Econometrica Supplementary Material

SUPPLEMENT TO "WALRAS–BOWLEY LECTURE: MARKET POWER AND WAGE INEQUALITY"

(*Econometrica*, Vol. 92, No. 3, May 2024, 603–636)

SHUBHDEEP DEB
Department of Economics, UPF Barcelona

JAN EECKHOUT
UPF Barcelona and ICREA-BSE-CREI

ASEEM PATEL
Department of Economics, University of Essex

LAWRENCE WARREN
US Census Bureau

APPENDIX A: DERIVATIONS

A.1. *Household's Optimization*

Optimum Consumption Functions

REPRESENTATIVE HOUSEHOLD MAXIMIZES THE FOLLOWING utility function subject to the budget constraint:

$$\max_{C_{inj},L_{inj},H_{inj}} C - \frac{1}{\bar{\phi}_L^{\frac{1}{\phi_L}}} \frac{L^{\frac{\phi_L+1}{\phi_L}}}{\frac{\phi_L+1}{\phi_L}} - \frac{1}{\bar{\phi}_H^{\frac{1}{\phi_H}}} \frac{H^{\frac{\phi_H+1}{\phi_H}}}{\frac{\phi_H+1}{\phi_H}}, \quad \text{s.t.} \quad PC = LW_L + HW_H + \Pi.$$

We solve the problem in two steps. First, we derive the household's market-level demand function and then we derive the establishment-level demand function. The solution to household's market-level demand function is a solution to

$$\max_{Y_j} \left(\int_j J^{\frac{-1}{\theta}} Y_j^{\frac{\theta-1}{\theta}} \, dj \right)^{\frac{\theta}{\theta-1}}, \quad \text{s.t.} \quad \int_J P_j Y_j \, dj \le Z. \tag{A1}$$

Then the optimal allocation is given by

$$\frac{\theta}{\theta-1} \left(\int_j J^{\frac{-1}{\theta}} Y_j^{\frac{\theta-1}{\theta}} \, dj \right)^{\frac{\theta}{\theta-1}-1} J^{\frac{-1}{\theta}} \frac{\theta-1}{\theta} Y_j^{\frac{\theta-1}{\theta}-1} = \lambda P_j. \tag{A2}$$

This can be simplified as $J^{\frac{-1}{\theta}} Y^{\frac{1}{\theta}} Y_j^{\frac{-1}{\theta}} = \lambda P_j$. Next, multiply each side by Y_j and integrate across J to get $Y = \lambda \int_j P_j Y_j \, dj$. We define the market price index P such that

Shubhdeep Deb: shubhdeep.deb@upf.edu
Jan Eeckhout: jan.eeckhout@upf.edu
Aseem Patel: aseem.patel@essex.ac.uk
Lawrence Warren: lawrence.fujio.warren@census.gov

$PY = \int_j P_j Y_j \, dj$, which would imply that $\lambda = P^{-1}$. Then, plugging this into the first-order condition delivers the market-specific demand function:

$$Y_j = \left(\frac{1}{J}\right)\left(\frac{P_j}{P}\right)^{-\theta} Y. \tag{A3}$$

The aggregate price index can be recovered by multiplying both sides by P_j and integrating across markets:

$$P = \left[\frac{1}{J}\int_J P_j^{1-\theta}\, dj\right]^{\frac{1}{1-\theta}}. \tag{A4}$$

We can apply a similar formulation to derive the establishment-specific demand function, $Y_{inj} = \frac{1}{I}(\frac{P_{inj}}{P_j})^{-\eta} Y_j$, and the market price index, $P_j = (\frac{1}{I}\sum_i P_{inj}^{1-\eta})^{\frac{1}{1-\eta}}$. Then, the establishment-specific demand function is given by

$$Y_{inj} = \frac{1}{J}\frac{1}{I}\left(\frac{P_{inj}}{P_j}\right)^{-\eta}\left(\frac{P_j}{P}\right)^{-\theta} Y. \tag{A5}$$

To derive the market-specific inverse demand function, we can write $P_j = J^{-\frac{1}{\theta}}(\frac{Y_j}{Y})^{-\frac{1}{\theta}} P$, and similarly at the establishment level as $P_{inj} = I^{-\frac{1}{\eta}}(\frac{Y_{inj}}{Y_j})^{-\frac{1}{\eta}} P_j$. Combining the last two equations, we can get the establishment-specific inverse demand curve as

$$P_{inj} = \left(\frac{1}{J}\right)^{\frac{1}{\theta}}\left(\frac{1}{I}\right)^{\frac{1}{\eta}} Y_{inj}^{-\frac{1}{\eta}} Y_j^{\frac{1}{\eta}-\frac{1}{\theta}} Y^{\frac{1}{\theta}} P. \tag{A6}$$

Optimum Labor Supply Functions

To derive equation (6), we follow Berger, Herkenhoff, and Mongey (2022) and adjust for the love for variety by scaling the utility function. The household's aggregate labor supply function for each skill $S \in \{H, L\}$ can be derived from

$$\max_S C - \frac{1}{\bar{\phi}_L^{\frac{1}{\phi_L}}}\frac{L^{\frac{\phi_L+1}{\phi_L}}}{\frac{\phi_L+1}{\phi_L}} - \frac{1}{\bar{\phi}_H^{\frac{1}{\phi_H}}}\frac{H^{\frac{\phi_H+1}{\phi_H}}}{\frac{\phi_H+1}{\phi_H}}, \quad \text{s.t.} \quad PC = LW_L + HW_H + \Pi.$$

Then, the first-order condition for $S \in \{H, L\}$ is

$$\frac{W_S}{P} = \overline{\phi}_S^{-\frac{1}{\phi_S}} S^{\frac{1}{\phi_S}} \quad \iff \quad S = \overline{\phi}_S\left(\frac{W_S}{P}\right)^{\phi_S},$$

which gives the aggregate labor supply function. The household's optimum choice of allocation of labor across markets can be written as the solution to

$$\min_{S_j}\left[\int_j \left(\frac{1}{J}\right)^{\frac{-1}{\hat{\theta}_S}} S_j^{\frac{\hat{\theta}_S+1}{\hat{\theta}_S}}\, dj\right]^{\frac{\hat{\theta}_S}{\hat{\theta}_S+1}}, \quad \text{s.t.} \quad \int_J W_{Sj} S_j \, dj \geq Z. \tag{A7}$$

Then, the optimal allocation is given by

$$\frac{\hat{\theta}_S}{\hat{\theta}_S+1}\left(\int_j\left(\frac{1}{J}\right)^{\frac{-1}{\hat{\theta}_S}} S_j^{\frac{\hat{\theta}_S+1}{\hat{\theta}_S}}\,dj\right)^{\frac{\hat{\theta}_S}{\hat{\theta}_S+1}-1}\left(\frac{1}{J}\right)^{\frac{-1}{\hat{\theta}_S}}\frac{\hat{\theta}_S+1}{\hat{\theta}_S} S_j^{\frac{\hat{\theta}_S+1}{\hat{\theta}_S}-1} = \lambda W_{Sj}. \tag{A8}$$

This can be simplified as $\frac{1}{J}^{\frac{-1}{\hat{\theta}_S}} S^{\frac{-1}{\hat{\theta}_S}} S_j^{\frac{1}{\hat{\theta}_S}} = \lambda W_{Sj}$. Next, multiply each side by S_j and integrate across J to get $S = \lambda \int_j W_{Sj} S_j\, dj$. We define the aggregate wage index W such that $WS = \int_j W_j S_j\, dj$, which would imply that $\lambda = W^{-1}$. Then, plugging this into the first-order condition delivers the market-specific labor supply equation as a function of wage levels and aggregate labor supply:

$$S_j = \left(\frac{1}{J}\right)\left(\frac{W_{Sj}}{W_S}\right)^{\hat{\theta}_S} S. \tag{A9}$$

The aggregate wage index can be recovered by multiplying both sides by W_j and integrating across markets:

$$W_S = \left[\frac{1}{J}\int_J W_{Sj}^{1+\hat{\theta}_S}\,dj\right]^{\frac{1}{1+\hat{\theta}_S}}. \tag{A10}$$

We can apply a similar formulation to derive the establishment-level labor supply, $S_{inj} = (\frac{1}{I})(\frac{W_{Sinj}}{W_{Sj}})^{\hat{\eta}_S} S_j$, and the market-specific wage index is $W_{Sj} = [(\frac{1}{I})\sum_i W_{Sinj}^{1+\hat{\eta}_S}]^{\frac{1}{1+\hat{\eta}_S}}$. Then, the establishment-level labor supply curve is given by

$$S_{inj} = \left(\frac{1}{J}\right)\left(\frac{1}{I}\right)\left(\frac{W_{Sinj}}{W_{Sj}}\right)^{\hat{\eta}_S}\left(\frac{W_{Sj}}{W_S}\right)^{\hat{\theta}_S} S. \tag{A11}$$

To derive the market-specific inverse labor supply function, write $W_{Sj} = (\frac{1}{J})^{-\frac{1}{\hat{\theta}_S}}(\frac{S_j}{S})^{\frac{1}{\hat{\theta}_S}} W_S$, and similarly at the establishment level as $W_{inj} = (\frac{1}{I})^{-\frac{1}{\hat{\eta}_S}}(\frac{S_{inj}}{S_j})^{\frac{1}{\hat{\eta}_S}} W_{Sj}$. Combining these two equations, we can get the establishment-level inverse labor supply curve as

$$W_{Sinj} = \left(\frac{1}{J}\right)^{-\frac{1}{\hat{\theta}_S}}\left(\frac{1}{I}\right)^{-\frac{1}{\hat{\eta}_S}} S_{inj}^{\frac{1}{\hat{\eta}_S}} S_j^{\frac{1}{\hat{\theta}_S}-\frac{1}{\hat{\eta}_S}} S^{-\frac{1}{\hat{\theta}_S}} W_S. \tag{A12}$$

A.2. *Solving the Equilibrium*

Optimal Firm Solution

There are N firms indexed by n in each market. A firm owns I/N establishments. An establishment's sales share and wage bill share are denoted by s_{inj} and e_{Linj}, e_{Hinj}, respectively. As a result, the firm's sales share and wage bill share can be expressed as $s_{nj} = \sum_{i\in\mathcal{I}_{nj}} s_{inj}$ and $e_{Lnj} = \sum_{i\in\mathcal{I}_{nj}} e_{Linj}$ for the low-skilled and $e_{Hnj} = \sum_{i\in\mathcal{I}_{nj}} e_{Hinj}$ for the high-skilled, respectively. Firm's problem here is to choose an employment level L_{inj}, H_{inj} for each establishment i simultaneously to maximize its profit. The FOC for input L_{inj} is derived below:

$$\left[P_{inj} + \frac{\partial P_{inj}}{\partial Y_{inj}} Y_{inj} + \sum_{i' \in \mathcal{I}_{nj} \setminus i} \left(\frac{\partial P_{i'nj}}{\partial Y_{inj}} Y_{i'nj} \right)\right] \frac{\partial Y_{inj}}{\partial L_{inj}}$$
$$= \left[W_{Linj} + \frac{\partial W_{Linj}}{\partial L_{inj}} L_{inj} + \sum_{i' \in \mathcal{I}_{nj} \setminus i} \left(\frac{\partial W_{Li'nj}}{\partial L_{inj}} L_{i'nj} \right)\right]. \tag{A13}$$

Note that $\frac{\partial P_{inj}}{\partial Y_{inj}} Y_{inj} = [-1/\eta + (1/\eta - 1/\theta)s_{inj}]P_{inj}$, and

$$\begin{aligned} \frac{\partial P_{i'nj}}{\partial Y_{inj}} Y_{i'nj} &= \frac{\partial P_{i'nj}/P_{i'nj}}{\partial Y_{inj}/Y_{inj}} \frac{P_{i'nj} Y_{i'nj}}{P_{inj} Y_{inj}} P_{inj} \\ &= \frac{\partial \log P_{i'nj}}{\partial \log Y_{inj}} \frac{s_{i'nj}}{s_{inj}} P_{inj} \\ &= \left[\left(\frac{1}{\eta} - \frac{1}{\theta}\right) s_{inj}\right] \frac{s_{i'nj}}{s_{inj}} P_{inj} \\ &= \left(\frac{1}{\eta} - \frac{1}{\theta}\right) s_{i'nj} P_{inj}, \end{aligned} \tag{A14}$$

and similarly, $\frac{\partial W_{Linj}}{\partial L_{inj}} L_{inj} = [1/\hat{\eta}_L + (1/\hat{\theta}_L - 1/\hat{\eta}_L)e_{Linj}]W_{Linj}$, and

$$\frac{\partial W_{Li'nj}}{\partial L_{inj}} L_{i'nj} = \left(\frac{1}{\hat{\theta}_L} - \frac{1}{\hat{\eta}_L}\right) e_{Li'nj} W_{Linj}. \tag{A15}$$

Combining these, the FOC can be rewritten into

$$\left[1 - \frac{1}{\theta} s_{nj} - \frac{1}{\eta}(1 - s_{nj})\right] P_{inj} \frac{\partial Y_{inj}}{\partial L_{inj}} = \left[1 + \frac{1}{\hat{\theta}_L} e_{Lnj} + \frac{1}{\hat{\eta}_L}(1 - e_{Lnj})\right] W_{Linj}, \tag{A16}$$

where markup and markdown are defined as

$$\begin{aligned} \mu_{inj} &= \frac{1}{1 + \varepsilon^P_{inj}} = \left[1 - \frac{1}{\theta} s_{nj} - \frac{1}{\eta}(1 - s_{nj})\right]^{-1}, \\ \delta_{Linj} &= 1 + \varepsilon^L_{inj} = \left[1 + \frac{1}{\hat{\theta}_L} e_{Lnj} + \frac{1}{\hat{\eta}_L}(1 - e_{Lnj})\right], \end{aligned} \tag{A17}$$

$$\delta_{Hinj} = 1 + \varepsilon^H_{inj} = \left[1 + \frac{1}{\hat{\theta}_H} e_{Hnj} + \frac{1}{\hat{\eta}_H}(1 - e_{Hnj})\right]. \tag{A18}$$

Solving the Model

Start from the first-order condition for a low-skilled worker:

$$Y_{inj}^{\frac{1}{\sigma}} A_{Linj}^{\frac{\sigma-1}{\sigma}} L_{inj}^{-\frac{1}{\sigma}} \left[1 - \frac{1}{\theta} s_{nj} - \frac{1}{\eta}(1 - s_{nj})\right] P_{inj} = \left[1 + \frac{1}{\hat{\theta}_L} e_{L,nj} + \frac{1}{\hat{\eta}_L}(1 - e_{Lnj})\right] W_{Linj}. \tag{A19}$$

We have a similar equation for a high-skilled worker:

$$
\begin{aligned}
&Y_{inj}^{\frac{1}{\sigma}} A_{Hinj}^{\frac{\sigma-1}{\sigma}} H_{inj}^{-\frac{1}{\sigma}} \left[1 - \frac{1}{\theta} s_{nj} - \frac{1}{\eta}(1 - s_{nj})\right] P_{inj} \\
&\quad = \left[1 + \frac{1}{\hat{\theta}_H} e_{Hnj} + \frac{1}{\hat{\eta}_H}(1 - e_{Hnj})\right] W_{Hinj}.
\end{aligned}
\tag{A20}
$$

By plugging into the inverse labor supply and inverse demand functions, we can rewrite each of these two conditions into

$$
\begin{aligned}
&\frac{1}{J}^{\frac{1}{\theta}} \frac{1}{I}^{\frac{1}{\eta}} (Y_{inj})^{-\frac{1}{\eta}} \left[\left(\frac{1}{I}^{\frac{1}{\eta}} \sum_i (Y_{inj})^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1} \frac{(\theta-\eta)}{\eta\theta}}\right] \\
&\quad \times \left[1 - \frac{1}{\theta} \frac{\sum_{i\in\mathcal{I}_{nj}} (Y_{inj})^{\frac{\eta-1}{\eta}}}{\sum_i (Y_{inj})^{\frac{\eta-1}{\hat{\eta}}}} - \frac{1}{\eta}\left(1 - \frac{\sum_{i\in\mathcal{I}_{nj}} (Y_{inj})^{\frac{\eta-1}{\eta}}}{\sum_i (Y_{inj})^{\frac{\eta-1}{\eta}}}\right)\right] \frac{\partial Y_{inj}}{\partial S_{inj}} Z_S \\
&= \frac{1}{J}^{\frac{-1}{\hat{\theta}_S}} \frac{1}{I}^{\frac{-1}{\hat{\eta}_S}} (S_{inj})^{\frac{1}{\hat{\eta}_S}} \left[\left(\frac{1}{I}^{\frac{-1}{\hat{\eta}_S}} \sum_i (S_{inj})^{\frac{\hat{\eta}_S+1}{\hat{\eta}_S}}\right)^{\frac{\hat{\eta}_S}{\hat{\eta}_S+1} \frac{(\hat{\eta}_S-\hat{\theta}_S)}{\hat{\eta}_S\hat{\theta}_S}}\right] \\
&\quad \times \left[1 + \frac{1}{\hat{\theta}_S} \frac{\sum_{i\in\mathcal{I}_{nj}} (S_{inj})^{\frac{\hat{\eta}_S+1}{\hat{\eta}_S}}}{\sum_i (S_{inj})^{\frac{\hat{\eta}_S+1}{\hat{\eta}_S}}} + \frac{1}{\hat{\eta}_S}\left(1 - \frac{\sum_{i\in\mathcal{I}_{nj}} (S_{inj})^{\frac{\hat{\eta}_S+1}{\hat{\eta}_S}}}{\sum_i (S_{inj})^{\frac{\hat{\eta}_S+1}{\hat{\eta}_S}}}\right)\right],
\end{aligned}
\tag{A21}
$$

where $S \in \{H, L\}$, $Z_S = W_S^{-1} S^{1/\hat{\theta}_S} Y^{1/\theta}$ is the skill-specific aggregate and the aggregate price P is normalized to 1. Finally, we replace Y_{inj} in the above expression by the production function, which gives us two first-order conditions that are functions of H_{inj} and L_{inj}. We use these two equations to solve the model computationally using the following algorithm.

A.3. *Algorithm to Solve the Model*

Given model primitives outlined in Table I, we proceed to compute the equilibrium of our economy using the following algorithm:

1. Guess three aggregates: $\{W_H^k, W_L^k, Y^k\}$, where k is the index of iteration.
2. Given those three initial values, solve the $2 \times I$ first-order conditions, market-by-market, and calculate H_{inj}, L_{inj}, and Y_{inj} for each establishment.
3. Compute $W_{H,inj}$, $W_{L,inj}$, and P_{inj} for each establishment using the inverse labor supply function for each skill and inverse demand function. Then, aggregate the establishment wages W_{Hinj}, W_{Linj} into W_H^{k+1}, W_L^{k+1} and establishment output Y_{inj} to Y^{k+1} using the respective CES aggregators.
4. Update the initial guess and iterate until all three aggregates converge $W_H^{k+1} = W_H^k$, $W_L^{k+1} = W_L^k$, and $Y^{k+1} = Y^k$ to get the equilibrium aggregates W_H^*, W_L^*, and Y^*.

A.4. *Algorithm to Back out Technology Shocks*

In order to back out the A_{Hinj} and A_{Linj} from the microdata, we proceed as follows:

1. Given that we can express the two first-order conditions for each establishment only as a function of $A_{Sinj}, S_{inj}, \forall i \in j$ in equation (A21) of Section (A.2), we begin by solving for $Z_S = W_S^{-1} S^{1/\hat{\theta}_S} Y^{1/\theta}$. We first use the aggregate labor supply function to substitute out W_S as a function of S using $W_S = \frac{S^{1/\varphi}}{\overline{\varphi}_S}$.[1]
2. Given our estimation of the labor supply function from Steps 1 and 2 in Section 4, we have estimates of $\hat{\eta}_S$, $\hat{\theta}_S$, $\overline{\varphi}_S$. Now $Z_S = S^{1/\hat{\theta}-1/\varphi} Y^{1/\theta} \overline{\varphi}_S$, where we only need to solve for Y. To do so, we use a two-step procedure.
   (a) **Step 1:** We guess $Y = \widetilde{Y}$ and solve for the A_{Sinj}, $\forall i$. At this stage, we identify the μ^*_{inj}, δ^*_{Sinj}, W^*_{Sinj}, and S^*_{inj}, where * denotes the equilibrium value of these quantities. S^*_{inj} is establishment-level skill-specific employment which we use from the data, W^*_{Sinj} are model wages from the labor supply function, and μ^*_{inj}, δ^*_{Sinj} are independent of aggregate Y as they only depend on the relative A_{Sinj} within a market.
   (b) **Step 2:** In Step 1, we identify $Y^* = \int_j \sum_i P_{inj} Y_{inj}\, dj$, as the establishment-level revenues are independent of the guess $\widetilde{Y}$. Therefore, we can solve the model a second time using Y^* to retrieve the estimated A^*_{Sinj} distribution.[2]

A.5. *Proofs*

Proof of Proposition 1

In homogeneous establishment case, the skill premium is given by

$$\kappa = \left[\left(\frac{A_H}{A_L} \right)^{\frac{\sigma-1}{\sigma+\phi}} \times \left(\frac{\bar{\phi}_L}{\bar{\phi}_H} \right)^{\frac{1}{\sigma+\phi}} \right] \times \left[\frac{1 + \frac{1}{\hat{\theta}_L}\frac{1}{N} + \frac{1}{\hat{\eta}_L}\left(1 - \frac{1}{N}\right)}{1 + \frac{1}{\hat{\theta}_H}\frac{1}{N} + \frac{1}{\hat{\eta}_H}\left(1 - \frac{1}{N}\right)} \right]^{\frac{\sigma}{\sigma+\phi}}. \tag{A22}$$

Then the skill premium elasticity is decreasing, that is, $\frac{\partial \kappa}{\partial N} / (\frac{\kappa}{N}) < 0$, iff $(1 + \frac{1}{\hat{\eta}_L})(\frac{1}{\hat{\theta}_H} - \frac{1}{\hat{\eta}_H}) < (1 + \frac{1}{\hat{\eta}_H})(\frac{1}{\hat{\theta}_L} - \frac{1}{\hat{\eta}_L})$.

PROOF: From first-order conditions, we know:

$$\begin{aligned} \kappa &\equiv \kappa_{ij} \\ &= \frac{A_{H,ij}^{\frac{\sigma-1}{\sigma}} H_{ij}^{-\frac{1}{\sigma}} \delta_{L,ij}}{A_{L,ij}^{\frac{\sigma-1}{\sigma}} L_{ij}^{-\frac{1}{\sigma}} \delta_{H,ij}} \end{aligned}$$

[1] Note that an equivalent way would be to compute the aggregate CES wage index (W_S) directly from the data, since the labor supply function holds in both the data and the model. This is because $\overline{\varphi}_S$ is estimated so that the aggregate labor supply function holds.

[2] An alternate way to solve for the aggregate Y^* would be to loop over guess $\widetilde{Y}$ until the goods market is in equilibrium.

$$= \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \cdot \left[\frac{1+\frac{1}{\hat{\theta}_L}e_{L,nj}+\frac{1}{\hat{\eta}_L}(1-e_{L,nj})}{1+\frac{1}{\hat{\theta}_H}e_{H,nj}+\frac{1}{\hat{\eta}_H}(1-e_{H,nj})}\right] \cdot \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}}$$

$$= \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \cdot \left[\frac{1+\frac{1}{\hat{\theta}_L}e_{L,nj}+\frac{1}{\hat{\eta}_L}(1-e_{L,nj})}{1+\frac{1}{\hat{\theta}_H}e_{H,nj}+\frac{1}{\hat{\eta}_H}(1-e_{H,nj})}\right] \cdot \left(\frac{\bar{\phi}_L}{\bar{\phi}_H}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{W_L}{W_H}\right)^{\frac{\phi}{\sigma}}.$$

By rearranging, we get the aforementioned expression.[3] Now we have following properties:

1. From equation (A22), when $N > 1$, it is clear that $\partial\kappa/\partial\hat{\theta}_L < 0$, $\partial\kappa/\partial\hat{\eta}_L < 0$, $\partial\kappa/\partial\hat{\theta}_H > 0$, and $\partial\kappa/\partial\hat{\eta}_H > 0$. In addition, it can be shown that $\partial\kappa/\partial A_H > 0$ and $\partial\kappa/\partial A_L < 0$.
2. With respect to the change in skill premium when changing N, we have

$$\frac{\partial\kappa}{\partial N}\Big/\left(\frac{\kappa}{N}\right) = \frac{\sigma}{\sigma+\phi}\frac{N\left[\left(1+\frac{1}{\hat{\eta}_L}\right)\left(\frac{1}{\hat{\theta}_H}-\frac{1}{\hat{\eta}_H}\right)-\left(1+\frac{1}{\hat{\eta}_H}\right)\left(\frac{1}{\hat{\theta}_L}-\frac{1}{\hat{\eta}_L}\right)\right]}{\left[N\left(1+\frac{1}{\hat{\eta}_H}\right)+\frac{1}{\hat{\theta}_H}-\frac{1}{\hat{\eta}_H}\right]\left[N\left(1+\frac{1}{\hat{\eta}_L}\right)+\frac{1}{\hat{\theta}_L}-\frac{1}{\hat{\eta}_L}\right]}.$$

A sufficient condition for this term to be negative is

$$\left(1+\frac{1}{\hat{\eta}_L}\right)\left(\frac{1}{\hat{\theta}_H}-\frac{1}{\hat{\eta}_H}\right) < \left(1+\frac{1}{\hat{\eta}_H}\right)\left(\frac{1}{\hat{\theta}_L}-\frac{1}{\hat{\eta}_L}\right).$$ *Q.E.D.*

APPENDIX B: DATA APPENDIX

In this section, we discuss the steps we took in the creation and cleaning of our data. We first outline the broad overview of our data cleaning and construction. We discuss some of the quality and coverage issues we face with our data and provide some insight into the different decisions we made in constructing our data for the analysis. Then we discuss the mapping of our model to the data.

Longitudinal Business Database. The data we use to estimate our model combine establishment-level data from the Longitudinal Business Database (LBD) with characteristics of the workers at these establishments from Longitudinal Employer-Household Dynamics (LEHD) data. The frame of the LBD comes from the Census Business Register, which is populated from the quinquennial economic census and from administrative sources. LBD is an establishment-level data set containing information on payroll, employment, revenue, ownership structure, geography (MSA), and industry classification (NAICS). We consider the LBD to be the frame of our sample, and augment this frame with information on worker composition.

[3]We denote $\kappa = \frac{W_H}{W_L} = \frac{\mathcal{W}_H}{\mathcal{W}_L}$ and assume that $\phi_L = \phi_H = \phi$.

Longitudinal Employer-Household Dynamics. The Longitudinal Employer-Household Dynamics (LEHD) data provide information on workers and firms in each state at quarterly frequency from unemployment insurance records. These data allow us to observe about 96% of workers and the identities of their employers (via tax identifiers) for a sample of 20 states, going back to 1997.[4] The LEHD infrastructure files include demographic information on workers from decennial censuses and the American Community Survey as well as administrative records, including age, sex, race and ethnicity, and education. We use the LEHD to construct measures of the education composition of each firm in our data.

B.1. *Education*

Skill Definition. For our exercise, we use the LEHD to derive measures of the composition of skill types and wages within each firm. We label individuals with some college education or greater as "high-skilled" and we label individuals with a high school diploma or less as "low-skilled" workers. The concept of a firm in LEHD is the State Employer Identification Number (SEIN) under which a firm typically reports its employment and payroll for all of its employees at all establishments within the state.

Earnings. Since we observe only earnings rather than wages in our data and only at quarterly frequency, we limit our measurement of earnings and employment to full-quarter observations where for time t, we require the worker is also employed at the firm in quarters $t-1$ and $t+1$ so we know the job existed for the duration of the entire quarter. To further limit marginal employment and outlier observations, we drop any earnings from workers at the firm below an earnings threshold equivalent to 130 hours worked (averaging 10 hours/week) at the federal minimum wage for that quarter. We also truncate worker earnings at the 99th percentile and restrict our earnings observations to prime age workers, between the ages of 25 and 65.

We aggregate these employment and earnings observations at the firm level by high- (and low-)skilled workers' share of employment, that is, their skill ratio. Similarly, we measure the high- (and low-)skilled workers' ratio of payroll per worker (skill premium) for each SEIN. Using the linkage of workers to employers in LEHD, we split the establishment-level payroll and employment in LBD by using the firm-level ratio of high- to low-skilled workers and payrolls we observed in LEHD. This provides us with our high- and low-skilled employment, H_{inj}, L_{inj}, and wages (payroll per worker) W_{Hinj}, W_{Linj}.

Define the skill ratio of the firm in LEHD as $\mathrm{SR}_{\mathrm{LEHD}} = \frac{H_{\mathrm{LEHD}}}{H_{\mathrm{LEHD}}+L_{\mathrm{LEHD}}}$, where H_{LEHD} and L_{LEHD} are full-quarter employment by skill type in LEHD. Then, skill-specific employment in LBD is $H_{inj} = \mathrm{SR}_{\mathrm{LEHD}}\,\mathrm{Emp}_{\mathrm{LBD}}$ and $L_{inj} = (1-\mathrm{SR}_{\mathrm{LEHD}})\,\mathrm{Emp}_{\mathrm{LBD}}$. Similarly, define skill premium as $\mathrm{SP}_{\mathrm{LEHD}} = W_{H,\mathrm{LEHD}}/W_{L,\mathrm{LEHD}}$, where $W_{S,\mathrm{LEHD}}$ is payroll per worker by skill type for full-quarter employment in each firm in LEHD. Using the definition of payroll as $\mathrm{Payroll}_{\mathrm{LBD}} = W_{Hinj}H_{inj} + W_{Linj}L_{inj}$ and the skill premium, the formula for the wage (payroll per worker) in our sample is

$$W_{Hinj} = \frac{\mathrm{Payroll}_{\mathrm{LBD}}}{H_{inj} + \frac{L_{inj}}{\mathrm{SP}_{\mathrm{LEHD}}}}, \qquad W_{Linj} = \frac{\mathrm{Payroll}_{\mathrm{LBD}}}{L_{inj} + (\mathrm{SP}_{\mathrm{LEHD}}\,H_{inj})}.$$

[4]Our sample includes CA, CO, CT, ID, IL, KS, LA, ME, MD, MN, MO, MT, NJ, NM, NC, OR, RI, TX, WA, and WI.

Given our formula for H_{inj} and L_{inj}, we can write W_{Sinj} in terms of our skill ratio, skill premium, employment, and payroll as

$$W_{Hinj} = \frac{\text{Payroll}_{\text{LBD}}}{\text{SR}_{\text{LEHD}}\,\text{Emp}_{\text{LBD}} + (1 - \text{SR}_{\text{LEHD}})\,\text{Emp}_{\text{LBD}} / \text{SP}_{\text{LEHD}}},$$

$$W_{Linj} = \frac{\text{Payroll}_{\text{LBD}}}{(1 - \text{SR}_{\text{LEHD}})\,\text{Emp}_{\text{LBD}} + (\text{SP}_{\text{LEHD}}\,\text{SR}_{\text{LEHD}}\,\text{Emp}_{\text{LBD}})}.$$

Coverage. While coverage of demographic information such as age and race is high in LEHD, the coverage of educational attainment data is lower than that of other individual characteristics. Education is available for workers in LEHD who were at least 25 years of age when surveyed in the 2000 decennial long form survey or the American Community Survey (ACS), and covers 27.6% of workers in our sample in 1997 and 16.9% in 2016. Because of the higher coverage of the 2000 decennial long form survey relative to ACS, education is observed for more workers in our sample in 1997 than in 2016. For workers without observed education, this value is imputed. The education imputation in LEHD is stationary, however, and poorly matches the time trends in educational attainment and skill premium observed in other data sets such as ACS and CPS. We would like to limit our use of education in LEHD to observed cases; however, this causes another issue of biasing our sample to only larger firms. To balance the representativeness of our establishment sample and also retain the trends in college attainment and skill premium in our sample, we use only observed data for any firm with at least one linked high-skilled and low-skilled worker with full-quarter earnings. For firms where we cannot observe at least one high-skilled and one low-skilled worker using observed education values, we use the imputed education values of the firm's full-quarter workers to get payroll and employment by skill level for the employer.

We merge the LBD to the LEHD by firm identifier, which provides each establishment within that firm with the same measure of skill ratio and payroll ratio from LEHD. For each establishment within the firm in a state, we split the LBD employment and payroll for that establishment by the skill ratio and payroll ratios we measured for the linked employer in LEHD. This approach preserves the establishment employment and payroll distribution within LBD. We show that our resulting measures using full-quarter earnings and employment, restricting our use of imputed education information, and applying the measures of skill ratio and payroll ratio to the LBD establishments gives us a sample which accurately reflects establishment counts and size as well as the trends in educational attainment and relative wages by skill.[5]

Matching Trends and the Size Distribution. Table AI shows the summary statistics of our baseline sample in comparison to our same sample construction if we used only observed worker information without any imputations, and our sample if we used all worker information including the observed and imputed worker education for all observations. We can see that using some imputed information maintains the coverage of our sample and the average establishment size, while restricting our use of imputed workers in large firms helps to preserve the skill composition and skill premium trends of the observed sample.

[5]Our estimated elasticities are qualitatively similar when we restrict to only using observed educational attainment. We have also established robustness of our elasticity estimates with different categorizations of skills.

TABLE AI

SAMPLE SUMMARY STATISTICS BY EDUCATION IMPUTE USAGE.

	Hybrid		Observed		All Workers	
	1997	2016	1997	2016	1997	2016
Total Employment	52.2	70.6	70.3	90.3	52.5	70.9
Skill Ratio	0.493	0.609	0.492	0.618	0.509	0.560
Skill Premium	1.52	1.73	1.52	1.77	1.45	1.46
$\mathcal{W}_H$	\$46,960	\$67,100	\$47,960	\$69,340	\$45,880	\$64,970
$\mathcal{W}_L$	\$30,980	\$38,690	\$31,590	\$39,290	\$31,630	\$44,560
Establishment Count	72,000	27,000	47,000	17,000	72,500	27,000

Note: Hybrid refers to our sample methodology where imputed workers are only used in the absence of at least one observed high- and low-skilled worker. The skill ratio is the establishment-level mean of the share of employment with high skill (some college education or more), weighted by employment. Total Employment refers to the average establishment size in the sample. $\mathcal{W}_H$ and $\mathcal{W}_L$ denote the employment-weighted mean of establishment payroll per worker for high- and low-skilled workers, respectively. Note that the skill premium is slightly different from the data values in Table VI as the samples in this table are constructed in the same manner as our estimation of labor supply elasticities, however it includes establishments with missing revenue information.

B.2. *Revenues*

Allocating Firm Revenue to the Establishments. As outlined in Section 4, the last step in our estimation procedure requires us to estimate the market structure. To do so, we need aggregate moments of the distribution of establishment-level revenue and payroll. Of course, our measure of sales which we can link to LBD is a firm-level measure derived from administrative tax data. To get at an establishment distribution, we follow Tanaka, Warren, and Wiczer (2023) and impute the revenue to the establishment by using the establishment's share of payroll within its firm. While imputing revenue to the establishment based on payroll shares is imperfect relative to a direct establishment-level measurement, we only use an aggregate moment for our market structure estimation.

Validation Using CMF. It is impossible to get the exact establishment-level distribution of revenues for all sectors in our data set, but it is possible to assess our imputation method by making a comparison between our payroll-share imputed revenue and a direct measure of establishment sales for the manufacturing sector in Economic Census years. We take the Census of Manufactures for the years 1997 and 2017 and apply similar restrictions to the payroll and revenue variables for establishments in our sample (non-missing and strictly positive payroll and revenue, and truncation of revenue at the 99th percentile). To check the quality of our impute, we focus on the sample of establishments within multi-establishment firms, as these are the units which we impute based on our payroll share. Figure A1 shows that the establishment-level payroll and revenue share distribution is nearly identical.[6]

We can further assess the revenue impute by comparing the difference between the directly measured establishment-level revenue to our payroll-imputed revenue measure. In Figure A2, we plot the distribution of the difference in logged imputed revenue minus the observed log of establishment revenue. The distribution of errors is symmetric and centered at 0. Looking at the revenue-weighted difference relative to the unweighted difference, the weighted distribution has a thicker left tail, suggesting that the imputed revenue is lower than the observed revenue especially for high-revenue establishments.

[6]We deflate revenue to 2002 dollars.

(a) 1997 (b) 2017

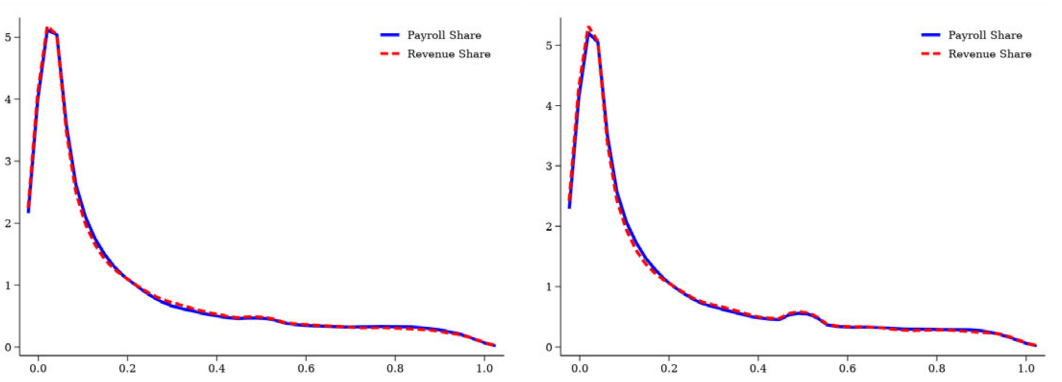
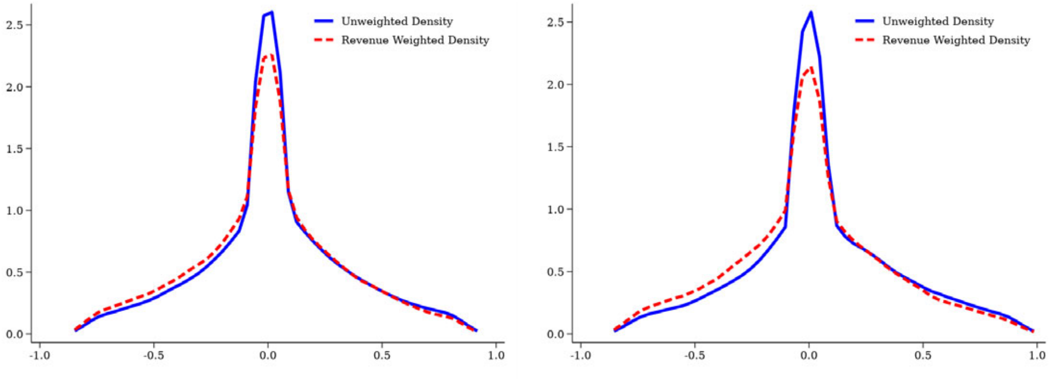


FIGURE A1.—Payroll shares and revenue shares of establishments in multi-unit firms. Notes: Kernel density plot of establishment payroll share and revenue share for establishments in multi-unit firms in CMF. Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.

Figure A3 displays the unweighted distribution of revenue over payroll using imputed revenue and observed revenue at the establishment level. When we look at our moment of interest, the sales-weighted distribution of revenue over payroll in Figure A4, we see that the distribution of observed revenue over payroll has a fatter tail than the imputed measure.[7] However, this error does not seem to affect the trends in our measure over time. The change in the sales-weighted establishment-level mean of revenue over payroll from 1997 to 2017 is nearly identical when using our imputed measure or the direct establishment measure, as can be seen in the last row of Table AII.



(a) 1997 (b) 2017

FIGURE A2.—Log difference of imputed revenue—observed revenue. Notes: Distributions of log differences in the imputed revenue and observed revenue at the establishment level for multi-unit establishments in CMF. This figure plots the differences for multi-unit firms, as these are the establishments which require imputation. Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.

[7]The heavier tail is not so obvious in the plotted distributions due to the truncation of the kernel densities at the 5th and 95th percentiles. However, the comparison of the weighted means in Table AII is consistent with more skewed distribution for the observed versus imputed measure.

(a) 1997 (b) 2017

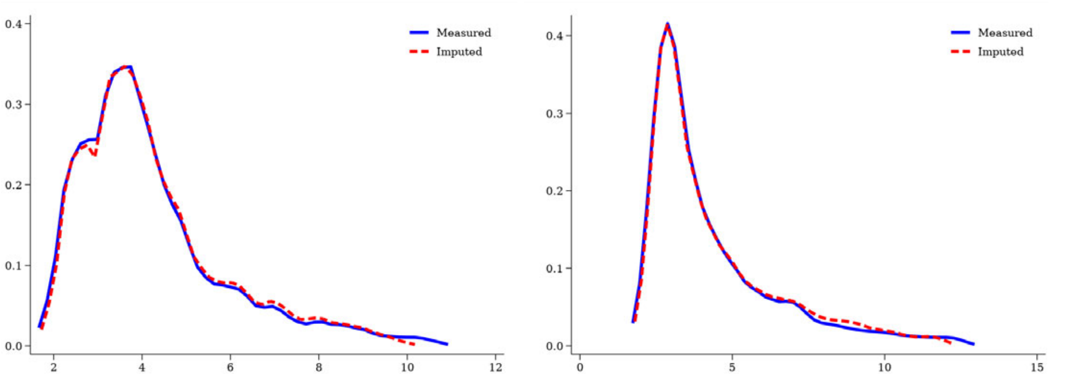


FIGURE A3.—Distribution of revenue over payroll. Notes: Unweighted distributions of revenue over payroll using imputed revenue and observed revenue at the establishment level for multi-unit establishments in CMF. Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.

(a) 1997 (b) 2017

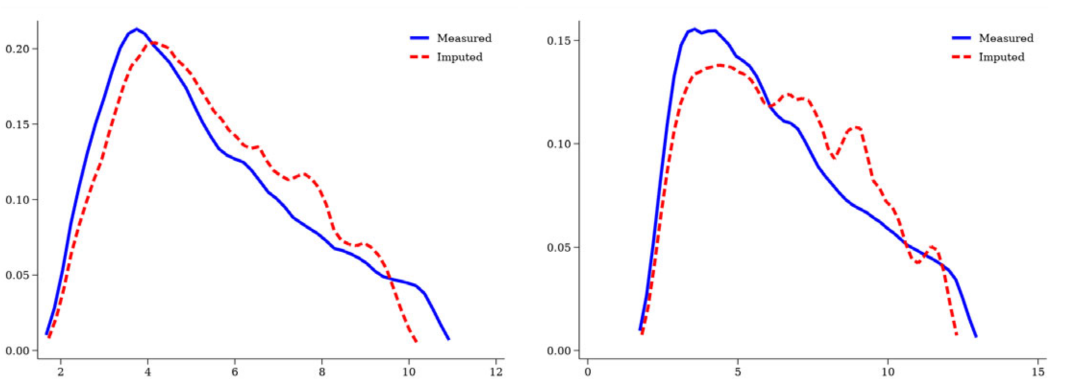


FIGURE A4.—Distribution of revenue over payroll (sales-weighted). Notes: Revenue-weighted distributions of revenue over payroll using imputed revenue and observed revenue at the establishment level for multi-unit establishments in CMF. Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.

TABLE AII

REVENUE OVER PAYROLL.

	Revenue over Payroll						
	Measured Mean	Imputed Mean	Measured Wgt Mean	Imputed Wgt Mean	MU_N	SU_N	Estab. Count
1997	4.87	4.70	8.87	7.49	68,000	326,000	394,000
2017	5.27	5.13	12.39	11.04	59,000	233,000	292,000
Change	0.40	0.43	3.52	3.55			

Note: The weighted mean in this table is weighted by observed establishment revenue. MU_N and SU_N are the rounded counts of establishments in multi-unit and single-unit firms, respectively. Imputations are only necessary for establishments within multi-unit firms.

B.3. *Market Definition*

In order to estimate the model, we need to define a market. Our approach is to stochastically define markets and use the structure of our model to estimate the scope of our markets. Practically, we start by defining a broad set of potential competitors as a NAICS 6 industry.[8] In order to define a market within each NAICS 6 industry, we first randomly assign establishments to markets of size I. Once we select those I establishments that form a market, thereafter we randomly establish the identity of the firms that compete, and how many firms N are active within a market by randomly assigning these I establishments into N subsets of size I/N. We drop the remainder of establishments in each industry that cannot be assigned to a full market of I establishments. In our exercise, we choose I to be 32.

Our baseline estimation uses NAICS 6 industry as the basis for our random market assignment to best match the features of the product market. Since our tax variation is at the state level, markets within a state will not have any variation in tax rates, which makes it difficult for us to condition on geography. Therefore, we use Tradeables and narrowly defined national industries (NAICS 6) as our baseline.[9] We could alternatively choose to match on characteristics of the labor market; however, we lack information such as occupation to satisfactorily define labor markets. Since our model assumes identical product and labor markets, our choice to match on product market characteristics implies that our labor markets are also national. In Appendix D, we perform robustness exercises where we eliminate random assignment of establishments to markets and where we segment our industries by geography (MSA) so that we are likely to be closer to the relevant boundary of the market for labor at the expense of the product market.

Ownership Assignment. Note that we do not use the ownership structure of firms and establishments from the data in our exercise. This discards some useful information about changes in the distribution of establishments and firms. Firm growth, especially at the tail, was documented by Cao, Hyatt, Mukoyama, and Sager (2022) to be largely driven by increases in the number of establishments a firm operates. However, we remain agnostic about the process of establishment birth, death, and consolidation. The primary reasons for our use of a stochastic ownership structure is that this allows a level of symmetry which is useful in our counterfactual analysis.

B.4. *Summary of Data Cleaning*

We have two samples which we use in our estimation process. For our backing out of technology and estimation of the market structure, we use cross-sections of establishments for the years 1997 and 2016. For this sample, we assign establishments to markets (and firms) stochastically as described above. We also require revenue information so we restrict the sample to establishments with non-missing revenue and truncate revenue at the 99th percentile.

For our estimation of labor supply elasticities, we use a panel of establishments from 1997 to 2011 as we have state-level corporate tax rates through 2011 from Giroud and Rauh (2019). In order to stochastically assign establishments to markets and retain a panel structure, we first randomly assign establishments to markets, conditional on NAICS 6,

[8]In Appendix D, we condition on geography and we define the broad set of competitors as those within NAICS 3 industry x MSA.

[9]Using Tradeables also helps our results to be comparable to Berger, Herkenhoff, and Mongey (2022).

in 1997 such that there are at most 32 establishments in each market. Once assigned to a market, the establishment always remains in that market as long as we observe it in the data. For every subsequent year starting from 1997, we again randomly assign the establishment unobserved previously (i.e., the new entrants) to one of the existing markets created in 1997. As a result, the size and the composition of the markets evolve randomly over time given the entry and exit of establishments from markets. Our baseline elasticity estimates are based on this sample. Since we want to estimate labor supply elasticities using the entire wage and earnings distribution, we do not restrict the sample based on revenue as we do when estimating market structure.

Our final data cleaning steps are common to both samples. Our sample is the subset of our LBD sample of establishments where the firm links to at least one SEIN in our 20-state LEHD sample. We drop firms in LBD where they account for less than 5 percent of the employment when measured at the linked firm in LEHD to drop some outlier firms in our linkage. We drop establishments with missing county. We keep only establishments of C-corporation firms for our tax instrument in the elasticity estimation. We use establishments in tradeable sectors (11, 21, 31, 32, 33, and 55) as defined in Delgado, Bryden, and Zyontz (2014). We drop establishments with five or fewer total employees, and for which we do not have at least one high- and one low-skilled employee and positive payroll for each skill type. We winsorize establishment employment and average high- and low-skilled payroll per worker, W_{Hinj}, W_{Linj}, at the 1st and 99th percentile.

APPENDIX C: IDENTIFICATION

C.1. *Derivation of Equation (21)*

To derive equation (21) in the main text, we proceed as follows. We start from the labor supply equation (rewritten below for convenience):

$$\ln W^*_{Sinjt} = k_{jt} + \left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_{jt} + \frac{1}{\hat{\eta}_S}\ln S_{injt} + \varepsilon_{Sinjt},$$

where $\ln W^*_{Sinjt} = \ln W_{Sinjt} + \varepsilon_{Sinjt}$ and $k_{jt} = \ln J_t^{\frac{1}{\hat{\theta}_S}} I_{jt}^{\frac{1}{\hat{\eta}_S}} S_t^{-\frac{1}{\hat{\theta}_S}} W_t$.

We construct sector-time average of the labor supply function to remove the sector-time fixed terms from the labor supply equation:

$$\overline{\ln W}^*_{Sjt} = k_{jt} + \left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_{jt} + \frac{1}{\hat{\eta}_S}\overline{\ln S}_{jt} + \overline{\varepsilon}_{Sjt},$$

where $\overline{\ln W}^*_{Sjt} = \frac{1}{I_j}\sum_{i\in j}\ln W^*_{Sinjt}$, $\overline{\ln S}_{jt} = \frac{1}{I_j}\sum_{i\in j}\ln S_{injt}$, and $\overline{\varepsilon}_{Sjt} = \frac{1}{I_j}\sum_{i\in j}\varepsilon_{injt}$.

Getting rid of sector-time components from the labor supply equation, we get

$$\ln W^*_{Sinjt} - \overline{\ln W}^*_{Sjt} = \frac{1}{\hat{\eta}_S}(\ln S_{injt} - \overline{\ln S}_{jt}) + (\varepsilon_{Sinjt} - \overline{\varepsilon}_{Sjt}).$$

Finally, we rely on the following moment conditions implied by Assumption 3 to get our equation of interest for $\hat{\eta}_S$:

$$0 = \mathbb{E}\left[(\varepsilon_{Sinjt} - \overline{\varepsilon}_{Sjt}) \times \tau_{X(i)t}\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\left\{\left(\ln W^*_{Sinjt} - \overline{\ln W}^*_{Sjt}\right) - \frac{1}{\hat{\eta}_S}\left(\ln S_{injt} - \overline{\ln S}_{jt}\right)\right\} \times \tau_{X(i)t}\right], \\
\mathbb{E}\left[\left(\ln W^*_{Sinjt} - \overline{\ln W}^*_{Sjt}\right) \times \tau_{X(i)t}\right] &= \frac{1}{\hat{\eta}_S}\mathbb{E}\left[\left(\ln S_{injt} - \overline{\ln S}_{jt}\right) \times \tau_{X(i)t}\right], \\
\hat{\eta}_S &= \frac{\mathbb{E}\left(\widetilde{S}_{injt} \times \tau_{X(i)t}\right)}{\mathbb{E}\left(\widetilde{W}^*{}_{Sinjt} \times \tau_{X(i)t}\right)},
\end{aligned}
$$

where $\widetilde{S}_{injt} = \ln S_{injt} - \overline{\ln S}_{jt}$ and $\widetilde{W}^*{}_{Sinjt} = \ln W^*_{Sinjt} - \overline{\ln W}^*_{Sjt}$.

In order to derive the expression for $\hat{\theta}_S$ in equation (21), we proceed as follows. Equipped with the estimate of $\hat{\eta}_S$, we rewrite the labor supply function as follows:

$$
\ln W^*_{Sinjt} - \frac{1}{\hat{\eta}_S}\ln S_{injt} \equiv \Omega_{Sinjt} = k_{jt} + \left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_{jt} + \varepsilon_{Sinjt}.
$$

Taking sector-time average on both sides, we get

$$
\overline{\Omega}_{Sjt} = k_j + k_t + \left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_{jt} + \nu_{jt} + \overline{\varepsilon}_{Sjt},
$$

where $k_{jt} = k_j + k_t + \nu_{jt}$ and $\overline{\Omega}_{Sjt} = \frac{1}{I_j}\sum_{i\in j}\Omega_{Sinjt}$.

Using the following moment implied by Assumption 3, we can calculate our expression of interest for $\hat{\theta}_S$:

$$
\begin{aligned}
\mathbb{E}(\bar{\varepsilon}_{Sjt} \times \bar{\tau}_{jt}) &= \mathbb{E}\left[\left(\overline{\Omega}_{Sjt} - k_{jt} - \left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_{jt}\right) \times \bar{\tau}_{jt}\right] = 0, \\
\mathbb{E}\left[(\overline{\Omega}_{Sjt} - k_{jt}) \times \bar{\tau}_{jt}\right] &= \mathbb{E}\left[\left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_{jt} \times \bar{\tau}_{jt}\right], \\
\hat{\theta}_S &= \left[\frac{\mathbb{E}\left(\{\overline{\Omega}_{Sjt} - k_{jt}\} \times \bar{\tau}_{jt}\right)}{\mathbb{E}(\ln S_{jt} \times \bar{\tau}_{jt})} + \frac{\mathbb{E}\left(\widetilde{W}^*{}_{Sinj} \times \tau_{X(i)t}\right)}{\mathbb{E}(\widetilde{S}_{inj} \times \tau_{X(i)t})}\right]^{-1}, \\
\hat{\theta}_S &= \left[\frac{\mathbb{E}\left(\left\{\overline{\Omega}_{Sjt} - (k_j + k_t + \upsilon_{jt})\right\} \times \bar{\tau}_{jt}\right)}{\mathbb{E}(\ln S_{jt} \times \bar{\tau}_{jt})} + \frac{\mathbb{E}\left(\widetilde{W}^*{}_{Sinj} \times \tau_{X(i)t}\right)}{\mathbb{E}(\widetilde{S}_{inj} \times \tau_{X(i)t})}\right]^{-1}, \\
\hat{\theta}_S &= \left[\frac{\mathbb{E}\left(\left\{\overline{\Omega}_{Sjt} - (k_j + k_t)\right\} \times \bar{\tau}_{jt}\right)}{\mathbb{E}(\ln S_{jt} \times \bar{\tau}_{jt})} + \frac{\mathbb{E}\left(\widetilde{W}^*{}_{Sinj} \times \tau_{X(i)t}\right)}{\mathbb{E}(\widetilde{S}_{inj} \times \tau_{X(i)t})}\right]^{-1},
\end{aligned}
$$

where $\bar{\tau}_{jt} = \frac{1}{I_j}\sum_{i\in j}\tau_{X(i)t}$. To go from line 3 to line 4, we rely on the fact that $k_{jt} = k_j + k_t + \nu_{jt}$. Finally, to go from line 4 to line 5, we rely on Assumption 3 outlined in the main text which implies $\mathbb{E}(\nu_{jt} \times \bar{\tau}_{jt}) = 0$.

C.2. *Identification Without Endogeneity*

In this section, we show that, under the assumption that the error term is *uncorrelated* with employment, we can identify $\hat{\eta}_S$ and $\hat{\theta}_S$ using the following moments:

$$\hat{\eta}_S = \left(\frac{\mathbb{C}ov(\widetilde{S}_{inj}, \widetilde{W^*}_{Sinj})}{\mathbb{V}ar(\widetilde{S}_{inj})}\right)^{-1}, \tag{A23}$$

$$\hat{\theta}_S = \left[\left(\frac{\mathbb{C}ov(\ln S_j, \overline{\Omega}_{Sj})}{\mathbb{V}ar(\ln S_j)}\right) + \left(\frac{\mathbb{C}ov(\widetilde{S}_{inj}, \widetilde{W^*}_{Sinj})}{\mathbb{V}ar(\widetilde{S}_{inj})}\right)\right]^{-1}, \tag{A24}$$

where we denote

$$\widetilde{S}_{inj} = \ln S_{inj} - \overline{\ln S_j}, \qquad \widetilde{W^*}_{Sinj} = \ln W^*_{Sinj} - \overline{\ln W^*_{Sj}}, \qquad \overline{\ln X_j} = \frac{1}{I}\sum_{i\in j} \ln X_{inj},$$

$$\Omega_{Sinj} = \ln W^*_{Sinj} - \frac{1}{\hat{\eta}_S}\ln S_{inj}, \qquad \overline{\Omega}_{Sj} = \frac{1}{I}\sum_{i\in j}\Omega_{Sinj}.$$

The moment condition in equation (A23) is equivalent to regressing the difference of log employment from the mean of market-level log employment on difference of log wages from the mean of market log wages. The moment condition in equation (A24) is equivalent to regressing the market-level employment CES index on average market-level wages (after removing the effect of average sectoral employment). Given that the sectoral CES index is a function of $\hat{\eta}_S$, we need to construct moments in equation (A23) and equation (A24) sequentially, starting with first retrieving the estimate of $\hat{\eta}_S$.

Deriving the Moment Conditions. To derive the moment conditions in equation (A23) and equation (A24), start by differencing out the market-specific mean wages and mean employment from equation (20) to get the following expression:

$$\ln W^*_{Sinj} - \overline{\ln W^*_{Sj}} = \frac{1}{\hat{\eta}_S}(\ln S_{inj} - \overline{\ln S_j}) + (\varepsilon_{Sinj} - \overline{\varepsilon}_{Sj}). \tag{A25}$$

An OLS regression of equation (A25) helps us retrieve $\hat{\eta}_S$ and equation (A23) specifies the moments that help us pin it down. Equipped with the estimate of $\hat{\eta}_S$, we can construct S_j, the CES index of market-level employment. In the second step, we can then estimate the between-market substitution parameter $\hat{\theta}_S$ by relying on equation (20) and subtracting $\frac{1}{\hat{\eta}_S}\ln S_{inj}$ from $\ln W^*_{Sinj}$:

$$\ln W^*_{Sinj} - \frac{1}{\hat{\eta}_S}\ln S_{inj} \equiv \Omega_{Sinj} = k + \left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_j + \varepsilon_{Sinj}, \tag{A26}$$

where $k = \ln J^{\frac{1}{\theta_S}} I^{\frac{1}{\eta_S}} S^{-\frac{1}{\theta_S}} W$.

To construct the moment in equation (A24), take market-specific averages of both sides on equation (A26) and regress $\ln S_j$ on $\overline{\Omega}_{Sj}$ to retrieve the estimate of θ:

$$\overline{\Omega}_{Sj} = k + \left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)\ln S_j + \overline{\varepsilon}_{Sj}. \tag{A27}$$

C.2.1. *Monte Carlo Simulation*

To see if our proposed estimator is able to recover the true structural parameters, we perform the following Monte Carlo simulation. First, we simulate the labor supply equation as follows:[10]

$$
\begin{aligned}
\ln W^*_{Sinj} &= k + \underbrace{\left(\frac{1}{\hat{\theta}_S} - \frac{1}{\hat{\eta}_S}\right)}_{\gamma_S} \ln S^*_j + \underbrace{\frac{1}{\hat{\eta}_S}}_{\beta_S} \ln S^*_{inj} + \underbrace{\epsilon^W_{inj} + \epsilon^S_{inj}}_{\epsilon_{inj}}, \\
\ln W^*_{Sinj} &= \ln W_{Sinj} + \epsilon^W_{inj}, \\
\ln W_{Sinj} &= k + \gamma_S \ln S_j + \beta_S \ln S_{inj}, \\
\ln S_{inj} &\sim \mathbb{N}(0,1), \\
S_j &= \left(\sum_i I^{\frac{1}{\hat{\eta}_S}} S_{inj}^{\frac{\hat{\eta}_S+1}{\hat{\eta}_S}}\right)^{\frac{\hat{\eta}_S}{\hat{\eta}_S+1}}, \\
\ln S^*_{inj} &= \ln S_{inj} + \rho \times \epsilon^S_{inj}, \qquad S^*_j = \left(\sum_i I^{\frac{1}{\hat{\eta}_S}} \left(S^*_{inj}\right)^{\frac{\hat{\eta}_S+1}{\hat{\eta}_S}}\right)^{\frac{\hat{\eta}_S}{\hat{\eta}_S+1}}, \\
\epsilon^S_{inj} &\sim \mathbb{N}(0,1), \qquad \epsilon^W_{inj} \sim \mathbb{N}(0,1),
\end{aligned}
\tag{A28}
$$

where ϵ^W_{inj} denotes the measurement error in wages and ϵ^S_{inj} denotes the measurement error in employment if $\rho \neq 0$. We assume that ϵ^W_{inj} and ϵ^S_{inj} are independent. Finally, we also assume that $\ln S_{inj}$ is independent of ϵ^W_{inj}.

Given this data generating process, we first verify that the estimator is able to recover the true structural parameters if $\rho = 0$. This implies that there is a zero correlation between ϵ_{inj} and $\ln S_{inj}$. The results of this exercise are provided in Table AIII. We find that, under this assumption, OLS can retrieve the true structural parameters $\hat{\eta}_S$ and $\hat{\theta}_S$ using cross-sectional data on employment and wages as outlined in equation (A25) and equation (A27).

TABLE AIII

MONTE CARLO SIMULATION.

		$\hat{\eta}_S$	$\hat{\theta}_S$	$\bar{\phi}_S$
True Value		3.00	1.50	10.00
$\rho = 0$	Mean	3.00	1.50	10.00
	Std. Dev	0.07	0.07	0.11
$\rho = 0.5$	Mean	3.75	2.02	16.80
	Std. Dev	0.10	0.12	0.70
$\rho = 1.5$	Mean	9.78	6.80	968.56
	Std. Dev	0.44	0.72	205.37

Note: In simulation, we assumed that $J = 500$ and $I = 32$ and ran 1000 trials for each value of ρ.

[10]We simulate only a cross-section and assume that each market has I establishments. We omit the time notation as we work with a cross-section.

In order to understand the role of endogeneity bias, we perform additional simulations where we assume that $\rho \neq 0$. This implies that $\ln S_{inj}$ is correlated with ϵ_{inj} in equation (A28). In practice, we pick $\rho \in \{0.5, 1.5\}$. As before, we try to recover the estimates of $\hat{\eta}_S$, $\hat{\theta}_S$, and $\bar{\phi}_S$ using OLS. The results of this exercise are also presented in Table AIII. We find that as ρ deviates from 0, it leads to an upward bias in the estimates of $\hat{\eta}_S$, $\hat{\theta}_S$, and $\bar{\phi}_S$, with the bias increasing as ρ increases.

APPENDIX D: ROBUSTNESS OF THE ESTIMATES OF LABOR SUBSTITUTABILITY PARAMETERS

This appendix presents two cases where we deviate from our baseline estimates in which we randomly assigned establishments to markets within NAICS 6. The main aim is to eliminate the influence of the random assignment of establishments into markets and explore the robustness of our estimates to potential misspecification of the labor market definition. Under each scenario in the robustness, the market size is allowed to vary based on the total number of establishments within each market and is entirely determined by the fixed market definition and the underlying microdata.

In Table AIV, the results of the parameters for labor substitutability are presented when establishments are no longer randomly assigned to markets and when product and labor markets are defined as NAICS 6. In Table AV, the results of labor substitutability are shown when we redefine our product and labor markets to be NAICS 3 x MSA and did not randomly assign establishments to markets.

In Table AIV, the estimates of the $\hat{\eta}_H$ and $\hat{\eta}_L$ are 2.70 and 2.50, respectively, as compared to our baseline values of 2.53 and 2.42. On the other hand, the estimates of $\hat{\theta}_H$ and $\hat{\theta}_L$ are 1.93 and 1.87, respectively, as compared to 2.02 and 1.85. These estimates are statistically significant at 1% when we cluster the standard errors at the state level.

In Table AV, we find that the value of the substitutability parameters for both high- and low-skilled workers increases (relative to the benchmark) and the difference between $\hat{\eta}_S - \hat{\theta}_S$ widens. For instance, the estimates of $\hat{\eta}_H$ and $\hat{\eta}_L$ are 5.40 and 6.41, respectively, and those of $\hat{\theta}_H$ and $\hat{\theta}_L$ are 2.92 and 3.43. However, we find that the second-stage estimates of $\beta_S = 1/\hat{\eta}_S$ are no longer statistically significant when we cluster at the state level.[11] In conclusion, we find that the baseline results are robust if the random assignment of establishments to markets is eliminated and the market is defined as NAICS 6. However, the estimate of $\hat{\eta}_S$ loses statistical significance if the market is defined as NAICS 3 x MSA.

APPENDIX E: ADDITIONAL RESULTS

E.1. *Distributions*

Figure A5 plots the distributions of log employment by skill level, $\ln H_{inj}$ and $\ln L_{inj}$. The employment distribution increases in variance, especially for high-skilled workers.

Figure A6 plots the distributions of log technology by skill level, $\ln A_{Hinj}$ and $\ln A_{Linj}$. It is worthwhile to note that while employment is the primary source of establishment-level heterogeneity in the model inputs, the distribution of technology reflects the model

[11]The only difference with regard to the baseline specification is that we do not include establishment fixed effects in our regression. When we include the establishment fixed effect, we find that the estimates of labor substitutability parameters are theory inconsistent.

TABLE AIV

ESTIMATES OF LABOR SUBSTITUTABILITY PARAMETERS: NAICS 6, TRADEABLES, WITHOUT RANDOM SAMPLING.

	OLS (1)	IV (2)		OLS (3)	IV (4)
A. OLS and Second-Stage IV Estimates					
β_H	−0.177	0.371	γ_H	0.146	0.148
SE	0.0007	0.057	SE	0.0002	0.001
State-level SE	(0.002)	(0.113)	Market SE	(0.023)	(0.043)
β_L	−0.108	0.399	γ_L	0.123	0.136
SE	0.0007	0.051	SE	0.0002	0.001
State-level SE	(0.003)	(0.097)	Market SE	(0.025)	(0.041)
Market x Year FE	Yes	Yes	Market FE	Yes	Yes
Establishment FE	Yes	Yes	Year FE	Yes	Yes
B. Structural Parameters					
$\hat{\eta}_H$	−5.64	2.70	$\hat{\theta}_H$	−31.37	1.93
$\hat{\eta}_L$	−9.30	2.50	$\hat{\theta}_L$	64.2	1.87
C. First-Stage Regressions for the IV					
$\tau^H_{X(i)t}$	–	−0.013	$\bar{\tau}^H_{jt}$	–	−0.015
SE		0.0008	SE		0.0009
State-level SE		(0.004)	Market SE		(0.001)
$\tau^L_{X(i)t}$	–	−0.015	$\bar{\tau}^L_{jt}$	–	−0.276
SE		0.0009	SE		0.0008
State-level SE		(0.006)	Market SE		(0.059)
Market x Year FE	–	Yes	Market FE	–	Yes
Establishment FE	–	Yes	Year FE	–	Yes
No. of obs (High-Skilled)	1,166,000	1,166,000		5900	5900
No. of obs (Low-Skilled)	1,166,000	1,166,000		5900	5900

Note: Non-clustered standard errors are reported without parentheses, while clustered standard errors are reported with parentheses. The significance stars correspond to clustered standard errors. Estimates of γ_S in columns 3 and 4 are conditional on the estimates of columns 1 and 2, respectively. Number of observations is common for both the first and the second stage. The number of observations reflects rounding for disclosure avoidance. $\tau^S_{X(i)t}$ denotes the co-efficient in front of taxes in the first-stage regression for the estimate of β_S. The same instrument is used separately, first to estimate β_H and then to estimate β_L.

structure, key parameters such as N and elasticities, as well as the market assignment.

Figure A7 shows that our estimated model matches the establishment-level distribution of skill premium remarkably well. Not only do we replicate the change in skill premium, but our model generates the substantial heterogeneity in the establishment-level skill premia we observe in the data.

Figure A8 plots the unweighted establishment-level distributions of the markup and skill-specific markdowns from our estimated model. Note that we observe a shift in all three distributions from 1997 to 2016. While we observe an increase in variance for both markdowns and markups, the upper bound for markdowns moves relatively little compared to the upper bound for markups. There is a much larger increase in the variance of markups over time.

TABLE AV

ESTIMATES OF LABOR SUBSTITUTABILITY PARAMETERS: NAICS 3 X MSA, WITHOUT RANDOM SAMPLING.

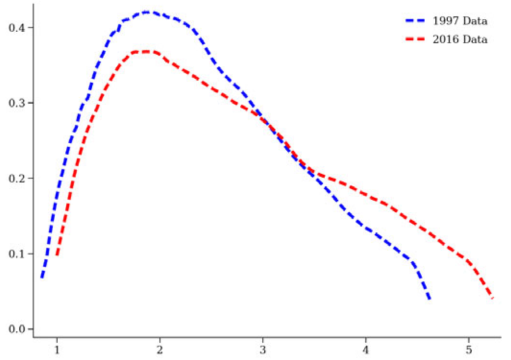
	OLS (1)	IV (2)		OLS (3)	IV (4)
A. OLS and Second-Stage IV Estimates					
β_H	0.079	0.185	γ_H	0.063	0.157
SE	0.0006	0.063	SE	0.0004	0.002
State-level SE	(0.003)	(0.189)	Market SE	(0.013)	(0.044)
β_L	0.029	0.156	γ_L	0.080	0.136
SE	0.0007	0.086	SE	0.0004	0.001
State-level SE	(0.005)	(0.310)	Market SE	(0.013)	(0.044)
Market x Year FE	Yes	Yes	Market FE	Yes	Yes
Establishment FE	Yes	Yes	Year FE	Yes	Yes
B. Structural Parameters					
$\hat{\eta}_H$	12.62	5.40	$\hat{\theta}_H$	7.05	2.92
$\hat{\eta}_L$	34.98	6.41	$\hat{\theta}_L$	9.23	3.43
C. First-Stage Regressions for the IV					
$\tau^H_{X(i)t}$	–	0.031	$\bar{\tau}^H_{jt}$	–	−0.110
SE		0.004	SE		0.0005
State-level SE		(0.008)	Market SE		(0.022)
$\tau^L_{X(i)t}$	–	0.024	$\bar{\tau}^L_{jt}$	–	−0.127
SE		0.004	SE		0.0005
State-level SE		(0.011)	Market SE		(0.023)
Market x Year FE	–	Yes	Market FE	–	Yes
Establishment FE	–	Yes	Year FE	–	Yes
No. of obs (High-Skilled)	497,000	497,000		5800	5800
No. of obs (Low-Skilled)	497,000	497,000		5800	5800

Note: Non-clustered standard errors are reported without parentheses, while clustered standard errors are reported with parentheses. The significance stars correspond to clustered standard errors. Estimates of γ_S in columns 3 and 4 are conditional on the estimates of columns 1 and 2, respectively. Number of observations is common for both the first and the second stage. The number of observations reflects rounding for disclosure avoidance. $\tau^S_{X(i)t}$ denotes the co-efficient in front of taxes in the first-stage regression for the estimate of β_S. The same instrument is used separately, first to estimate β_H and then to estimate β_L.

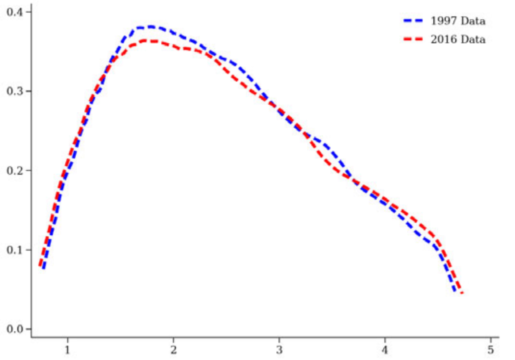
E.2. *Decomposing Estimated Productivity*

In Table IV, we decomposed the total variance in $\ln A_{Linj}$ and $\ln A_{Hinj}$ into within and between NAICS 6 industries. To do so, we denote industry as $k \in \{1, \dots, K\}$, the total number of establishments in a given industry k as $\tilde{I}_k$, and the total number of establishments in the economy as $\tilde{I} = \sum_{k=1}^{K} \tilde{I}_k$. Additionally, we denote $\ln A_{Sik} = a_{Sik}$, $S \in \{H, L\}$. We can then decompose the $\mathbb{V}ar(a_{Sik})$ as follows:

$$\mathbb{V}ar(a_{ik}) = \underbrace{\frac{1}{\tilde{I}} \sum_{i=1}^{\tilde{I}} (a_{ik} - \overline{a}_k)^2}_{\text{Within NAICS 6}} + \underbrace{\frac{1}{\tilde{I}} \sum_{k=1}^{K} \tilde{I}_k (\overline{a}_k - \overline{a})^2}_{\text{Between NAICS 6}}.$$

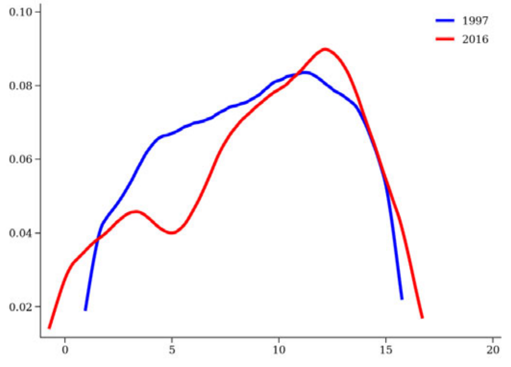


(a) PDF of $\ln H_{inj}$

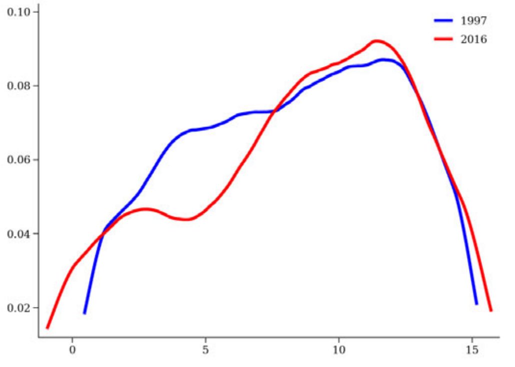


(b) PDF of $\ln L_{inj}$

FIGURE A5.—Distribution of employment by skill. Notes: Panels (a) and (b) show the probability density function of productivities of $\ln A_{Hinj}$ and A_{Linj}, respectively, for 1997 and 2016. Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.



(a) PDF of $\ln A_{Hinj}$

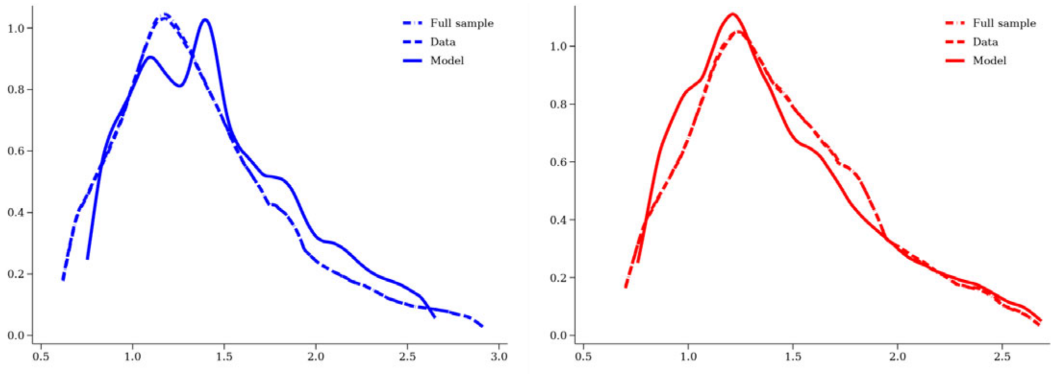


(b) PDF of $\ln A_{Linj}$

FIGURE A6.—Estimated distribution of skill-specific technology. Notes: Panels (a) and (b) show the probability density function of productivities of $\ln A_{Hinj}$ and $\ln A_{Linj}$, respectively, for 1997 and 2016. Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.

APPENDIX F: ADDITIONAL TABLES PERTAINING TO RANDOMIZATION

In order to deepen our understanding of the influence of randomness in our main findings presented in Tables VII and VIII, we present additional evidence in this appendix. As highlighted in the main text, to conduct counterfactual simulations, we randomized establishment assignments to firms a total of 41 times. For each counterfactual scenario, we provide estimates of the 5th and 95th percentiles in Tables AVI and AVII to capture the range of possible outcomes.



(a) 1997 (b) 2016

FIGURE A7.—Distribution of skill premium. Notes: Full sample corresponds to the set of establishments in the data where we observe high- and low-skilled wage. Data refer to the subset of full sample after the assignment of establishments to markets of size I. Model corresponds to the model-predicted skill premium for the same set of establishments in the Data sample. Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.

(a) Distribution of δ_{Hinj} (b) Distribution of δ_{Linj} (c) Distribution of μ_{inj}

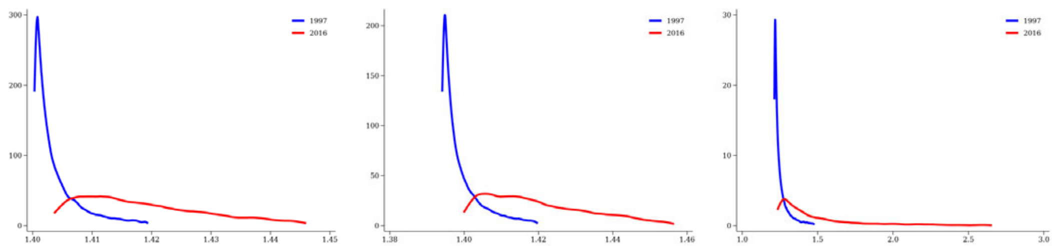


FIGURE A8.—Estimated markup and markdown distribution. Notes: Variables are truncated at the 5th and 95th percentiles before plotting kernel densities.

TABLE AVI

CONFIDENCE INTERVALS FOR COUNTERFACTUAL RESULTS OF TABLE VII.

	Level (1)	5th Percentile (2)	95th Percentile (3)
N	1.480	1.475	1.482
A_{Hinj}, A_{Linj}	1.934	1.931	1.940
$\bar{\phi}_H$, $\bar{\phi}_L$	1.242	1.241	1.243
A_{Hinj}, A_{Linj} and N	1.956	1.942	1.959
A_{Hinj}, A_{Linj} and $\bar{\phi}_H$, $\bar{\phi}_L$	1.631	1.628	1.635
N and $\bar{\phi}_H$, $\bar{\phi}_L$	1.255	1.250	1.256

Note: Column 1 denotes the level of the skill premium for each of the counterfactuals presented in Table VII. These values correspond to the seed that produces the median change in the total variance of wage inequality in Table VIII. Columns 2 and 3 provide the estimates of the 5th and the 95th percentiles for each of the counterfactuals we performed.

TABLE AVII

CONFIDENCE INTERVALS FOR COUNTERFACTUAL RESULTS OF TABLE VIII.

	Level (1)	5th Percentile (2)	95th Percentile (3)
		Total	
N	0.329	0.327	0.330
A_{Hinj}, A_{Linj}	0.400	0.401	0.403
$\bar{\phi}_H$, $\bar{\phi}_L$	0.293	0.293	0.294
A_{Hinj}, A_{Linj} and N	0.416	0.413	0.418
A_{Hinj}, A_{Linj} and $\bar{\phi}_H$, $\bar{\phi}_L$	0.356	0.356	0.358
N and $\bar{\phi}_H$, $\bar{\phi}_L$	0.308	0.306	0.308
		Within	
N	0.048	0.048	0.048
A_{Hinj}, A_{Linj}	0.087	0.087	0.087
$\bar{\phi}_H$, $\bar{\phi}_L$	0.027	0.027	0.027
A_{Hinj}, A_{Linj} and N	0.089	0.088	0.089
A_{Hinj}, A_{Linj} and $\bar{\phi}_H$, $\bar{\phi}_L$	0.050	0.050	0.050
N and $\bar{\phi}_H$, $\bar{\phi}_L$	0.028	0.028	0.028
		Between	
N	0.282	0.280	0.282
A_{Hinj}, A_{Linj}	0.313	0.314	0.315
$\bar{\phi}_H$, $\bar{\phi}_L$	0.266	0.266	0.267
A_{Hinj}, A_{Linj} and N	0.327	0.324	0.329
A_{Hinj}, A_{Linj} and $\bar{\phi}_H$, $\bar{\phi}_L$	0.306	0.306	0.308
N and $\bar{\phi}_H$, $\bar{\phi}_L$	0.280	0.278	0.281

Note: Column 1 (titled Level) denotes the level of total, within-, or between-establishment wage inequality for each of the counterfactuals presented in Table VIII. These values correspond to the seed that produces the median change in the total variance of wage inequality in Table VIII. Columns 2 and 3 provide the estimates of the 5th and the 95th percentiles over all seeds for each of the counterfactuals we performed.

REFERENCES

BERGER, DAVID, KYLE HERKENHOFF, AND SIMON MONGEY (2022): "Labor Market Power," *American Economic Review*, 112 (4), 1147–1193. [2,13]

CAO, DAN, HENRY HYATT, TOSHIHIKO MUKOYAMA, AND ERICK SAGER (2022): "Firm Growth Through New Establishments," Unpublished Working Paper. [13]

DELGADO, MERCEDES, RICHARD BRYDEN, AND SAMANTHA ZYONTZ (2014): "Categorization of Traded and Local Industries in the US Economy," Unpublished Working Paper. [14]

GIROUD, XAVIER, AND JOSHUA RAUH (2019): "State Taxation and the Reallocation of Business Activity: Evidence From Establishment-Level Data," *Journal of Political Economy*, 127 (3), 1262–1316. [13]

TANAKA, SATOSHI, LAWRENCE WARREN, AND DAVID WICZER (2023): "Earnings Growth, Job Flows and Churn," *Journal of Monetary Economics*, 135, 86–98. [10]

Co-editor Charles I. Jones handled this manuscript.

Manuscript received 13 September, 2022; final version accepted 11 February, 2024; available online 13 February, 2024.