report regression coefficients respectively with and without population weighting. In the regressions reported in columns 3 and 4 an additional slope for the medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.

Regressions show that the relationship between initial exposure to the most routine and offshorable jobs at the city level in 1994 is not a good predictor of increases of the shares of high- or low-paid or declines of middle-paid jobs. In population-weighted regressions, larger initial exposure to the most routine and offshorable jobs leads to a *lower* increase of high-paid occupations' share, no tendency for low-paid jobs, and a *lower* destruction of middle-paid jobs. At the same time, non-population weighted regressions (column 2) do not show any relationship between this exposure and the changes in the studied occupational shares. Allowing for a separate slope for large cities reveals that labor market responses in these metropolitan areas are always different from those of small ones.

Table F.39: Employment share changes 1994-2015 and exposure to 4 most routine/offshorable occupations (MRO) in 1994, Part II.

Δ *Emp. sh., MRO jobs*								
emp. sh., MRO jobs in 1994	-0.24*** (0.05)	-0.39*** (0.04)	-0.29*** (0.05)	-0.40*** (0.04)	-0.28*** (0.05)	-0.38*** (0.04)	-0.28*** (0.05)	-0.36*** (0.04)
middle × emp. sh., MRO jobs in 1994			-0.01 (0.01)	-0.01 (0.01)	-0.02 (0.01)	-0.02 (0.01)	-0.01 (0.01)	-0.02 (0.01)
large × emp. sh., MRO jobs in 1994			-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)	-0.06*** (0.02)
constant	-0.02 (0.02)	0.05*** (0.02)	0.01 (0.02)	0.06*** (0.02)	0.01 (0.02)	0.05*** (0.02)	0.01 (0.02)	0.04*** (0.01)
R^2	0.43	0.55	0.53	0.58	0.52	0.55	0.53	0.54
Δ *Emp. sh., OMP jobs*								
emp. sh., MRO jobs in 1994	0.52*** (0.05)	0.38*** (0.03)	0.48*** (0.05)	0.38*** (0.03)	0.48*** (0.05)	0.36*** (0.03)	0.48*** (0.05)	0.35*** (0.02)
middle × emp. sh., MRO jobs in 1994			-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
large × emp. sh., MRO jobs in 1994			-0.05*** (0.02)	-0.04** (0.02)	-0.05*** (0.02)	-0.04* (0.02)	-0.05*** (0.02)	-0.04** (0.02)
constant	-0.24*** (0.02)	-0.17*** (0.01)	-0.21*** (0.02)	-0.17*** (0.01)	-0.21*** (0.02)	-0.16*** (0.01)	-0.21*** (0.02)	-0.16*** (0.01)
R^2	0.84	0.71	0.87	0.73	0.86	0.71	0.87	0.73
Δ *Emp. sh., middle-paid > median wage*								
emp. sh., MRO jobs in 1994	0.18** (0.09)	0.01 (0.07)	0.13* (0.07)	0 (0.06)	0.13*(0.07)	0 (0.07)	0.14** (0.06)	0.01 (0.04)
middle × emp. sh., MRO jobs in 1994			-0.05*** (0.02)	-0.06*** (0.02)	-0.05*** (0.02)	-0.06*** (0.02)	-0.04** (0.02)	-0.04** (0.02)
large × emp. sh., MRO jobs in 1994			-0.13*** (0.02)	-0.11*** (0.02)	-0.13*** (0.02)	-0.11*** (0.02)	-0.12*** (0.02)	-0.10*** (0.02)
constant	-0.18*** (0.04)	-0.09*** (0.03)	-0.13*** (0.03)	-0.07*** (0.03)	-0.13*** (0.03)	-0.07** (0.03)	-0.14*** (0.03)	-0.08*** (0.02)
R^2	0.17	0.00	0.36	0.14	0.36	0.15	0.41	0.13
Δ *Emp. sh., middle-paid < median wage*								
emp. sh., MRO jobs in 1994	0.10** (0.05)	-0.01 (0.06)	0.06 (0.07)	-0.02 (0.06)	0.06 (0.08)	-0.02 (0.07)	0.05 (0.07)	-0.04 (0.05)
middle × emp. sh., MRO jobs in 1994			0.03** (0.02)	0.03* (0.02)	0.03* (0.02)	0.03* (0.02)	0.02 (0.02)	0.02 (0.02)
large × emp. sh., MRO jobs in 1994			0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)
constant	-0.08*** (0.02)	-0.04 (0.02)	-0.07** (0.03)	-0.04 (0.03)	-0.07** (0.03)	-0.04 (0.03)	-0.07** (0.03)	-0.03 (0.02)
R^2	0.09	0.00	0.12	0.04	0.12	0.04	0.10	0.02
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in MRO share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: MRO jobs are the 4 most routine or offshorable occupations with the highest RTI or OFF-GMS indexes (CS 48, 54, 62 and 67). Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants as of 2015. Population figures from 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of different middle-paid occupational categories in total hours worked over the period 1994-2015 on the initial employment share of MRO jobs at the individual city level in 1994. The first two columns report regression coefficients respectively with and without population weighting. In the regressions reported in columns 3 and 4 an additional slope for medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.
It can be observed in the top panel that initial exposure to MRO jobs is negatively correlated with the change in their employment shares. The interaction of a dummy for large cities with exposure is robustly negatively different from zero. Routine and offshorable jobs in large cities are destroyed at a higher rate than in small cities.
There is a strong and significant positive relation between the exposure to MRO group of jobs and the employment share of OMP jobs. We see that the largest decline in OMP occupations occurs in cities *least* exposed to MRO jobs. The simple linear fit indicates that in cities with largest initial MRO job exposure (above 0.5), the employment share of the OMP jobs increases over the period 1994-2015 (cf. Figure F.11).
We cannot observe any robustly significant relationship between exposure to MRO jobs and the changes in the employment share of middle-paid jobs with wages below the median average wage in 1994, regardless of the city size.

Table F.40: Employment share changes of the most offshorable jobs over the period 1994-2015 and exposure to 6 most offshorable occupations (OFF6) in 1994.

Employment share change of OFF6 jobs								
employment share of OFF6 jobs in 1994	-0.03	-0.29***	-0.17**	-0.31***	-0.15*	-0.30***	-0.14*	-0.25***
	(0.12)	(0.06)	(0.08)	(0.06)	(0.08)	(0.06)	(0.08)	(0.05)
middle × employment share of OFF6 jobs in 1994			-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
large × employment share of OFF6 jobs in 1994			-0.08***	-0.06***	-0.08***	-0.06***	-0.07***	-0.05***
			(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
constant	-0.15**	0.02	-0.05	0.03	-0.06	0.03	-0.07	0
	(0.07)	(0.03)	(0.05)	(0.03)	(0.05)	(0.04)	(0.05)	(0.03)
R^2	0.00	0.28	0.33	0.34	0.33	0.32	0.34	0.30
Observations	117	117	117	117	115	115	115	115
population weighted?	y	n	y	n	y	n	y	n
no outliers in OFF6 share	n	n	n	n	y	y	n	n
no outliers with employment share change	n	n	n	n	n	n	y	y

Notes: The category of 6 most offshorable occupations (OFF6) are those with highest OFF-GMS index (CS 46, 48, 54, 62, 65, 67). Robust standard errors. N=117; 11 cities > 0.5m, 44 cities between 0.1-0.5m and 62 cities between 0.05-0.1m inhabitants in 2015. Weighting by population as of 1990. ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

This Table shows the results of OLS regressions of the change in the employment share of 6 most offshorable occupations in total hours worked over the period 1994-2015 on their initial employment share at the individual city level in 1994. The first two columns report regression coefficients respectively with and without population weighting. In the regressions reported in columns 3 and 4 an additional slope for medium size (0.1-0.5m) and largest cities (above 0.5m inhabitants) is allowed. The last four columns show robustness checks, without outlier observations either in the initial job share or the change of employment shares over 1994-2015.
The relationship between the exposure to the most offshorable jobs at the city level and their share change over the period 1994-2015 is negative (but not significant in the population weighted regressions). However, this relationship is revealed to be significantly (at the 1% level) negative for largest cities once we allow for a separate slope. The 6 most offshorable jobs in large cities are destroyed at a higher rate than in small cities conditioning on the same level of exposure.

F.6 Pre-1994 Labor market developments

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

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

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

There are two data sources that allow to track occupations at the 2-digit PCS level back to 1982 when the PCS classification was introduced: the French Labor Market Survey (yearly data) and the Census (1982, 1990 and 1999). The publicly available Labor Market Survey gives data at the department but not at the commune level, hence it is impossible to precisely characterize city-level labor markets. The Census, on the other hand, gives the commune location of respondents but does not give data about hours worked or wages. We use the Census as we are interested in the shares of employment in cities, but in contrast to data presented in main text the patterns will refer to shares of people employed and not actual hours worked. We use the publicly available individual data for the 1982, 1990 and 1999 censuses (covering 1/4th for 1982 and 1990, and 1/20th for 1999 of the entire population respectively).

The 1982-1999 counterparts to Table 2 using Census data are in Table F.41.

Table F.41: Share of 4 highest-paid occupations per metropolitan area size, Census data 1982-1999.

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

The conclusions from this exercise are as follows. First of all, exposure to the most routine and offshorable jobs is much higher in 1982 for large cities above 0.5m inhabitants than in 1994 in the DADS data, and the discrepancies in terms of shares of high- middle- and MRO jobs across city sizes are lower. Employment shares of the MRO and, more generally, middle-paid jobs indeed decline faster

Table F.42: Share of 10 middle-paid occupations per metropolitan area size, Census data 1982-1999.

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

Table F.43: Share of 4 lowest-paid occupations per metropolitan area size, Census data 1982-1999.

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

in larger cities whether in 1982-1990 or in the entire 1982-1999 period. The labor market polarization across cities manifests itself as our theory predicts: high-paid jobs' shares increase most in largest cities, as found in the exhaustive DADS data for 1994-2015. Low-paid jobs do not increase at all in largest cities in 1982-1990 and increase less in terms of percentage points over 1990-1999 and the entire 1982-1999 period.

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

For individual occupational categories, the routine/offshorable job categories whose employment declines most in the studied years 1994-2015 in the DADS data in large cities are in particular PCS 46 and 54 (mid-level professionals and office workers respectively), whereas it is 62 and 67 (skilled and unskilled industrial workers respectively) for small cities (cf. the patterns in Figure F.7).[59] A part of the answer of such a differential evolution may lay in the fact that large cities had different shares of these jobs at the beginning of the 1990s (see Table F.45) than small cities, and such a discrepancy existed already in 1982. In particular, the share of mid-level professionals and office workers in employment was higher than that of industrial workers in 1982 in the largest cities above .75m inhabitants while the opposite is true for smaller cities. Therefore, the additional adjustment in terms of percentage points we observe in these blue-collar categories over the period 1994-2015 may be less pronounced as well. This feature of data may be explained by different deindustrialization across time and geographies as

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

Table F.44: Share of the 4 most routine and offshorable occupations (CS 48, 54, 62 and 67) per metropolitan area size, Census data 1982-1999.

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

shown in Table F.46 where the share of industry employment at Census years is given for the period 1968-2015. Already over the period 1968-1982 large cities experienced faster deindustrialization than small cities. Reports from research bodies as the INSEE or DATAR (Délégation interministérielle à l'aménagement du territoire et à l'attractivité régionale) indicate the following. Internal offshoring of manufacturing tasks within France might have played a part due to the reduction in internal transport costs (both because of highway and railway construction), environmental regulations to keep polluting industries out of high density areas or a deliberate government policy to decentralize economic activity across France (e.g. moving public engineering schools outside Paris), and hence not related to automation and offshoring shocks. For 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.

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

Table F.45: Employment share of selected middle-paid occupations per metropolitan area size, Census data 1982-1999.

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

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

Table F.46: Employment share in industry per metropolitan area size, Census data 1968-2015.

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

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

F.7 Public sector employment 1990-2015.

DADS-Postes data does not contain public sector employee data prior to the end of the 2000s. We can, however, assess the evolution of public sector jobs (number of positions but not hours) using the harmonized Census SAPHIR. In this exercise, we cannot restrict the private sector only to incorporated firms (catégorie juridique "5" in INSEE's classification) as in the main sample.

We restrict attention to all 2-digit CS categories between 23 and 68 apart from CS 44 (clergy) among employed in cities considered in our sample within the age range 25-64 for 1990 and 2015. We classify CS 33 (category A: public administration managers) and 34 (higher education professors, scientists in public employment) as high-paid, CS 42 (teachers) and CS 45 (category B: intermediate public administration workers) as middle-paid, CS 52 (category C and D employees of public administration) and the public employment portion of CS 53 (policemen, military) as low-paid.

Public sector employment grew nationally by 0.5 pp in total from 23.3% to 23.8% of total employment between 1990 and 2015. In 1990, small cities with populations below 0.1m were relatively more abundant in public sector jobs with a 24.8% of total versus 21.3% in Paris. The share of high-paid public sector jobs increased in the period 1990-2015 by 1.2 pp (to 6.8% of total jobs).

The breakdown of the change within cities is given in Table F.47. There was growth of public sector employment in cities with population below 0.5m (between 1.9 pp in cities between 0.2-0.5m and 2.7 pp in cities between 0.1-0.2m). High-paid public sector jobs increased in cities of all sizes (from 0.7 pp in cities with a population between 0.5m-0.75m to 1.7 pp in cities between 0.1-0.2m). Low-paid jobs increased in smallest cities by 1.2 pp and decreased by 2.1 pp in Paris, with intermediate values for other cities. Changes in the shares of middle-paid public sector jobs are negligible.

Based on this data, we see that there was no comparable shock to middle-paid jobs in the public sector as in the private sector discussed in the paper. The public sector does not seem to play an important role in the evolution of jobs nationally, within (or across, not shown) cities, and does not exhibit any spatially interesting patterns. In particular, the decrease in middle-paid jobs observed in the private sector was not absorbed by middle-paid or low-paid public sector jobs, either nationally or at the city level. Moreover, the changes in low-paid public sector jobs (decreases in larger cities and increases in smaller cities) do not "compensate" the lower growth of low-paid jobs in larger cities and lead to a higher growth of low-paid occupations there in the aggregate but actually exacerbate the differential growth in these types of jobs across cities overall.

Table F.47: Change in the share of private and public sector jobs within city groups and in the aggregate 1990-2015 from Census data.

job type	Paris	$>$ 0.75M	0.5-0.75M	0.2-0.5M	0.1-0.2M	0.05-0.1M	All cities
private sector	0.007	-0.001	0.002	-0.019	-0.027	-0.025	-0.005
public sector high-paid	0.011	0.011	0.007	0.016	0.017	0.009	0.012
public sector middle-paid	0.003	0.000	-0.003	0.002	0.005	0.004	0.002
public sector low-paid	-0.021	-0.010	-0.006	0.001	0.005	0.012	-0.009

Notes: Aggregation to the city-level based on metropolitan areas as of 2015. Includes employed individuals between 25-64 of age for all 2-digit CS categories 23-68 with the exception of CS category 44 (clergy).