Online Appendix:
Job Preferences, Labor Market Power, and Inequality

Oscar Volpe

November 11, 2025

This online appendix accompanies the paper "Job Preferences, Labor Market Power, and Inequality". It provides derivations of key economic quantities, proofs of equilibrium properties and identification results, model extensions, details about data preparation and estimation, and robustness analyses. All notation is consistent with the rest of the paper.

A. *Derivation of Equilibrium Quantities*

A.1. *Firm Labor Supply*

Given any set of wage offers $\mathbf{W}(X) = \{W_k(X)\}_k$, a worker with skills X and marginal utility of (log) earnings β would choose to work at firm j with probability $\mathrm{P}(j(i) = j|\beta, X)$, where:

$$
\begin{aligned}
\mathrm{P}(j(i) = j|\beta, X) &= \mathrm{P}\bigg(u_{ij}\big(W_j(X_i), a_j(X_i)\big) > \max_{k \neq j} \big\{u_{ik}\big(W_k(X_i), a_k(X_i)\big)\big\} \,\bigg|\, \beta_i = \beta, X_i = X\bigg) \\
&= \mathrm{P}\bigg(\beta \log W_j(X) + a_j(X) + \epsilon_{ij} > \beta \log W_k(X) + a_k(X) + \epsilon_{ik}, \forall k \neq j\bigg) \\
&= \mathrm{P}\bigg(\epsilon_{ik} < \beta \log\left(\frac{W_j(X)}{W_k(X)}\right) + a_j(X) - a_k(X) + \epsilon_{ij}, \forall k \neq j, \forall k \neq j\bigg) \\
&= \int_{-\infty}^{\infty} \mathrm{P}\bigg(\epsilon_{ik} < \beta \log\left(\frac{W_j(X)}{W_k(X)}\right) + a_j(X) - a_k(X) + \epsilon_{ij}, \forall k \neq j \,\bigg|\, \epsilon_{ij} = \tilde{\epsilon}\bigg) f_{\epsilon_{ij}}(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= \int_{-\infty}^{\infty} \prod_{k \neq j} \mathrm{P}\bigg(\epsilon_{ik} < \beta \log\left(\frac{W_j(X)}{W_k(X)}\right) + a_j(X) - a_k(X) + \epsilon_{ij} \,\bigg|\, \epsilon_{ij} = \tilde{\epsilon}\bigg) f_{\epsilon_{ij}}(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= \int_{-\infty}^{\infty} \prod_{k \neq j} F_{\epsilon}\bigg(\beta \log\left(\frac{W_j(X)}{W_k(X)}\right) + a_j(X) - a_k(X) + \tilde{\epsilon}\bigg) f_{\epsilon}(\tilde{\epsilon}) d\tilde{\epsilon}.
\end{aligned}
$$

Here, the second-to-last equality holds because $\{\epsilon_{ij}\}_{i,j}$ are independent, and the final equality holds because $\{\epsilon_{ij}\}_{i,j}$ are identically distributed. The mass of workers with skills X at firm j is:

$$
S_j(X) = \int \left(\int_{-\infty}^{\infty} \prod_{k \neq j} F_{\epsilon} \left(\beta \log\left(\frac{W_j(X)}{W_k(X)}\right) + a_j(X) - a_k(X) + \tilde{\epsilon}\right) f_{\epsilon}(\tilde{\epsilon}) d\tilde{\epsilon} \right) f_{\beta,X}(\beta, X) d\beta.
$$

Under a logit error structure, where $F_\epsilon(\epsilon) = \exp(-\exp(-\epsilon))$, the choice probability becomes:

$$\mathrm{P}(j(i) = j|\beta, X) = \frac{\exp\big(\beta \log W_j(X) + a_j(X)\big)}{\sum_{k=1}^{J} \exp\big(\beta \log W_k(X) + a_k(X)\big)}.$$

Defining $I(\beta, X) = \sum_{k=1}^{J} \exp\big(\beta \log W_k(X) + a_k(X)\big)$ to be the worker's wage index, I obtain:

$$S_j(X) = \int \frac{1}{I(\beta, X)} \exp\big(\beta \log W_j(X) + a_j(X)\big) f_{\beta,X}(\beta, X) d\beta.$$

Using this parameterization for F_ε, I will now derive the firm-specific labor supply elasticity $\frac{\partial \log S_j(X)}{\partial \log W_j(X)}$, as well as the higher-order derivatives of $\log S_j(X)$. Consider the following property.

Property A.1. *Assume $F_\epsilon(\epsilon) = \exp(-\exp(-\epsilon))$. From the perspective of a strategically small firm, the first five derivatives of (log) labor supply curve with respect to the (log) wage are:*

$$\begin{aligned}
\frac{\partial \log S_j(X)}{\partial \log W_j(X)} &= \mathrm{E}\big(\beta|X, j(i) = j\big) \\
\frac{\partial^2 \log S_j(X)}{\partial \left[\log W_j(X)\right]^2} &= \mathrm{Var}\big(\beta|X, j(i) = j\big) \\
\frac{\partial^3 \log S_j(X)}{\partial \left[\log W_j(X)\right]^3} &= \mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i) = j)]^3 \big| X, j(i) = j\big) \\
\frac{\partial^4 \log S_j(X)}{\partial \left[\log W_j(X)\right]^4} &= \mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i) = j)]^4 \big| X, j(i) = j\big) - 3\,\mathrm{Var}\big(\beta|X, j(i) = j\big)^2 \\
\frac{\partial^5 \log S_j(X)}{\partial \left[\log W_j(X)\right]^5} &= \mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i) = j)]^5 \big| X, j(i) = j\big) \\
&\quad - 10\,\mathrm{Var}\big(\beta|X, j(i) = j\big)\,\mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i) = j)]^3 \big| X, j(i) = j\big).
\end{aligned}$$

Proof. If the firm views itself as strategically small, then it does not internalize the impact of its own wage on each worker's wage index. Specifically, it sets $\partial I(\beta, X)/\partial W_j(X) = 0$ for all β and X. Under this assumption, the following relationships hold for any function $g : \mathbb{R} \to \mathbb{R}$:

$$\begin{aligned}
\frac{\partial\, \mathrm{E}(g(\beta)\, \mathrm{P}(j(i) = j|\beta, X) \mid X)}{\partial \log W_j(X)} &= \int \frac{\partial}{\partial \log W_j(X)} \left[\frac{g(\beta) \exp\big(\beta \log W_j(X) + a_j(X)\big)}{I(\beta, X)}\right] f_{\beta|X}(\beta|X) d\beta \\
&= \int \frac{\beta g(\beta) \exp\big(\beta \log W_j(X) + a_j(X)\big)}{I(\beta, X)} f_{\beta|X}(\beta|X) d\beta \\
&= \mathrm{E}\left(\beta g(\beta)\, \mathrm{P}(j(i) = j|\beta, X) \mid X\right). \qquad \text{(A.1)}
\end{aligned}$$

$$
\begin{aligned}
\frac{\mathrm{E}(g(\beta)\,\mathrm{P}(j(i)=j|\beta,X)\mid X)}{\mathrm{P}(j(i)=j|X)} &= \frac{\sum_{k=1}^{J}\mathrm{E}(g(\beta)\,\mathrm{P}(j(i)=j|\beta,X)\mid X, j(i)=k)\times \mathrm{P}(j(i)=k|X)}{\mathrm{P}(j(i)=j|X)}\\
&= \frac{\mathrm{E}(g(\beta)\mid X, j(i)=j)\times \mathrm{P}(j(i)=j|X)+\sum_{k\neq j}0\times \mathrm{P}(j(i)=k|X)}{\mathrm{P}(j(i)=j|X)}\\
&= \mathrm{E}(g(\beta)\mid X, j(i)=j)\,. \qquad \text{(A.2)}
\end{aligned}
$$

To ease notation, let $s_j(X) = \log S_j(X)$ and $w_j(X) = \log W_j(X)$. Also, define $\tau_{jX,s}$ such that:

$$
\tau_{jX,s}(w) = \mathrm{E}\big(\beta^s \exp\big(\beta w + a_j(X)\big)/I(\beta,X)|X\big).
$$

Equation (A.1) implies $\partial\tau_{jX,s}(w_j(X))/\partial w_j(X) = \tau_{jX,s+1}(w_j(X))$, while equation (A.2) implies $\tau_{jX,s}(w_j(X))/\tau_{jX,0}(w_j(X)) = \mathrm{E}(\beta^s|X, j(i)=j)$. Using these properties, I derive the following:

<u>*First Derivative*</u>. The first derivative of the labor supply curve, $\partial s_j(X)/\partial w_j(X)$, is:

$$
\frac{\partial s_j(X)}{\partial w_j(X)} = \frac{\partial \log\big(\tau_{jX,0}(w_j(X))\big)}{\partial w_j(X)} = \frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))} = \mathrm{E}\big(\beta|X, j(i)=j\big).
$$

<u>*Second Derivative*</u>. The second derivative of the labor supply curve, $\partial^2 s_j(X)/\partial w_j^2(X)$, is:

$$
\begin{aligned}
\frac{\partial^2 s_j(X)}{\partial w_j^2(X)} &= \frac{\partial\big(\tau_{jX,1}(w_j(X))/\tau_{jX,0}(w_j(X))\big)}{\partial w_j(X)}\\
&= \frac{\tau_{jX,2}(w_j(X))\tau_{jX,0}(w_j(X)) - \tau_{jX,1}^2(w_j(X))}{\tau_{jX,0}^2(w_j(X))} = \frac{\tau_{jX,2}(w_j(X))}{\tau_{jX,0}(w_j(X))} - \left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^2\\
&= \mathrm{E}\big(\beta^2|X, j(i)=j\big) - \mathrm{E}\big(\beta|X, j(i)=j\big)^2 = \mathrm{Var}\big(\beta|X, j(i)=j\big).
\end{aligned}
$$

<u>*Third Derivative*</u>. The third derivative of the labor supply curve, $\partial^3 s_j(X)/\partial w_j^3(X)$, is:

$$
\begin{aligned}
\frac{\partial^3 s_j(X)}{\partial w_j^3(X)} &= \frac{\partial}{\partial w_j(X)}\left(\frac{\tau_{jX,2}(w_j(X))\tau_{jX,0}(w_j(X)) - \tau_{jX,1}^2(w_j(X))}{\tau_{jX,0}^2(w_j(X))}\right)\\
&= \frac{\tau_{jX,3}(w_j(X))\tau_{jX,0}^3(w_j(X)) - 3\tau_{jX,2}(w_j(X))\tau_{jX,1}(w_j(X))\tau_{jX,0}^2(w_j(X)) + 2\tau_{jX,1}^3(w_j(X))\tau_{jX,0}(w_j(X))}{\tau_{jX,0}^4(w_j(X))}\\
&= \frac{\tau_{jX,3}(w_j(X))}{\tau_{jX,0}(w_j(X))} - 3\left(\frac{\tau_{jX,2}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right) + 2\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^3\\
&= \mathrm{E}\big(\beta^3|X, j(i)=j\big) - 3\,\mathrm{E}\big(\beta|X, j(i)=j\big)\,\mathrm{E}\big(\beta^2|X, j(i)=j\big) + 2\,\mathrm{E}\big(\beta|X, j(i)=j\big)^3\\
&= \mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i)=j)]^3\big|X, j(i)=j\big).
\end{aligned}
$$

Fourth Derivative. The fourth derivative of the labor supply curve, $\partial^4 s_j(X)/\partial w_j^4(X)$, is:

$$
\begin{aligned}
\frac{\partial^4 s_j(X)}{\partial w_j^4(X)} &= \frac{\partial}{\partial w_j(X)}\left(\frac{\tau_{jX,3}(w_j(X))\tau_{jX,0}^3(w_j(X)) - 3\tau_{jX,2}(w_j(X))\tau_{jX,1}(w_j(X))\tau_{jX,0}^2(w_j(X)) + 2\tau_{jX,1}^3(w_j(X))\tau_{jX,0}(w_j(X))}{\tau_{jX,0}^4(w_j(X))}\right) \\
&= \frac{\tau_{jX,4}(w_j(X))\tau_{jX,0}^7(w_j(X)) - 4\tau_{jX,3}(w_j(X))\tau_{jX,1}(w_j(X))\tau_{jX,0}^6(w_j(X)) - 3\tau_{jX,2}^2(w_j(X))\tau_{jX,0}^6(w_j(X))}{\tau_{jX,0}^8(w_j(X))} \\
&\quad + \frac{12\tau_{jX,2}(w_j(X))\tau_{jX,1}^2(w_j(X))\tau_{jX,0}^5(w_j(X)) - 6\tau_{jX,1}^4(w_j(X))\tau_{jX,0}^4(w_j(X))}{\tau_{jX,0}^8(w_j(X))} \\
&= \frac{\tau_{jX,4}(w_j(X))}{\tau_{jX,0}(w_j(X))} - 4\left(\frac{\tau_{jX,3}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right) - 3\left(\frac{\tau_{jX,2}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^2 \\
&\quad + 12\left(\frac{\tau_{jX,2}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^2 - 6\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^4 \\
&= \mathrm{E}\big(\beta^4|X, j(i)=j\big) - 4\,\mathrm{E}\big(\beta|X, j(i)=j\big)\,\mathrm{E}\big(\beta^3|X, j(i)=j\big) - 3\,\mathrm{E}\big(\beta^2|X, j(i)=j\big)^2 \\
&\quad + 12\,\mathrm{E}\big(\beta|X, j(i)=j\big)^2\,\mathrm{E}\big(\beta^2|X, j(i)=j\big) - 6\,\mathrm{E}\big(\beta|X, j(i)=j\big)^4 \\
&= \mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i)=j)]^4\big|X, j(i)=j\big) - 3\,\mathrm{Var}\big(\beta|X, j(i)=j\big)^2.
\end{aligned}
$$

Fifth Derivative. The fifth derivative of the labor supply curve, $\partial^5 s_j(X)/\partial w_j^5(X)$, is:

$$
\begin{aligned}
\frac{\partial^5 s_j(X)}{\partial w_j^5(X)} &= \frac{\partial}{\partial w_j(X)}\left(\frac{\tau_{jX,4}(w_j(X))\tau_{jX,0}^7(w_j(X)) - 4\tau_{jX,3}(w_j(X))\tau_{jX,1}(w_j(X))\tau_{jX,0}^6(w_j(X)) - 3\tau_{jX,2}^2(w_j(X))\tau_{jX,0}^6(w_j(X))}{\tau_{jX,0}^8(w_j(X))}\right. \\
&\qquad\left. + \frac{12\tau_{jX,2}(w_j(X))\tau_{jX,1}^2(w_j(X))\tau_{jX,0}^5(w_j(X)) - 6\tau_{jX,1}^4(w_j(X))\tau_{jX,0}^4(w_j(X))}{\tau_{jX,0}^8(w_j(X))}\right) \\
&= \frac{\tau_{jX,5}(w_j(X))\tau_{jX,0}^{15}(w_j(X))}{\tau_{jX,0}^{16}(w_j(X))} - \frac{60\tau_{jX,2}(w_j(X))\tau_{jX,1}^3(w_j(X))\tau_{jX,0}^{12}(w_j(X)) - 24\tau_{jX,1}^5(w_j(X))\tau_{jX,0}^{11}(w_j(X))}{\tau_{jX,0}^{16}(w_j(X))} \\
&\quad - \frac{5\tau_{jX,4}(w_j(X))\tau_{jX,1}(w_j(X))\tau_{jX,0}^{14}(w_j(X)) - 20\tau_{jX,3}(w_j(X))\tau_{jX,1}^2(w_j(X))\tau_{jX,0}^{13}(w_j(X))}{\tau_{jX,0}^{16}(w_j(X))} \\
&\quad - \frac{10\tau_{jX,3}(w_j(X))\tau_{jX,2}(w_j(X))\tau_{jX,0}^{14}(w_j(X)) - 30\tau_{jX,2}^2(w_j(X))\tau_{jX,1}(w_j(X))\tau_{jX,0}^{13}(w_j(X))}{\tau_{jX,0}^{16}(w_j(X))} \\
&= \frac{\tau_{jX,5}(w_j(X))}{\tau_{jX,0}(w_j(X))} - 5\left(\frac{\tau_{jX,4}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right) + 20\left(\frac{\tau_{jX,3}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^2 + 24\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^5 \\
&\quad - 10\left(\frac{\tau_{jX,3}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)\left(\frac{\tau_{jX,2}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right) + 30\left(\frac{\tau_{jX,2}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^2\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right) - 60\left(\frac{\tau_{jX,2}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^3 \\
&= \mathrm{E}\big(\beta^5|X, j(i)=j\big) - 5\,\mathrm{E}\big(\beta|X, j(i)=j\big)\,\mathrm{E}\big(\beta^4|X, j(i)=j\big) + 20\,\mathrm{E}\big(\beta|X, j(i)=j\big)^2\,\mathrm{E}\big(\beta^3|X, j(i)=j\big) \\
&\quad - 10\,\mathrm{E}\big(\beta^2|X, j(i)=j\big)\,\mathrm{E}\big(\beta^3|X, j(i)=j\big) + 30\,\mathrm{E}\big(\beta|X, j(i)=j\big)\,\mathrm{E}\big(\beta^2|X, j(i)=j\big)^2 \\
&\quad - 60\,\mathrm{E}\big(\beta|X, j(i)=j\big)^3\,\mathrm{E}\big(\beta^2|X, j(i)=j\big) + 24\,\mathrm{E}\big(\beta|X, j(i)=j\big)^5 \\
&= \mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i)=j)]^5\big|X, j(i)=j\big) - 10\,\mathrm{Var}\big(\beta|X, j(i)=j\big)\,\mathrm{E}\big([\beta - \mathrm{E}(\beta|X, j(i)=j)]^3\big|X, j(i)=j\big).
\end{aligned}
$$

□

A.2. *Equilibrium Wage Equation*

In equilibrium, each firm chooses a set of wages $\{W_j(\chi,\varphi)\}_{\chi,\varphi}$ to maximize profit Π_j, where:

$$\Pi_j = T_j\left(\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(\int \varphi D_j(\chi,\varphi)d\varphi\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}} - \sum_{\chi\in\mathcal{X}}\left(\int W_j(\chi,\varphi)D_j(\chi,\varphi)d\varphi\right).$$

Let $L_j(\chi,\varphi)$ denote the equilibrium employment of skill group $X=(\chi,\varphi)$ at firm j, defined as the amount of labor hired when $D_j(\chi,\varphi)=S_j(\chi,\varphi)$. The firm's first order condition is:

$$\begin{aligned}\frac{\partial\Pi_j}{\partial W_j(\chi,\varphi)} =&\varphi T_j(1-\alpha_j)\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j-1}\left(\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-1}\frac{\partial L_j(\chi,\varphi)}{\partial W_j(\varphi,\chi)}\\ &- L_j(\chi,\varphi) - W_j(\chi,\varphi)\frac{\partial L_j(\chi,\varphi)}{\partial W_j(\varphi,\chi)} = 0, \quad \text{for all } (\chi,\varphi)\in\mathcal{X}\times\mathbb{R},\end{aligned}$$

where $L_j^{\text{eff}}(\chi)=\int\varphi L_j(\chi,\varphi)d\varphi$ denotes the efficiency units of labor for a given skill type χ. Defining $\boldsymbol{\varepsilon}_j(X)$ as the labor supply elasticity $\partial\log L_j(X)/\partial\log W_j(X)$, this condition becomes:

$$W_j(\chi,\varphi) = \frac{\boldsymbol{\varepsilon}_j(\chi,\varphi)}{1+\boldsymbol{\varepsilon}_j(\chi,\varphi)}\times\varphi T_j(1-\alpha_j)\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j-1}\left(\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-1}.$$

The wage markdown at firm j for workers with skills X is: $\frac{W_j(X)}{\partial Y_j/\partial L_j(X)} = \boldsymbol{\varepsilon}_j(X)/(1+\boldsymbol{\varepsilon}_j(X))$.

Lemmas 1 and 2 establish the existence of an equilibrium corresponding to a unique set of profit-maximizing wages $[W_j(X)]_X$ for each firm j. In this equilibrium, the firm's problem has an interior solution, which satisfies the first-order condition. Taking logs, the condition is:

$$\begin{aligned}w_j(\chi,\varphi) =& \log\varphi + \log T_j + \log(1-\alpha_j) + \log\theta_{j\chi} - \log\left(1+\boldsymbol{\varepsilon}_j^{-1}(\chi,\varphi)\right)\\ &- (1-\rho_j)\log L_j^{\text{eff}}(\chi) + \frac{1-\alpha_j-\rho_j}{\rho_j}\log\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}.\end{aligned}$$

A.3. *Firm Production*

After plugging in the labor supply constraint, a firm j's value added from hiring workers is:

$$Y_j = T_j\left(\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(\int\varphi L_j(\chi,\varphi)d\varphi\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}}.$$

For workers with skills $X = (\chi, \varphi)$, the marginal product of labor at firm j is $\frac{\partial Y_j}{\partial L_j(\chi,\varphi)}$, where:

$$\frac{\partial Y_j}{\partial L_j(\chi,\varphi)} = \varphi T_j(1-\alpha_j)\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j - 1}\left(\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-1}.$$

To better understand the relationship between skill types χ and skill levels φ in the production function, it will be useful to define the standardized unit of labor at the firm to be N_j, where:

$$N_j = \left(\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(\int\varphi L_j(\chi,\varphi)d\varphi\right)^{\rho_j}\right)^{1/\rho_j}.$$

For any skill profile $X = (\chi, \varphi)$, the elasticity of N_j with respect to $L_j(\chi, \varphi)$ is defined by:

$$\frac{\partial \log N_j}{\partial \log L_j(\chi,\varphi)} = \underbrace{\frac{\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}}{\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}}}_{\partial\log N_j/\partial\log L_j^{\text{eff}}(\chi)} \times \underbrace{\frac{\varphi L_j(\chi,\varphi)}{\int\varphi' L_j(\chi,\varphi')d\varphi'}}_{\partial\log L_j^{\text{eff}}(\chi)/\partial\log L_j(\chi,\varphi)}.$$

Observe that these elasticities sum up to one: $\sum_{\chi\in\mathcal{X}}\int\frac{\partial\log N_j}{\partial\log L_j(\chi,\varphi)}d\varphi = 1$. They can therefore be interpreted as the share of effective labor contributed to firm j by workers with skills (χ, φ).

Property A.2. *Suppose that $\alpha_j \in (0,1)$ and $\rho_j < 1$. Then a firm's value added Y_j is concave in labor inputs. Additionally, Y_j exhibits decreasing marginal returns within each skill type:*

$$\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0, \quad \textit{for all } \chi\in\mathcal{X} \textit{ and } \varphi,\varphi'\in\mathbb{R}_{++},$$

and complementarity (or substitutability) between skill types depending on the value of $1-\rho_j$*:*

$$\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi',\varphi')}\begin{cases} > 0 & \textit{if } \alpha_j < 1-\rho_j \\ < 0 & \textit{if } \alpha_j > 1-\rho_j, \end{cases} \quad \textit{for all } \chi\neq\chi' \textit{ and } \varphi,\varphi'\in\mathbb{R}_{++}.$$

Proof. To show that Y_j is concave, first note that it is continuous, strictly increasing, and homogeneous of degree $1-\alpha_j \in (0,1]$. Therefore, by Jehle & Reny (2011), Theorem 3.1, it is sufficient to show that Y_j is strictly quasiconcave. Note that Y_j can be written as a composition of two functions, $Y_j = h\circ g(\{L_j^{\text{eff}}(\chi)\}_\chi)$, where $g : \mathbb{R}_{++}^{|\mathcal{X}|}\to\mathbb{R}$ and $h:\mathbb{R}\to\mathbb{R}$ are defined by:

$$g(\{L_j^{\text{eff}}(\chi)\}_\chi) = \sum_{\chi\in\mathcal{X}}\theta_{j\chi}\big(L_j^{\text{eff}}(\chi)\big)^{\rho_j} \quad\text{and}\quad h(x) = T_j x^{(1-\alpha_j)/\rho_j}.$$

The function g has a diagonal Hessian matrix with entries: $\frac{\partial^2 g}{\partial[L_j^{\text{eff}}(\chi)]^2} = -\rho_j(1-\rho_j)\theta_{j\chi}\big(L_j^{\text{eff}}(\chi)\big)^{\rho_j-2}$.

I now analyze three cases. First, if $\rho_j \in (0,1)$, then h is strictly increasing and g is strictly concave, because its Hessian is negative definite. It follows that Y_j is strictly quasiconcave in $\{L_j^{\text{eff}}(\chi)\}_\chi$, as it is a strictly increasing transformation of a strictly concave function. Second, if $\rho_j < 0$, then h is strictly decreasing and g is strictly convex, because its Hessian is positive definite. Therefore, Y_j is still strictly quasiconcave in $\{L_j^{\text{eff}}(\chi)\}_\chi$, as it is a strictly decreasing transformation of a strictly convex function. Third, if ρ_j approaches zero, then Y_j reduces to a Cobb-Douglas production function that is strictly concave in $\{L_j^{\text{eff}}(\chi)\}_\chi$. In all three cases, Y_j is shown to be strictly quasiconcave in $\{L_j^{\text{eff}}(\chi)\}_\chi$. By Jehle & Reny (2011), Theorem 3.1, this implies that Y_j is concave in $\{L_j^{\text{eff}}(\chi)\}_\chi$. Furthermore, since $L_j^{\text{eff}}(\chi)$ is a linear function of $\{L_j(\chi,\varphi)\}_\varphi$, it follows that a firm's value added Y_j is a concave function of all labor inputs.

To show that $\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0$, fix some $\chi \in \mathcal{X}$. For any skill levels φ and φ', I write:

$$
\begin{aligned}
\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} &= \varphi T_j(1-\alpha_j)\theta_{j\chi}\Bigg[\varphi'(\rho_j-1)\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j-2}\left(\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-1} \\
&\quad + \varphi'(1-\alpha_j-\rho_j)\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{2\rho_j-2}\left(\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-2}\Bigg] \\
&= \varphi'\left(L_j^{\text{eff}}(\chi)\right)^{-1}\frac{\partial Y_j}{\partial L_j(\chi,\varphi)}\left[(\rho_j-1)+(1-\alpha_j-\rho_j)\frac{\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}}{\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}}\right].
\end{aligned}
$$

With positive employment, $\varphi'[L_j^{\text{eff}}(\chi)]^{-1}\frac{\partial Y_j}{\partial L_j(\chi,\varphi)} > 0$. So, $\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0$ if and only if:

$$
\begin{aligned}
0 &> (\rho_j-1)+(1-\alpha_j-\rho_j)\frac{\partial \log N_j}{\partial \log L_j^{\text{eff}}(\chi)} \\
&= (\rho_j-1)\left[1-\frac{\partial \log N_j}{\partial \log L_j^{\text{eff}}(\chi)}\right]-\alpha_j\frac{\partial \log N_j}{\partial \log L_j^{\text{eff}}(\chi)} > (\rho_j-1)-\alpha_j\frac{\partial \log N_j}{\partial \log L_j^{\text{eff}}(\chi)},
\end{aligned}
$$

where the final inequality follows because $\partial \log N_j/\partial \log L_j^{\text{eff}}(\chi)$ is bounded between 0 and 1, and $\rho_j < 1$ by assumption. Re-arranging terms, this inequality may be re-written as follows:

$$
\rho_j < 1+\alpha_j\frac{\partial \log N_j}{\partial \log L_j^{\text{eff}}(\chi)}.
$$

Since this holds trivially when $\alpha_j \in (0,1)$ and $\rho_j < 1$, I conclude that $\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0$.[1]

[1] Note that values of ρ_j above one can also satisfy this inequality, particularly if the returns to scale are small.

Finally, fix $\chi, \chi' \in \mathcal{X}$, where $\chi \neq \chi'$, and $\varphi, \varphi' \in \mathbb{R}$. The derivative $\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi',\varphi')}$ is:

$$\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi',\varphi')} = (1-\alpha_j-\rho_j)\underbrace{\varphi\varphi' T_j(1-\alpha_j)\theta_{j\chi}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi)L_j^{\text{eff}}(\chi')\right)^{\rho_j-1}\left(\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-2}}_{(*)}.$$

Observe that the term $(*)$ is strictly positive. Thus, the sign of the derivative is pinned down by $1-\alpha_j-\rho_j$. It is positive whenever $\rho_j < 1-\alpha_j$, and it is negative whenever $\rho_j > 1-\alpha_j$. □

A.4. Firm Profit

After plugging in the labor supply constraint, the profit function Π_j for a firm j is defined by:

$$\Pi_j = Y_j - \int W_j(X)L_j(X)dX.$$

In equilibrium, wages are determined by setting the derivatives $\partial\Pi_j/\partial W_j(X)$ to zero, where:

$$\frac{\partial \Pi_j}{\partial W_j(X)} = \left(\frac{\partial Y_j}{\partial L_j(X)} - W_j(X)\right)\frac{\partial L_j(X)}{\partial W_j(X)} - L_j(X).$$

To understand when profit is a concave function of wages, I evaluate the Hessian matrix of Π_j.

Property A.3. *The Hessian of the profit function Π_j is given by $H_{\Pi_j} = \mathbf{A}'(H_{Y_j})\mathbf{A} + \mathbf{B}$, where $H_{Y_j} = [\frac{\partial^2 Y_j}{\partial L_j(X)\partial L_j(X')}]_{X,X'}$ is the Hessian of Y_j, and $\mathbf{A}$ and $\mathbf{B}$ are diagonal matrices with entries:*

$$\mathbf{A}[X,X] = \frac{\partial L_j(X)}{\partial W_j(X)} \quad \textit{and} \quad \mathbf{B}[X,X] = \frac{L_j(X)}{W_j(X)}\left(\frac{\partial^2 \log L_j(X)}{\partial\left[\log W_j(X)\right]^2}\boldsymbol{\varepsilon}_j^{-1}(X) - (\boldsymbol{\varepsilon}_j(X)+1)\right).$$

Proof. The diagonal entries of the matrix H_{Π_j}, denoted by $\mathrm{H}_{\Pi_j}[X,X] = \frac{\partial^2 \Pi_j}{\partial W_j^2(X)}$, are given by:

$$\begin{aligned}\frac{\partial^2 \Pi_j}{\partial W_j^2(X)} &= \frac{\partial^2 Y_j}{\partial L_j^2(X)} \times \left(\frac{\partial L_j(X)}{\partial W_j(X)}\right)^2 + \frac{\partial Y_j}{\partial L_j(X)} \times \frac{\partial^2 L_j(X)}{\partial W_j^2(X)} - \frac{\partial L_j(X)}{\partial W_j(X)} - \frac{\partial L_j(X)}{\partial W_j(X)} - W_j(X)\frac{\partial^2 L_j(X)}{\partial W_j^2(X)} \\ &= \frac{\partial^2 Y_j}{\partial L_j^2(X)} \times \left(\frac{\partial L_j(X)}{\partial W_j(X)}\right)^2 + \left(\frac{\partial Y_j}{\partial L_j(X)} - W_j(X)\right)\frac{\partial^2 L_j(X)}{\partial W_j^2(X)} - 2\frac{\partial L_j(X)}{\partial W_j(X)}.\end{aligned}$$

If the first-order condition binds, then $\frac{\partial Y_j}{\partial L_j(X)} - W_j(X) = W_j(X)/\boldsymbol{\varepsilon}_j(X)$, which implies that:

$$\begin{aligned}
\frac{\partial^2 \Pi_j}{\partial W_j^2(X)} &= \frac{\partial^2 Y_j}{\partial L_j^2(X)} \times \left(\frac{\partial L_j(X)}{\partial W_j(X)}\right)^2 + \frac{W_j(X)}{\boldsymbol{\varepsilon}_j(X)} \times \frac{\partial^2 L_j(X)}{\partial W_j^2(X)} - 2\frac{\partial L_j(X)}{\partial W_j(X)} \\
&= \frac{\partial^2 Y_j}{\partial L_j^2(X)} \times \left(\frac{\partial L_j(X)}{\partial W_j(X)}\right)^2 + \frac{L_j(X)}{W_j(X)} \left(\frac{W_j^2(X)}{\boldsymbol{\varepsilon}_j(X) L_j(X)} \times \frac{\partial^2 L_j(X)}{\partial W_j^2(X)} - 2\,\boldsymbol{\varepsilon}_j(X)\right) \\
&= \frac{\partial^2 Y_j}{\partial L_j^2(X)} \times \left(\frac{\partial L_j(X)}{\partial W_j(X)}\right)^2 + \frac{L_j(X)}{W_j(X)} \left(\frac{\partial^2 \log L_j(X)}{\partial \left[\log W_j(X)\right]^2}\, \boldsymbol{\varepsilon}_j^{-1}(X) - (\boldsymbol{\varepsilon}_j(X) + 1)\right),
\end{aligned}$$

where the final equality relies on the observation that $\partial^2 \log L_j(X)/\partial[\log W_j(X)]^2$ equals:

$$\begin{aligned}
\frac{\partial^2 \log L_j(X)}{\partial \left[\log W_j(X)\right]^2} &= \frac{W_j^2(X)}{L_j(X)} \times \frac{\partial^2 L_j(X)}{\partial W_j^2(X)} - \left(\frac{\partial L_j(X)}{\partial W_j(X)}\right)^2 \times \frac{W_j^2(X)}{L_j^2(X)} + \frac{\partial L_j(X)}{\partial W_j(X)} \times \frac{W_j(X)}{L_j(X)} \\
&= \frac{W_j^2(X)}{L_j(X)} \times \frac{\partial^2 L_j(X)}{\partial W_j^2(X)} - \boldsymbol{\varepsilon}_j^2(X) + \boldsymbol{\varepsilon}_j(X).
\end{aligned}$$

The off-diagonal entries of H_{Π_j}, denoted by $\mathrm{H}_{\Pi_j}[X, X'] = \frac{\partial^2 \Pi_j}{\partial W_j(X) \partial W_j(X')}$ for $X \neq X'$, equal:

$$\frac{\partial^2 \Pi_j}{\partial W_j(X) \partial W_j(X')} = \frac{\partial^2 Y_j}{\partial L_j(X) \partial L_j(X')} \times \frac{\partial L_j(X)}{\partial W_j(X)} \times \frac{\partial L_j(X')}{\partial W_j(X')}.$$

Taken together, I can write $\mathrm{H}_{\Pi_j} = \mathbf{A}'(\mathrm{H}_{Y_j})\mathbf{A} + \mathbf{B}$, where H_{Y_j}, $\mathbf{A}$, and $\mathbf{B}$ are all defined above.

□

I now highlight two notable special cases where Π_j is a concave function of wages $\{W_j(X)\}_X$.

Example (Homogeneous β). Suppose β is constant within each skill group: $\mathrm{Var}(\beta_i | X_i = X) = 0$. Then by Property A.1, $\partial^2 \log L_j(X)/\partial[\log W_j(X)]^2 = \mathrm{Var}(\beta_i | X_i = X, j(i) = j) = 0$. In this case, $\mathbf{B}$ is a negative definite matrix with entries $\mathbf{B}[X, X] = -[L_j(X)/W_j(X)] \times [\boldsymbol{\varepsilon}_j(X) + 1]$. Since H_{Y_j} is negative semidefinite by Property A.2, H_{Π_j} is also negative semidefinite, because:

$$\begin{aligned}
v'\mathrm{H}_{\Pi_j} v &= v'\left(\mathbf{A}'(\mathrm{H}_{Y_j})\mathbf{A} + \mathbf{B}\right)v \\
&= v'\mathbf{A}'(\mathrm{H}_{Y_j})\mathbf{A}v + v'\mathbf{B}v \\
&= (\mathbf{A}v)'(\mathrm{H}_{Y_j})(\mathbf{A}v) + v'\mathbf{B}v \leq 0,
\end{aligned}$$

for any vector v. Therefore, the firm's profit function Π_j must be concave in wages $\{W_j(X)\}_X$.

Example (Firm is a Price-taker). Suppose that the firm does not exercise any wage-setting power, setting $W_j(X) = \frac{\partial Y_j}{\partial L_j(X)}$ for all skill groups X. In this case, $\mathbf{B}$ will reduce to a negative definite matrix with entries $\mathbf{B}[X, X] = -2[\partial L_j(X)/\partial W_j(X)]$. Just as in the previous case, this implies that H_{Π_j} is negative semidefinite. So, the firm's profit function must be concave.

A.5. Worker Rents

Worker rents R_i^w are defined so that $u_{ij}\big(W_j(X_i) - R_i^w, a_j(X_i)\big) = \max_{k \neq j(i)} u_{ik}\big(W_k(X_i), a_k(X_i)\big)$. As a share of a worker's earnings, these rents can be expressed as a strictly increasing function of the utility difference between a worker i's chosen firm $j(i)$ and the next-best alternative:

$$\frac{R_i^w}{W_{j(i)}(X_i)} = 1 - \exp\left(-\frac{u_{ij(i)}\big(W_{j(i)}(X_i), a_{j(i)}(X_i)\big) - \max_{k \neq j(i)} u_{ik}\big(W_k(X_i), a_k(X_i)\big)}{\beta}\right).$$

I now derive the distribution of these rent shares among workers with skills $X_i = X$ at firm j. I show that, if β is homogeneous, then these shares are independent of a worker's skills and firm.

Property A.4. *For workers with skills X at firm j, the density of $\omega_i = R_i^w / W_{j(i)}(X_i)$ is given by:*

$$f_{\omega|X,j}(\omega | X_i = X, j(i) = j) = \mathrm{E}\Big(\beta_i (1-\omega)^{\beta_i - 1} \;\Big|\; X_i = X, j(i) = j\Big).$$

Whenever β is homogeneous, ω_i will be independent of both a worker's skills X_i and firm $j(i)$.

Proof. Let $W_j(X)$ be the wage that a firm j pays workers with skills X. For any $W \leq W_j(X)$, the density of these workers who would be willing to accept their current firm at wage W is:

$$\frac{\partial L_j(X, W)}{\partial W} \times \frac{1}{L_j(X, W_j(X))}.$$

By the change-of-variables formula, the density of ω_i given $X_i = X$ and $j(i) = j$ is equal to:

$$\begin{aligned}
f_{\omega|X,j}(\omega | X_i = X, j(i) = j) &= -\frac{\partial L_j(X, (1-\omega) W_j(X))}{\partial \omega} \times \frac{1}{L_j(X, W_j(X))} \\
&= -\frac{\partial}{\partial \omega}\left(\int \frac{\exp\big(\beta \log\big((1-\omega)W_j(X)\big) + a_j(X)\big) f_{\beta|X}(\beta|X) f_X(X)}{I(\beta, X)} d\beta\right) \times \frac{1}{L_j(X, W_j(X))} \\
&= \left(\int \beta (1-\omega)^{\beta - 1} \frac{\exp\big(\beta \log W_j(X) + a_j(X)\big) f_{\beta|X}(\beta|X) f_X(X)}{I(\beta, X)} d\beta\right) \times \frac{1}{L_j(X, W_j(X))} \\
&= \frac{\mathrm{E}\Big(\beta (1-\omega)^{\beta-1} \times \exp\big(\beta \log W_j(X) + a_j(X)\big) / I(\beta, X) \;\Big|\; X\Big)}{\mathrm{E}\Big(\exp\big(\beta \log W_j(X) + a_j(X)\big) / I(\beta, X) \;\Big|\; X\Big)} \\
&= \mathrm{E}\Big(\beta_i (1-\omega)^{\beta_i - 1} \;\Big|\; X_i = X, j(i) = j\Big).
\end{aligned}$$

If β is homogeneous, then this density simplifies to $f_{\omega|X,j}(\omega | X_i = X, j(i) = j) = \beta(1-\omega)^{\beta-1}$. Since this function does not depend on a worker's skills $X_i = X$ or firm $j(i) = j$, I can write:

$$f_{\omega|X,j}(\omega | X_i = X, j(i) = j) = f_\omega(\omega), \quad \text{for all } X \text{ and } j.$$

□

When β is heterogeneous, the share of wages that workers earn as rents systematically varies across firms and skill groups. To characterize this variation, I compute the average rent share, $\mathrm{E}(R_i^w/W_{j(i)}(X_i)|X_i = X, j(i) = j)$, for workers with skills X at firm j. I also show that (fixing skills) workers who sort into higher paying firms receive a smaller share of their wage as rents.

Property A.5. *Among workers with skills X employed at firm j, the expected rent share is:*

$$\mathrm{E}\left(\frac{R_i^w}{W_{j(i)}(X_i)} \;\Big|\; X_i = X, j(i) = j\right) = \mathrm{E}\left(\frac{1}{1+\beta_i} \;\Big|\; X_i = X, j(i) = j\right).$$

If $\mathrm{Var}(\beta_i|X_i = X) > 0$, *then* $\mathrm{E}(R_i^w/W_{j(i)}(X_i)|X_i = X, j(i) = j)$ *is strictly decreasing in* $W_j(X)$.

Proof. Average worker rents $\mathrm{E}(R_i^w|X, j(i) = j)$ are derived by integrating $W_j(X) - W$ with respect to the density of workers who would be willing to accept their current job at wage W.

$$\begin{aligned}
\mathrm{E}(R_i^w|X, j(i) = j) &= \int_0^{W_j(X)} \left(W_j(X) - W\right)\left(\frac{\partial L_j(X,W)}{\partial W} \times \frac{1}{L_j(X, W_j(X))}\right) dW \\
&= -\frac{W_j(X)}{L_j(X, W_j(X))} \times \int_0^1 \omega \left(\frac{\partial}{\partial \omega} \int \frac{\exp\big(\beta \log\left((1-\omega)W_j(X)\right) + a_j(X)\big) f_{\beta|X}(\beta|X) f_X(X)}{I(\beta, X)} d\beta\right) d\omega \\
&= \frac{W_j(X)}{L_j(X, W_j(X))} \times \int_0^1 \omega \left(\int \frac{\beta(1-\omega)^{\beta-1} \exp\big(\beta \log W_j(X) + a_j(X)\big) f_{\beta|X}(\beta|X) f_X(X)}{I(\beta, X)} d\beta\right) d\omega.
\end{aligned}$$

The final equality relies on the assumption that firms view themselves to be strategically small in the market. Changing the order of integration, the average worker rents can be re-written as:

$$\begin{aligned}
\mathrm{E}(R_i^w|X, j(i) = j) &= \frac{W_j(X)}{L_j(X, W_j(X))} \times \int \frac{1}{I(\beta, X)} \exp\big(\beta \log W_j(X) + a_j(X)\big) f_{\beta|X}(\beta|X) f_X(X) \left(\int_0^1 \omega\beta(1-\omega)^{\beta-1} d\omega\right) d\beta \\
&= \frac{W_j(X)}{L_j(X, W_j(X))} \times \int \frac{1}{I(\beta, X)} \left(\frac{1}{1+\beta}\right) \exp\big(\beta \log W_j(X) + a_j(X)\big) f_{\beta|X}(\beta|X) f_X(X) d\beta \\
&= W_j(X) \times \frac{\mathrm{E}\Big(\frac{1}{1+\beta} \times \exp\big(\beta \log W_j(X) + a_j(X)\big)/I(\beta, X) \mid X\Big)}{\mathrm{E}\big(\exp\big(\beta \log W_j(X) + a_j(X)\big)/I(\beta, X) \mid X\big)} \\
&= W_j(X) \times \mathrm{E}\left(\frac{1}{1+\beta_i} \;\Big|\; X_i = X, j(i) = j\right).
\end{aligned}$$

To see how $\mathrm{E}(R_i^w/W_{j(i)}(X_i)|X_i = X, j(i) = j)$ depends on $W_j(X)$, I evaluate the elasticity:

$$\frac{\partial \log \mathrm{E}(R_i^w/W_{j(i)}(X_i)|X_i = X, j(i) = j)}{\partial \log W_j(X)} = \frac{\frac{\partial \mathrm{E}\left(\frac{1}{1+\beta_i}\middle|X_i = X, j(i) = j\right)}{\partial \log W_j(X)}}{\mathrm{E}\left(\frac{1}{1+\beta_i} \mid X_i = X, j(i) = j\right)} = \frac{\mathrm{Cov}\left(\beta_i, \frac{1}{1+\beta_i} \mid X_i = X, j(i) = j\right)}{\mathrm{E}\left(\frac{1}{1+\beta_i} \mid X_i = X, j(i) = j\right)}.$$

As long as $\mathrm{Var}(\beta_i|X_i = X) > 0$, the covariance term is strictly negative. Hence, fixing skills, workers at higher paying firms tend to receive a smaller share of their wage in the form of rents.

□

A.6. Employer Rents

Employer rents come in the form of excess profits that firms obtain by exploiting their wage-setting power. To calculate these rents, I consider a counterfactual setting where firms act as price-takers in the market, facing perfectly-elastic labor supply curves. I then define the rents at firm j to be the difference between the realized and counterfactual profits $R_j^e = \Pi_j - \Pi_j^{\text{pt}}$.

Property A.6. *Let Y_j^{pt} be a firm j's value added when it is a price taker. The firm's rents are:*

$$R_j^e = Y_j \times \left[\sum_{\chi \in \mathcal{X}} \int \left(\frac{1 + \alpha_j \boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)} \right) \frac{\partial \log N_j}{\partial \log L_j(\chi, \varphi)} d\varphi - \alpha_j \left(\frac{Y_j^{pt}}{Y_j} \right) \right].$$

If β is homogeneous, then all firms receive the same rents as a share of their profits, provided that they have the same returns to scale, α_j, and that labor is perfectly substitutable, $\rho_j = 1$.

Proof. For any firm j, the profit Π_j that the firm obtains in the monopsonistic equilibrium is:

$$\begin{aligned}
\Pi_j &= Y_j - \sum_{\chi \in \mathcal{X}} \left(\int \left(\frac{\boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)} \right) \left(\frac{\partial Y_j}{\partial L_j(\chi, \varphi)} \right) L_j(\chi, \varphi) d\varphi \right) \\
&= Y_j - \sum_{\chi \in \mathcal{X}} \left(\int \left(\frac{\boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)} \right) \varphi T_j (1 - \alpha_j) \theta_{j\chi} \left(L_j^{\text{eff}}(\chi) \right)^{\rho_j - 1} \left(\sum_{\chi' \in \mathcal{X}} \theta_{j\chi'} \left(L_j^{\text{eff}}(\chi') \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j} - 1} L_j(\chi, \varphi) d\varphi \right) \\
&= Y_j \times \left(1 - (1 - \alpha_j) \sum_{\chi \in \mathcal{X}} \left(\frac{\theta_{j\chi} \left(L_j^{\text{eff}}(\chi) \right)^{\rho_j}}{\sum_{\chi' \in \mathcal{X}} \theta_{j\chi'} \left(L_j^{\text{eff}}(\chi') \right)^{\rho_j}} \right) \left(\int \left(\frac{\boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)} \right) \left(\frac{\varphi L_j(\chi, \varphi)}{\int \varphi' L_j(\chi, \varphi') d\varphi'} \right) d\varphi \right) \right). \\
&= Y_j \times \left(1 - (1 - \alpha_j) \sum_{\chi \in \mathcal{X}} \int \left(\frac{\boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)} \right) \frac{\partial \log N_j}{\partial \log L_j(\chi, \varphi)} d\varphi \right) \\
&= Y_j \times \sum_{\chi \in \mathcal{X}} \int \left(\frac{1 + \alpha_j \boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)} \right) \frac{\partial \log N_j}{\partial \log L_j(\chi, \varphi)} d\varphi.
\end{aligned}$$

If the firm is a price-taker in the market, then its profit Π_j^{pt} is defined in the following way:

$$\Pi_j^{\text{pt}} = \max_{\{D_j^{\text{pt}}(\chi, \varphi)\}_{\chi, \varphi}} T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi D_j^{\text{pt}}(\chi, \varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}} - \sum_{\chi \in \mathcal{X}} \left(\int W_j^{\text{pt}}(\chi, \varphi) D_j^{\text{pt}}(\chi, \varphi) d\varphi \right).$$

Taking first-order conditions with respect to labor demand yields the following wage equation:

$$W_j^{\text{pt}}(\chi, \varphi) = \varphi T_j (1 - \alpha_j) \theta_{j\chi} \left(\int \varphi D_j^{\text{pt}}(\chi, \varphi) d\varphi \right)^{\rho_j - 1} \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi D_j^{\text{pt}}(\chi, \varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j} - 1}.$$

In equilibrium, the firm's labor demand equals its labor supply evaluated at the wage $W_j^{\text{pt}}(X)$. Let $L_j^{\text{pt}}(X)$ be the market-clearing quantity of labor, so that $D_j^{\text{pt}}(X) = S_j\big(X, W_j^{\text{pt}}(X)\big)$. I write:

$$\begin{aligned}
\Pi_j^{\text{pt}} &= T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi L_j^{\text{pt}}(\chi,\varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}} \times \left(1 - (1-\alpha_j) \sum_{\chi \in \mathcal{X}} \int \frac{\partial \log N_j}{\partial \log L_j(\chi,\varphi)} d\varphi \right) \\
&= \alpha_j T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi L_j^{\text{pt}}(\chi,\varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}} \\
&= \alpha_j Y_j^{\text{pt}},
\end{aligned}$$

where the second equality above follows because the elasticities $\frac{\partial \log N_j}{\partial \log L_j(\chi,\varphi)}$ aggregate to one. Given these expressions for the firm's realized and counterfactual profits, Π_j and Π_j^{pt}, I obtain:

$$R_j^e = Y_j \times \left[\sum_{\chi \in \mathcal{X}} \int \left(\frac{1 + \alpha_j \boldsymbol{\varepsilon}_j(\chi,\varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi,\varphi)} \right) \frac{\partial \log N_j}{\partial \log L_j(\chi,\varphi)} d\varphi - \alpha_j \left(\frac{Y_j^{\text{pt}}}{Y_j} \right) \right].$$

If β is homogeneous across workers, then the labor supply elasticity $\boldsymbol{\varepsilon}_j(X)$ is constant across all firms j and skill groups X. In this case, a firm j's rents as a share of its profits are given by:

$$\begin{aligned}
\frac{R_j^e}{\Pi_j} = 1 - \frac{\alpha_j(1+\boldsymbol{\varepsilon})}{1+\alpha_j \boldsymbol{\varepsilon}} \times \frac{Y_j^{\text{pt}}}{Y_j} &= 1 - \frac{\alpha_j(1+\boldsymbol{\varepsilon})}{1+\alpha_j \boldsymbol{\varepsilon}} \times \left(\frac{T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi L_j^{\text{pt}}(\chi,\varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}}}{T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi L_j(\chi,\varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}}} \right) \\
&= 1 - \frac{\alpha_j(1+\boldsymbol{\varepsilon})}{1+\alpha_j \boldsymbol{\varepsilon}} \times \left(\frac{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi \left(\frac{W_j^{\text{pt}}(\chi,\varphi)}{W_j(\chi,\varphi)} \right)^{\beta} L_j(\chi,\varphi) d\varphi \right)^{\rho_j}}{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi L_j(\chi,\varphi) d\varphi \right)^{\rho_j}} \right)^{\frac{1-\alpha_j}{\rho_j}}.
\end{aligned}$$

If labor is perfectly substitutable, so that $\rho_j = 1$, then the wage ratio $W_j^{\text{pt}}(\chi,\varphi)/W_j(\chi,\varphi)$ is:

$$\begin{aligned}
\frac{W_j^{\text{pt}}(\chi,\varphi)}{W_j(\chi,\varphi)} &= \frac{1+\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}} \left(\frac{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \int \varphi L_j^{\text{pt}}(\chi,\varphi) d\varphi}{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \int \varphi L_j(\chi,\varphi) d\varphi} \right)^{-\alpha_j} \\
&= \frac{1+\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}} \left(\frac{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \int \varphi \left(\frac{W_j^{\text{pt}}(\chi,\varphi)}{W_j(\chi,\varphi)} \right)^{\beta} L_j(\chi,\varphi) d\varphi}{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \int \varphi L_j(\chi,\varphi) d\varphi} \right)^{-\alpha_j} \\
&= \frac{1+\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}} \left(\frac{W_j^{\text{pt}}(\chi,\varphi)}{W_j(\chi,\varphi)} \right)^{\beta} \left(\frac{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \int \varphi L_j(\chi,\varphi) d\varphi}{\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \int \varphi L_j(\chi,\varphi) d\varphi} \right)^{-\alpha_j} = \frac{1+\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}} \left(\frac{W_j^{\text{pt}}(\chi,\varphi)}{W_j(\chi,\varphi)} \right)^{\beta} = \left(\frac{1+\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}} \right)^{\frac{1}{1-\beta}},
\end{aligned}$$

where the third equality holds since $W_j^{\text{pt}}(\chi,\varphi)/W_j(\chi,\varphi)$ is constant within each firm j. Substituting this expression, I can write R_j^e/Π_j as a deterministic function of the returns to scale α_j.

$$\begin{aligned}\frac{R_j^e}{\Pi_j} &= 1 - \frac{\alpha_j(1+\boldsymbol{\varepsilon})}{1+\alpha_j\,\boldsymbol{\varepsilon}} \times \left(\frac{\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\int\varphi\left(\frac{1+\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}}\right)^{\frac{1}{1-\beta}}L_j(\beta,\chi)d\varphi}{\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\int\varphi L_j(\chi,\varphi)d\varphi}\right)^{1-\alpha_j}\\ &= 1 - \frac{\alpha_j(1+\boldsymbol{\varepsilon})}{1+\alpha_j\,\boldsymbol{\varepsilon}} \times \left(\frac{1+\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}}\right)^{\frac{1-\alpha_j}{1-\beta}}.\end{aligned}$$

□

A.7. *Pass-through of TFP Shocks to Wages*

In the next property, I derive an expression for the pass-through of a hypothetical shock to firm productivity (specifically, the TFP parameter T_j) on the wages $\{W_j(X)\}_X$ of different workers.

Property A.7. *The vector of elasticities of firm j's wages $\{W_j(X)\}_X$ with respect its TFP T_j is $\left[\frac{\partial\log W_j(X)}{\partial\log T_j}\right]_X = (I-\boldsymbol{\Gamma}_j-\mathbf{K}_j)^{-1}\mathbf{1}$, where $\boldsymbol{\Gamma}_j$ and $\mathbf{K}_j$ are both matrices in $\mathbb{R}^{|X|\times|X|}$, with entries:*

$$\begin{aligned}\boldsymbol{\Gamma}_j[X,X'] &= \mathbb{1}\{X=X'\}\frac{\frac{\partial^2\log L_j(X)}{\partial[\log W_j(X)]^2}}{\boldsymbol{\varepsilon}_j(X)[1+\boldsymbol{\varepsilon}_j(X)]}\\ \mathbf{K}_j[X,X'] &= \left((1-\alpha_j-\rho_j)\frac{\partial\log N_j}{\log L_j(\chi',\varphi')} - \mathbb{1}\{\chi=\chi'\}(1-\rho_j)\frac{\partial\log L_j^{\mathit{eff}}(\chi')}{\partial\log L_j(\chi',\varphi')}\right)\boldsymbol{\varepsilon}_j(\chi',\varphi').\end{aligned}$$

If β is homogeneous, then $\frac{\partial\log W_j(X)}{\partial\log T_j}$ is constant across all firms j and skill groups X, provided that firms have the same returns to scale, α_j, and that labor is perfectly substitutable, $\rho_j=1$.

Proof. For any skill group X, the elasticity of $W_j(X)$ with respect to firm TFP is given by:

$$\begin{aligned}\frac{\partial\log W_j(\chi,\varphi)}{\partial\log T_j} &= 1 + \frac{\partial\left(\frac{\boldsymbol{\varepsilon}_j(\chi,\varphi)}{1+\boldsymbol{\varepsilon}_j(\chi,\varphi)}\right)}{\partial\log T_j} - (1-\rho_j)\frac{\partial\log L_j^{\text{eff}}(\chi)}{\partial\log T_j} + (1-\alpha_j-\rho_j)\frac{\partial\log N_j}{\partial\log T_j}\\ &= 1 + \frac{\partial\left(\frac{\boldsymbol{\varepsilon}_j(\chi,\varphi)}{1+\boldsymbol{\varepsilon}_j(\chi,\varphi)}\right)}{\partial\log W_j(\chi,\varphi)}\frac{\partial\log W_j(\chi,\varphi)}{\partial\log T_j} - (1-\rho_j)\int\frac{\partial\log L_j^{\text{eff}}(\chi)}{\partial\log L_j(\chi,\varphi')}\boldsymbol{\varepsilon}_j(\chi,\varphi')\frac{\partial\log W_j(\chi,\varphi')}{\partial\log T_j}d\varphi'\\ &\quad + (1-\alpha_j-\rho_j)\sum_{\chi'\in\mathcal{X}}\left(\int\frac{\partial\log N_j}{\partial\log L_j(\chi',\varphi')}\boldsymbol{\varepsilon}_j(\chi',\varphi')\frac{\partial\log W_j(\chi',\varphi')}{\partial\log T_j}d\varphi'\right)\\ &= 1 + \left(\frac{1}{\boldsymbol{\varepsilon}_j(\chi,\varphi)[1+\boldsymbol{\varepsilon}_j(\chi,\varphi)]}\right)\frac{\partial^2\log L_j(\chi,\varphi)}{\partial[\log W_j(\chi,\varphi)]^2}\frac{\partial\log W_j(\chi,\varphi)}{\partial\log T_j}\\ &\quad + \sum_{\chi'\in\mathcal{X}}\left[(1-\alpha_j-\rho_j)\frac{\partial\log N_j}{\log L_j^{\text{eff}}(\chi')} - (1-\rho_j)\mathbb{1}\{\chi=\chi'\}\right]\int\frac{\partial\log L_j^{\text{eff}}(\chi')}{\partial\log L_j(\chi',\varphi')}\boldsymbol{\varepsilon}_j(\chi',\varphi')\frac{\partial\log W_j(\chi',\varphi')}{\partial\log T_j}d\varphi'.\end{aligned}$$

So, in each firm j, the elasticities solve a linear system, which can be written in matrix form as:

$$\left[\frac{\partial \log W_j(X)}{\partial \log T_j}\right]_X = \mathbf{1} + \mathbf{\Gamma}_j \left[\frac{\partial \log W_j(X)}{\partial \log T_j}\right]_X + \mathbf{K}_j \left[\frac{\partial \log W_j(X)}{\partial \log T_j}\right]_X,$$

with $\mathbf{\Gamma}_j$ and $\mathbf{K}_j$ defined above. If $(I-\mathbf{\Gamma}_j-\mathbf{K}_j)^{-1}$ exists, then $\left[\frac{\partial \log W_j(X)}{\partial \log T_j}\right]_X = (I-\mathbf{\Gamma}_j-\mathbf{K}_j)^{-1}\mathbf{1}$.

If β is homogeneous across workers, the labor supply elasticity $\boldsymbol{\varepsilon}_j(X)$ is constant across firms j and skill groups X, and the second derivative $\partial^2 \log L_j(X)/\partial[\log W_j(X)]^2$ equals zero. So, if workers are perfectly substitutable in a firm's production function, so that $\rho_j = 1$, then:

$$\begin{aligned}
\frac{\partial \log W_j(\chi,\varphi)}{\partial \log T_j} &= 1 - \alpha_j\,\boldsymbol{\varepsilon} \sum_{\chi' \in \mathcal{X}} \int \frac{\partial \log N_j}{\log L_j(\chi',\varphi')} \frac{\partial \log W_j(\chi',\varphi')}{\partial \log T_j} d\varphi' \\
&= 1 - \alpha_j\,\boldsymbol{\varepsilon}\, \frac{\partial \log W_j(\chi,\varphi)}{\partial \log T_j} = \frac{1}{1-\alpha_j\,\boldsymbol{\varepsilon}}.
\end{aligned}$$

So, $\frac{\partial \log W_j(X)}{\partial \log T_j}$ is constant across firms and skill groups if firms have the same returns to scale.

□

A.8. Allocative Inefficiency

Social welfare is defined as the sum of workers' utilities (net of amenities) and firms' profits:

$$\mathcal{W} = \mathrm{E}\left(\max_j \left\{u_{ij}\big(W_j(X_i), a_j(X_i)\big)\right\}\right) + \log \sum_{j=1}^{J} \Pi_j.$$

Using the formula for the expectation of a maximum over T1EV random variables, I write:

$$\begin{aligned}
\mathcal{W} &= \mathrm{E}\left(\max_j \left\{\beta_i \log W_j(X_i) + a_j(X_i) + \epsilon_{ij}\right\}\right) + \log \sum_{j=1}^{J} \Pi_j \\
&= \int \mathrm{E}\left[\max_j \left\{\beta_i \log W_j(X_i) + a_j(X_i) + \epsilon_{ij}\right\}\middle| \beta_i = \beta, X_i = X\right] f_{\beta,X}(\beta_i, X) d(\beta_i, X) + \log \sum_{j=1}^{J} \Pi_j \\
&= \int \left(\log \sum_{j=1}^{J} \exp\left[\beta \log W_j(X) + a_j(X)\right] + \gamma\right) f_{\beta,X}(\beta, X) d(\beta, X) + \log \sum_{j=1}^{J} \Pi_j \\
&= \int \log I(\beta, X) f_{\beta,X}(\beta, X) d(\beta, X) + \gamma + \log \sum_{j=1}^{J} \Pi_j,
\end{aligned}$$

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.[2] The social planner seeks to maximize welfare by solving $\mathcal{W}^* = \max_{\{j(i)\}_i} \mathcal{W}$. The first order condition of the planner's problem is:

$$\frac{\partial \mathcal{W}}{\partial L_j(X)} = 0,$$

for all X and j. The following property characterizes when the optimal allocation is achieved.

Property A.8. *The allocation of labor that solves the social planner's problem, denoted by $\{L_j^*(X)\}_{j,X}$, coincides with the equilibrium allocation in a competitive (Walrasian) economy.*

Proof. The planner's solution is characterized by setting $\frac{\partial \mathcal{W}}{\partial L_j(X)}$ to zero for all X and j, where:

$$\begin{aligned}
\frac{\partial \mathcal{W}}{\partial L_j(X)} &= \frac{\partial\left[\int \log I(\beta,X) f_{\beta|X}(\beta|X) f_X(X) d\beta\right]}{\partial L_j(X)} + \frac{\partial \log \sum_{j=1}^J \Pi_j}{\partial L_j(X)} \\
&= \int \frac{\partial \log I(\beta,X))}{\partial L_j(X)} f_{\beta|X}(\beta|X) f_X(X) d\beta + \frac{\partial \Pi_j}{\partial L_j(X)} \left(\sum_{j=1}^J \Pi_j\right)^{-1} \\
&= \int \frac{\beta \exp\left(\beta \log W_j(X) + a_j(X)\right)}{I(\beta,X)} \left(\frac{\partial \log W_j(X)}{\partial L_j(X)}\right) f_{\beta|X}(\beta|X) f_X(X) d\beta + \frac{\partial \Pi_j}{\partial L_j(X)} \left(\sum_{j=1}^J \Pi_j\right)^{-1}. \\
&= \left(\frac{\partial \log W_j(X)}{\partial \log L_j(X)}\right) \mathrm{E}(\beta|X, j(i)=j) + \frac{\partial \Pi_j}{\partial L_j(X)} \left(\sum_{j=1}^J \Pi_j\right)^{-1} \\
&= \left(\frac{\partial \log W_j(X)}{\partial \log L_j(X)}\right) \mathrm{E}(\beta|X, j(i)=j) + \frac{\partial\left[Y_j - \int W_j(X')L_j(X')dX'\right]}{\partial L_j(X)} \left(\sum_{j=1}^J \Pi_j\right)^{-1} \\
&= \left(\frac{\partial \log W_j(X)}{\partial \log L_j(X)}\right) \mathrm{E}(\beta|X, j(i)=j) + \frac{\frac{\partial Y_j}{\partial L_j(X)} - W_j(X)}{\sum_{j=1}^J \Pi_j} - \frac{\int \frac{\partial \log W_j(X')}{\partial \log L_j(X)} W_j(X')dX'}{\sum_{j=1}^J \Pi_j}.
\end{aligned}$$

If all firms are price-takers, then they take wages as given, so: $\partial \log W_j(X')/\partial \log L_j(X) = 0$ for all $X, X' \in \mathcal{X} \times \mathbb{R}_+$. In this case, the solution to the social planner's problem reduces to:

[2]This property is proven in Small & Rosen (1981). Even without T1EV errors, expected maximal utility is:

$$\begin{aligned}
\mathrm{E}\left(\max_j \left\{u_{ij}\left(W_j(X_i), a_j(X_i)\right)\right\}\right) &= \sum_{j=1}^J \left(\int \mathrm{E}\left[\beta_i \log W_j(X_i) + a_j(X_i) + \epsilon_{ij} \middle| j(i)=j, X_i = X\right] L_j(X)dX\right) \\
&= \sum_{j=1}^J \left(\int \left(\boldsymbol{\varepsilon}_j(X) \log W_j(X) + a_j(X)\right) L_j(X)dX\right) + \mathrm{E}\left[\epsilon_{ij} \middle| j(i)=j\right].
\end{aligned}$$

$$\frac{\partial \mathcal{W}}{\partial L_j(X)} = 0 \quad \Longleftrightarrow \quad W_j(X) = \frac{\partial Y_j}{\partial L_j(X)},$$

for all X and j, which coincides with the equilibrium wage condition in a Walrasian economy. □

Property A.5 implies that the first-best allocation is implemented in a competitive (Walrasian) economy. To compute the optimal welfare $\mathcal{W}^*$, I therefore consider a counterfactual setting where all firms are price takers in the labor market. In this setting, social welfare is given by:

$$\mathcal{W}^* = \int \log I^*(\beta, X) f_{\beta,X}(\beta, X) d(\beta, X) + \gamma + \log \sum_{j=1}^{J} \alpha_j Y_j^*.$$

I measure allocative inefficiency by comparing $\mathcal{W}^*$ to welfare in a monopsonistic equilibrium:

$$\mathcal{W} = \int \log I(\beta, X) f_{\beta,X}(\beta, X) d(\beta, X) + \gamma + \log \sum_{j=1}^{J} Y_j \left(\int \left(\frac{1 + \alpha_j \boldsymbol{\varepsilon}_j(X)}{1 + \boldsymbol{\varepsilon}_j(X)} \right) \omega_j(X) dX \right).$$

To implement the first-best optimal allocation, a planner can give wage-specific wage subsidies to workers, where the shape of the subsidy curve depends on the distribution of preferences.

A.9. Micro-foundation for the Worker's Indirect Utility Function

I now present a simple micro-foundation for the indirect utility function of the worker, where each worker i chooses a firm j to maximize utility from consuming goods and leisure. Let C_{ij} denote the worker's expected consumption from working at the firm and let H_{ij} denote the worker's expected time spent working. A worker i's utility from choosing a firm j equals:

$$u_{ij} = f_i(C_{ij}, H_{ij}).$$

Let $\bar{H}_j(X_i)$ denote the scheduled work hours for employees with skills X_i at firm j, accounting for paid time off, overtime, vacation leave, and other benefits. Let $\tilde{H}_{ij}$ denote the idiosyncratic component of hours worked, accounting for commuting time and worker-firm match factors. Assume that $f_i : \mathbb{R}_+^2 \to \mathbb{R}$ is log-additive in consumption and scheduled hours, such that:

$$f_i(C_{ij}, H_{ij}) = a_i + \kappa_i \log C_{ij} - \eta \log \bar{H}_j(X_i) - \tilde{f}(\tilde{H}_{ij}).$$

The marginal rate of substitution between log consumption and log work hours equals $-\kappa_i/\eta$. Define the worker's budget constraint as $C_{ij} = W_j(X_i)$, where $W_j(X_i)$ denotes total earnings. Additionally, define $a_j(X) = -\log \bar{H}_j(X_i)$, $\beta_i = \kappa_i/\eta$, and $\epsilon_{ij} = -\eta^{-1}\tilde{f}(\tilde{H}_{ij})$. It follows that:

$$u_{ij} = \frac{a_i}{\eta} + \beta_i \log W_j(X_i) + a_j(X_i) + \epsilon_{ij},$$

which corresponds to utility specification (1) in the paper.[3] Note that this utility function can be extended to include non-labor income. For example, let V_i denote the worker's other forms of income. The budget constraint becomes $C_{ij} = W_j(X_i) + V_i$. If V_i is observed in data, then I can re-define earnings to be $\tilde{W}_{ij}(X_i) = W_j(X_i) + V_i$, and the same utility specification applies.

A.10. Capital and Monopolistic Competition in the Product Market

I now give an extension of the model that includes capital and monopolistic competition in the product market. Consider a monopolistic firm j with the following production function:

$$Q_j = T_j K_j^{\eta_j} N_j^{1-\alpha_j},$$

where K_j denotes capital. In a monopolistic product market, the revenue curve is $Y_j = Q_j^{1-\kappa_j}$. For each skill vector X, labor is hired according to the labor supply curve $L_j(X)$ and capital is rented at some fixed price r_j. The firm's profit function has the following representation:

$$\begin{aligned}
\Pi_j &= Q_j^{1-\kappa_j} - \sum_{\chi \in \mathcal{X}} \left(\int W_j(\chi, \varphi) L_j(\chi, \varphi) d\varphi \right) - r_j K_j \\
&= \tilde{T}_j K_j^{\tilde{\eta}_j} N_j^{1-\tilde{\alpha}_j} - \sum_{\chi \in \mathcal{X}} \left(\int W_j(\chi, \varphi) L_j(\chi, \varphi) d\varphi \right) - r_j K_j,
\end{aligned}$$

where $\tilde{T}_j = T_j^{1-\kappa_j}$, $\tilde{\eta}_j = \eta_j(1-\kappa_j)$, and $\tilde{\alpha}_j = \alpha_j + \kappa_j(1-\alpha_j)$. I now show that both perfect and monopolistic competition in the product market yield the same profit function. As a first step, I derive the first order condition of the firm's problem with respect to capital K_j. I write:

$$K_j = \left(\frac{r_j}{\tilde{\eta}_j \tilde{T}_j N_j^{1-\tilde{\alpha}_j}} \right)^{\frac{1}{\tilde{\eta}_j - 1}}.$$

Plugging this condition into the firm's profit function, I obtain the following:

$$\begin{aligned}
\Pi_j &= \tilde{T}_j \left(\frac{r}{\tilde{\eta}_j \tilde{T}_j N_j^{1-\tilde{\alpha}_j}} \right)^{\frac{\tilde{\eta}_j}{\tilde{\eta}_j - 1}} N_j^{1-\tilde{\alpha}_j} - \sum_{\chi \in \mathcal{X}} \left(\int W_j(\chi, \varphi) L_j(\chi, \varphi) d\varphi \right) - r \left(\frac{r}{\tilde{\eta}_j \tilde{T}_j N_j^{1-\tilde{\alpha}_j}} \right)^{\frac{1}{\tilde{\eta}_j - 1}} \\
&= \left[\tilde{T}_j \left(\frac{r}{\tilde{\eta}_j \tilde{T}_j} \right)^{\frac{\tilde{\eta}_j}{\tilde{\eta}_j - 1}} - r \left(\frac{r}{\tilde{\eta}_j \tilde{T}_j} \right)^{\frac{1}{\tilde{\eta}_j - 1}} \right] N_j^{-\frac{1-\tilde{\alpha}_j}{\tilde{\eta}_j - 1}} - \sum_{\chi \in \mathcal{X}} \left(\int W_j(\chi, \varphi) L_j(\chi, \varphi) d\varphi \right).
\end{aligned}$$

Note that this is just a reinterpretation of the original problem, where the firm's profit equals:

$$\Pi_j = \hat{T}_j N_j^{1-\hat{\alpha}_j} - \sum_{\chi \in \mathcal{X}} \left(\int W_j(\chi, \varphi) L_j(\chi, \varphi) d\varphi \right), \quad \text{where:} \quad \begin{cases} \hat{T}_j &= \tilde{T}_j \left(\frac{r}{\tilde{\eta}_j \tilde{T}_j} \right)^{\frac{\tilde{\eta}_j}{\tilde{\eta}_j - 1}} - r \left(\frac{r}{\tilde{\eta}_j \tilde{T}_j} \right)^{\frac{1}{\tilde{\eta}_j - 1}} \\ \hat{\alpha}_j &= \frac{\tilde{\eta}_j - \tilde{\alpha}_j}{\tilde{\eta}_j - 1}. \end{cases}.$$

[3] The term $\frac{a_i}{\eta}$ does not vary across firms. Therefore, it does not impact the worker's employment decisions.

B. Proofs of Main Results

B.1. Equilibrium Properties of the Model

This appendix section establishes equilibrium properties relating to existence and uniqueness, which are necessary for the identification arguments. I begin by proving an equilibrium exists.

Lemma 1. There exists an equilibrium involving strictly positive wages and employment.

Proof. The taste shocks ϵ_{ij} have strictly positive density, so $\mathrm{P}(j(i) = j|\beta, X)$ is strictly positive for all β, X and j. Therefore, an equilibrium (if it exists) involves strictly positive employment.

I restrict attention to equilibria with non-negative wages, so that $W_j(X) \geq 0$ for all X and j. For a firm to be profitable, its wages cannot exceed the revenue that it earns per unit of labor. Since $\partial^2 Y_j/\partial L_j^2(X) < 0$ for all X and j, this means that there exists a strict upper bound on the wage vectors $\{W_j(\mathbf{X})\}_j$. Thus, the set of feasible wage vectors $\{W_j(\mathbf{X})\}_j$ is contained in a convex, compact subset of Euclidean space. Every equilibrium must lie within the interior of this subspace, since $\lim_{W_j(X)\to 0} \partial \Pi_j/\partial W_j(X) > 0$ and $\lim_{W_j(X)\to\infty} \partial \Pi_j/\partial W_j(X) < 0$ for all X and j, implying that firms can always raise their profit by deviating from a corner solution. Given these properties, I can restrict attention to wages that satisfy the first-order conditions:

$$W_j(X) = \frac{\boldsymbol{\varepsilon}_j(X)}{1 + \boldsymbol{\varepsilon}_j(X)} \times \frac{\partial Y_j}{\partial L_j(X)}, \quad \text{for all } X \text{ and } j.$$

The right-hand-side is a continuously differentiable function of $W_j(X)$ that is bounded within the set of feasible wages, since $0 < [\boldsymbol{\varepsilon}_j(X)/(1 + \boldsymbol{\varepsilon}_j(X))] \times \partial Y_j/\partial L_j(X) < \partial Y_j/\partial L_j(X)$, where $\partial Y_j/\partial L_j(X)$ is finite for all $L_j(X) > 0$. By Brouwer's fixed point theorem, there is a solution to this system of equations, which corresponds to an equilibrium with strictly positive wages. □

Identification of firm-specific labor supply elasticities, as laid out in Section 4.A, requires that an infinitesimal, isolated productivity shock to one firm does not generate discrete jumps in that firm's wages and employment. One way to guarantee this is to show that, conditional on the wages at rival firms, $k \neq j$, each firm j's profit maximization problem admits a unique, locally stable solution.[4] In the following lemma, I show that this property holds almost surely.

[4]This property is immediate if profit Π_j is concave in wages $\{W_j(X)\}_X$, as this guarantees a unique solution to the firm's first-order condition. As shown in appendix section A.4, concavity holds under two special cases: the competitive (Walrasian) benchmark and a model with constant wage–amenity tradeoffs, such that $\beta_i = \beta$. However, the property does not apply in general, which means that it cannot be leveraged for the proof. From inspection of the Hessian matrix of the profit function, H_{Π_j}, I can conclude that Π_j is concave in wages whenever $\partial^2 \log L_j(X)/\partial[\log W_j(X)]^2 < \boldsymbol{\varepsilon}_j(X)(\boldsymbol{\varepsilon}_j(X) + 1)$ for all X. In a logit model, $\boldsymbol{\varepsilon}_j(X)$ equals $\mathrm{E}(\beta|X, j(i) = j)$ and $\partial^2 \log L_j(X)/\partial[\log W_j(X)]^2$ equals $\mathrm{Var}(\beta|X, j(i) = j)$. Hence, this condition is more likely to hold if workers have higher marginal utilities of (log) earnings and/or if the dispersion of these marginal utilities is low.

Lemma 2. Conditional on the wages of rival firms, $\{W_k(\mathbf{X})\}_{k \neq j}$, firm j's profit-maximization problem will have a unique optimal wage schedule $W_j(\mathbf{X})$ for almost every value of (α_j, ρ_j).

Proof. Let $W_j(\mathbf{X})$ be firm j's wage vector. The first-order condition requires $\partial \Pi_j / \partial W_j(X) = 0$ for every X. If $W_j(\mathbf{X})$ satisfies this condition, then the Hessian $[\partial^2 \Pi_j / \partial W_j(X) \partial W_j(X')]_{X,X'}$ is equal to $\mathrm{H}_{\Pi_j} = \mathbf{A}'(\mathrm{H}_{Y_j})\mathbf{A} + \mathbf{B}$, where H_{Y_j}, $\mathbf{A}$, and $\mathbf{B}$ are derived in Properties A.2 and A.3. This Hessian matrix satisfies two properties. First, $\mathbf{A}'(\mathrm{H}_{Y_j})\mathbf{A}$ and $\mathbf{B}$ are both nonsingular, as $\mathbf{A}'(\mathrm{H}_{Y_j})\mathbf{A}$ is symmetric, negative definite and $\mathbf{B}$ is diagonal. Second, given rival firms' wages $\{W_k(\mathbf{X})\}_{k \neq j}$, amenities $\{a_k(\mathbf{X})\}_k$, and worker distributions $(F_X, \{F_{\beta|X}\}_X, F_\epsilon)$, the matrices $\mathbf{A}$ and $\mathbf{B}$ are deterministic functions of $W_j(\mathbf{X})$. So, for any wage vector $W_j(\mathbf{X})$, H_{Π_j} depends on the firm's technology parameters only through H_{Y_j}. It follows that, for almost all (α_j, ρ_j) in $(0, 1) \times (-\infty, 1)$, the matrix H_{Π_j}, evaluated at $W_j(\mathbf{X})$, has a nonzero determinant and is therefore nonsingular. By the inverse function theorem, any solution $W_j(\mathbf{X})$ to the first-order condition is locally unique (and hence isolated) with probability one. Because all fixed points are isolated and the set of fixed points is compact, there can only exist finitely many of them.

Not every critical point $W_j(\mathbf{X})$ of Π_j is necessarily a global maximizer of the firm's profit function. I now show that, for almost all parameter values, there is a unique global maximizer, corresponding to a unique optimal wage schedule for the firm. Suppose that two distinct wage vectors both maximized Π_j. Since these vectors are (almost surely) locally stable fixed points of the first-order condition, the implicit function theorem delivers well-defined marginal effects of $\frac{1-\alpha_j}{\rho_j}$ on Π_j at each one. As these marginal effects differ, any arbitrarily small change in $\frac{1-\alpha_j}{\rho_j}$ yields different profit levels at these two wage vectors, so that they can no longer both be global maximizers. Thus, for almost all (α_j, ρ_j), the firm's problem has a unique solution. $\square$

Throughout the paper, I consider two key counterfactuals. First, I analyze the competitive (Walrasian) equilibrium, defined as the equilibrium in which firms are price-takers, acting as if they face perfectly elastic labor supply curves. Second, I analyze a version of the model in which workers' wage-amenity tradeoffs are homogeneous, so $\beta_i = \beta$ for all i. In both cases, it is important to show that the equilibrium is unique; otherwise, multiple counterfactual equilibria could exist, complicating the interpretation of the estimates. In the following lemma, I show that the equilibrium is indeed unique under either of these counterfactual specifications. For simplicity, I only provide a proof for the case where there is a single skill type, i.e., $|\chi| = 1$.

Lemma 3. Let $|\chi| = 1$. There exists a unique equilibrium under each of the two counterfactual specifications of the model: the Walrasian benchmark and the benchmark where β is constant.

Proof. Under each counterfactual specification, the wage markdown is constant: $\frac{\varepsilon_j(X)}{1+\varepsilon_j(X)} = \frac{\varepsilon}{1+\varepsilon}$.[5] Thus, letting $|\chi| = 1$, I can define the effective wage at firm j as $W_j^{\text{eff}}(\varphi) = W_j(\varphi)/\varphi$, where:

$$
\begin{aligned}
W_j^{\text{eff}}(\varphi) &= \frac{\varepsilon}{1+\varepsilon} T_j(1-\alpha_j)\left(\int \varphi L_j\big(\varphi, \mathbf{W}(\varphi)\big)d\varphi\right)^{-\alpha_j} \\
&= \frac{\varepsilon}{1+\varepsilon} T_j(1-\alpha_j)\left(\int \varphi \int \frac{\exp\big(\beta \log W_j(\varphi) + a_j(\varphi)\big)}{\sum_{k=1}^J \exp\big(\beta \log W_k(\varphi) + a_k(\varphi)\big)} f_{\beta,\varphi}(\beta,\varphi)d\beta d\varphi\right)^{-\alpha_j} \\
&= \frac{\varepsilon}{1+\varepsilon} T_j(1-\alpha_j)\left(\int \varphi \int \frac{\exp\big(\beta \log W_j^{\text{eff}}(\varphi) + a_j(\varphi)\big)}{\sum_{k=1}^J \exp\big(\beta \log W_k^{\text{eff}}(\varphi) + a_k(\varphi)\big)} f_{\beta,\varphi}(\beta,\varphi)d\beta d\varphi\right)^{-\alpha_j} \\
&= \frac{\varepsilon}{1+\varepsilon} T_j(1-\alpha_j)\left(\int \varphi L_j\big(\varphi, \mathbf{W}^{\text{eff}}(\varphi)\big)d\varphi\right)^{-\alpha_j}.
\end{aligned}
$$

Because the right-hand-side is constant within each firm, I can write $W_j^{\text{eff}}(\varphi) = W_j^{\text{eff}}$ for all j. Hence, an equilibrium is characterized by a solution to $g(\mathbf{W}^{\text{eff}}) = \mathbf{0}$, where I define g so that:

$$
g_j(\mathbf{W}^{\text{eff}}) = \frac{\varepsilon}{1+\varepsilon} T_j(1-\alpha_j)[L_j^{\text{eff}}(\mathbf{W}^{\text{eff}})]^{-\alpha_j} - W_j^{\text{eff}}, \quad \text{for } j \in \{1, \dots, J\}.
$$

To prove that this equilibrium is unique, I apply Theorem 3.1 in Kennan (2001), which implies that there is at most one equilibrium if g is strictly radially quasi-concave and quasi-increasing. To show that g is strictly radially quasi-concave, fix any $\mathbf{W}^{\text{eff}} > \mathbf{0}$ where $g(\mathbf{W}^{\text{eff}}) = \mathbf{0}$. I write:

$$
\begin{aligned}
g_j(\lambda \mathbf{W}^{\text{eff}}) &= \frac{\varepsilon}{1+\varepsilon} T_j(1-\alpha_j)[L_j^{\text{eff}}(\lambda \mathbf{W}^{\text{eff}})]^{-\alpha_j} - \lambda W_j^{\text{eff}} \\
&= \frac{\varepsilon}{1+\varepsilon} T_j(1-\alpha_j)[L_j^{\text{eff}}(\mathbf{W}^{\text{eff}})]^{-\alpha_j} - \lambda W_j^{\text{eff}} \\
&= (1-\lambda) W_j^{\text{eff}} > 0,
\end{aligned}
$$

for any $\lambda \in (0,1)$ and $j \in \{1, \dots, J\}$. The second equality above holds since the labor supply curves $L_j(\varphi, \mathbf{W})$ are homogeneous of degree zero. To show that g is quasi-increasing, notice that, for every firm j, $g_j(\lambda \mathbf{W}^{\text{eff}})$ is a decreasing function of effective labor $L_j^{\text{eff}}(\mathbf{W}^{\text{eff}})$, which is itself decreasing in the effective wages of j's rival firms, $\{W_k^{\text{eff}}\}_{k \neq j}$. Therefore, if $\tilde{W}_j^{\text{eff}} = W_j^{\text{eff}}$ and $\tilde{W}_k^{\text{eff}} \geq W_k^{\text{eff}}$ for all $k \neq j$, it must be that $g_j(\tilde{\mathbf{W}}^{\text{eff}}) \geq g_j(\mathbf{W}^{\text{eff}})$. The fact that g is radially quasi-concave, together with monotonicity, allows me to apply Kennan (2001), Theorem 3.1. So, I conclude that there is a unique equilibrium wage vector in each counterfactual economy. □

[5] For the Walrasian benchmark, this claim holds trivially: firms do not markdown wages, so $\frac{\varepsilon_j(X)}{1+\varepsilon_j(X)} = 1$. For the benchmark with constant β, it follows from Property A.1, which demonstrates $\varepsilon_j(X) = \mathrm{E}(\beta | X, j(i) = j)$.

B.2. *Properties of the Labor Supply Curve*

To understand how $S_j(X)$ depends on the wage $W_j(X)$, I first establish the following property.

Lemma 4. The own-wage elasticity of labor supply, $\frac{\partial \log S_j(X)}{\partial \log W_j(X)}$, takes the following form:

$$\frac{\partial \log S_j(X)}{\partial \log W_j(X)} = \mathrm{E}\left(\frac{\partial \log \mathrm{P}(j(i) = j|\beta, X)}{\partial \log W_j(X)} \,\middle|\, X, j(i) = j \right).$$

Proof. The derivative of a firm's labor supply $\log S_j(X)$ with respect to $\log W_j(X)$ is given by:

$$\begin{aligned}
\frac{\partial \log S_j(X)}{\partial \log W_j(X)} &= \frac{W_j(X)}{S_j(X)} \times \frac{\partial}{\partial W_j(X)} \int \mathrm{P}(j(i) = j|\beta, X) f_{\beta,X}(\beta, X) d\beta \\
&= S_j^{-1}(X) \int \frac{\partial \log \mathrm{P}(j(i) = j|\beta, X)}{\partial \log W_j(X)} \mathrm{P}(j(i) = j|\beta, X) f_{\beta,X}(\beta, X) d\beta \\
&= \frac{\int \frac{\partial \log \mathrm{P}(j(i)=j|\beta,X)}{\partial \log W_j(X)} \mathrm{P}(j(i) = j|\beta, X) f_{\beta,X}(\beta|X) d\beta}{\mathrm{P}(j(i) = j|X)} \\
&= \frac{\mathrm{E}\left(\frac{\partial \log \mathrm{P}(j(i)=j|\beta,X)}{\partial \log W_j(X)} \mathrm{P}(j(i) = j|\beta, X)|X \right)}{\mathrm{P}(j(i) = j|X)}.
\end{aligned}$$

By the Law of Iterated Expectations, this labor supply elasticity can be re-written as follows:

$$\begin{aligned}
\frac{\partial \log S_j(X)}{\partial \log W_j(X)} &= \frac{\sum_{k=1}^{J} \mathrm{E}\left(\frac{\partial \log \mathrm{P}(j(i)=j|\beta,X)}{\partial \log W_j(X)} \mathrm{P}(j(i) = j|\beta, X)|X, j(i) = k \right) \times \mathrm{P}(j(i) = k|X)}{\mathrm{P}(j(i) = j|X)} \\
&= \frac{\mathrm{E}\left(\frac{\partial \log \mathrm{P}(j(i)=j|\beta,X)}{\partial \log W_j(X)} |X, j(i) = j \right) \times \mathrm{P}(j(i) = j|X) + \sum_{k \neq j} 0 \times \mathrm{P}(j(i) = k|X)}{\mathrm{P}(j(i) = j|X)} \\
&= \mathrm{E}\left(\frac{\partial \log \mathrm{P}(j(i) = j|\beta, X)}{\partial \log W_j(X)} |X, j(i) = j \right).
\end{aligned}$$

□

This lemma shows that the elasticity of labor supply at firm j can be expressed as the average β-specific elasticity among incumbent workers at that firm. This representation is quite general, and it does not rely on assumptions about firm conduct, such as monopsonistic, Bertrand, or Cournot competition. To see how $\frac{\partial \log S_j(X)}{\partial \log W_j(X)}$ depends on $W_j(X)$, consider the next property.

Lemma 5. The second derivative of $\log S_j(X)$ with respect to $\log W_j(X)$ takes the form:

$$\frac{\partial^2 \log S_j(X)}{\partial \left[\log W_j(X)\right]^2} = \mathrm{E}\left(\frac{\partial^2 \log \mathrm{P}(j(i) = j|\beta, X)}{\partial \left[\log W_j(X)\right]^2} \middle| X, j(i) = j \right) + \mathrm{Var}\left(\frac{\partial \log \mathrm{P}(j(i) = j|\beta, X)}{\partial \log W_j(X)} \middle| X, j(i) = j \right).$$

Proof. To simplify notation, let $\boldsymbol{\varepsilon}_j(X) = \frac{\partial \log S_j(X)}{\partial \log W_j(X)}$ and $\boldsymbol{\varepsilon}_j(\beta, X) = \frac{\partial \log \mathrm{P}(j(i)=j|\beta,X)}{\partial \log W_j(X)}$. I write:

$$
\begin{aligned}
\frac{\partial^2 \log S_j(X)}{\partial \left[\log W_j(X)\right]^2} &= \frac{\partial}{\partial \log W_j(X)} \left[\frac{\int \boldsymbol{\varepsilon}_j(\beta, X)\, \mathrm{P}(j(i) = j|\beta, X) f_{\beta,X}(\beta|X) d\beta}{\mathrm{P}(j(i) = j|X)} \right] \\
&= \frac{\int \frac{\partial\left[\boldsymbol{\varepsilon}_j(\beta,X)\, \mathrm{P}(j(i)=j|\beta,X)\right]}{\partial \log W_j(X)} f_{\beta|x}(\beta|X) d\beta - \boldsymbol{\varepsilon}_j(X) \int \boldsymbol{\varepsilon}_j(\beta, X)\, \mathrm{P}(j(i) = j|\beta, X) f_{\beta|x}(\beta|X) d\beta}{\mathrm{P}(j(i) = j|X)} \\
&= \frac{\int \frac{\partial\left[\boldsymbol{\varepsilon}_j(\beta,X)\, \mathrm{P}(j(i)=j|\beta,X)\right]}{\partial \log W_j(X)} f_{\beta|x}(\beta|X) d\beta}{\mathrm{P}(j(i) = j|X)} - \boldsymbol{\varepsilon}_j^2(X) \\
&= \frac{\int \frac{\partial \boldsymbol{\varepsilon}_j(\beta,X)}{\partial \log W_j(X)}\, \mathrm{P}(j(i) = j|\beta, X) f_{\beta|x}(\beta|X) d\beta}{\mathrm{P}(j(i) = j|X)} + \frac{\int \boldsymbol{\varepsilon}_j^2(\beta, X)\, \mathrm{P}(j(i) = j|\beta, X) f_{\beta|x}(\beta|X) d\beta}{\mathrm{P}(j(i) = j|X)} - \boldsymbol{\varepsilon}_j^2(X) \\
&= \mathrm{E}\left(\frac{\partial\, \boldsymbol{\varepsilon}_j(\beta, X)}{\partial \log W_j(X)} \middle| X, j(i) = j \right) + \mathrm{E}\left(\boldsymbol{\varepsilon}_j^2(\beta, X) \middle| X, j(i) = j \right) - \mathrm{E}\left(\boldsymbol{\varepsilon}_j(\beta, X) \middle| X, j(i) = j \right)^2 \\
&= \mathrm{E}\left(\frac{\partial\, \boldsymbol{\varepsilon}_j(\beta, X)}{\partial \log W_j(X)} \middle| X, j(i) = j \right) + \mathrm{Var}\left(\boldsymbol{\varepsilon}_j(\beta, X) \middle| X, j(i) = j \right).
\end{aligned}
$$

□

B.3. *Identification of Worker Skills*

Proposition 1. If Assumption II holds, then a worker's log earnings may be decomposed as:

$$\log W_{i,\tau} = \log \varphi_{i,\tau} + \psi_{j(i,\tau),\chi(i,\tau),\tau}.$$

Proof. Let $W_{j\tau}^{\mathrm{eff}}(\chi, \varphi) = W_{j\tau}(\chi, \varphi)/\varphi$ denote the effective wage for workers with skill type χ at firm j in period τ. If Assumption II holds, then the equilibrium wage condition can be written as $W_{j\tau}^{\mathrm{eff}}(\chi, \varphi) = g_{j\chi\tau}(\{W_{k\tau}^{\mathrm{eff}}(\chi, \varphi)\}_k)$ for all (j, χ, τ), where $g_{j\chi\tau}$ does not depend on φ. I write:

$$W_{j\tau}^{\mathrm{eff}}(\chi, \varphi) = \frac{\boldsymbol{\varepsilon}_{j\tau}(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_{j\tau}(\chi, \varphi)} T_{j\tau}(1 - \alpha_j)\theta_{j\chi} \left(L_{j\tau}^{\mathrm{eff}}(\chi) \right)^{\rho_j - 1} \left(\sum_{\chi' \in X} \theta_{j\chi'} \left(L_{j\tau}^{\mathrm{eff}}(\chi') \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j} - 1},$$

where the labor supply elasticity $\boldsymbol{\varepsilon}_{j\tau}(\chi, \varphi)$ equals a (χ, τ)-specific function of $\left\{W_{k\tau}^{\mathrm{eff}}(\chi, \varphi)\right\}_k$:

$$
\begin{aligned}
\boldsymbol{\varepsilon}_{j\tau}(\chi, \varphi) &= \int \beta \times \frac{\frac{\exp\left(\beta \log W_{j\tau}(\chi,\varphi) + a_j(\chi,\varphi)\right)}{\sum_{k=1}^{J} \exp(\beta \log W_{k\tau}(\chi,\varphi) + a_k(\chi,\varphi))} f_{\beta|\chi,\tau}(\beta|\chi) f_{\chi,\varphi|\tau}(\chi, \varphi)}{\int \frac{\exp\left(\beta' \log W_{j\tau}(\chi,\varphi) + a_j(\chi,\varphi)\right)}{\sum_{k=1}^{J} \exp(\beta' \log W_{k\tau}(\chi,\varphi) + a_k(\chi,\varphi))} f_{\beta|\chi,\tau}(\beta'|\chi) f_{\chi,\varphi|\tau}(\chi, \varphi) d\beta'} d\beta \\
&= \int \beta \times \frac{\frac{\exp\left(\beta \log W_{j\tau}^{\mathrm{eff}}(\chi,\varphi)\right)}{\sum_{k=1}^{J} \exp\left(\beta \log W_{k\tau}^{\mathrm{eff}}(\chi,\varphi) + a_{k\chi}\right)} f_{\beta|\chi,\tau}(\beta|\chi)}{\int \frac{\exp\left(\beta' \log W_{j\tau}^{\mathrm{eff}}(\chi,\varphi)\right)}{\sum_{k=1}^{J} \exp\left(\beta' \log W_{k\tau}^{\mathrm{eff}}(\chi,\varphi) + a_{k\chi}\right)} f_{\beta|\chi,\tau}(\beta'|\chi) d\beta'} d\beta.
\end{aligned}
$$

As shown above, the equilibrium is fully characterized in terms of effective wages and labor, where φ does not enter any of the determining functions. Therefore, the effective wages do not depend on φ, so that $W^{\text{eff}}_{j\tau}(\chi,\varphi) = W^{\text{eff}}_{j\tau}(\chi)$, which delivers the desired log-additive separability. $\square$

B.4. Identification of Labor Supply Elasticities

Proposition 2. Under Assumptions I and II, the following moment conditions are satisfied:

$$\mathrm{E}\left[\ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi)\middle|Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] = \mathrm{E}\left[\ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi)\middle|Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right]$$
$$\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,0}(\chi) - w^{\text{eff}}_{j\tau_0,0}(\chi)\middle|Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] = \mathrm{E}\left[w^{\text{eff}}_{j\tau_1,0}(\chi) - w^{\text{eff}}_{j\tau_0,0}(\chi)\middle|Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right].$$

Proof. Let $\Delta w^{\text{eff}}_{j0}(\chi) = w^{\text{eff}}_{j\tau_1,0}(\chi) - w^{\text{eff}}_{j\tau_0,0}(\chi)$ and $\Delta\ell_{j0}(\chi) = \ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi)$ denote the change in untreated potential wage and labor outcomes from τ_0 and τ_1 for skill type χ at firm j. In the model, these untreated potential outcomes evolve due to spillover effects that result from the TFP shocks at treated firms. These spillovers operate through two channels: workers' wage indices $\{I(\beta,X)\}_{\beta,X}$ and the joint distribution of worker types $F_{\beta,X}$. So, for the common trends assumption to hold, I must show that treated and untreated firms would be affected in the same way, on average, by any change to $\{I(\beta,X)\}_{\beta,X}$ and $F_{\beta,X}$. I prove this in two steps.

<u>Step 1.</u> *Derive a sufficient statistic for $\Delta w^{eff}_{j0}(\chi)$ and $\Delta\ell_{j0}(\chi)$.*

First, I introduce some new notation. For $\Gamma_j = \left(\rho_j, \alpha_j, \{\theta_{j\chi}\}_\chi, \{a_{j\chi} - a_{j\chi'}\}_{\chi,\chi'}\right)'$, I define:

$$g_{\chi,\tau}(w|\Gamma_j) = \log(1-\alpha_j) + \log\theta_{j\chi} - (1-\rho_j)\log h_{\chi,\tau}(w|\Gamma_j) + \frac{1-\alpha_j-\rho_j}{\rho_j}\log\sum_{\chi'\in X}\theta_{j\chi'}\left(h_{\chi',\tau}(w|\Gamma_j)\right)^{\rho_j}$$
$$h_{\chi,\tau}(w|\Gamma_j) = \int\int\varphi\left(\frac{\exp\left(\beta w + a_{j\chi} - a_{j\chi^*}\right)}{I^{\text{eff}}_\tau(\beta,\chi)}\right) f_{\beta|\chi,\tau}(\beta|\chi) f_{\varphi|\chi,\tau}(\varphi) f_{\chi|\tau}(\chi)\, d\beta\, d\varphi,$$

for $\tau \in \{\tau_0,\tau_1\}$, where χ^* denotes some "reference" skill type. Since each firm j views itself as strategically small, it takes the wage index $I^{\text{eff}}_\tau(\beta,\chi)$ as given. Note that $g_{\chi,\tau}(w|\Gamma_j)$ can be interpreted as the equilibrium wage equation for a price-taking firm (with zero markdowns) where $\log(T_j) = 0$ and $a_{j\chi}$ is normalized relative to the reference type, i.e., $a_{j\chi} = a_{j\chi} - a_{j\chi^*}$. As shown in Appendix B.1, there is a unique fixed point solution to this system. Thus, I can write $g_{\chi,\tau}(w^*|\Gamma_j) = g^*_{\chi,\tau}(\Gamma_j)$. Given this property, the untreated potential wage $w^{\text{eff}}_{j,\tau_1,0}(\chi)$ is:

$$\begin{aligned} w^{\text{eff}}_{j\tau_1,0}(\chi) &= \log T_j + \log(1-\alpha_j) + \log\theta_{j\chi} - \log\left(1 + \frac{1}{\boldsymbol{\varepsilon}_{j\tau_1,0}(\chi)}\right) \\ &\qquad - (1-\rho_j)\ell^{\text{eff}}_{j\tau_1,0}(\chi) + \frac{1-\alpha_j-\rho_j}{\rho_j}\log\sum_{\chi'\in X}\theta_{j\chi'}\left(\exp\ell^{\text{eff}}_{j\tau_1,0}(\chi)\right)^{\rho_j} \\ &= \log T_j - \alpha_j a_{j\chi^*} - \log\left(1 + \frac{1}{\boldsymbol{\varepsilon}_{j\tau_1,0}(\chi)}\right) + g^*_{\chi,\tau_1}(\Gamma_j), \end{aligned}$$

where the second equality follows from the fact that $\ell^{\text{eff}}_{j\tau_1,0}(\chi) = a_{j\chi^*} + \log h_{\chi,\tau}(w^{\text{eff}}_{j\tau_1,0}(\chi)|\Gamma_j)$. Because the wage changes are infinitesimal, I can write $\boldsymbol{\varepsilon}_{j\tau_1,0}(\chi) = \boldsymbol{\varepsilon}_{\tau_1,0}(\chi, w^{\text{eff}}_{j\tau_0,0}(\chi))$, where $\boldsymbol{\varepsilon}_{j\tau_1,0}(\chi, w^{\text{eff}}_{j\tau_0,0}(\chi))$ represents the labor supply elasticity faced by an untreated firm j at time period τ_1, evaluated at the pre-period effective wage $w^{\text{eff}}_{j\tau_0,0}(\chi)$. By this relationship, I obtain:

$$\Delta w^{\text{eff}}_{j0}(\chi) = \underbrace{\log\left(1 + \frac{1}{\boldsymbol{\varepsilon}_{\tau_1,0}(\chi, w^{\text{eff}}_{j\tau_0,0}(\chi))}\right) - \log\left(1 + \frac{1}{\boldsymbol{\varepsilon}_{\tau_0,0}(\chi, w^{\text{eff}}_{j\tau_0,0}(\chi))}\right) + g^*_{\chi,\tau_1}(\Gamma_j) - g^*_{\chi,\tau_0}(\Gamma_j)}_{q_w(w^{\text{eff}}_{j\tau_0,0}(\chi)|\Gamma_j,\tau_0,\tau_1,\chi)}.$$

In addition, I can write the change in untreated potential labor outcomes, $\Delta\ell_{j0}(\chi)$, as follows:

$$\Delta\ell_{j0}(\chi) = \underbrace{\log\left(\frac{\int \frac{\exp[\beta q_w(w^{\text{eff}}_{j\tau_0,0}(\chi)|\Gamma_j,\tau_0,\tau_1,\chi) + \beta w^{\text{eff}}_{j\tau_0,0}(\chi) + a_{j\chi} - a_{j\chi^*}]}{I_{\tau_1}(\beta, X)} f_{\beta|\chi,\tau_1}(\beta|\chi) f_{\chi,\varphi|\tau_1}(\chi,\varphi)d\beta}{\int \frac{\exp[\beta' w^{\text{eff}}_{j\tau_0,0}(\chi) + a_{j\chi} - a_{j\chi^*}]}{I_{\tau_0}(\beta', X)} f_{\beta|\chi,\tau_0}(\beta'|\chi) f_{\chi,\varphi|\tau_0}(\chi,\varphi)d\beta'}\right)}_{q_\ell(w^{\text{eff}}_{j\tau_0,0}(\chi)|\Gamma_j,\tau_0,\tau_1,\chi)}.$$

<u>Step 2.</u> *Show that Assumption I implies common trends.*

Plugging in the functions $q_w(w^{\text{eff}}_{j\tau_0,0}(\chi)|\Gamma_j, \tau_0, \tau_1, \chi)$ and $q_\ell(w^{\text{eff}}_{j\tau_0,0}(\chi)|\Gamma_j, \tau_0, \tau_1, \chi)$, I obtain:

$$\begin{aligned}
\mathrm{E}\left[\Delta w^{\text{eff}}_{j0}(\chi)\middle|Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] &= \mathrm{E}\left[q_w(w^{\text{eff}}_{j\tau_0}(\chi)|\Gamma_j, \tau_0, \tau_1, \chi)\middle|Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] \\
&= \mathrm{E}\left[q_w(w^{\text{eff}}_{j\tau_0}(\chi)|\Gamma_j, \tau_0, \tau_1, \chi)\middle|Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] \\
&= \mathrm{E}\left[\Delta w^{\text{eff}}_{j0}(\chi)\middle|Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right]
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left[\Delta\ell_{j0}(\chi)\middle|Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] &= \mathrm{E}\left[q_\ell(w^{\text{eff}}_{j\tau_0}(\chi)|\Gamma_j, \tau_0, \tau_1, \chi)\middle|Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] \\
&= \mathrm{E}\left[q_\ell(w^{\text{eff}}_{j\tau_0}(\chi)|\Gamma_j, \tau_0, \tau_1, \chi)\middle|Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] \\
&= \mathrm{E}\left[\Delta\ell_{j0}(\chi)\middle|Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right].
\end{aligned}$$

In both equations, the second equality holds by Assumption I. So, I can write $\mathrm{DiD}_{\tau_0,\tau_1}(w|\chi)$ as:

$$
\begin{aligned}
\mathrm{DiD}_{\tau_0,\tau_1}(w|\chi) &= \frac{\mathrm{E}\left[\ell_{j\tau_1}(\chi)-\ell_{j\tau_0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]-\mathrm{E}\left[\ell_{j\tau_1}(\chi)-\ell_{j\tau_0}(\chi)\middle|Z_j=0, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]}{\mathrm{E}\left[w^{\text{eff}}_{j\tau_1}(\chi)-w^{\text{eff}}_{j\tau_0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]-\mathrm{E}\left[w^{\text{eff}}_{j\tau_1}(\chi)-w^{\text{eff}}_{j\tau_0}(\chi)\middle|Z_j=0, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]} \\
&= \frac{\mathrm{E}\left[\ell_{j\tau_1,1}(\chi)-\ell_{j\tau_0,1}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]-\mathrm{E}\left[\ell_{j\tau_1,0}(\chi)-\ell_{j\tau_0,0}(\chi)\middle|Z_j=0, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]}{\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,1}(\chi)-w^{\text{eff}}_{j\tau_0,1}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]-\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,0}(\chi)-w^{\text{eff}}_{j\tau_0,0}(\chi)\middle|Z_j=0, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]} \\
&= \frac{\mathrm{E}\left[\ell_{j\tau_1,1}(\chi)-\ell_{j\tau_0,1}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]-\mathrm{E}\left[\ell_{j\tau_1,0}(\chi)-\ell_{j\tau_0,0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]}{\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,1}(\chi)-w^{\text{eff}}_{j\tau_0,1}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]-\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,0}(\chi)-w^{\text{eff}}_{j\tau_0,0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]} \\
&= \frac{\mathrm{E}\left[\ell_{j\tau_1,1}(\chi)-\ell_{j\tau_1,0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]}{\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,1}(\chi)-w^{\text{eff}}_{j\tau_1,0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right]}.
\end{aligned}
$$

Since the TFP shocks are infinitesimal, the difference $w^{\text{eff}}_{j\tau_1,1}(\chi)-w^{\text{eff}}_{j\tau_0,1}(\chi)$ is also infinitesimal for any firm j. Thus, I can conclude that $\boldsymbol{\varepsilon}_{\tau_1}(\chi, w) = \mathrm{DiD}_{\tau_0,\tau_1}(w|\chi)$ for any wage w, because:

$$
\begin{aligned}
\mathrm{E}\left[\ell_{j\tau_1,1}(\chi)-\ell_{j\tau_1,0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right] &= \mathrm{E}\left[\boldsymbol{\varepsilon}_{j\tau_1}(\chi)\times\left(w^{\text{eff}}_{j\tau_1,1}(\chi)-w^{\text{eff}}_{j\tau_1,0}(\chi)\right)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right] \\
&= \mathrm{E}\left[\boldsymbol{\varepsilon}_{\tau_1}\left(\chi, w^{\text{eff}}_{j\tau_0,1}(\chi)\right)\times\left(w^{\text{eff}}_{j\tau_1,1}(\chi)-w^{\text{eff}}_{j\tau_1,0}(\chi)\right)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right] \\
&= \boldsymbol{\varepsilon}_{\tau_1}(\chi, w)\times\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,1}(\chi)-w^{\text{eff}}_{j\tau_1,0}(\chi)\middle|Z_j=1, w^{\text{eff}}_{j\tau_0}(\chi)=w\right].
\end{aligned}
$$

□

B.5. Identification of Technology

Proof of Proposition 3.

Proof. Consider two skill types χ and χ'. For each time period $\tilde{\tau} \in \{\tau, \tau'\}$, it must be that:

$$
\log\left(\frac{\mathrm{MPL}^{\text{eff}}_{j\tilde{\tau}}(\chi)}{\mathrm{MPL}^{\text{eff}}_{j\tilde{\tau}}(\chi')}\right) = \log\theta_{j\chi} - \log\theta_{j\chi'} - (1-\rho_j)\left[\log L^{\text{eff}}_{j\tilde{\tau}}(\chi) - \log L^{\text{eff}}_{j\tilde{\tau}}(\chi')\right].
$$

Because ρ_j and $\{\theta_{j\chi}\}_\chi$ are constant over time, the elasticity of substitution may be recovered by computing inter-temporal shifts in the relative marginal products and effective labor shares:

$$
(1-\rho_j)^{-1} = \frac{\log\left(\frac{L^{\text{eff}}_{j\tau}(\chi)}{L^{\text{eff}}_{j\tau}(\chi')}\right) - \log\left(\frac{L^{\text{eff}}_{j\tau'}(\chi)}{L^{\text{eff}}_{j\tau'}(\chi')}\right)}{\log\left(\frac{\mathrm{MPL}^{\text{eff}}_{j\tau}(\chi)}{\mathrm{MPL}^{\text{eff}}_{j\tau}(\chi')}\right) - \log\left(\frac{\mathrm{MPL}^{\text{eff}}_{j\tau'}(\chi)}{\mathrm{MPL}^{\text{eff}}_{j\tau'}(\chi')}\right)}.
$$

□

Proof of Proposition 4.

Proof. I normalize the firm-specific efficiencies $\{\theta_{j\chi}\}_\chi$ so that $\theta_{j\chi^*} = 1$ for skill type $\chi^* \in \mathcal{X}$. Under this normalization, and given knowledge of ρ_j, these parameters may be computed as:

$$\theta_{j\chi} = \exp\left[\log\left(\frac{\text{MPL}^{\text{eff}}_{j\tau}(\chi)}{\text{MPL}^{\text{eff}}_{j\tau}(\chi^*)}\right) + (1-\rho_j)\log\left(\frac{L^{\text{eff}}_{j\tau}(\chi)}{L^{\text{eff}}_{j\tau}(\chi^*)}\right)\right].$$

A firm's returns to scale and total factor productivity can then be recovered from the effective marginal products, the effective labor shares, and (log) value added. Specifically, I can write:

$$1-\alpha_j = \exp\left[\log \text{MPL}^{\text{eff}}_{j\tau}(\chi) - y_{j\tau} - \log(\theta_{j\chi}) + (1-\rho_j)\log L^{\text{eff}}_{j\tau}(\chi) + \log\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L^{\text{eff}}_{j\tau}(\chi')\right)^{\rho_j}\right]$$

$$T_{j\tau} = \exp\left[y_{j\tau} - \frac{1-\alpha_j}{\rho_j}\log\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(L^{\text{eff}}_{j\tau}(\chi)\right)^{\rho_j}\right].$$

□

B.6. Identification of Non-Wage Amenities

Proof of Proposition 5.

Proof. For some firm j^*, set $a_{j^*\chi} = 0$. Under this normalization, the amenities $\{a_{j\chi}\}_{j\neq j^*}$ are:

$$a_{j\chi} = \log L_{j\tau}\big(\chi, w^{\text{eff}}_{k\tau}(\chi)\big) - \log L_{j^*\tau}(\chi),$$

where $L_{j^*\tau}(\chi)$ is the labor supplied to firm j^* by skill type χ at time τ, and $L_{j\tau}\big(\chi, w^{\text{eff}}_{k\tau}(\chi)\big)$ is the labor supplied to firm j if it posts the same log effective wage as firm k. If the elasticity $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$ is known to the researcher for all wages w, then $\log L_j(\chi, w^{\text{eff}}_{k\chi})$ is recovered from:

$$a_{j\chi} = \log L_{j\tau}(\chi) + \int_{w^{\text{eff}}_{j\tau}(\chi)}^{w^{\text{eff}}_{j^*\tau}(\chi)} \boldsymbol{\varepsilon}_{j\tau}(\chi, w)\,dw - \log L_{j^*\tau}(\chi).$$

□

B.7. Identification of Worker Preferences

Proof of Proposition 6.

Proof. Suppose that the elasticity curve $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$ is known to the researcher for skill type χ.

Then the firm-specific labor supply curves $\{L_{j\tau}(\chi, w)\}_j$ can be recovered through integration:

$$\log L_{j\tau}(\chi, w) = \log L_{j\tau}(\chi) + \int_{w_{j\tau}^{\text{eff}}(\chi)}^{w} \boldsymbol{\varepsilon}_{j\tau}(\chi, \tilde{w}) d\tilde{w}.$$

Each labor supply curve $L_{j\tau}(\chi, w)$ may be expressed as a Laplace transform $\mathcal{L}\{g\}(s)$, where:

$$\mathcal{L}\{g\}(s) = \int_0^{\infty} g(t)\exp(-st)dt, \quad \text{such that} \quad \begin{aligned} s &= -w \\ t &= \beta \\ g(t) &= \frac{\exp(a_{j\chi})}{\sum_{k=1}^{J}\exp\left(tw_{k\tau}^{\text{eff}}(\chi) + a_{k\chi}\right)} f_{\beta|\chi,\tau}(t|\chi) f_{\chi|\tau}(\chi). \end{aligned}$$

The transform $\mathcal{L}\{g\}(s)$ is a one-to-one mapping of $g(t)$. Specifically, any two functions $g(t)$ can only share the same Laplace transform if they differ on a set of Lebesgue measure zero. Therefore, when the elasticity $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$ is point-identified, so is the function $g(\beta)$, where:

$$g(\beta) = \frac{\exp(a_{j\chi})}{\sum_{k=1}^{J}\exp\left(\beta w_{k\tau}^{\text{eff}}(\chi) + a_{k\chi}\right)} f_{\beta|\chi,\tau}(\beta|\chi) f_{\chi|\tau}(\chi).$$

If the amenities $\{a_{j\chi}\}_j$ are identified up-to-scale, i.e., relative to some reference amenity $a_{j^*\chi}$, then the density $f_{\beta|\chi}(\beta|\chi)$ is point-identified for any β-value from the following relationship:

$$f_{\beta|\chi}(\beta|\chi) = g(\beta) \times \frac{\sum_{k=1}^{J}\exp\left(\beta w_{k}^{\text{eff}}(\chi) + \tilde{a}_{k\chi} - \tilde{a}_{j^*\chi}\right)}{\exp\left(a_{j\chi} - a_{j^*\chi}\right) f_{\chi|\tau}(\chi)}.$$

□

B.8. Details on the Estimation of Firm-Specific Labor Supply Elasticities

In my estimation procedure, I consider the following nonparametric Kernel estimator:

$$\widehat{\text{DiD}}_{\tau_0,\tau_1}(w|\chi) = \frac{\sum_j K_{1,j}(w)\mathbf{1}\{Z_j = 1\}\left(\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi)\right) - \sum_j K_{0,j}(w)\mathbf{1}\{Z_j = 0\}\left(\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi)\right)}{\sum_j K_{1,j}(w)\mathbf{1}\{Z_j = 1\}\left(w_{j\tau_1}^{\text{eff}}(\chi) - w_{j\tau_0}^{\text{eff}}(\chi)\right) - \sum_j K_{0,j}(w)\mathbf{1}\{Z_j = 0\}\left(w_{j\tau_1}^{\text{eff}}(\chi) - w_{j\tau_0}^{\text{eff}}(\chi)\right)},$$

where I define $K_{z,j}(w)$ to be the Kernel weight for firm j with treatment status $z \in \{0, 1\}$. Examples of kernel functions include the Gaussian and Uniform kernel, defined as follows:

$$\textit{Gaussian:} \quad K_{z,j}(w) = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{w_{j\tau_0}^{\text{eff}}(\chi) - w}{h}\right)^2\right].$$

$$\textit{Uniform:} \quad K_{z,j}(w) = \frac{\mathbb{1}\left\{w - h \le w_{j\tau_0}^{\text{eff}}(\chi) \le w + h\right\}}{\sum_j \mathbb{1}\left\{w - h \le w_{j\tau_0}^{\text{eff}}(\chi) \le w + h\right\}}.$$

In each case, the tuning parameter h determines the bandwidth. As $h \to 0$, I find that:

$$\lim_{h\to 0} \widehat{\mathrm{DiD}}_{\tau_0,\tau_1}(w|\chi) = \frac{\frac{1}{N_1}\sum_j \mathbf{1}\{Z_j = 1\}\big(\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi)\big) - \frac{1}{N_0}\sum_j \mathbf{1}\{Z_j = 0\}\big(\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi)\big)}{\frac{1}{N_1}\sum_j \mathbf{1}\{Z_j = 1\}\big(w^{\text{eff}}_{j\tau_1}(\chi) - w^{\text{eff}}_{j\tau_0}(\chi)\big) - \frac{1}{N_0}\sum_j \mathbf{1}\{Z_j = 0\}\big(w^{\text{eff}}_{j\tau_1}(\chi) - w^{\text{eff}}_{j\tau_0}(\chi)\big)},$$

where $N_0 = \sum_j \mathbf{1}\{Z_j = 0\}$ and $N_1 = \sum_j \mathbf{1}\{Z_j = 1\}$. Also, by the weak law of large numbers:

$$\frac{1}{N_z}\sum_j \mathbf{1}\{Z_j = z\}\big(\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi)\big) \xrightarrow{p} \mathrm{E}\left[\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi)\middle| Z_j = z, w^{\text{eff}}_{j\tau_0}(\chi) = w\right], \quad \text{and:}$$

$$\frac{1}{N_z}\sum_j \mathbf{1}\{Z_j = z\}\big(w^{\text{eff}}_{j\tau_1}(\chi) - w^{\text{eff}}_{j\tau_0}(\chi)\big) \xrightarrow{p} \mathrm{E}\left[w^{\text{eff}}_{j\tau_1}(\chi) - w^{\text{eff}}_{j\tau_0}(\chi)\middle| Z_j = z, w^{\text{eff}}_{j\tau_0}(\chi) = w\right],$$

for $z \in \{0, 1\}$. Furthermore, using the continuous mapping theorem, I obtain the property:

$$\lim_{h\to 0} \widehat{\mathrm{DiD}}_{\tau_0,\tau_1}(w|\chi) \xrightarrow{p} \mathrm{DiD}_{\tau_0,\tau_1}(w|\chi) \quad \text{as } J \to \infty.$$

D. Data Preparation and Robustness Analyses

D.1. Details about the Estimation Sample

Figure A.1 plots the empirical distributions of log wages, log labor, log value added, and log profits for the main estimation sample. Note that the distribution of log labor is truncated, as the estimation sample is restricted to firms that maintain at least five full-time employees.

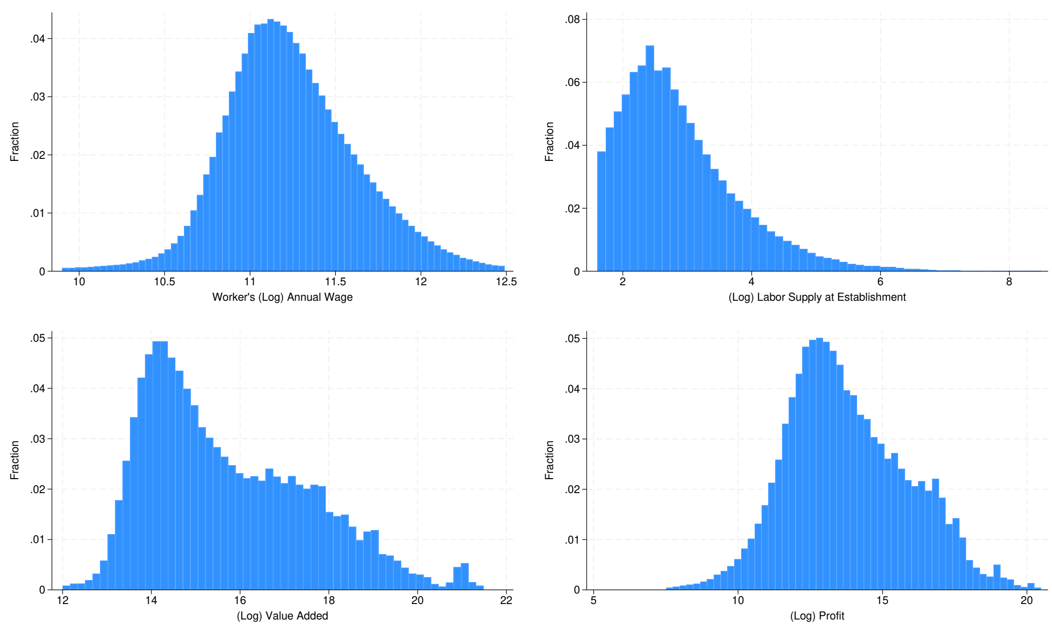
Additional Data Sources: External Instruments

Public procurement contracts in Norway are enforced by the Public Procurement Act and associated regulations. The laws apply to procurement of products, services, and construction contracts valued at over NOK 100,000, and they cover all public sector entities in Norway. All public sector procurement is competitive, and procurement auctions are publicly announced on Doffin.no. The number and scale of announcements depends on the size of each contract.

H. de Frahan et al. (2024) collect data on contract award announcements from Doffin.no during the period from 2003 to 2018, encompassing over 60,000 announcements. The data collection process involves three main steps. First, the researchers obtain HTML files through the Doffin IT department, which contain the full history of contract awards for the specified period. Second, they employ standard scraping techniques to extract information from these HTML files. For each contract, various characteristics are collected, including the name of the winning firm, the size and date of the contract, and the product specifications. Notably, the Doffin files do not provide information on the names of losing bidders. Also, for 90% of the contract award announcements, the entire tax identifier of the winning firm is missing.

In the third step, the researchers utilize a fuzzy string matching procedure, following Raffo and Lhuillery (2009), to link firm names in the Doffin data with government registers that include both the names and tax identifiers of all Norwegian firms. This matching process yields results for approximately 30,000 contracts. The researchers assess the accuracy of the fuzzy matching procedure by analyzing the 10% of contract awards for which both the name and tax identifier of the winning firm are available. They find that the algorithm correctly assigns the tax identifier in 93% of cases. The accuracy of this matching procedure does not appear to be correlated with contract characteristics, such as the value or number of bidders.

Figure A.1: Empirical Distributions of Wages, Labor, Value Added, and Profits



Notes. This figure plots histograms of wages, labor, value added, and profits (in logs) for the estimation sample.

The estimation sample is defined using the same data and following similar protocols as those used to analyze the internal instrument. Firms are classified as "treated" when they win their first procurement contracts through Doffin. All firms that receive treatment within the same calendar year τ are categorized into a "cohort". The control group for a given cohort consists of firms that win their first procurement contract in year $\tau' > \tau + 3$ or never. Future winner control firms are drawn from the Doffin data, and never winners are drawn from the full firm sample. Just as with the main sample, the data is restricted to full-time workers and firms that remain operational, maintaining at least five employees, for nine consecutive years. In total, the Doffin sample is composed of 25,714 unique firms and 901,811 unique workers.

D.2. Robustness Analyses

The first set of robustness analyses replicates the main IV estimates presented in Figure 4 of the paper using alternative specifications. In Figure A.2, I show a version of the estimates that exclude market fixed effects, effectively treating Norway as one single labor market. In Figure A.3, I report estimates based on a restricted sample where firms only operate a single establishment. In Figures A.4 and A.5, I show estimates computed using different types of Kernel estimators—Uniform and Epanechnikov kernel functions, respectively. In Figure A.6, I provide estimates for a restricted sample of firms with market shares below the median in their local labor markets. Note that these estimates aim to address potential concerns about firms leveraging their market shares to compete strategically for workers. Finally, in Figure A.7, I present estimates calibrated using the external instruments described in Appendix D.1.

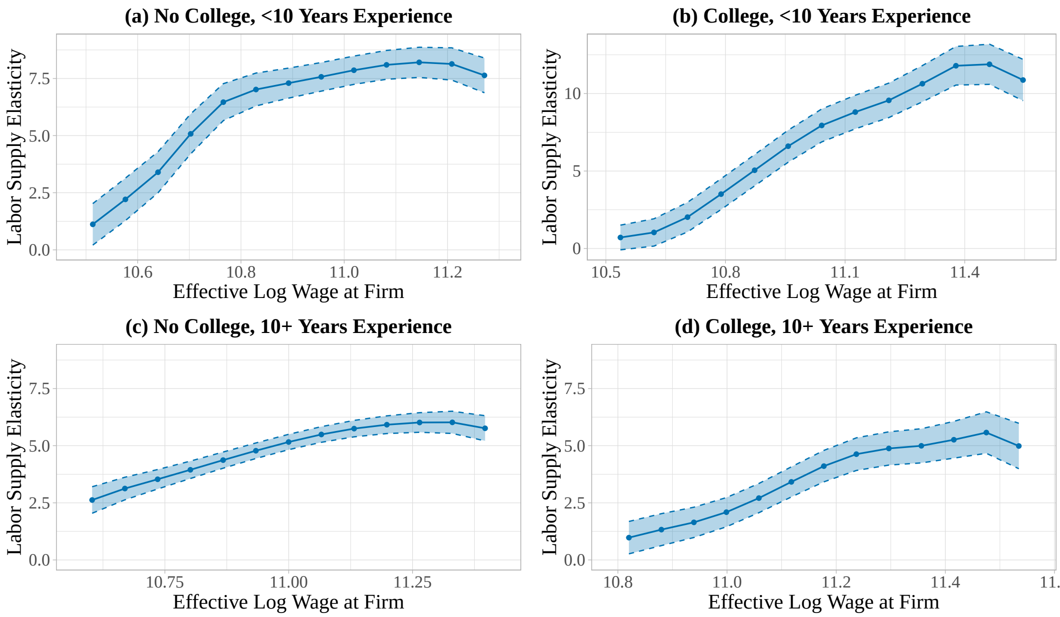
These exercises reveal that the main IV estimates are robust to specification changes. In particular, restricting the sample to firms with a single establishment or below-median market shares does not significantly affect the results. Additionally, alternative types of estimators, including the nonparametric binning estimator with a Uniform Kernel, yield similar estimates. Furthermore, all the estimates remain relatively stable with and without market fixed effects.

The estimates calibrated using external instruments are not directly comparable to those presented in Figure 4 due to differences in specification choices between the two papers. In addition to imposing different sample restrictions, H. de Frahan et al. (2024) do not account for skill differences among workers. Rather, they estimate wage and employment responses to the instrument using aggregate firm-level measures, such as average wages and total labor. Additionally, these wage and labor responses are estimated at the firm level rather than at the establishment level. Perhaps even more importantly, their DiD estimators for computing wage and employment responses differ from mine. Whereas I control for a firm's initial effective wages, the estimator used by H. de Frahan et al. (2024) controls for a firm's initial labor share. This approach would be invalid under my framework, as I allow for the possibility that two firms with identical labor shares face different labor supply elasticities. Taken together, these differences in methodology could lead to notable variations in the IV estimates.

Despite these differences, when I calibrate my model using the wage and labor responses estimated by H. de Frahan et al. (2024), I observe a pattern that closely aligns with the main IV estimates in my paper. Specifically, I find that labor supply elasticities vary significantly across firms, with higher elasticities being associated with higher average wages at the firm. Any discrepancies between the two sets of estimates may attributed to the specification differences outlined above. Yet, it is reassuring that the same patterns emerge in both cases.

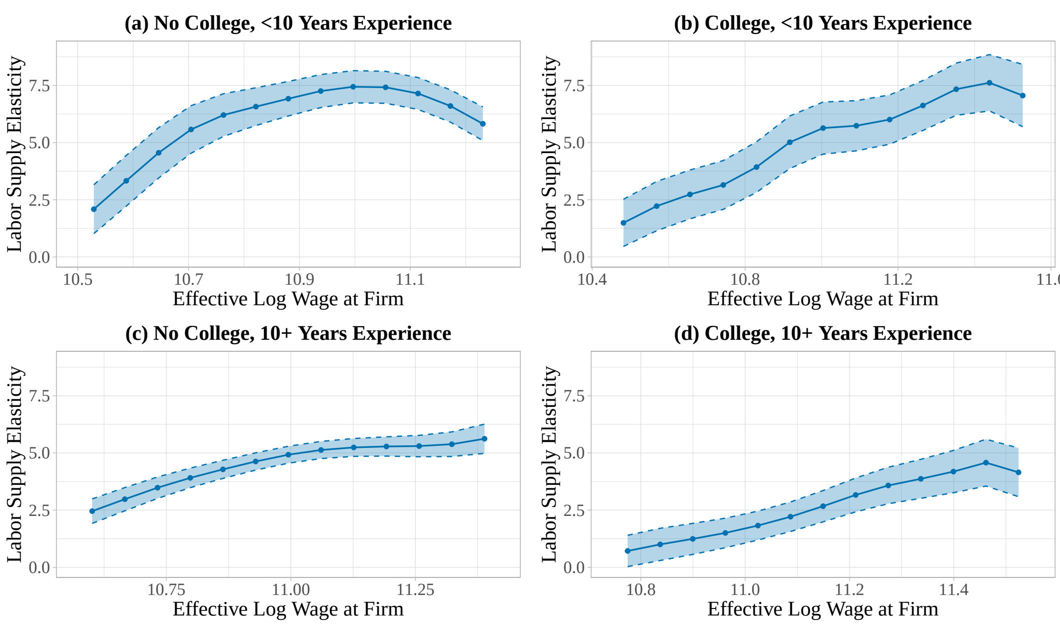
The second set of robustness analyses are misspecification tests. First, I test whether the internal instrument Z impacts labor on the intensive margin by influencing work hours. I find that there is no significant effect. Next, I test whether the IV estimates (averaged across years, firms, and skill types) differs between firms with below-median and above-median market shares in their local labor market. I find no significant market share effect, which supports my assumption that firms do not internalize their market shares when setting workers' wages.

Figure A.2: IV Estimates of Labor Supply Elasticities—No Market Fixed Effects



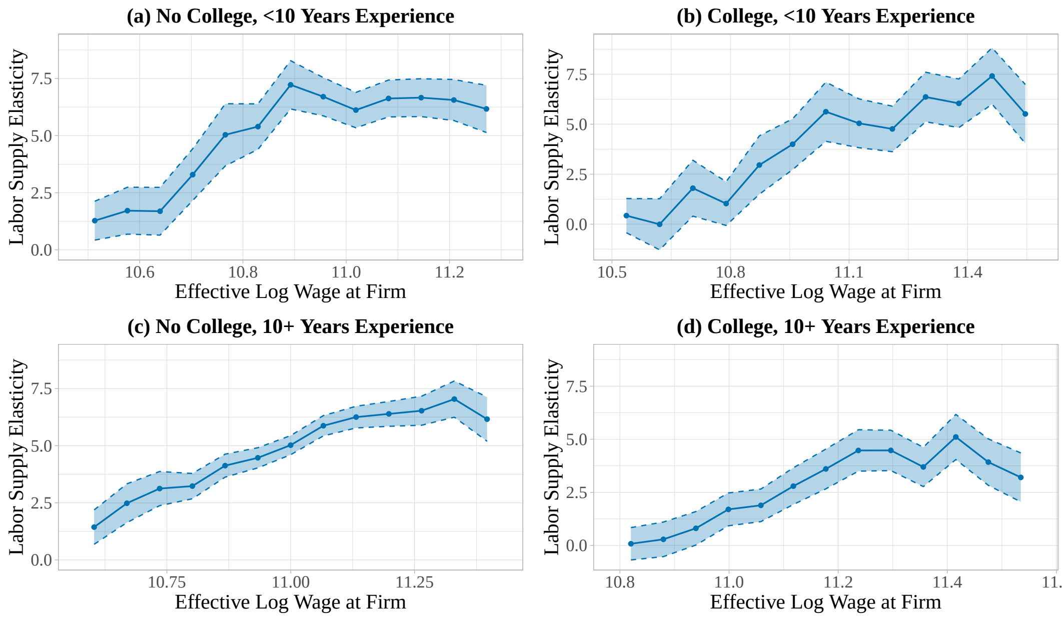
Notes. This figure presents IV estimates of firm-specific labor supply elasticities, excluding market fixed effects. Otherwise, the estimates are based on the same specification choices used to generate Figure 4 in the main text.

Figure A.3: IV Estimates of Labor Supply Elasticities—Single-Establishment Firms



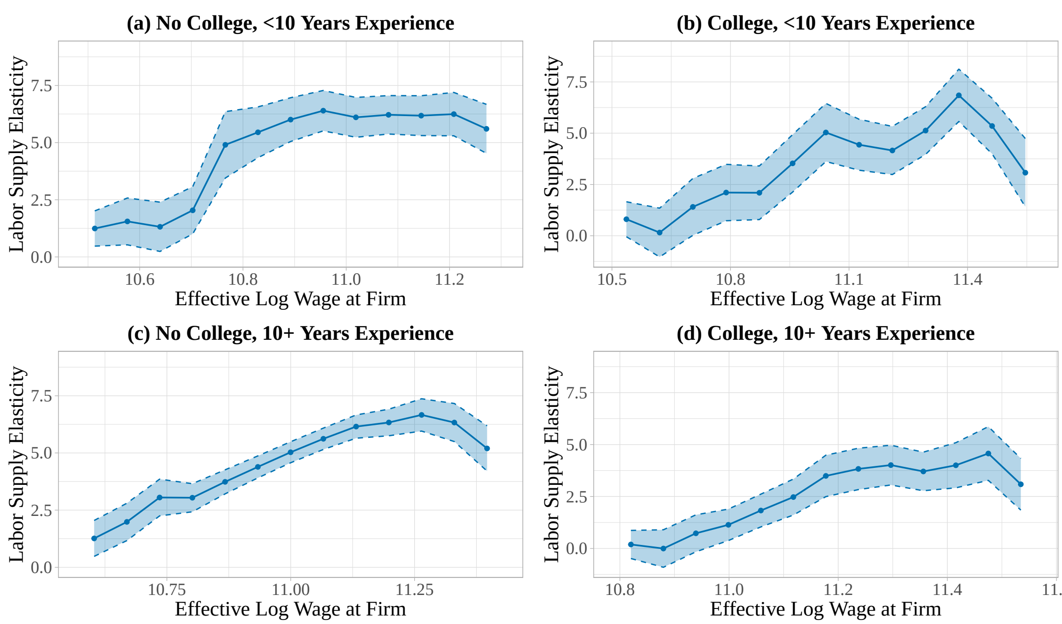
Notes. This figure plots IV estimates of firm-specific labor supply elasticities with the sample restricted to firms with one establishment. Aside from this change, all other specifications are unchanged, aligning with Figure 4.

Figure A.4: IV Estimates of Labor Supply Elasticities—Uniform Kernel Estimator



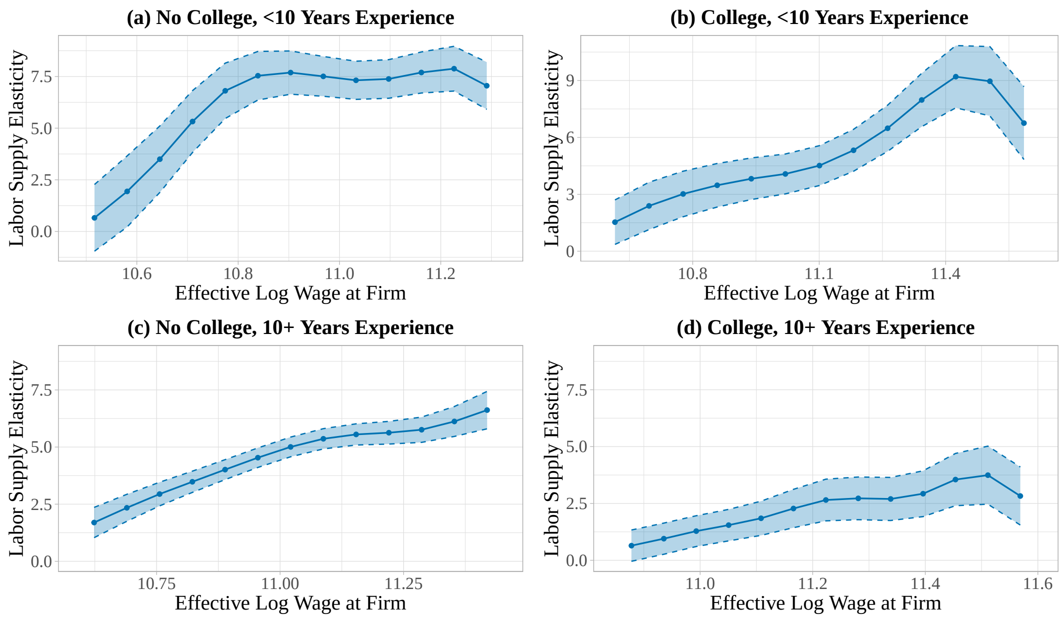
Notes. This figure plots IV estimates of firm-specific labor supply elasticities using a Uniform Kernel estimator. Otherwise, the estimates are based on the same specification choices used to generate Figure 4 in the main text.

Figure A.5: IV Estimates of Labor Supply Elasticities—Epanechnikov Kernel Estimator



Notes. This figure plots IV estimates of firm-specific labor supply elasticities using an Epanechnikov Kernel estimator. Otherwise, the estimates rely on the same specifications used to generate Figure 4 in the main text.

Figure A.6: IV Estimates of Labor Supply Elasticities—Small Market Shares



Notes. This figure plots IV estimates of firm-specific labor supply elasticities with the sample restricted to firms with below-median market shares. Otherwise, all other specifications are unchanged and align with Figure 4.

Figure A.7: Calibrated Estimates for External Instruments

Notes. This figure plots the reduced form estimates for the external instrument, averaged across skill types and years. The point estimates and standard errors in the first two figures come from H. de Frahan et al. (2024). The bottom two figures show the estimated firm-specific labor supply elasticities calibrated to my framework. Note that elasticities are measured at the firm level, and they do not account for skill differences between workers.

Table A.1: Two Misspecification Tests

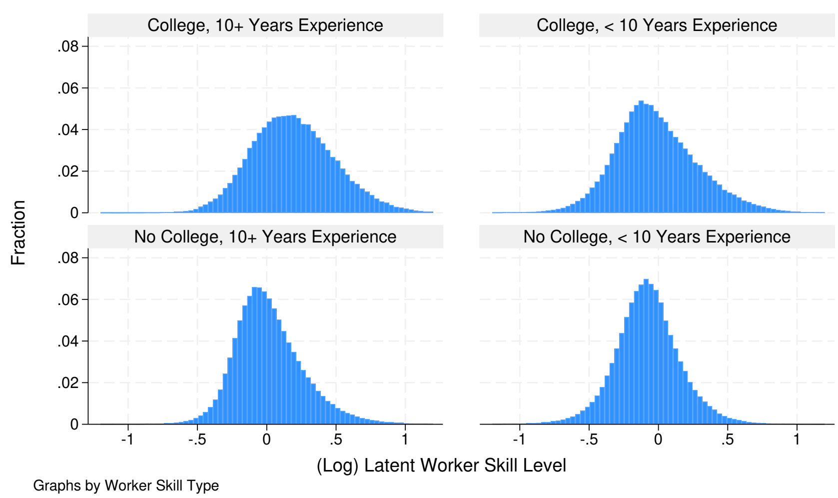
	Effect of Z on Work Hours		Market Share Effect	
	Estimated Effect	p-value	Estimated Effect	p-value
Market FEs	-0.011	0.105	-0.380	0.311
	(0.007)		(0.773)	
No Market FEs	-0.003	0.624	-0.073	0.462
	(0.007)		(0.774)	

Notes. This table presents two misspecification tests. The first two columns report the estimated effect of Z on worker's scheduled hours. The third and fourth columns report the effect of having an above-median local market share on the labor supply elasticities. Both tests are conducted with and without market fixed effects.

D.3. Additional Tables and Figures

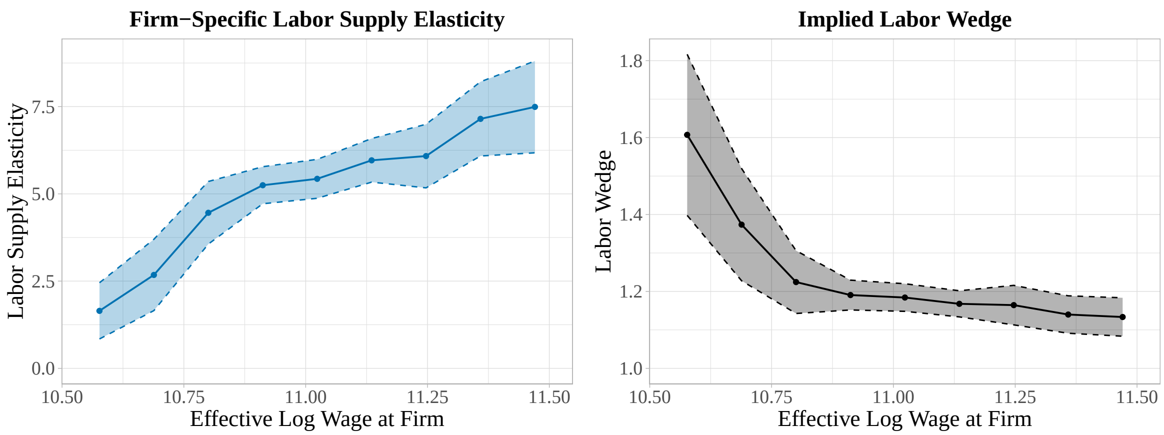
This Appendix section provides additional tables and figures that are not included in the main text of the paper. In Figure A.8, I plot the estimated distributions of worker skill levels, disaggregated by college attainment and experience. In Figure A.9, I plot the IV estimates for my main specification, averaged across all years and skill types. In Figure A.10, I plot the estimated distributions of worker and firm rents across firms in the main estimation sample.

Figure A.8: Estimated Distributions of Worker Skill Levels



Notes. This figure plots the the estimated distributions of worker skill levels by college attainment and experience.

Figure A.9: Average IV Estimates by Effective Wage

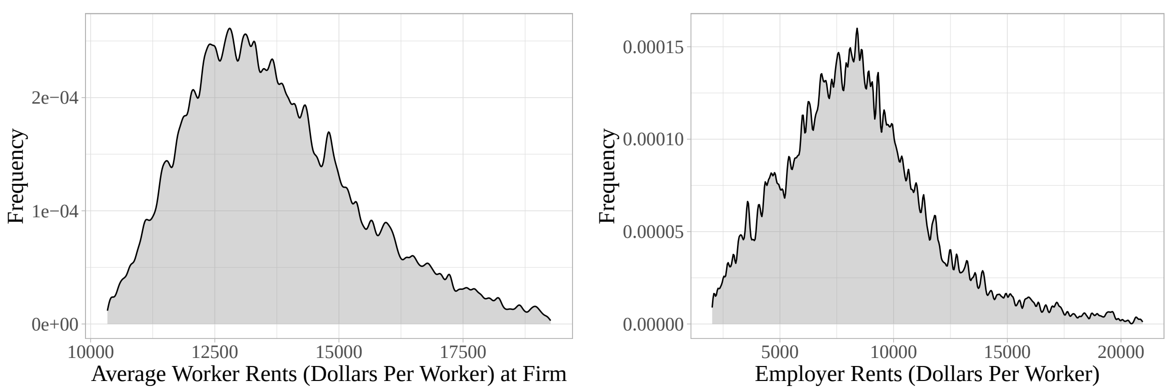


Notes. This figure plots firm-specific labor supply elasticities and labor wedges, averaged across all subgroups and years. Market fixed effects are included. All standard errors are bootstrapped using 500 bootstrap samples.

Figure A.10: Estimated Distributions of Worker and Firm Rents

(a) Average Worker Rents at Firm

(b) Firm Rents



Notes. This figure plots the distributions of estimated firm rents and estimated average worker rents at a firm.