Supplemental Material - Not For Publication

Union and Firm Labor Market Power

Miren Azkarate-Askasua and Miguel Zerecero

This Supplemental Material is organized as follows. Section I presents additional derivations, which includes different extensions of the model and derivations with an alternative production function. Section II presents additional proofs and results on the model aggregation. Section III contains additional estimation results. Section IV contains additional counterfactual. Section VI presents details on sample selection and variable construction and Section VII presents summary statistics.

I Additional derivations

I.1 Extension: Endogenous participation

We showed in the proof of Proposition 4 that the solution of sector prices $\mathbf{P}$ is homogeneous of degree zero with respect to total employment level which we denote here as L_e. We have that,

$$L_{io}(w_{io}) = \frac{T_{io}w_{io}^{\varepsilon_b}}{\Phi_m}\frac{\Phi_m^{\eta/\varepsilon_b}\Gamma_b^{\eta}}{\Phi}L = \frac{T_{io}w_{io}^{\varepsilon_b}}{\Phi_m}\frac{\Phi_m^{\eta/\varepsilon_b}\Gamma_b^{\eta}}{\Phi_e}L_e.$$

We have that $L_e = \frac{\Phi_e}{\Phi}L$ with $\Phi_e \equiv \sum_{m\in\mathcal{M}}\Phi_m^{\eta/\varepsilon_b}\Gamma_b^{\eta}$ is the part of Φ that comes from the employed and $\Phi_u \equiv \sum_{uo\in\mathcal{U}}(T_{uo}w_{uo}^{\varepsilon_b})^{\eta/\varepsilon_b}\Gamma_b^{\eta}$ is the part from the out of the labor force as in the main text.

The model aggregation steps are the same as in Section A with the exception that L_b now is $L_{b,e}$. We normalize all the reservation wages w_{uo} to 1. We recover the out-of-the-labor-force amenities T_{uo} to match the observed unemployment rate and we can compute Φ_u. There are no markdowns for the OTLF and we set the productivities of the fictitious OTLF establishments to zero such that they do not contribute to aggregate output.

Aggregating from (B5) from the Online Appendix,

$$\Phi_{b,e} = \left(\frac{\Phi}{L}\right)^{\frac{\eta\delta}{1+\eta\delta}}\sum_{m\in\mathcal{M}_b}\widetilde{\Phi}_m^{\frac{\eta(1+\varepsilon_b)}{\varepsilon_b(1+\eta\delta)}}P_b^{\frac{\eta}{(1-\alpha_b)(1+\eta\delta)}} = \left(\frac{\Phi}{L}\right)^{\frac{\eta\delta}{1+\eta\delta}}\widetilde{\Phi}_{b,e}P_b^{\frac{\eta}{(1-\alpha_b)(1+\eta\delta)}} \tag{I1}$$

$$\widetilde{\Phi}_{b,e} \equiv \sum_{m\in\mathcal{M}_b}\widetilde{\Phi}_m^{\frac{\eta(1+\varepsilon_b)}{\varepsilon_b(1+\eta\delta)}}, \qquad \Phi \equiv \Phi_e + \Phi_u,$$

and,

$$\Phi_e \equiv \left(\frac{\Phi}{L}\right)^{\frac{\eta\delta}{1+\eta\delta}} \sum_{b\in\mathcal{B}} \widetilde{\Phi}_{b,e} P_b^{\frac{\eta}{(1-\alpha_b)(1+\eta\delta)}} \Gamma_b^{\eta} = \left(\frac{\Phi}{L}\right)^{\frac{\eta\delta}{1+\eta\delta}} \widetilde{\Phi}_e \tag{I2}$$

$$\widetilde{\Phi}_e \equiv \sum_{b\in\mathcal{B}} \widetilde{\Phi}_{b,e} P_b^{\frac{\eta}{(1-\alpha_b)(1+\eta\delta)}} \Gamma_b^{\eta}$$

$$L_{b,e} = \frac{\Phi_{b,e}\Gamma_b^{\eta}}{\Phi_e} L_e = \frac{\widetilde{\Phi}_{b,e}\Gamma_b^{\eta} P_b^{\frac{\eta}{(1-\alpha_b)(1+\eta\delta)}}}{\widetilde{\Phi}_e} L_e. \tag{I3}$$

We can solve for the prices without knowing total employment level L_e. Total employment level is $L_e = \frac{\Phi_e}{\Phi} L$, where L is total labor supply (employed and out-of-the-labor-force) that will determine the level of aggregate output. We can find it by solving for Φ_e in equation (I2),

$$\Phi_e^{\frac{1+\eta\delta}{\eta\delta}} L = (\Phi_e + \Phi_u) \widetilde{\Phi}_e^{\frac{1+\eta\delta}{\eta\delta}}.$$

The solution is obviously unique as the left hand side is convex and the right hand side linear. With the solution for Φ_e one can construct all the aggregates back.

I.2 Extension: Agglomeration

Plugging the labor supply into (29), the wage in the baseline economy is:

$$w_{io} = \left(\beta_b \lambda(\mu_{io}, \varphi_b) \frac{Z_{io}}{(T_{io}\Gamma_b^{\eta})^{\delta}}\right)^{\frac{1}{1+\varepsilon_b\delta}} \Phi_m^{\nu_b - \frac{\eta}{\varepsilon_b}\widetilde{\nu_b}} \left(\frac{\Phi}{L}\right)^{\widetilde{\nu_b}}, \quad \nu_b = \frac{\delta}{1+\varepsilon_b\delta}, \quad \widetilde{\nu_b} = \frac{\delta - \gamma}{1+\varepsilon_b\delta}.$$

The baseline wage can be written as: $w_{io} = \widetilde{w}_{io} \Phi_m^{\nu_b - \frac{\eta}{\varepsilon_b}\widetilde{\nu_b}} \left(\frac{\Phi}{L}\right)^{\widetilde{\nu_b}}$. Analogously, the counterfactual wage is: $w_{io} = \omega_{io} \widehat{P}_b^{\frac{1}{(1-\alpha_b)(1+\varepsilon_b\delta)}} \Phi_m^{\nu_b - \frac{\eta}{\varepsilon_b}\widetilde{\nu_b}} \left(\frac{\Phi}{L}\right)^{\widetilde{\nu_b}}$. Aggregating to generate Φ_m,

$$\Phi_m = \widetilde{\Phi}_m^{\frac{1+\varepsilon_b\delta}{1+\eta(\delta-\gamma)}} \left(\frac{\Phi}{L}\right)^{\frac{\varepsilon_b(\delta-\gamma)}{1+\eta(\delta-\gamma)}}. \tag{I4}$$

The counterfactual Φ'_m is analogously $\Phi'_m = \widetilde{\Phi'}_m^{\frac{1+\varepsilon_b\delta}{1+\eta(\delta-\gamma)}} \widehat{P}_b^{\frac{\varepsilon_b}{(1-\alpha_b)(1+\eta(\delta-\gamma))}} \left(\frac{\Phi}{L}\right)^{\frac{\varepsilon_b(\delta-\gamma)}{1+\eta(\delta-\gamma)}}$.

In order to be able to find a solution to the model, we need that the exponents are bounded. This is equivalent to requiring $\gamma \neq \frac{1}{\eta} + \delta$. The parameter γ governs the strength of agglomeration forces within a local labor market, and δ and $\frac{1}{\eta}$ are related with dispersion forces. Those come from the decreasing returns to scale (δ) and from the variance of taste shocks ($\frac{1}{\eta}$). When the latter is high, the mass of workers having extreme taste shocks is higher. This implies that agglomeration forces will impact less as workers would be more inelastic to changes in wages. The standard condition for uniqueness of the equilibrium with agglomeration would be that it is sufficiently weak ($\gamma < \frac{1}{\eta} + \delta$). We instead find the weaker condition $\gamma \neq \frac{1}{\eta} + \delta$.

The counterfactual industry labor supply is:

$$L'_b = \frac{\widehat{P}_b^{\frac{\eta}{(1-\alpha_b)(1+\eta(\delta-\gamma))}} \widetilde{\Phi}'_b \Gamma_b^{\eta}}{\sum_{b \in \mathcal{B}} \widehat{P}_{b'}^{\frac{\eta}{(1-\alpha_{b'})(1+\eta(\delta-\gamma))}} \widetilde{\Phi}'_{b'} \Gamma_{b'}^{\eta}}, \quad \widetilde{\Phi}'_b \equiv \sum_{m \in \mathcal{M}_b} \widetilde{\Phi'}_m^{\frac{\eta(1+\varepsilon_b \delta)}{\varepsilon_b(1+\eta(\delta-\gamma))}}$$

The counterfactual establishment-occupation output y'_{io} and sector output Y'_b are:

$$y'_{io} = \frac{\widehat{P}_b^{\frac{\alpha_b}{1-\alpha_b}}}{P_b} Z_{io} {L'_{io}}^{1-\delta} {L'_m}^{\gamma}, \quad Y'_b = \frac{\widehat{P}_b^{\frac{\alpha_b}{1-\alpha_b}}}{P_b} \overline{Z}_b \overline{\Omega}'_b {L'_b}^{1-\delta+\gamma},$$

where γ changed the returns to scale and the aggregations. We define:

$$\begin{aligned}
\overline{\Omega}'_b &\equiv \sum_{io \in \mathcal{I}_b} \frac{A_{io}}{A_b} {s'_{io|m}}^{1-\delta} {s'_{m|b}}^{1-\delta+\gamma} \\
\overline{Z}_b &\equiv \sum_{io \in \mathcal{I}_b} Z_{io} \tilde{s}_{io|m}^{1-\delta} \tilde{s}_{m|b}^{1-\delta+\gamma},
\end{aligned}$$

where $\overline{Z}_b$ is a measure of sector productivity under the efficient allocation with agglomeration (that is the same in the baseline and in the counterfactual) and $\overline{\Omega}'_b$ is a measure of misallocation in the counterfactual.

The expressions for the baseline are analogous but setting $\widehat{P}_b = 1$, and defining the above with baseline employment shares, $Y'_b = \frac{1}{P_b} \overline{Z}_b \overline{\Omega}_b {L_b}^{1-\delta+\gamma}$.

The intermediate good demand in the counterfactual relative to the baseline is:

$$\begin{aligned}
\widehat{P}_b^{\frac{1}{1-\alpha_b}} \widehat{\overline{\Omega}}_b \left(\frac{L'_b(\widehat{\mathbf{P}})}{L_b} \right)^{1-\delta+\gamma} &= \prod_{b' \in \mathcal{B}} \widehat{P}_b^{\frac{\alpha_{b'}}{1-\alpha_{b'}}} \widehat{\overline{\Omega}}_{b'} \left(\frac{L'_{b'}(\widehat{\mathbf{P}})}{L_{b'}} \right)^{1-\delta+\gamma} \\
\Leftrightarrow \widehat{P}_b^{\frac{1+\eta}{(1-\alpha_b)(1+\eta(\delta-\gamma))}} \widehat{\overline{\Omega}}_b \left(\frac{\widetilde{\Phi}'_b \Gamma_b^{\eta}}{L_b} \right)^{1-\delta+\gamma} &= \prod_{b' \in \mathcal{B}} \widehat{P}_{b'}^{\alpha_{b'}(1+\eta(\delta-\gamma))+\eta(1-\delta+\gamma)} \widehat{\overline{\Omega}}_{b'} \left(\frac{\widetilde{\Phi}'_{b'} \Gamma_{b'}^{\eta}}{L_{b'}} \right)^{1-\delta+\gamma}.
\end{aligned}$$

Uniqueness of the solution to this system of equations is guaranteed by $\sum_{b \in \mathcal{B}} \alpha_b \theta_b < 1$. This condition being the same as for Proposition 4, uniqueness of the equilibrium with agglomeration forces only needs the additional requirement of $\gamma \neq \frac{1}{\eta} + \delta$.

I.3 Alternative production function

For completeness, in this section we lay out a model with an alternative Cobb-Douglas production function with generic capital and a labor composite that is at odds with the data.

Suppose that establishment i produces using some generic capital K_i and a labor composite H_i of different occupations:

$$y_i = \widetilde{A}_i K_i^{\alpha_b} H_i^{\beta_b} = \widetilde{A}_i K_i^{\alpha_b} \left(\prod_{o \in \mathcal{O}} L_{io}^{\gamma_o} \right)^{\beta_b}, \quad \sum_o \gamma_o = 1, \quad \alpha_b + \beta_b \leq 1. \tag{I5}$$

The first order conditions with respect to capital and the bargained wage are:

$$w_{io} = \beta_b \gamma_o \lambda(\mu_{io}, \varphi_b) P_b \frac{y_i}{L_{io}},$$
$$R_b = \alpha_b \widetilde{A}_i K_i^{\alpha_b - 1} H_i^{\beta_b}.$$

Substituting the first order condition for capital into the production function, the wage first order condition becomes:

$$w_{io} = \beta_b \gamma_o \lambda(\mu_{io}, \varphi_b) A_i H_i^{1-\delta} L_{io}^{-1} P_b^{\frac{1}{1-\alpha_b}},$$

where we plugged the labor supply and used the definition of $\delta = 1 - \frac{\beta_b}{1-\alpha_b}$ from the main text and $A_i = \widetilde{A}_i^{\frac{1}{1-\alpha_b}} \left(\frac{\alpha_b}{R_b}\right)^{\frac{\alpha_b}{1-\alpha_b}}$ as in the main text. Using those and solving for L_{io}, we can write the labor composite H_i as function of wages:

$$H_i^{\delta} = P_b^{\frac{1}{1-\alpha_b}} \prod_{o \in \mathcal{O}} \beta_b \gamma_o \lambda(\mu_{io}, \varphi_b) w_{io}^{-1}$$

Substituting the wage equation with the labor supply (12) into the expression above, we get:

$$\begin{aligned} H_i^{1+\varepsilon_b \delta} &= P_b^{\frac{\varepsilon_b}{1-\alpha_b}} \prod_{o \in \mathcal{O}} \left(\beta_b \gamma_o \lambda(\mu_{io}, \varphi_b) A_i (T_{io} \Gamma_b^{\eta})^{1/\varepsilon_b}\right)^{\varepsilon_b \gamma_o} \prod_{o \in \mathcal{O}} \left(\Phi_m^{1-\eta/\varepsilon_b} \frac{\Phi}{L}\right)^{-\gamma_o} \\ &= P_b^{\frac{\varepsilon_b}{1-\alpha_b}} (\beta_b \Upsilon A_i)^{\varepsilon_b} T_i \Gamma \prod_{o \in \mathcal{O}} \lambda(\mu_{io}, \varphi_b)^{\varepsilon_b \gamma_o} \prod_{o \in \mathcal{O}} \left(\Phi_m^{1-\eta/\varepsilon_b} \frac{\Phi}{L}\right)^{-\gamma_o}, \end{aligned}$$

where $\Upsilon \equiv \prod_{o \in \mathcal{O}} \gamma_o$, $\Gamma \equiv \prod_{o \in \mathcal{O}} \Gamma_b^{\eta}$ and $T_i \equiv \prod_{o \in \mathcal{O}} T_{io}$. Plugging back into the wage equation and rearranging, we get:

$$\begin{aligned} w_{io} = \Bigg[& \lambda(\mu_{io}, \varphi_b) \frac{\gamma_o}{T_{io} \Gamma_b^{\eta}} (\beta_b A_i)^{\frac{1+\varepsilon_b}{1+\varepsilon_b \delta}} (\Upsilon (T_i \Gamma)^{1/\varepsilon_b})^{\frac{\varepsilon_b (1-\delta)}{1+\varepsilon_b \delta}} \\ & \times \left(\prod_{o' \in \mathcal{O}} \lambda(\mu_{io'}, \varphi_b)^{\varepsilon_b \gamma'_o}\right)^{\frac{1-\delta}{1+\varepsilon_b \delta}} \left(\prod_{o' \in \mathcal{O}} \Phi_{m'}^{(\eta/\varepsilon_b - 1)\gamma'_o}\right)^{\frac{1-\delta}{1+\varepsilon_b \delta}} \Phi_m^{1-\eta/\varepsilon_b} \Bigg]^{\frac{1}{1+\varepsilon_b}} \left(\frac{\Phi}{L}\right)^{\frac{1}{1+\varepsilon_b}} P_b^{1/\chi_b}, \end{aligned} \tag{I6}$$

with $\chi_b = (1-\alpha_b)(1+\varepsilon_b \delta)$. Define the following:

$$\begin{aligned} c_{io} &\equiv \frac{\gamma_o}{T_{io} \Gamma_b^{\eta}} (\beta_b A_i)^{\frac{1+\varepsilon_b}{1+\varepsilon_b \delta}} (\Upsilon (T_i \Gamma)^{1/\varepsilon_b})^{\frac{\varepsilon_b (1-\delta)}{1+\varepsilon_b \delta}}, \\ C_l &\equiv \prod_{o' \in \mathcal{O}} \left(\Phi_{m'}^{(\eta/\varepsilon_b - 1)\gamma_o}\right)^{\frac{\delta}{1+\varepsilon_b \delta}} \left(\frac{\Phi}{L}\right)^{\frac{1}{1+\varepsilon_b}}, \\ F_b &\equiv P_b^{1/\chi_b}, \end{aligned}$$

where C_l is a location constant with $l = n \times h$. Rearranging we have that:

$$w_{io} = \left[\lambda(\mu_{io}, \varphi_b) c_{io} \left(\prod_{o' \in \mathcal{O}} \lambda(\mu_{io'}, \varphi_b)^{\varepsilon_b \gamma'_o}\right)^{\frac{1-\delta}{1+\varepsilon_b \delta}} \frac{\Phi_m^{1-\eta/\varepsilon_b}}{\prod_{o' \in \mathcal{O}} \Phi_{m'}^{(1-\eta/\varepsilon_b)\gamma'_o}}\right]^{\frac{1}{1+\varepsilon_b}} C_l F_b. \tag{I7}$$

The last system is equivalent to the one in (I6) and has the benefit to being able to write the wages as $w_{io} = \widetilde{w}_{io} C_m F_b$, where we want $\widetilde{w}_{io}$ to be homogeneous of degree zero with respect constants to m level. Note that the last term inside the brackets is homogeneous of degree zero with respect to location l constants shared by all the occupations of a establishments. Then, defining $\widetilde{\Phi}_m \equiv \sum_{i \in \mathcal{I}_m} T_{io} w_{io}^{\varepsilon_b}$, the establishment-occupation or normalized wage is:

$$\widetilde{w}_{io} \equiv \left[\lambda(\mu_{io}, \varphi_b) c_{io} \left(\prod_{o' \in \mathcal{O}} \lambda(\mu_{io'}, \varphi_b)^{\varepsilon_b \gamma'_o} \right)^{\frac{1-\delta}{1+\varepsilon_b \delta}} \frac{\widetilde{\Phi}_m^{1-\eta/\varepsilon_b}}{\prod_{o' \in \mathcal{O}} \widetilde{\Phi}_{m'}^{(1-\eta/\varepsilon_b)\gamma'_o}} \right]^{\frac{1}{1+\varepsilon_b}}. \tag{I8}$$

$\widetilde{w}_{io}$ is homogeneous of degree zero with respect to location l constants shared by all occupations. This property, makes the model with the alternative production function also block recursive. That is, it allows solving for the normalized wages of every location l (combinations of commuting zone n and sub-industry h combinations) independently and then recover the aggregate constants. Aggregating (I8) and solving for $\widetilde{\Phi}_m$, we have:

$$\widetilde{\Phi}_m = \left[\frac{\sum_{i \in I_m} \left(\lambda(\mu_{io}, \varphi_b) c_{io} T_{io}^{\frac{1+\varepsilon_b}{\varepsilon_b}} \prod_{o' \in \mathcal{O}} \lambda(\mu_{io'}, \varphi_b)^{\varepsilon_b \gamma'_o} \right)^{\frac{1-\delta}{1+\varepsilon_b \delta}}}{\prod_{o' \in \mathcal{O}} \widetilde{\Phi}_{m'}^{(1-\eta/\varepsilon_b)\gamma'_o}} \right]^{\frac{\varepsilon_b}{1+\eta}}.$$

Taking everything to the power $(1-\eta/\varepsilon_b)\gamma_o$ and taking the product,

$$\mathcal{L}_l \equiv \prod_{o' \in \mathcal{O}} \widetilde{\Phi}_{m'}^{(1-\eta/\varepsilon_b)\gamma'_o} = \prod_{o' \in \mathcal{O}} \left[\sum_{i \in I_m} \left(\lambda(\mu_{io}, \varphi_b) c_{io} T_{io}^{\frac{1+\varepsilon_b}{\varepsilon_b}} \prod_{o' \in \mathcal{O}} \lambda(\mu_{io'}, \varphi_b)^{\varepsilon_b \gamma'_o} \right)^{\frac{1-\delta}{1+\varepsilon_b \delta}} \right]^{\gamma_{o'} \frac{\varepsilon_b - \eta}{1+\varepsilon_b - \eta}},$$

which recovers all the local labor market constants inside $\widetilde{w}_{io}$.

In order to prove the existence and uniqueness of the solution of the system (I8), define $\widehat{w}_{io}$ as:

$$\widehat{w}_{io} = \left[\lambda(\mu_{io}, \varphi_b) \left(\prod_{o' \in \mathcal{O}} \lambda(\mu_{io'}, \varphi_b)^{\varepsilon_b \gamma'_o} \right)^{\frac{1-\delta}{1+\varepsilon_b \delta}} \right]^{\frac{1}{1+\varepsilon_b}} c_{io}^{\frac{1}{1+\varepsilon_b}}$$

$$w_{io} = \widehat{w}_{io} \left[\frac{\widetilde{\Phi}_m^{1-\eta/\varepsilon_b}}{\mathcal{L}_l} \right]^{\frac{1}{1+\varepsilon_b}} C_l F_b = \widehat{w}_{io} z_l = \widetilde{w}_{io} C_l F_b. \tag{I9}$$

We can show that the system formed by (I9) has a solution and is unique.

Proposition 5. *For given parameters $0 \leq \alpha_b, \beta_b < 1$, $1 < \eta < \varepsilon_b$, $0 \leq \delta \leq 1$, transformed price F_b, constants C_l, $\widetilde{\Phi}_m$, $\mathcal{L}_l$ and non-negative vectors of productivities $\{A_i\}_{i \in m}$ and amenities $\{T_{io}\}_{io \in m}$, there exists a unique vector of wages $\{w_{io}\}_{io \in I_m}$ for every location l (combination of commuting zone n and sub-industry h) that solves the system formed by (I9).*

Proof. For existence, first note that $\lambda(\mu_{io}, \varphi_b) \in \left[(1-\varphi_b)\frac{\eta}{1+\eta} + \varphi_b \frac{1}{1-\delta}, (1-\varphi_b)\frac{\varepsilon_b}{1+\varepsilon_b} + \varphi_b \frac{1}{1-\delta} \right], \forall i, o$. Define a vector $\mathbf{w}$ with wage of all the establishment-occupations at location l, $\mathbf{w} \equiv \{w_{11}, w_{12}, ..., w_{1O}, ..., w_{I1}, ..., w_{IO}\}$. Taking for now the elements of z_l as constants. The system to solve is: $f_{io}(\mathbf{w}) = \widehat{w}_{io} z_l$.

We have that

$$\mathbf{w} \in \mathcal{C} \equiv \left[\left((1-\varphi_b)\frac{\eta}{1+\eta}+\varphi_b\frac{1}{1-\delta}\right)^{\frac{1}{1+\eta\delta}} c_{11}^{\frac{1}{1+\varepsilon_b}} z_{I1}, \left((1-\varphi_b)\frac{\varepsilon_b}{1+\varepsilon_b}+\varphi_b\frac{1}{1-\delta}\right)^{\frac{1}{1+\eta\delta}} c_{11}^{\frac{1}{1+\varepsilon_b}} z_{I1}\right]$$
$$\times \ldots \times \left[\left((1-\varphi_b)\frac{\eta}{1+\eta}+\varphi_b\frac{1}{1-\delta}\right)^{\frac{1}{1+\eta\delta}} c_{IO}^{\frac{1}{1+\varepsilon_b}} z_{IO}, \left((1-\varphi_b)\frac{\varepsilon_b}{1+\varepsilon_b}+\varphi_b\frac{1}{1-\delta}\right)^{\frac{1}{1+\eta\delta}} c_{IO}^{\frac{1}{1+\varepsilon_b}} z_{IO}\right].$$

The system f_{io} is continuous on wages and maps into itself on $\mathcal{C}$. The last set being a compact set we can apply Brower's fixed point theorem.

For uniqueness, once the product of the wedges is substituted, $\widehat{w}_{io}$ is:

$$\widehat{w}_{io} = \left[\lambda(\mu_{io}, \varphi_b) c_{io} \prod_{o' \in \mathcal{O}} (w_{io'} c_{io}^{-\frac{1}{1+\varepsilon_b}})^{\gamma'_o \varepsilon_b (1-\delta)}\right]^{\frac{1}{1+\varepsilon_b}}$$

Define the function $g_{io}(\mathbf{w}) = f_{io}(\mathbf{w}) - w_{io}$. Gross substitution is fulfilled if $\frac{\partial g_{io}(\mathbf{w})}{\partial w_{jo}} > 0, \forall j \neq i$ with $j \in \mathcal{I}_l$ and $\frac{\partial g_{io}(\mathbf{w})}{\partial w_{io'}}, \forall o'$. Gross substitution resumes to taking the partial derivatives of $\widehat{w}_{io}$ which are positive by similar reasoning as in the main proof. Finally, $\widehat{w}_{io}$ is homogeneous of degree $\frac{\varepsilon_b}{1+\varepsilon_b}(1-\delta) < 1$. Therefore the solution to the system (I9) exists and is unique. □

Finally, the model can be aggregated up to the industry level following similar steps as in Proposition 3.

II Additional results and proofs

We use the following Theorem and Corollary to establish uniqueness in our proofs. These are taken from Allen, Arkolakis, and Li (2016) as they are not present any more in the current version of their paper Allen, Arkolakis, and Li (2021). Of course, any error should be attributed to us.

Theorem 1. *Consider $g : \mathbb{R}^n_{++} \times \mathbb{R}^m_{++}$ for some $n \in \{1, ..., N\}$ and $m \in \{1, ..., M\}$ such that:*

(i) homogeneity of any degree: $g(tx, ty) = t^k g(x, y)$, $t \in \mathbb{R}_{++}$ and $k \in \mathbb{R}$,

(ii) gross-substitution property: $\frac{\partial g_i}{\partial x_j} > 0$ for all $i \neq j$,

(iii) monotonicity with respect to the joint variable: $\frac{\partial g_i}{\partial y_k} \geq 0$, for all i, k.

Then, for any given $y^0 \in \mathbb{R}^M_{++}$ there exists at most one solution satisfying $g(x, y^0) = 0$.

Proof. We proceed by contradiction. Suppose there are two different up-to-scale, solutions, x^1, x^2, such that $f(x^1) = f(x^2) = 0$ i.e. $g(x^1, y^0) = g(x^2, y^0) = 0$. Without loss of generality, suppose there exists some $t > 1$ such that $tx^1_j \geq x^2_j$ for all $j \in \{1, ..., n\}$ and the equality holds for at least one $j = \bar{j}$. Then the inequality must strictly hold since x^1 and x^2 are different up-to-scale. Condition (iii) $\frac{\partial g_i}{\partial y_k} \geq 0$, for all i, k implies that $g(tx^1, y^0) \leq g(tx^1, ty^0) = 0$ where $g(tx^1, ty^0) = 0$ is from condition (i) (and also $g(tx^2, ty^0) = 0$ because x^1 and x^2 are solutions). However, condition (ii) implies $g_j(tx^1, y^0) > g_j(x^2, y^0) = 0$, thus a contradiction. □

Corollary 1. *Assume (i) $f(x)$ satisfies gross-substitution and (ii) $f(x)$ can be decomposed as $f(x) = \sum_{j=1}^{\nu_f} g^j(x) - \sum_{k=1}^{\nu_g} h^k(x)$, where $g^j(x), h^k(x)$ are non-negative vector functions and, respectively, homogeneous of degree α_j and β_k, with $\bar{\alpha} = \max \alpha_j \leq \min \beta_k$.*

1. *Then there is at most one up-to-scale solution of $f(x) = 0$.*

2. *In particular, if for some j,k $\alpha_j \neq \beta_k$, then there is at most one solution.*

Proof. Define $m(x,y)$ as a vector function where $m_i(x,y) = \sum_{j=1}^{\nu_f} y^{\bar{\alpha}-\alpha_j} g_i^j(x) - \sum_{k=1}^{\nu_g} y^{\bar{\alpha}-\beta_k} h_i^k(x)$. Obviously, $m(x,y)$ is of homogenous degree $\bar{\alpha}$ and $\frac{\partial m_i}{\partial y} \geq 0$. Also we have $f(x) = m(x,y^0)$ where $y^0 = 1$, thus the above theorem applies.

Furthermore, if $f_i(x)$ is not homogeneous of some degree because $\alpha_j \neq \beta_k$, there is at most one solution. Suppose not, if tx^1 and x^1 are the solutions, then $f_i(x^1) > t^{-min(\beta_k)} f_i(tx^1) = 0$, also a contradiction. □

II.1 Existence and uniqueness of local market equilibrium in Berger, Herkenhoff, and Mongey (2022)

Our existence and uniqueness proof extends easily to the case consider by Berger et al. (2022), where instead of using shares of employment $s_{io|m}$, they use wage bill shares $s_{io|m}^w = \frac{w_{io}L_{io}}{\sum_{j\in\mathcal{I}_m} w_{jo}Ljo}$, and no bargaining power. i.e. $\varphi_b = 0$. The existence proof is exactly the same. For uniqueness and to establish gross-substitution of a similar function $g_{io}(\mathbf{w})$, we can follow all the steps of the previous proof and note that:

$$s_{io|m}^w = \frac{T_{io} w_{io}^{1+\varepsilon_b}}{\sum_{j\in\mathcal{I}_m} T_{jo} w_{jo}^{1+\varepsilon_b}}.$$

Thus, clearly, $\frac{\partial s_{io|m}}{\partial w_{jo}} < 0$ for any $i \neq j$ and $g_{io}(\mathbf{w})$ also satisfies the gross-substitution property. Then we can conclude that the local labor market equilibrium of Berger et al. (2022) also exists and is unique.

II.2 Aggregation of the model

Here we provide additional details that correspond to the aggregation of the model. Parts of these were already covered in the proof of Proposition 3 in the Online Appendix. We include them in the interest of clarity.

Aggregating establishment-occupation output (6) and using the restriction $\frac{\beta_b}{1-\alpha_b} = 1 - \delta \in [0,1]$, the local labor market output is:

$$Y_m = \sum_{i\in\mathcal{I}_m} y_{io} = P_b^{\frac{\alpha_b}{1-\alpha_b}} \sum_{i\in\mathcal{I}_m} A_{io} L_{io}^{1-\delta} = P_b^{\frac{\alpha_b}{1-\alpha_b}} \sum_{i\in\mathcal{I}_m} A_{io} s_{io|m}^{1-\delta} L_m^{1-\delta} = P_b^{\frac{\alpha_b}{1-\alpha_b}} \Omega_m A_m L_m^{1-\delta},$$

where the local labor market productivity and misallocation are measured as:

$$\Omega_m \equiv \sum_{i \in \mathcal{I}_m} \frac{A_{io}}{A_m} s_{io|m}^{1-\delta}, \quad A_m \equiv \sum_{i \in \mathcal{I}_m} A_{io} \tilde{s}_{io|m}^{1-\delta}, \quad \tilde{s}_{io|m} = \frac{\left(T_{io}^{1/\varepsilon_b} A_{io}\right)^{\frac{\varepsilon_b}{1+\varepsilon_b \delta}}}{\sum_{j \in \mathcal{I}_m} \left(T_{jo}^{1/\varepsilon_b} A_{jo}\right)^{\frac{\varepsilon_b}{1+\varepsilon_b \delta}}}.$$

The definition of $\tilde{s}_{io|m}$ comes from Proposition 1 with constant labor wedges.

Further aggregating to sector level according to (4):

$$Y_b = \sum_{m \in \mathcal{M}_b} Y_m = P_b^{\frac{\alpha_b}{1-\alpha_b}} \sum_{m \in \mathcal{M}_b} \Omega_m A_m L_m^{1-\delta} = P_b^{\frac{\alpha_b}{1-\alpha_b}} \Omega_b A_b L_b^{1-\delta}. \tag{II1}$$

The sector level measures of productivity and misallocation are:

$$\begin{aligned}
\Omega_b &\equiv \sum_{m \in \mathcal{M}_b} \Omega_m \frac{A_m}{A_b} s_{m|b}^{1-\delta} = \sum_{m \in \mathcal{M}_b} \sum_{io \in \mathcal{I}_m} \frac{A_{io}}{A_b} s_{io|m}^{1-\delta} s_{m|b}^{1-\delta}, \\
A_b &\equiv \sum_{m \in \mathcal{M}_b} A_m \tilde{s}_{m|b}^{1-\delta} = \sum_{m \in \mathcal{M}_b} \sum_{io \in \mathcal{I}_m} A_{io} \tilde{s}_{io|m}^{1-\delta} \tilde{s}_{m|b}^{1-\delta}, \\
\tilde{s}_{m|b} &= \frac{\left[\sum_{j \in \mathcal{I}_m} \left(T_{jo}^{1/\varepsilon_b} A_{jo}\right)^{\frac{\varepsilon_b}{1+\varepsilon_b \delta}}\right]^{\frac{\eta(1+\varepsilon_b \delta)}{\varepsilon_b(1+\eta)}}}{\sum_{m' \in \mathcal{M}_b} \left[\sum_{j' \in \mathcal{I}_{m'}} \left(T_{j'o}^{1/\varepsilon_b} A_{j'o}\right)^{\frac{\varepsilon_b}{1+\varepsilon_b \delta}}\right]^{\frac{\eta(1+\varepsilon_b \delta)}{\varepsilon_b(1+\eta)}}}.
\end{aligned}$$

A_b is an employment weighted industry productivity with the employment shares that would arise with constant labor wedges. Similary, Ω_b is an employment weighted sum of productivities where employment shares incorporate the labor wedge normalized by A_b. The covariance between productivities and employment shares is key in order to determine sector productivity. As long as market power distorts the employment distribution making more productive firms to constrain their size, the covariance between productivity and employment is lower than in the case with constant wedges.

Turning to wages, from (15), the establishment wage bill is:

$$w_{io} L_{io} = \beta_b P_b^{\frac{1}{1-\alpha_b}} \lambda_{io} A_{io} L_{io}^{1-\delta} = \beta_b \lambda_{io} P_b y_{io},$$

where we used the production function (6). The local labor market wage bill is,

$$\begin{aligned}
\sum_{i \in \mathcal{I}_m} w_{io} L_{io} &= \beta_b \sum_{i \in \mathcal{I}_m} \lambda_{io} P_b y_{io} = \beta_b \sum_{i \in \mathcal{I}_m} \lambda_{io} \frac{P_b y_{io}}{P_b Y_m} P_b Y_m = \beta_b \lambda_m P_b Y_m, \\
\lambda_m &\equiv \sum_{i \in \mathcal{I}_m} \lambda_{io} \frac{A_{io}}{\Omega_m A_m} s_{io|m}^{1-\delta},
\end{aligned}$$

where λ_m is a value added weighted sum of establishment labor wedges. Aggregating to the sector,

$$\sum_{m\in\mathcal{M}_b}\sum_{i\in\mathcal{I}_m} w_{io}L_{io} = \beta_b \sum_{m\in\mathcal{M}_b} \lambda_m \frac{P_b Y_m}{P_b Y_b} P_b Y_b = \beta_b \lambda_b P_b Y_b,$$

$$\lambda_b \equiv \sum_{m\in\mathcal{M}_b} \lambda_m \frac{A_m\Omega_m}{\Omega_b A_b} s_{m|b}^{1-\delta} = \sum_{m\in\mathcal{M}_b}\sum_{i\in\mathcal{I}_m} \lambda_{io}\frac{A_{io}}{\Omega_b A_b} s_{io|m}^{1-\delta} s_{m|b}^{1-\delta}.$$

Using the sectoral production function (III1) and the final good production function (2) we have that:

$$Y = \prod_{b\in\mathcal{B}} \left(P_b^{\frac{\alpha_b}{1-\alpha_b}} A_b \Omega_b L_b^{1-\delta} \right)^{\theta_b} = \prod_{b\in\mathcal{B}} P_b^{\frac{\alpha_b\theta_b}{1-\alpha_b}} \prod_{b\in\mathcal{B}} \left(A_b\Omega_b s_b^{1-\delta}\right)^{\theta_b} L^{1-\delta}$$

$$= \prod_{b\in\mathcal{B}} \bar{P}_b^{\frac{\alpha_b\theta_b}{1-\alpha_b}} \prod_{b\in\mathcal{B}} \left[\Omega_b \frac{A_b}{A} s_b^{1-\delta} \left(\frac{P_b}{\bar{P}_b}\right)^{\frac{\alpha_b}{1-\alpha_b}} \right]^{\theta_b} A L^{1-\delta} = \bar{P}\Omega A L^{1-\delta},$$

where:

$$\bar{P} \equiv \prod_{b\in\mathcal{B}} \bar{P}_b^{\frac{\alpha_b\theta_b}{1-\alpha_b}}, \quad \Omega \equiv \prod_{b\in\mathcal{B}} \left[\Omega_b \frac{A_b}{A} s_b^{1-\delta} \left(\frac{P_b}{\bar{P}_b}\right)^{\frac{\alpha_b}{1-\alpha_b}} \right]^{\theta_b}$$

$$A \equiv \sum_{b\in\mathcal{B}}\sum_{m\in\mathcal{M}_b}\sum_{io\in\mathcal{I}_m} A_{io}\tilde{s}_{io|m}^{1-\delta}\tilde{s}_{m|b}^{1-\delta}\tilde{s}_b^{1-\delta} = \sum_{b\in\mathcal{B}} A_b \tilde{s}_b^{1-\delta}$$

$$\tilde{s}_b = \frac{\sum_{m\in\mathcal{M}_b}\left[\sum_{j\in\mathcal{I}_m}\left(T_{jo}^{1/\varepsilon_b} A_{jo}\right)^{\frac{\varepsilon_b}{1+\varepsilon_b\delta}}\right]^{\frac{\eta(1+\varepsilon_b\delta)}{\varepsilon_b(1+\eta)}}}{\sum_{b'\in\mathcal{B}}\sum_{m'\in\mathcal{M}_{b'}}\left[\sum_{j'\in\mathcal{I}_{m'}}\left(T_{j'o}^{1/\varepsilon_{b'}} A_{j'o}\right)^{\frac{\varepsilon_{b'}}{1+\varepsilon_{b'}\delta}}\right]^{\frac{\eta(1+\varepsilon_{b'}\delta)}{\varepsilon_{b'}(1+\eta)}}}.$$

Ω represents an aggregate misallocation measure taking into account general equilibrium effects, $\bar{P}_b$ is the price of sector b good if all the labor wedges in the economy where constant and A is a measure of undistorted productivity.

Aggregate labor share. From the above, the sector labor share is,

$$LS_b = \beta_b\lambda_b, \quad LS = \frac{\sum_{b\in\mathcal{B}}\sum_{m\in\mathcal{M}_b}\sum_{i\in\mathcal{I}_m} w_{io}L_{io}}{\sum_{b\in\mathcal{B}}\sum_{m\in\mathcal{M}_b}\sum_{i\in\mathcal{I}_m} P_b Y_{io}}. \tag{II2}$$

Realizing that industry b expenditure share is equal to θ_b, the aggregate labor share is $LS = \sum_{b\in\mathcal{B}}\beta_b\lambda_b\theta_b$. For given parameters, knowing the industry wedges $\{\lambda_b\}_{b=1}^{B}$ is enough to compute the aggregate labor share.

III Additional estimation results

In Table III1 we provide additional estimation results to complement the ones from the Online Appendix C.6. We compare the baseline estimates of the within local labor market elasticities of substitution to estimates with two-period lagged instruments.

Table III1: Estimated Within Elasticities for Different Lags

Industry Code	Industry Name	1 Lag $\hat{\varepsilon}_b$	2 Lags $\hat{\varepsilon}_b$
15	Food	1.69	1.99
17	Textile	1.49	1.83
18	Clothing	1.41	1.69
19	Leather	2.09	2.50
20	Wood	1.51	1.77
21	Paper	3.06	3.39
22	Printing	1.52	1.79
24	Chemical	3.25	3.56
25	Plastic	2.51	3.04
26	Other Minerals	1.62	1.77
27	Metallurgy	3.77	4.35
28	Metals	1.22	1.48
29	Machines and Equipments	2.18	2.63
30	Office Machinery	3.33	3.72
31	Electrical Equipment	3.02	3.61
32	Telecommunications	3.54	4.08
33	Optical Equipment	1.91	2.36
34	Transport	4.05	4.56
35	Other Transport	3.49	4.05
36	Furniture	1.57	1.90

Notes: All the estimated parameters are 2-digit industry specific. *1 Lag* $\hat{\varepsilon}_b$ are the estimated within local labor market elasticities when we instrument for the wages with one lag and *2 Lags* $\hat{\varepsilon}_b$ present the analogous when we instrument with two lags.

IV Additional counterfactuals

We present additional results of the main counterfactual and other implications of labor market power on urban-rural differences.

IV.1 The effect of labor market power on urban-rural differences

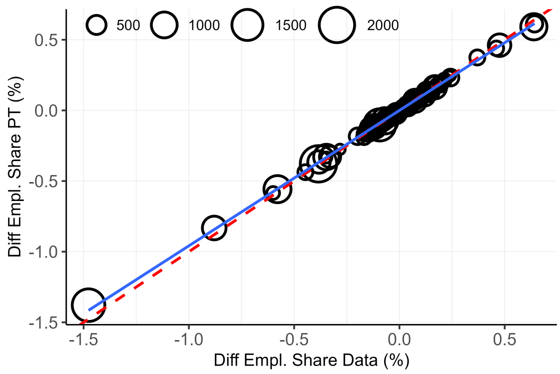
Figure IV2 suggests an important labor reallocation from cities to rural areas in the counterfactual without labor wedges. This section explores the impact of employer and union labor market power on the urban-rural mobility over time. We find that the importance of cities in manufacturing would have declined more slowly in absence of labor market power coming from firms and unions. A potential reason is that the closure of manufacturing establishments in cities would increase the labor concentration of urban areas, making small labor markets relatively more attractive.

Employment changes

We compare the urban-rural manufacturing employment changes over time observed in the data to the ones from yearly counterfactuals without union and firm labor market power. In the data, the de-industrialization or the reduction of manufacturing employment occurred primarily in cities leading to the gain in relative importance of rural areas within manufacturing. Figure IV1 compares the relative employment shares observed in the data to the one in a counterfactual without labor wedges for each commuting zone.

First, we perform the counterfactual where there are no labor wedges because establishments and unions act as price takers (PT) for the initial year 1994. Then we compute the commuting zone employment share out of total manufacturing employment for the initial and final years (1994

Figure IV1: De-industrialization differences



Notes: The x-axis shows the percentage differences of commuting zone employment shares out of manufacturing over time in the data ($\Delta^D = S^D_{07} - S^D_{94}$). The y-axis presents the analogous for the counterfactual without wedges ($\Delta^M = S^{PT}_{07} - S^{PT}_{94}$). The first year is 1994 and the last one is 2007. The bubble size represents the level of employment in thousands at the commuting zone for the first year. The blue line represents a fitted line from an OLS regression. A weighted least squares regression using initial employment as weights gives a similar result.

and 2007 respectively) and for the different scenarios. To compare mobility over time, we compute the differences over time of the commuting zone employment shares in the data ($\Delta^D = S^D_{07} - S^D_{94}$) and in the counterfactual ($\Delta^M = S^{PT}_{07} - S^{PT}_{94}$). Figure IV1 in presents this comparison. The x axis shows the time difference in the data Δ^D and the y axis shows the time difference in the model counterfactual without labor wedges Δ^M. The size of the dots is the initial level of manufacturing employment of the commuting zone. The counterfactual urban-rural mobility is very similar to the process observed in the data which is mostly guided by exogenous productivity and firm location decisions and not by labor market distortions.

The line generated by the largest population commuting zones in Figure IV1 is slightly flatter than the 45 degree line. Cities would loose their relative importance a bit slower in the counterfactual. A potential reason is the closure of manufacturing firms in the largest cities, which became more concentrated over time leading to distortions closer to the ones present in rural areas.

IV.2 Extensions

Table IV1 presents counterfactuals with agglomeration externalities under perfect competition where wages are equal to the marginal revenue product of labor. Baseline gains (1.62%) are amplified with agglomeration due to the productivity gains.

Figure IV3 shows productivity changes in the counterfactual with oligopsonistic competition relative to the baseline. The map shows that the biggest productivity losses happen outside big cities and some commuting zones increase overall productivity due to labor mobility across sectors.

V Additional empirical evidence

We present an additional robustness exercise related to the reduced form regression in Section 3 in the main text, where we change the definition of the local labor markets.

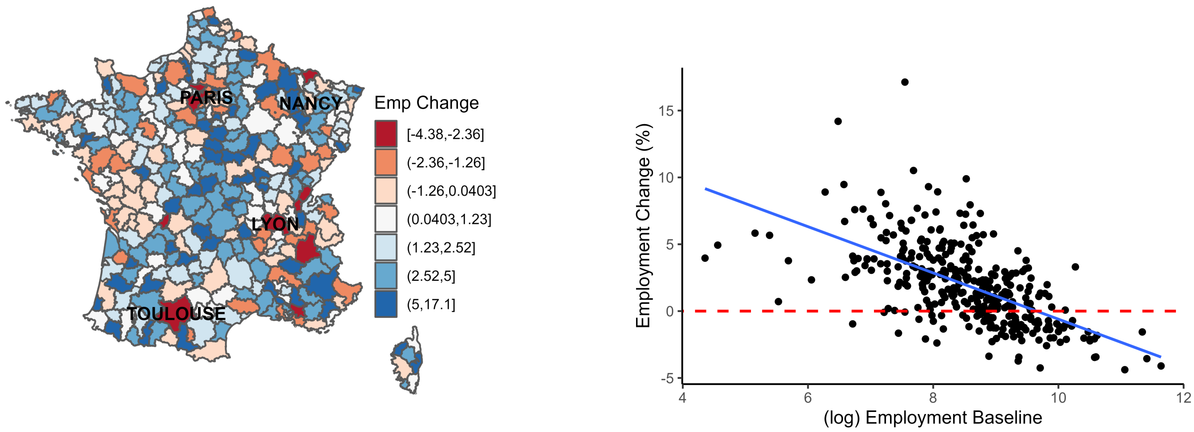
Local labor market. Figure V1 does the same exercise as in the main empirical strategy but chang-

Table IV1: Counterfactuals: Agglomeration. Perfect Competition

	ΔY (%)	Δ Prod (%)	Contribution ΔY (%) GE	Productivity	Labor
No Agglomeration	1.62	1.33	9	83	8
Agglomeration					
$\gamma = 0.05$	1.73	1.40	8	82	10
$\gamma = 0.1$	1.84	1.48	7	81	12
$\gamma = 0.2$	2.08	1.66	5	80	15
$\gamma = 0.25$	2.22	1.75	3	80	17
$\gamma = 0.3$	2.36	1.86	2	80	18

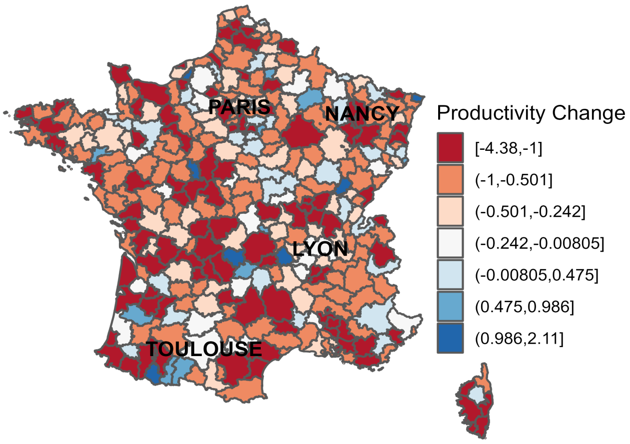
Notes: Results are in percentages. First column ΔY is the change of aggregate output with respect to the baseline, Δ *Prod* is the change in aggregate productivity from decomposition (28). Last three columns present the contribution of each of the elements of the decomposition (28) to output gains. *No Agglomeration* is the main counterfactual without wedges, under free mobility of labor, fixed total labor supply and no agglomeration forces. All the other counterfactuals in this table allow for agglomeration within the local labor market. Similarly to the main counterfactual, workers are freely mobile and total employment is fixed. We present different counterfactuals depending on the agglomeration parameter γ.

Figure IV2: Employment Change (%) in the Counterfactual: Perfect Competition



Notes: The map presents employment changes with respect to the baseline economy in percentages within commuting zones. The counterfactual without labor wedges is performed for the year 2007. The figure in the right plots the employment change in the counterfactual versus the log of employment in the baseline. The blue line is a fitted line from an OLS regression.

Figure IV3: Productivity Change (%) in the Counterfactual: Oligopsonistic Competition



Notes: The map presents productivity changes with respect to the baseline economy in percentages. Each block constitutes a commuting zone. Local labor markets are aggregated up to the commuting zone. Commuting zone productivity is an employment weighted average of individual productivities. Following the discussion in Section A.4, keeping fixed the baseline revenue productivities, any change in the counterfactual comes from changes in aggregate productivities from the reallocation of workers. Counterfactuals are performed for the year 2007.

ing the definition of local labor market. Local labor markets are here defined with 2-digit industries instead of 3-digit industries.[1] The specification includes commuting zone fixed effects as in Figure E1 Panel (b).

VI Data details

In this section we provide details about sample selection and variable construction.

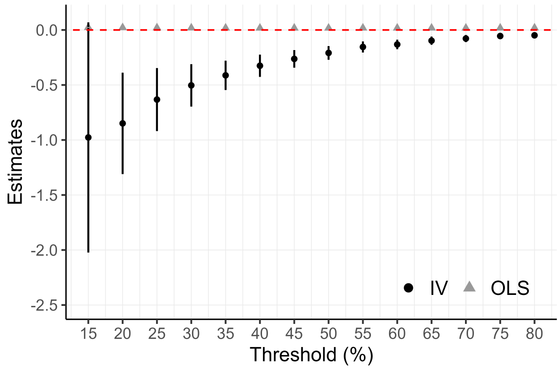
VI.1 Sample selection

Ficus/Fare. This data source comes from tax records therefore we observe yearly firm information. We exclude the source tables belonging to public firms.[2] Before 2000 we take table sources in euros and from 2001 onward we use data from consolidated economic units.[3] After excluding firms without a firm identifier, the raw data sample contains about 29 million firms, of which about 2.8 million are manufacturing firms.[4] Manufacturing sector (sector code equal to *D*) constitutes on average 10% of the observations, 19.2% of value added and 27.2% of employment.

Postes. *DADS Postes* covers all the employment spells of a salaried employee per year. If a worker has several spells in a year we would have multiple observations. The main benefit of this employer-employee data source is that we can know the establishment and employment location of the workers. We exclude workers in establishments with fictitious identifiers (SIREN starting by F) and in

[1] That is, a local labor market is defined as a combination between commuting zone, 2-digit industry and occupation.
[2] We only use the Financial units (*FIN*) and Other units (*TAB*) tables and exclude Public administration (*APU*).
[3] The profiling of big groups consolidates legal units into economic units. In 2001 the Peugeot-Citroën PSA was treated, Renault in 2003 and the group Accor in 2005. This implies the definition of new economic entities and would therefore lead to the creation of new firm identifiers. Given the potential impact of big establishments in local labor markets we opted to maintain them.
[4] We consider a missing firm identifier (SIREN) also if the identifier equals to zero for all the 9 digits.

Figure V1: Robustness. Local Labor Market at 2-digit Industry



Notes: This figure presents the point estimates and 95% confidence bands of the OLS and IV exercises on the y-axis. The x-axis presents different thresholds κ that define a mass layoff shock. We focus on non-affected competitors (not suffering a mass layoff shock). The instrument is the presence of a mass layoff shock firm in the local labor market. The definition of local labor market is a combination of commuting zone, 2-digit industry and occupation. The difference with respect to the figure in the main text is that the local labor market is at 2-digit rather than 3-digit industry.

public firms. For every establishment identifier (SIRET) we sum the wage bill and the full time equivalent number of employees.

Merged data. After merging both data sources, we end up with data that include yearly establishment observations. After the filters and merging the sample consists of 1.3 million firms and 1.6 million establishment observations. In the process of filtering and merging, about half of the original firms are lost. Wages and value added are deflated using the Consumer Price Index.[5]

Labor and wage data, coming from the balance sheets (at the firm level) and the one from employee records, needs to be consolidated. In order to be consistent with balance sheet information we assign labor and employment coming from *FICUS* to the establishments according to their respective shares. We proceed in several steps. First, we filter out observations with no wage or employment information from *Postes* from firms present at different commuting zones. Second, we get rid of observations with no labor, capital and wage bill information coming from *FICUS* and also observations with non existing or missing commuting zone. Third, we aggregate employee data to the firm times commuting zone level.[6] What we call establishment throughout the text is the entity aggregated at the commuting zone level. Then we compute the labor and wage shares of these entities out of the firm's aggregates. Finally, we split firm data from the balance sheet according to those shares. This procedure leaves the firms in a unique commuting zone with their balance sheet data but allows to split wage bill and employment data coming from the balance sheet for multi-location firms. Establishment wage is simply the average wage. That is, wage bill over total full time equivalent employees.

We further exclude Tobacco (2-digits industry code 16), Refineries & Nuclear industry (code 23) and Recycling (code 37). We finally get rid of the outliers reducing the sample 1.5% and finish with

[5]Nominal variables are expressed in constant 2015 euros.

[6]Data from 1994 and 1995 do not have commuting zone information. We therefore impute it using correspondence tables between city code and commuting zone. A city code has 5 digits coming from the department and city. Some commuting zone codes beyond the 2 missing years were modified or cleaned. City codes (*commune* codes) of Paris, Marseille and Lyon were divided into different *arrondissements*. We assign them codes 75056, 13055 and 69123 respectively. Then we proceed to the cleaning of commuting zones by assigning to the non existing codes the one corresponding to the city where the establishment is located. We get rid of non matched or missing commuting zone codes. We aggregate the data coming from *Postes* at the commuting zone level after this cleaning.

4,156,754 establishment-occupation-year observations that belong to 1.25 million firms.[7]

VI.2 Variable construction

Ficus:

- Value added: value added net of taxes (*VACBF*). We restrict to firms with strictly positive value added.[8]
- Capital: tangible and intangible capital without counting depreciation. It is the sum of the variables *IMMOCOR* and *IMMOINC*.
- Employment: full time equivalent employment at the firm (*EFFSALM*).
- Wage bill: gross total wage bills. Is the sum of wages (*SALTRAI*) and firm taxed (*CHARSOC*).[9]
- Industry: industry classification comes from *APE*. The sub-industries h are 3 digit industries and industries b are at two digits.

Postes:

- Occupation: original occupation categories come from the two digit occupations (*CS2*). We group occupations with first digits 2 and 3 into a unique occupation group.[10] This regrouping is done to avoid substantial changes in occupation groups between 1994 and 2007. Observations with missing occupation information are excluded.
- Employment: full time equivalent at the establishment-occupation level (*etp*).
- Wage: is the gross wage (per year) of individual worker (*sbrut*). If the spell is less than a year is the gross wage corresponding to the spell.
- Commuting zone: depending on the year, the commuting zone classification is taken from the variable *zemp* or *zempt*. Commuting zone information is missing for the years 1994 and 1995 and is imputed using the city codes.[11]

VI.3 Construction of required rates

In order to construct the required rates for the different sectors we follow the methodology proposed by Barkai (2020) using the Capital Input Data from the EU KLEMS database, December 2016 revision. In this dataset one can find, for a given industry, different depreciation rates and price indices for different types of capital. The types of capital that are present in the manufacturing sector are: Computing Equipment, Communications Equipment, Computer Software and Databases, Transport Equipment, Buildings and structures (non-residential), and Research and Development. We construct a required rate for each of the industries at the 2 digit level defined in the NAF classification. However, unlike the NAF classification, that we have up to 20 different industries,

[7] We get rid of wage per capita outliers by truncating the sample at the 0.5% below and 99.5%.
[8] We follow the advise of the French statistical instiute (INSEEE) in using net value added to perform comparisons across industries.
[9] For firms declaring at the BIC-BRN regime (*TYPIMPO*= 1) we only take *SALTRAI* .
[10] Occupations with first digit 1 and 7 are excluded. They constituted less than 0.05% of the matched sample.
[11] City codes are the concatenation of department (*DEP*) and city (*COM*).

there are only 11 industries classified within manufacturing within the EU KLEMS database. Any industry classification in EU KLEMS is just an aggregation of the 2 digit industry classification in NAF. Therefore, there are industries within the NAF classification that will have the same required rate of return on capital.

For a type of capital s and sector b, we define the the required rate of return R_{sb} as:

$$R_{sb} = \left(i^D - \mathbb{E}\left[\pi_{sb}\right] + \delta_{sb}\right),$$

where i^D is a the cost fo debt borrowing in financial markets, and π_{sb} and δ_{sb} are, respectively, the inflation and depreciation rates of capital type s in sector b.

Then we define the total expenditures on capital type s in sector b as:

$$E_{sb} = R_{sb} P^K_{sb} K_{sb},$$

where $P^K_{sb} K_{sb}$ is the nominal value of capital stock of type s. Summing over all types of capital within a sector we can obtain the total expenditures of capital of such sector:

$$E_b = \sum_{sb} R_{sb} P^K_{sb} K_{sb}.$$

Multiplying and dividing by the total nominal value of capital stock we obtain:

$$\sum_s R_{sb} P^K_{sb} K_{sb} = \underbrace{\sum_s \frac{P^K_{sb} K_{sb}}{\sum_{s'} P^K_{s'b} K_{s'b}} R_{sb}}_{R_b} \underbrace{\sum_s P^K_{sb} K_{sb}}_{P^{Kb} K_b},$$

where the first term R_b is the interest rate that we use in the model.

VII Summary statistics

Tables VII1, VII2 and VII3 contain summary statistics of sub-industries, local labor markets and commuting zones for the year 2007, which is the year we use for our counterfactuals. Table VII4 presents worker transition probabilities across occupations, industries and commuting zones.

Table VII1: Sub-industry Summary Statistics. Baseline Year

Variable	Mean	Pctl(25)	Median	Pctl(75)	St. Dev.
N_h	2,840	493	1,261	2,639	4,530.5
L_h	30,466	7,559	15,070	50,036	33,899.3
$\overline{w}_h$	34.6	29.6	33.0	37.531	6.9
LS_h	0.52	0.48	0.53	0.58	0.10
KS_h	0.26	0.17	0.23	0.32	0.13

Notes: There are 97 3-digit industries, or sub-industries, in the sample. N_h is the number of establishments per 3-digit industry h, L_h is total employment of h, $\overline{w}_h$ is the average establishment wage of h, LS_h is the labor share and KS_h is the capital share. We get the capital shares following Barkai (2020). All the nominal variables are in thousands of constant 2015 euros.

Table VII2: Local Labor Market Summary Statistics. Baseline Year

Variable	Mean	Pctl(25)	Median	Pctl(75)	St. Dev.
N_m	4.76	1	2	4	14.4
L_m	51.0	2.8	9.4	34.9	196.2
$\overline{w}_m$	36.6	24.3	30.2	42.5	36.1
$\widehat{w}_m$	36.2	24.1	30.0	42.2	25.6
$\text{HHI}(s_{io\|m})$	0.67	0.38	0.68	1.00	0.32
$\text{HHI}(s^w_{io\|m})$	0.68	0.39	0.70	1.00	0.32

Notes: There are 57,940 local labor markets in the year 2007. N_m is the number of competitors in the local labor market m, L_m is total employment in m, $\overline{w}_m$ is the mean w_{iot} of the establishment-occupations in m, $\widehat{w}_m$ is the weighted average wage at m with weights equal to employment shares, $\text{HHI}(s_{io|m})$ and $\text{HHI}(s^w_{io|m})$ are respectively the Herfindahls with employment and wage shares. All the nominal variables are in thousands of constant 2015 euros.

Table VII3: Commuting Zones Summary Statistics. Baseline Year

Variable	Mean	Pctl(25)	Median	Pctl(75)	St. Dev.
N_n	773.798	266.8	461	861.2	1,168.407
L_n	8,300.567	2,567.403	5,244.300	10,086.210	11,322.000
$\overline{L}_n$	11.389	8.148	10.878	13.547	6.043
$\overline{w}_n$	34.399	32.707	34.161	35.593	3.242

Notes: There are 356 commuting zones in the sample. N_n is the number of establishments at the CZ, L_n is full time equivalent employment at CZ, $\overline{L}_n$ is the average L_{iot} of establishment-occupations at n, $\overline{w}_n$ is the mean w_{iot} at n in thousands of constant 2015 euros.

Table VII4: Transition Probabilities

Occupation	Commuting Zone	Industry	Trans. Prob. FTE	Trans. Prob.
0	0	0	91.39	91.01
0	0	1	2.37	2.36
0	1	0	0.02	0.02
1	0	0	6.03	6.40
1	0	1	0.20	0.21
1	1	0	0.00	0.00
1	1	1	0.00	0.00

Notes: The transition rates are computed over the whole sample period 1994-2007. *Occupation* is an indicator function of occupational change, *Commuting Zone* is an indicator function of commuting zone change, *Industry* is an indicator function of 3-digit industry change, *Trans. Prob. FTE* are the unconditional transition probabilities based on full time equivalent units and *Trans. Prob.* are the unconditional transition probabilities based on counts of working spells independently of duration and part-time status.

References

Allen, T., C. Arkolakis, and X. Li (2016): "On the Existence and Uniqueness of Trade Equilibria," *Unpublished manuscript*.

——— (2021): "On the Equilibrium Properties of Network Models with Heterogeneous Agents," Tech. rep., National Bureau of Economic Research.

Barkai, S. (2020): "Declining Labor and Capital Shares," *The Journal of Finance*, 75, 2421–2463, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12909.

Berger, D., K. Herkenhoff, and S. Mongey (2022): "Labor Market Power," *American Economic Review*, 112, 1147–93.