Online Appendix

Oscar Volpe

[Access the **latest draft** and **online appendix**.]

These appendices accompany the paper "Job Preferences, Labor Market Power, and Inequality". It includes derivations of economic quantities, along with proofs and discussions related to equilibrium properties, identification, and estimation. Additionally, it provides details about data preparation, robustness analyses, and model extensions.

A. *Derivation of Equilibrium Quantities*

A.1. Firm Labor Supply

Given a set of wage offers $\mathbf{W}(X) = \{W_k(X)\}_{k=1}^{J}$, a worker with skills X whose marginal utility of (log) earnings equals β would choose to work for an employer j with probability:

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

Here, the second-to-last equality holds because $\{\epsilon_{ij}\}_{i,j}$ are independent and the final equality holds because ϵ_{ij} is 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:

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

Define $I(\beta, X) = \sum_{k=1}^{J} \exp\big(\beta \log W_k(X) + a_k(X)\big)$ as the wage index for this type of worker. For workers with skills $X_i = X$, the total mass of workers supplied to an employer j equals:

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

A.2. Equilibrium Wage Equation

Each employer j in the market chooses a set of wages $\{W_j(\chi, \varphi)\}_{\chi,\varphi}$ to maximize profit:

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

In this expression, $D_j(\chi, \varphi)$ is the labor demand for skills $X = (\chi, \varphi)$, which equals the labor supply curve $S_j(\chi, \varphi)$ in equilibrium. Plugging in these curves, the 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.

Let $\boldsymbol{\varepsilon}_j(X) = \partial \log L_j(X) / \partial \log W_j(X)$ be the labor supply elasticity for workers with skills $X = (\chi, \varphi)$. As $\boldsymbol{\varepsilon}_j(X) = \partial L_j(X) W_j(X) / \partial W_j(X) L_j(X)$, the first-order condition is:

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

For workers with skills X, the wage markdown at a firm j 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, so wages satisfy the first-order condition. In logs, this 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 Labor Supply Elasticity and Higher-Order Derivatives*

I now derive the firm-specific elasticity of labor supply, as well as higher derivatives of the labor supply curves, from the perspective of an employer that views itself as strategically small. In particular, I assume that the firm does not internalize the impact of its own wage on a worker i's wage index, i.e., $\partial I(\beta, X)/\partial W_j(X) = 0$. As a first step, I establish two claims.

Claim A.3.1. For any function $g : \mathbb{R}_+ \to \mathbb{R}$ taking values in the support of β, it follows that:

$$\frac{\partial \, \mathrm{E}_X\big(g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big)/I_i(X)\big)}{\partial w_j(X)} = \mathrm{E}_X\big(\beta g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big).$$

Proof. The derivative of $\mathrm{E}_X\big(g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)$ with respect to $w_j(X)$ is:

$$\begin{aligned}
\frac{\partial \, \mathrm{E}_X\big(g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big)/I_i(X)\big)}{\partial w_j(X)} &= \int \frac{\partial}{\partial w_j(X)}\left[\frac{g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big) f_{\beta|X}(\beta|X)}{I(\beta, X)}\right] d\beta \\
&= \int \frac{\beta g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big) f_{\beta|X}(\beta|X)}{I_i(\beta, X)} d\beta \\
&= \mathrm{E}_X\big(\beta g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big).
\end{aligned}$$

□

Claim A.3.2. For any function $g : \mathbb{R}_+ \to \mathbb{R}$ taking values in the support of β, it follows that:

$$\frac{\mathrm{E}_X\big(g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)}{\mathrm{E}_X\big(\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)} = \mathrm{E}_X\big(g(\beta)|j(i) = j\big).$$

Proof. Given that $\mathrm{P}(j(i) = j|\beta, X)$ is equal to $\exp\big(\beta \log W_j(X) + a_j(X)\big)/I(\beta, X)$, I write:

$$\frac{\mathrm{E}_X\big(g(\beta)\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)}{\mathrm{E}_X\big(\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)} = \frac{\mathrm{E}_X\big(g(\beta)\,\mathrm{P}(j(i) = j|\beta, X)\big)}{\mathrm{P}(j(i) = j|X)}.$$

Using the Law of Iterated Expectations, I can decompose this quantity in the following way:

$$\begin{aligned}
\frac{\mathrm{E}_X\big(g(\beta)\,\mathrm{P}(j(i) = j|\beta, X)\big)}{\mathrm{P}(j(i) = j|X)} &= \frac{\mathrm{E}_X\big(g(\beta)\,\mathrm{P}(j(i) = j|\beta, X)\big|j(i) = j\big)}{\mathrm{P}(j(i) = j|X)} \times \mathrm{P}(j(i) = j|X) \\
&\quad + \frac{\mathrm{E}_X\big(g(\beta)\,\mathrm{P}(j(i) = j|\beta, X)\big|j(i) \neq j\big)}{\mathrm{P}(j(i) = j|X)} \times \big[1 - \mathrm{P}(j(i) = j|X)\big] \\
&= \mathrm{E}_X\big(g(\beta) \times 1\big|j(i) = j\big) + \frac{\mathrm{E}_X\big(g(\beta) \times 0\big|j(i) \neq j\big)}{\mathrm{P}(j(i) = j|X)} \times \big[1 - \mathrm{P}(j(i) = j|X)\big] \\
&= \mathrm{E}_X\big(g(\beta)\big|j(i) = j\big).
\end{aligned}$$

□

To simplify notation going forward, I define the function $\tau_{jX,s} : \mathbb{R}_+ \to \mathbb{R}_+$ such that:

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

Claim A.3.1 implies that: $\partial \tau_{jX,s}(w_j(x))/\partial w_j(X) = \tau_{jX,s+1}(w_j(X))$. Claim A.3.2 implies that: $\tau_{jX,s}(w_j(X))/\tau_{jX,0}(w_j(X)) = \mathrm{E}_X(\beta^s | j(i) = j)$. Using these properties, I derive the first five derivatives of the (log) labor supply curve $\ell_j(X)$ with respect to (log) wage $w_j(X)$.

First Derivative. The labor supply elasticity for a firm is $\boldsymbol{\varepsilon}_j(X) = \partial \ell_j(X)/\partial w_j(X)$, where:

$$\begin{aligned}
\frac{\partial \ell_j(X)}{\partial w_j(X)} &= \frac{\partial \log\big(\tau_{jX,0}(w_j(X)) f_X(X)\big)}{\partial w_j(X)} \\
&= \frac{\partial \log\big(\tau_{jX,0}(w_j(X))\big)}{\partial w_j(X)} \\
&= \frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))} \\
&= \mathrm{E}_X\big(\beta_i | j(i) = j\big).
\end{aligned}$$

As a consequence of Claim A.3.2, this elasticity can be decomposed in the following way:

$$\begin{aligned}
\frac{\partial \ell_j(X)}{\partial w_j(X)} &= \frac{\mathrm{E}_X\big(\beta \exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)}{\mathrm{E}_X\big(\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)} \\
&= \mathrm{E}_X(\beta) + \frac{\mathrm{Cov}_X\big(\beta, \exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)}{\mathrm{E}_X\big(\exp\big(\beta w_j(X) + a_j(X)\big)/I(\beta, X)\big)}.
\end{aligned}$$

Second Derivative. The second derivative of the labor supply curve $\partial^2 \ell_j(X)/\partial w_j^2(X)$ is:

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

Third Derivative. The third derivative of the labor supply curve $\partial^3 \ell_j(X)/\partial w_j^3(X)$ is:

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

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

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

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

$$
\begin{aligned}
\frac{\partial^5 \ell_j(X)}{\partial w_j^5(X)} = \frac{\partial}{\partial w_j(X)} \Bigg(& \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))} \\
& + \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))} \Bigg) \\
= & \frac{\tau_{jX,5}(w_j(X))\tau_{jX,0}^{15}(w_j(X))}{\tau_{jX,0}^{16}(w_j(X))} \\
& - \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))} \\
& - \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{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))} \\
= & \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 \\
& - 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 + 24\left(\frac{\tau_{jX,1}(w_j(X))}{\tau_{jX,0}(w_j(X))}\right)^5 \\
= & \operatorname{E}_X\left(\beta_i^5 | j(i) = j\right) - 5\operatorname{E}_X\left(\beta_i | j(i) = j\right)\operatorname{E}_X\left(\beta_i^4 | j(i) = j\right) + 20\operatorname{E}_X\left(\beta_i | j(i) = j\right)^2\operatorname{E}_X\left(\beta_i^3 | j(i) = j\right) \\
& - 10\operatorname{E}_X\left(\beta_i^2 | j(i) = j\right)\operatorname{E}_X\left(\beta_i^3 | j(i) = j\right) + 30\operatorname{E}_X\left(\beta_i | j(i) = j\right)\operatorname{E}_X\left(\beta_i^2 | j(i) = j\right)^2 \\
& - 60\operatorname{E}_X\left(\beta_i | j(i) = j\right)^3\operatorname{E}_X\left(\beta_i^2 | j(i) = j\right) + 24\operatorname{E}_X\left(\beta_i | j(i) = j\right)^5 \\
= & \operatorname{E}_X\left([\beta_i - \operatorname{E}_X(\beta_i | j(i) = j)]^5 \middle| j(i) = j\right) - 10\operatorname{Var}_X\left(\beta_i | j(i) = j\right)\operatorname{E}_X\left([\beta_i - \operatorname{E}_X(\beta_i | j(i) = j)]^3 \middle| j(i) = j\right).
\end{aligned}
$$

A.4. *Worker Rents*

The average rents for workers with skills X at firm j are $R_{jX}^w = \operatorname{E}_X(R_i^w | j(i) = j)$, where:

$$u_{ij}\left(W_j(X_i) - R_i^w, a_j(X_i)\right) = \max_{j' \neq j(i)} u_{ij'}\left(W_{j'}(X_i), a_{j'}(X_i)\right).$$

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

$$L_j'(X, W) = \left.\frac{\partial L_j(X)}{\partial W_j(X)}\right|_{W_j(X)=W} \times \frac{1}{L_j(X)}.$$

Average worker rents are computed by integrating $W_j(X) - W$ with respect to this density:

$$\begin{aligned}
R^w_{jX} &= \int_0^{W_j(X)} \big(W_j(X) - W\big) L'_j(X, W) dW \\
&= \int_0^{W_j(X)} \big(W_j(X) - W\big) \left(\frac{\partial L_j(X)}{\partial W_j(X)}\bigg|_{W_j(X)=W} \times \frac{1}{L_j(x)} \right) dW \\
&= \frac{W_j(X)}{L_j(X)} \times \int_0^1 (1-\omega) \left(\frac{\partial}{\partial \omega} \int \frac{\exp\big(\beta_i \log\big(\omega W_j(X)\big) + a_j(X)\big) f_{\beta|X}(\beta_i|X) f_X(X)}{I_i(X)} d\beta_i \right) d\omega \\
&= \frac{W_j(X)}{L_j(X)} \times \int_0^1 (1-\omega) \left(\int \frac{\beta_i \omega^{\beta_i - 1} \exp\big(\beta_i \log W_j(X) + a_j(X)\big) f_{\beta|X}(\beta_i|X) f_X(X)}{I_i(X)} d\beta_i \right) d\omega.
\end{aligned}$$

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

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

By averaging over worker skills X, I can compute the mean rents for all workers at firm j:

$$R^w_j = \mathrm{E}(R^w_i | j(i) = j) = \int W_j(X)\, \mathrm{E}\bigg(\frac{1}{1+\beta_i} \;\Big|\; j(i) = j, X_i = X \bigg) f_X(X) dX.$$

Note that the conditional expectation $\mathrm{E}\big(\frac{1}{1+\beta_i} \big| j(i) = j, X_i = X\big)$ can be decomposed such that:

$$\mathrm{E}\bigg(\frac{1}{1+\beta_i} \Big| j(i) = j, X_i = X \bigg) = \mathrm{E}_X \bigg(\frac{1}{1+\beta_i} \bigg) + \frac{\mathrm{Cov}_X\Big(\frac{1}{1+\beta_i}, \exp\big(\beta_i \log W_j(X) + a_j(X)\big)/I_i(X) \Big)}{\mathrm{E}_X\big(\exp\big(\beta_i \log W_j(X) + a_j(X)\big)/I_i(X) \big)}.$$

The elasticity of average worker rents with respect to the wage $W_j(X)$ is $\frac{\partial \log R^w_{jX}}{\partial \log W_j(X)}$, where:

$$\frac{\partial \log R^w_{jX}}{\partial \log W_j(X)} = 1 + \frac{\frac{\partial \mathrm{E}\left(\frac{1}{1+\beta_i} \big| j(i)=j, X_i = X \right)}{\partial \log W_j(X)}}{\mathrm{E}\Big(\frac{1}{1+\beta_i} \;\big|\; j(i) = j, X_i = X \Big)} = 1 + \frac{\mathrm{Cov}\Big(\beta_i, \frac{1}{1+\beta_i} \;\big|\; j(i) = j, X_i = X \Big)}{\mathrm{E}\Big(\frac{1}{1+\beta_i} \;\big|\; j(i) = j, X_i = X \Big)}.$$

If $\mathrm{Var}(\beta_i | X_i = X) > 0$, then the second term is negative, which means that $\frac{\partial R^w_{jX} W_j(X)}{\partial W_j(X) R^w_{jX}} < 1$.

A.5. 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 are price-takers in the market, facing perfectly-elastic labor supply curves. I define the rents at firm j to be the difference between the true and counterfactual profits $R_j^e = \Pi_j - \Pi_j^{\text{price-taker}}$.

For any employer j, the profit Π_j that is realized in the monopsonistic labor market is:

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

To simplify the expression above, define $\omega_j(\chi,\varphi)$ as the share of effective labor that workers with skills $X = (\chi,\varphi)$ contribute to firm j. These effective labor shares are defined so that:

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

These shares aggregate to one, since: $\sum_{\chi \in \mathcal{X}} \int \omega_j(\chi,\varphi) d\varphi = 1$. Using this property, I write:

$$\begin{aligned}
\Pi_j &= Y_j \times \left(1 - (1-\alpha_j) \sum_{\chi \in \mathcal{X}} \int \left(\frac{\boldsymbol{\varepsilon}_j(\chi,\varphi)}{1+\boldsymbol{\varepsilon}_j(\chi,\varphi)} \right) \omega_j(\chi,\varphi) d\varphi \right) \\
&= Y_j \times \sum_{\chi \in \mathcal{X}} \int \left[1 - (1-\alpha_j) \left(\frac{\boldsymbol{\varepsilon}_j(\chi,\varphi)}{1+\boldsymbol{\varepsilon}_j(\chi,\varphi)} \right) \right] \omega_j(\chi,\varphi) d\varphi \\
&= Y_j \times \sum_{\chi \in \mathcal{X}} \int \left(\frac{1+\alpha_j \, \boldsymbol{\varepsilon}_j(\chi,\varphi)}{1+\boldsymbol{\varepsilon}_j(\chi,\varphi)} \right) \omega_j(\chi,\varphi) d\varphi.
\end{aligned}$$

If an employer j is a price-taker in the market, then its profit $\Pi_j^{\text{price-taker}}$ equals:

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

Taking first-order conditions with respect to labor demand yields the 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 labor demand $D_j^{\text{pt}}(X)$ equals the labor supply $L_j^{\text{pt}}(X)$, which is given by:

$$L_j^{\text{pt}}(X) = \int \frac{1}{I(\beta, X)} \exp\big(\beta \log W_j^{\text{pt}}(X) + a_j(X)\big) f_{\beta|X}(\beta|X) f_X(X) d\beta.$$

Given counterfactual wages and labor supply curves, the profit $\Pi_j^{\text{price-taker}}$ can be written as:

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

For any firm j, define $Y_j^{\text{pt}} = T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi L_j^{\text{pt}}(\chi, \varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}}$. The employer rents are:

$$\Pi_j^* - \Pi_j^{\text{price-taker}} = Y_j \times \left[\sum_{\chi \in \mathcal{X}} \int \left(\frac{1 + \alpha_j \boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)} \right) \omega_j(\chi, \varphi) d\varphi - \alpha_j \left(\frac{Y_j^{\text{pt}}}{Y_j} \right) \right].$$

A.6. Pass-through of TFP Shocks to Wages

The elasticity of the wage $W_j(X)$ with respect to a firm's total factor productivity T_j is:

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

A.7. Allocative Inefficiency

In this economy, aggregate social welfare is defined as $\mathcal{W}$, where:

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

Using the formula for the expectation of a maximum over T1EV random variables, I write:[1]

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

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

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

for all skills X and employers j. By evaluating these derivatives, I obtain the following:

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

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

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

Let $\boldsymbol{\varepsilon}_j^*(X)$ be the elasticity of wage w.r.t. labor in the solution to the social planner's problem.

$$\begin{aligned}
\frac{\partial \mathcal{W}}{\partial L_j(X)} &= \frac{\boldsymbol{\varepsilon}_j(X)}{\boldsymbol{\varepsilon}_j^*(X)} + \frac{\partial \Pi_j}{\partial L_j(X)}\left(\sum_{j=1}^{J}\Pi_j\right)^{-1} \\
&= \frac{\boldsymbol{\varepsilon}_j(X)}{\boldsymbol{\varepsilon}_j^*(X)} + \frac{\partial\left[Y_j - \int W_j(X)L_j(X)dX\right]}{\partial L_j(X)}\left(\sum_{j=1}^{J}\Pi_j\right)^{-1} \\
&= \frac{\boldsymbol{\varepsilon}_j(X)}{\boldsymbol{\varepsilon}_j^*(X)} + \left(\sum_{j=1}^{J}\Pi_j\right)^{-1}\left(\frac{\partial Y_j}{\partial L_j(X)} - \frac{\partial W_j(X)L_j(X)}{\partial L_j(X)} - W_j(X)\right) \\
&= \frac{\boldsymbol{\varepsilon}_j(X)}{\boldsymbol{\varepsilon}_j^*(X)} + \frac{\frac{\partial Y_j}{\partial L_j(X)} - W_j(X)}{\sum_{j=1}^{J}\Pi_j} - \frac{W_j(X)}{\boldsymbol{\varepsilon}_j^*(X)\times\sum_{j=1}^{J}\Pi_j}.
\end{aligned}$$

The solution to the social planner's problem is to set the elasticity of wages with respect to labor to zero for all X and j. This involves adjusting wages so that the markdowns are zero. To implement the first-best policy, a planner can give wage-specific wage subsidies to workers, where the shape of the subsidy curve depends on the distribution of preferences.

In the monopsonistic economy, without any policy intervention, welfare is given by:

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

Under the first-best optimal allocation that solves the planner's problem, welfare is given by:

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

The first-best optimal allocation is achieved in a competitive (Walrasian) economy. To compute optimal welfare $\mathcal{W}^*$, I consider a counterfactual setting where all firms are price-takers.

B. *Properties of the Model: Proofs and Discussion*

B.1. *Preference Heterogeneity and Substitution Patterns*

Consider a version of the utility specification (1) where the taste shocks $\{\epsilon_{ij}\}_{i,j}$ are i.i.d. and the coefficient β is constant across workers. In this case, preferences can be written as:

$$u_{ij}\big(W_j(X_i), a_j(X_i)\big) = \delta_{jX_i} + \epsilon_{ij},$$

where $\delta_{jX_i} = \beta\log W_j(X_i) + a_j(X_i)$ is a deterministic function of $\big(W_j(X_i), a_j(X_i)\big)$. Given posted wages $\{W_k(X)\}_{k=1}^{J}$, a worker with skills X will work for employer j with probability:

$$\mathrm{P}(j(i) = j|X_i = X) = \int_{-\infty}^{\infty}\prod_{k\neq j}F_\epsilon\Big(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\Big)f_\epsilon(\tilde{\epsilon})d\tilde{\epsilon}.$$

Proof of Property 1.

Define $P_{jX} = \mathrm{P}(j(i) = j|X_i = X)$ as the share of workers with skills X at firm j. If the terms $\{\epsilon_{ij}\}_{i,j}$ are i.i.d. and $\beta_i = \beta$ for all i, then the own-wage elasticity of labor supply is:

$$\begin{aligned}
\frac{\partial P_{jX} W_j(X)}{\partial W_j(X) P_{jX}} &= \frac{W_j(X)}{P_{jX}} \times \frac{\partial}{\partial W_j(X)} \int_{-\infty}^{\infty} \prod_{k \neq j} F_\epsilon\left(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= \frac{W_j(X)}{P_{jX}} \times \int_{-\infty}^{\infty} \frac{\partial}{\partial W_j(X)} \prod_{k \neq j} F_\epsilon\left(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= \frac{W_j(X)}{P_{jX}} \times \int_{-\infty}^{\infty} \left[\sum_{k \neq j} \left(\prod_{\ell \notin \{j,k\}} F_\epsilon\left(\delta_{jX} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\right) \frac{\partial \delta_{jX}}{\partial W_j(X)} \right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right] \\
&= \frac{W_j(X)}{P_{jX}} \times \frac{\partial \delta_{jX}}{\partial W_j(X)} \times \sum_{k \neq j} \left[\int_{-\infty}^{\infty} \prod_{\ell \notin \{j,k\}} F_\epsilon\left(\delta_{jX} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right] \\
&= \frac{\beta}{P_{jX}} \times \sum_{k \neq j} \left[\int_{-\infty}^{\infty} \prod_{\ell \notin \{j,k\}} F_\epsilon\left(\delta_{jX} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right].
\end{aligned}$$

In addition, the cross-wage elasticity of labor supply with respect to any firm $k \neq j$ equals:

$$\begin{aligned}
\frac{\partial P_{jX} W_k(X)}{\partial W_k(X) P_{jX}} &= \frac{W_k(X)}{P_{jX}} \times \frac{\partial}{\partial W_k(X)} \int_{-\infty}^{\infty} \prod_{\ell \neq j} F_\epsilon\left(\delta_{jX} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= \frac{W_k(X)}{P_{jX}} \times \int_{-\infty}^{\infty} \frac{\partial}{\partial W_k(X)} \prod_{\ell \neq j} F_\epsilon\left(\delta_{jX} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= -\frac{W_k(X)}{P_{jX}} \times \frac{\partial \delta_{kX}}{\partial W_k(X)} \times \int_{-\infty}^{\infty} \prod_{\ell \notin \{j,k\}} F_\epsilon\left(\delta_{jX} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= -\frac{\beta}{P_{jX}} \times \int_{-\infty}^{\infty} \prod_{\ell \notin \{j,k\}} F_\epsilon\left(\delta_{jX} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{jX} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon}.
\end{aligned}$$

Suppose that $P_{jX} = P_{j'X}$ for firms $j, j' \in \{1, \ldots, J\}$. Then $\delta_{jX} = \delta_{j'X}$, which means that:[2]

$$\begin{aligned}
\frac{\partial P_{jX} W_j(X)}{\partial W_j(X) P_{jX}} &= \frac{\beta}{P_{j'X}} \times \sum_{k \neq j} \left[\int_{-\infty}^{\infty} \prod_{\ell \notin \{j,k\}} F_\epsilon\left(\delta_{j'X} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{j'X} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right] \\
&= \frac{\beta}{P_{j'X}} \times \left[\int_{-\infty}^{\infty} \prod_{\ell \notin \{j,j'\}} F_\epsilon\left(\delta_{j'X} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon^2(\tilde{\epsilon}) d\tilde{\epsilon} \right. \\
&\quad \left. + \sum_{k \notin \{j,j'\}} \left[\int_{-\infty}^{\infty} \prod_{\ell \notin \{j,j',k\}} F_\epsilon\left(\delta_{j'X} - \delta_{\ell X} + \tilde{\epsilon}\right) F_\epsilon\left(\tilde{\epsilon}\right) f_\epsilon\left(\delta_{j'X} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right] \right] \\
&= \frac{\beta}{P_{j'X}} \times \sum_{k \neq j'} \left[\int_{-\infty}^{\infty} \prod_{\ell \notin \{j',k\}} F_\epsilon\left(\delta_{j'X} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{j'X} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right].
\end{aligned}$$

[2]The labor shares P_{jX} are strictly increasing in (and thus are uniquely determined by) the parameters δ_{jX}.

This final expression corresponds to $\frac{\partial P_{j'X} W'_j(X)}{\partial W'_j(X) P_{j'X}}$. Moreover, for any $k \notin \{j, j'\}$, I can write:

$$
\begin{aligned}
\frac{\partial P_{jX} W_k(X)}{\partial W_k(X) P_{jX}} &= -\frac{\beta}{P_{j'X}} \times \int_{-\infty}^{\infty} \prod_{\ell \notin \{j,k\}} F_\epsilon\left(\delta_{j'X} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{j'X} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= -\frac{\beta}{P_{j'X}} \times \int_{-\infty}^{\infty} \prod_{\ell \notin \{j,j',k\}} F_\epsilon\left(\delta_{j'X} - \delta_{\ell X} + \tilde{\epsilon}\right) F_\epsilon\left(\tilde{\epsilon}\right) f_\epsilon\left(\delta_{j'X} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \\
&= -\frac{\beta}{P_{j'X}} \times \int_{-\infty}^{\infty} \prod_{\ell \notin \{j',k\}} F_\epsilon\left(\delta_{j'X} - \delta_{\ell X} + \tilde{\epsilon}\right) f_\epsilon\left(\delta_{j'X} - \delta_{kX} + \tilde{\epsilon}\right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon}.
\end{aligned}
$$

This expression equals $\frac{\partial P_{j'X} W_k(X)}{\partial W_k(X) P_{j'X}}$, the cross-wage elasticity at firm j' with respect to firm k.

$\square$

Consider the special case where ϵ_{ij} follows a Type I extreme value distribution. The labor shares in this case have the following closed form expressions: $P_{jX} = \exp(\delta_{jX}) / \sum_{k=1}^{J} \exp(\delta_{kX})$. Under this specification, it is easy to see that Property 1 holds since the wage elasticities are:

$$
\frac{\partial P_{jX} W_k(X)}{\partial W_k(X) P_{jX}} = \begin{cases} \beta(1 - P_{jX}) & \text{if } j = k \\ -\beta P_{kX} & \text{if } j \neq k. \end{cases}
$$

Next, suppose that β_i is heterogeneous across workers. Then the choice probabilities are:

$$
\mathrm{P}(j(i) = j | X_i = X) = \int \mathrm{P}(j(i) = j | \beta_i = \beta, X_i = X) f_{\beta|x}(\beta|X) d\beta.
$$

Equivalently, this probability can be expressed as $\mathrm{E}[\mathrm{P}(j(i) = j | \beta_i, X_i) | X_i = X]$, which equals the conditional expectation of β-specific choice probabilities among workers with skills X.

Proof of Property 2.

To ease notation, let $P_{jX} = \mathrm{P}(j(i) = j | X_i = X)$ and $P_{jX}(\beta_i) = \mathrm{P}(j(i) = j | \beta_i, X_i = X)$ be the aggregate and β-specific labor shares, respectively, for a firm. The wage elasticities are:

$$
\begin{aligned}
\frac{\partial P_{jX} W_k(X)}{\partial W_k(X) P_{jX}} &= \frac{W_k(X)}{P_{jX}} \times \frac{\partial}{\partial W_k(X)} \left(\int P_{jX}(\beta) f_{\beta|x}(\beta|X) d\beta \right) \\
&= P_{jX}^{-1} \int \frac{\partial P_{jX}(\beta)}{\partial W_k(X)} W_k(X) f_{\beta|x}(\beta|X) d\beta \\
&= P_{jX}^{-1} \int \frac{\partial P_{jX}(\beta) W_k(X)}{\partial W_k(X) P_{jX}(\beta)} P_{jX}(\beta) f_{\beta|x}(\beta|X) d\beta,
\end{aligned}
$$

for any firm $k \in \{1, \ldots, J\}$. These elasticities can also be expressed in terms of expectations:

$$\begin{aligned}
\frac{\partial P_{jX} W_k(X)}{\partial W_k(X) P_{jX}} &= P_{jX}^{-1} \operatorname{E}\left(\frac{\partial P_{jX}(\beta_i) W_k(X)}{\partial W_k(X) P_{jX}(\beta_i)} P_{jX}(\beta_i)\right) \\
&= P_{jX}^{-1}\left[\operatorname{E}\left(\frac{\partial P_{jX}(\beta_i) W_k(X)}{\partial W_k(X) P_{jX}(\beta_i)} \times 1 \;\middle|\; j(i) = j\right) P_{jX} + \operatorname{E}\left(\frac{\partial P_{jX}(\beta_i) W_k(X)}{\partial W_k(X) P_{jX}(\beta_i)} \times 0 \;\middle|\; j(i) \neq j\right)(1 - P_{jX})\right] \\
&= \operatorname{E}\left(\frac{\partial P_{jX}(\beta_i) W_k(X)}{\partial W_k(X) P_{jX}(\beta_i)} \;\middle|\; j(i) = j\right).
\end{aligned}$$

So, the aggregate elasticity is the average of β-specific elasticities among workers at the firm.

□

<u>Proof of Property 3.</u>

To ease notation, let $\boldsymbol{\varepsilon}_j(X) = \frac{\partial P_{jX} W_j(X)}{\partial W_j(X) P_{jX}}$ and $\boldsymbol{\varepsilon}_j(\beta, X) = \frac{\partial P_{jX}(\beta_i) W_k(X)}{\partial W_k(X) P_{jX}(\beta_i)}$ be the aggregate and β-specific own-wage elasticities of labor supply, respectively, for a firm. It follows that:

$$\begin{aligned}
\frac{\partial\, \boldsymbol{\varepsilon}_j(X)}{\partial \log W_j(X)} &= \frac{\partial}{\partial \log W_j(X)}\left[P_{jX}^{-1} \int \boldsymbol{\varepsilon}_j(\beta, X) P_{jX}(\beta) f_{\beta|x}(\beta|X) d\beta\right] \\
&= P_{jX}^{-1} \int \frac{\partial\left[\boldsymbol{\varepsilon}_j(\beta, X) P_{jX}(\beta)\right]}{\partial \log W_j(X)} f_{\beta|x}(\beta|X) d\beta - P_{jX}^{-2}\left(\frac{\partial P_{jX} W_j(X)}{\partial W_j(X)}\right) \int \boldsymbol{\varepsilon}_j(\beta, X) P_{jX}(\beta) f_{\beta|x}(\beta|X) d\beta \\
&= P_{jX}^{-1} \int \frac{\partial\left[\boldsymbol{\varepsilon}_j(\beta, X) P_{jX}(\beta)\right]}{\partial \log W_j(X)} f_{\beta|x}(\beta|X) d\beta - \boldsymbol{\varepsilon}_j^2(X) \\
&= P_{jX}^{-1}\left[\int \frac{\partial\, \boldsymbol{\varepsilon}_j(\beta, X)}{\partial \log W_j(X)} P_{jX}(\beta) f_{\beta|x}(\beta|X) d\beta + \int \frac{\partial P_{jX}(\beta)}{\partial \log W_j(X)} \boldsymbol{\varepsilon}_j(\beta, X) f_{\beta|x}(\beta|X) d\beta\right] - \boldsymbol{\varepsilon}_j^2(X) \\
&= P_{jX}^{-1} \int \frac{\partial\, \boldsymbol{\varepsilon}_j(\beta, X)}{\partial \log W_j(X)} P_{jX}(\beta) f_{\beta|x}(\beta|X) d\beta + P_{jX}^{-1} \int \boldsymbol{\varepsilon}_j^2(\beta, X) P_{jX}(\beta) f_{\beta|x}(\beta|X) d\beta - \boldsymbol{\varepsilon}_j^2(X).
\end{aligned}$$

By the same reasoning that is used in the proof of Property 2, I can re-write this quantity as:

$$\begin{aligned}
\frac{\partial\, \boldsymbol{\varepsilon}_j(X)}{\partial \log W_j(X)} &= \operatorname{E}_X\left(\frac{\partial\, \boldsymbol{\varepsilon}_j(\beta_i, X)}{\partial \log W_j(X)} \;\middle|\; j(i) = j\right) + \operatorname{E}_X\left(\boldsymbol{\varepsilon}_j^2(\beta_i, X) \;\middle|\; j(i) = j\right) - \operatorname{E}_X\left(\boldsymbol{\varepsilon}_j(\beta_i, X) \;\middle|\; j(i) = j\right)^2 \\
&= \operatorname{E}\left(\frac{\partial\, \boldsymbol{\varepsilon}_j(\beta_i, X)}{\partial \log W_j(X)} \;\middle|\; X_i = X, j(i) = j\right) + \operatorname{Var}\left(\boldsymbol{\varepsilon}_j(\beta_i, X) \;\middle|\; X_i = X, j(i) = j\right).
\end{aligned}$$

The final expression above is additively separable into two terms. The first term is the average derivative of $\boldsymbol{\varepsilon}_j(\beta_i, X)$ taken with respect to $\log W_j(X)$ among workers i at the firm. The second term represents the conditional variance of $\boldsymbol{\varepsilon}_j(\beta_i, X)$ among workers i at the firm.

□

B.2. Existence and Uniqueness of an Equilibrium

In order to prove the existence and uniqueness of an equilibrium, it is first necessary to establish some basic properties about the firm-specific labor supply curves, as well as the production and profit functions. The properties that I discuss below will guide my analysis.

Firm-Specific Labor Supply Curves

The total mass of workers with skills X employed at a firm j is given by $S_j(X)$, where:

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

Throughout my analysis, I assume that firms perceive themselves to be strategically small within the economy. Thus, a firm j sets its wages $\{W_j(X)\}_X$ without considering the impact of changing its own wages on the labor that is supplied to other firms. By Property 2, I write:

$$\begin{aligned}
\frac{\partial \log S_j(X)}{\partial \log W_j(X)} &= \mathrm{E}\left(\frac{\partial \log \mathrm{P}(j(i) = j | \beta_i, X_i)}{\partial \log W_j(X_i)} \,\middle|\, X_i = X, j(i) = j \right) \\
&= \mathrm{E}\left(\frac{\beta_i}{P_{jX_i}} \times \sum_{k \neq j} \left[\int_{-\infty}^{\infty} \prod_{\ell \notin \{j,k\}} F_\epsilon \left(\delta_{jX_i} - \delta_{\ell X_i} + \tilde{\epsilon} \right) f_\epsilon \left(\delta_{jX_i} - \delta_{kX_i} + \tilde{\epsilon} \right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right] \,\middle|\, X_i = X, j(i) = j \right),
\end{aligned}$$

where I define $\delta_{jX_i} = \beta \log W_j(X_i) + a_j(X_i)$. Using this formula, I prove the following claim.

Claim B.2.1. The labor supplied to a firm $S_j(X)$ is strictly-increasing in the wage $W_j(X)$.

Proof. Because $\log(\cdot)$ is a strictly-increasing transformation, it is sufficient to show that the derivative $\partial \log S_j(X) / \partial \log W_j(X)$ is strictly positive for any wage $W_j(X) \in \mathbb{R}_{++}$. First, note that $0 < F_\epsilon(\epsilon_{ij}) < 1$ and $0 < f_\epsilon(\epsilon_{ij}) < 1$ for all $\epsilon_{ij} \in \mathbb{R}$ since the taste shocks ϵ_{ij} take positive density everywhere on $\mathbb{R}$. In addition, because $\mathrm{P}(\beta_i > 0 | X_i = X) > 0$ for every X, it must be that $\partial \log \mathrm{P}(j(i) = j | \beta_i, X_i) / \partial \log W_j(X_i)$ is strictly positive for all values of (β_i, X_i). From this property, I conclude that the derivative $\partial \log S_j(X) / \partial \log W_j(X)$ is strictly positive. □

This claim ensures that, for any values of $\{W_k(X)\}_{k \neq j}$, the labor supply $S_j(X)$ for firm j is uniquely defined by firm j's wage $W_j(X)$. Thus, any equilibrium is uniquely characterized by the wages that maximize the firms' profit function subject to the labor supply constraint. The goal of the rest of this section is to demonstrate that such wages do exist and are unique.

Firm Production Functions

In equilibrium, labor demand $D_j(X)$ equals labor supply $S_j(X)$. A firm j's output is:

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

Among workers with skills $X = (\chi, \varphi)$, the marginal product of labor at the firm is given by:

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

I prove the next three claims by deriving the Hessian matrix H_{Y_j} of the production function.

Claim B.2.2. If $\alpha \in (0,1)$ and $\rho_j < 1$, then $\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0$ for $\chi \in \mathcal{X}$ and $\varphi, \varphi' \in \mathbb{R}_{++}$.

Proof. Fix $\chi \in \mathcal{X}$ and $\varphi, \varphi' \in \mathbb{R}$. The derivative of $\frac{\partial Y_j}{\partial L_j(\chi,\varphi)}$ with respect to $L_j(\chi, \varphi')$ is:

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

Define $\zeta_{j\chi\varphi\varphi'} = \varphi\varphi' T_j(1-\alpha_j)\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j-2}\left(\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-1}$. This term is positive since $\varphi, \varphi', T_j, \theta_{j\chi} > 0$ and employment is positive. Using this notation, I write:

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

As $\zeta_{j\chi\varphi\varphi'} > 0$, it must be that $\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0$ if and only if $\frac{1}{\zeta_{j\chi\varphi\varphi'}} \times \frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0$. The effective labor share $\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}/\sum_{\chi\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}$ for skill type χ is bounded between 0 and 1. Given these properties, I conclude that $\frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} < 0$ if and only if:

$$0 > \frac{1}{\zeta_{j\chi\varphi\varphi'}} \times \frac{\partial^2 Y_j}{\partial L_j(\chi,\varphi)\partial L_j(\chi,\varphi')} = -(1-\rho_j)\left[1 - \frac{\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}}{\sum_{\chi\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}}\right] - \alpha_j \frac{\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}}{\sum_{\chi\in\mathcal{X}}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}}$$
$$= -(1-\rho_j) - \alpha_j \frac{\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}}{\sum_{\chi'\neq\chi}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}}.$$

Re-arranging terms, this inequality becomes: $\rho_j < 1 + \alpha_j \times \theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j} / \sum_{\chi'\neq\chi}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}$. Whenever $\alpha_j \in (0,1)$ and $\rho_j < 1$, this inequality holds trivially. It is worth noting that values of ρ_j above unity may also satisfy this inequality, particularly if the returns to scale parameter $1-\alpha_j$ is small and/or if the effective labor share for the skill type χ is large within this firm. □

Claim B.2.3. Suppose that $\alpha \in (0,1)$. For any $\chi, \chi' \in \mathcal{X}$, 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 positive when $\rho_j < 1-\alpha_j$ and negative when $\rho_j > 1-\alpha_j$.

Proof. 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:

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

The term $\zeta_{j\chi\chi'\varphi\varphi'} = \varphi\varphi' T_j(1-\alpha_j)\theta_{j\chi}\theta_{j\chi'}\left(L_j^{\text{eff}}(\chi)L_j^{\text{eff}}(\chi')\right)^{\rho_j-1}\left(\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}-2}$ is always greater than zero. Therefore, the sign of the derivative is pinned down by $1-\alpha_j-\rho_j$. This quantity will be positive whenever $\rho_j < 1-\alpha_j$ and will be negative whenever $\rho_j > 1-\alpha_j$. □

Claim B.2.4. Suppose that $\alpha \in (0,1)$ and $\rho_j < 1$. The production function Y_j is concave.

Proof. The effective labor $L_{j\chi}^{\text{eff}} = \int \varphi L_j(\chi,\varphi)d\varphi$ in each skill type χ is a concave, strictly increasing function of $\{L_j(\chi,\varphi)\}_\varphi$. Define the mappings $g : \mathbb{R}_+^{|\mathcal{X}|} \to \mathbb{R}_+$ and $h : \mathbb{R}_+ \to \mathbb{R}_+$ so that: $g(L_j^{\text{eff}}) = \sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j}$ and $h(x) = T_j x^{(1-\alpha_j)/\rho_j}$. If $0 < \rho_j \leq 1$, then $g(L_j^{\text{eff}})$ has a diagonal Hessian matrix with negative eigenvalues $-\rho_j(1-\rho_j)\theta_{j\chi}\left(L_j^{\text{eff}}(\chi)\right)^{\rho_j-2}$. Also, since $h(x)$ is monotonically increasing, $Y_j = h(g(L_j^{\text{eff}}))$ is quasiconcave in L_j^{eff}. Now suppose that $\rho_j < 0$. Then the Hessian matrix of $g(L_j^{\text{eff}})$ is positive definite, which means that $g(L_j^{\text{eff}})$ is convex. In addition, the function $h(x)$ is monotonically decreasing, so $Y_j = h(g(L_j^{\text{eff}}))$ is quasiconcave in L_j^{eff}. Any positive, quasiconcave function is concave if it is homogeneous of degree $k \in (0,1]$. It follows that Y_j is concave in L_j^{eff}. So, it is also concave in $\{L_j(\chi,\varphi)\}_{\chi,\varphi}$. □

Claim B.2.2 establishes conditions under which firms face decreasing marginal returns to hiring labor of the same skill type. Claim B.2.3 presents conditions under which different skill types are treated as substitutes (or complements) in the firm's production. Claim B.2.4 demonstrates that the output of a firm is a concave function of its labor inputs $\{L_j(\chi, \varphi)\}_{\chi,\varphi}$.

Firm Profit Functions

After plugging in the labor supply constraint, the profit function for any firm j equals:

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

The derivative of this function with respect to the wage $W_j(X)$ for workers with skills X is:

$$\begin{aligned}
\frac{\partial \Pi_j}{\partial W_j(X)} &= \frac{\partial Y_j}{\partial L_j(X)} \times \frac{\partial L_j(X)}{\partial W_j(X)} - L_j(X) - W_j(X) \frac{\partial L_j(X)}{\partial W_j(X)} \\
&= L_j(X) \times \left(\frac{\partial Y_j}{\partial L_j(X)} \times \frac{\boldsymbol{\varepsilon}_j(X)}{W_j(X)} - 1 - \boldsymbol{\varepsilon}_j(X) \right).
\end{aligned}$$

Next, I derive the Hessian matrix H_{Π_j} of the firm's profit function. The diagonal entries are:

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

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:

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

where the second-to-last equality above relies on the observation that $\frac{\partial \boldsymbol{\varepsilon}_j(X)}{\partial \log W_j(X)}$ is equal to:

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

The off-diagonal entries of the Hessian matrix are given by $\frac{\partial^2 \Pi_j}{\partial W_j(X) \partial W_j(X')}$, for $X \neq X'$, where:

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

Taken together, the Hessian of Π_j has the form $\mathrm{H}_{\Pi_j} = \mathrm{A}^\intercal (\mathrm{H}_{Y_j}) \mathrm{A} + \mathrm{B}$, where H_{Y_j} is the Hessian matrix of the firm's production function, and where A and B are both diagonal matrices with entries $\mathrm{A}_{XX} = \frac{\partial L_j(X)}{\partial W_j(X)}$ and $\mathrm{B}_{XX} = \frac{L_j(X)}{W_j(X)} \big[(\frac{\partial \boldsymbol{\varepsilon}_j(X)}{\partial \log W_j(X)}) / \boldsymbol{\varepsilon}_j(X) - (\boldsymbol{\varepsilon}_j(X) + 1) \big]$, respectively.[3] To interpret the properties of the Hessian matrix H_{Π_j}, I will consider two special cases:

Special Case 1. Assume β is constant for workers with skills X, i.e., $\mathrm{Var}(\beta_i | X_i = X) = 0$. In this case, $\partial \boldsymbol{\varepsilon}_j(X) / \partial \log W_j(X) = \mathrm{Var}(\beta_i | X_i = X, j(i) = j) = 0$. It follows that H_{Π_j} equals:

$$\mathrm{H}_{\Pi_j} = \mathrm{A}^\intercal (\mathrm{H}_{Y_j}) \mathrm{A} + \mathrm{B}, \quad \text{where} \quad \mathrm{B} = -\operatorname{diag}\left(\left[\frac{L_j(X) \big(\boldsymbol{\varepsilon}_j(X) + 1 \big)}{W_j(X)} \right]_X \right).$$

This matrix is negative definite, which means that the profit function Π_j is concave. To see why, note that B has strictly negative eigenvalues and that H_{Y_j} is always negative definite when $\alpha_j \in (0,1)$ and $\rho_j < 1$. Thus, H_{Π_j} must also be negative definite, since, for any $v \neq \mathbf{0}$:

$$\begin{aligned}
v^\intercal \mathrm{H}_{\Pi_j} v &= v^\intercal \big(\mathrm{A}^\intercal (\mathrm{H}_{Y_j}) \mathrm{A} + \mathrm{B} \big) v \\
&= v^\intercal \mathrm{A}^\intercal (\mathrm{H}_{Y_j}) \mathrm{A} v + v^\intercal \mathrm{B} v \\
&= (\mathrm{A} v)^\intercal (\mathrm{H}_{Y_j}) (\mathrm{A} v) + v^\intercal \mathrm{B} v < 0.
\end{aligned}$$

Special Case 2. Assume that firms do not exercise wage-setting power, i.e., $W_j(X) = \frac{\partial Y_j}{\partial L_j(X)}$. In this setting, the Hessian matrix of the firm's profit function equals $\mathrm{H}_{\Pi_j} = \mathrm{A}^\intercal (\mathrm{H}_{Y_j}) \mathrm{A} + \mathrm{B}$, where $\mathrm{B} = -2 \operatorname{diag} \big(\big[\frac{\partial L_j(X)}{\partial W_j(X)} \big]_X \big)$. Just as in the previous case, this simplification ensures that this matrix H_{Π_j} is negative definite. Therefore, the firm's profit function must be concave.

Both special cases lead to concavity of the profit function, which is a useful property for proving that a unique equilibrium exists. However, this property does not apply in general,

[3] Since $\frac{\partial L_j(X)}{\partial W_j(X)} > 0$ for all X, the entries of the matrix $\mathrm{A}^\intercal (\mathrm{H}_{Y_j}) \mathrm{A}$ share the same signs as the entries of H_{Y_j}.

which means that I cannot use it in the proof. To understand why, consider the matrix B. The diagonal entries of B, which also correspond to its eigenvalues, are negative as long as:

$$\frac{\partial\, \boldsymbol{\varepsilon}_j(X)}{\partial \log W_j(X)} < \boldsymbol{\varepsilon}_j(X)\big(\boldsymbol{\varepsilon}_j(X)+1\big), \quad \text{for all } X.$$

Under the logit model, $\boldsymbol{\varepsilon}_j(X)$ equals $\mathrm{E}(\beta_i|X_i = X, j(i) = j)$ and $\partial\, \boldsymbol{\varepsilon}_j(X)/\partial \log W_j(X)$ equals $\mathrm{Var}(\beta_i|X_i = X, j(i) = j)$. So, the inequality above is more likely to hold if workers at firm j have higher marginal utilities of (log) earnings and/or if the dispersion of these marginal utilities is low. Since the elasticity $\boldsymbol{\varepsilon}_j(X)$ increases in a firm's wage $W_j(X)$, the inequality is more likely to hold when firms offer higher wages. Note that it is difficult to draw general conclusions about the signs of the entries of B without placing restrictions on the distribution of β. This ambiguity makes it hard to establish when H_{Π_j} is negative definite.

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

Proof. I only consider equilibria where wages are positive, i.e., $W_j(X) > 0$ for every X. Also, since the taste shocks ϵ_{ij} take positive density on $\mathbb{R}$, it follows that $\mathrm{P}(j(i) = j|\beta_i, X_i) > 0$ for all (β_i, X_i). Thus, $L_j(X) = f_X(X) \times \mathrm{E}\left[\mathrm{P}(j(i) = j|\beta_i, X_i)|X_i = X\right] > 0$ for any X, which means that any equilibrium involves strictly positive employment. For a firm to be profitable, its wage cannot exceed the revenue that it receives per unit of labor. Because $\frac{\partial^2 Y_j}{\partial L_j^2(X)} < 0$, this restriction guarantees that there exists a strict upper bound on the wage $W_j(X)$ at each firm. Therefore, for any firm j, the set of feasible wages $\{W_j(X)\}_X$ is contained within a convex, compact subset of the Euclidean space. Moreover, any equilibrium must lie in 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, implying that a firm is always able to increase profit by deviating from a corner solution.

Given this reasoning, I restrict attention to wages that satisfy the first-order condition:

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

This condition takes the form of a continuously differentiable system of equations, where the right-hand-side is bounded within the set of feasible wages.[4] By Brouwer's fixed point theorem, there is a solution to this system of equations, which corresponds to an equilibrium. □

Note. Not every critical point of the profit function is necessarily an equilibrium. For the wages $\{W_j(X)\}_X$ to exist in equilibrium, they must be a global maximizer of the firm's profit function Π_j. I now prove that there is almost always a unique global maximizer of Π_j.

[4]Specifically, $0 < \left(\frac{\varepsilon_j(X)}{1+\varepsilon_j(X)}\right)\left(\frac{\partial Y_j}{\partial L_j(X)}\right) < \frac{\partial Y_j}{\partial L_j(X)}$, where $\frac{\partial Y_j}{\partial L_j(X)}$ is finite for all $L_j(X) > 0$. The labor supply curve $L_j(X)$ is bounded from below by 0: $L_j(X|W_j(X) = 0) = f_X(X) \int \frac{1}{I_i(X)} \exp(a_j(X)) f_{\beta|x}(\beta_i|X) d\beta_i > 0$.

Lemma 2. There is a unique solution to the firm's problem for almost all values of (α_j, ρ_j).

Proof. Let $W_j(\mathbf{X}) = [W_j(X)]_X$ denote the vector of wages at firm j. The first-order condition requires that $g_j\big(W_j(\mathbf{X})\big)$ equals zero, where $g_j\big(W_j(\mathbf{X})\big)$ is a multi-valued function satisfying:

$$g_{jX}\big(W_j(\mathbf{X})\big) = W_j(X) - \frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)} \times \frac{\partial Y_j\big(W_j(\mathbf{X})\big)}{\partial L_j(X)}, \quad \text{for all } X.$$

I examine the properties of g_j by deriving the Jacobian J_{g_j}. This matrix has diagonal entries:

$$\begin{aligned}\frac{\partial g_X\big(W_j(\mathbf{X})\big)}{\partial W_j(X)} &= 1 - \big(1+\boldsymbol{\varepsilon}_j(X)\big)^{-2} \times \frac{\partial\, \boldsymbol{\varepsilon}_j(X)}{\partial W_j(X)} \times \frac{\partial Y_j\big(W_j(\mathbf{X})\big)}{\partial L_j(X)} - \frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)} \times \frac{\partial^2 Y_j\big(W_j(\mathbf{X})\big)}{\partial L_j^2(X)} \times \frac{\partial L_j(X)}{\partial W_j(X)} \\ &= 1 - \frac{1}{\boldsymbol{\varepsilon}_j(X)\big(1+\boldsymbol{\varepsilon}_j(X)\big)} \times \frac{\partial\, \boldsymbol{\varepsilon}_j(X)}{\partial \log W_j(X)} - \frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)} \times \frac{\partial^2 Y_j\big(W_j(\mathbf{X})\big)}{\partial L_j^2(X)} \times \frac{\partial L_j(X)}{\partial W_j(X)},\end{aligned}$$

where the second equality uses the first-order condition. The off-diagonal entries of J_{g_j} are:

$$\frac{\partial g_X\big(W_j(\mathbf{X})\big)}{\partial W_j(X')} = -\frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)} \times \frac{\partial^2 Y_j\big(W_j(\mathbf{X})\big)}{\partial L_j(X)\partial L_j(X')} \times \frac{\partial L_j(X')}{\partial W_j(X')}, \quad \text{for } X \neq X'.$$

Thus, this Jacobian matrix has the form $\mathrm{J}_{g_j} = I - \big(\operatorname{diag}\big(\frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)}\big)(\mathrm{H}_{Y_j})\mathrm{A} + \mathrm{C}\big)$, where H_{Y_j} is the Hessian of the production function, A is a diagonal matrix with entries $\mathrm{A}_{XX} = \frac{\partial L_j(X)}{\partial W_j(X)}$, C is a diagonal matrix with entries $\mathrm{C}_{XX} = \boldsymbol{\varepsilon}_j^{-1}(X)\big(1+\boldsymbol{\varepsilon}_j(X)\big)^{-1} \times \big[\partial\, \boldsymbol{\varepsilon}_j(X)/\partial \log W_j(X)\big]$, I is the identity matrix, and $\operatorname{diag}\big(\frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)}\big)$ is a diagonal matrix of skill-specific wage markdowns.

Consider any vector of wages $W_j(\mathbf{X})$ that satisfies the first-order condition. Given firm amenities and the distribution of workers' preference parameters, the matrices $\operatorname{diag}\big(\frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)}\big)$, A, and C are fully determined by these wages. Moreover, for any $\alpha_j \in (0,1)$ and $\rho_j < 1$, the Hessian matrix H_{Y_j} of the production function is always nonsingular for any wage vector that is realized. Given this property, the determinant of the Jacobian matrix J_{g_j}, when evaluated at the wage vector $W_j(\mathbf{X})$, is nonzero for almost all values of $(\alpha_j, \rho_j) \in (0,1) \times (-\infty, 1)$. Therefore, the matrix J_{g_j} will be nonsingular with probability one at the wage vector $W_j(\mathbf{X})$. By the inverse function theorem, it is further guaranteed that $W_j(\mathbf{X})$ is a locally unique fixed point solution to the first-order condition with probability one. Furthermore, given that the set of fixed points is compact, it must have finitely-many elements with probability one.

Lastly, suppose that there are two equilibria. Since they are almost always locally stable, we may use the implicit function theorem to define the marginal effect of $\frac{1-\alpha_j}{\rho_j}$ on the firm's profit at each equilibrium. Since these marginal effects are different, any slight change in $\frac{1-\alpha_j}{\rho_j}$ causes profit to differ at the resulting equilibria. So, there is almost always one equilibrium. $\square$

C. Identification Proofs and Estimation Details

C.1. Identification of Worker Skills

Proof of Proposition 1.

Suppose that $\varphi \perp \beta | \chi, \tau$ and $a_j(\chi, \varphi) = a_{j\chi} + a_{\chi\varphi}$. The labor supplied to a firm j is:

$$L_{j\tau}(\chi, \varphi) = \int \frac{\exp\left(\beta \log W^{\text{eff}}_{j\tau}(\chi, \varphi) + a_{j\chi}\right)}{\sum_{k=1}^{J} \exp\left(\beta \log W^{\text{eff}}_{k\tau}(\chi, \varphi) + a_{k\chi}\right)} f_{\beta|\chi,\tau}(\beta|\chi) f_{\chi,\varphi|\tau}(\chi, \varphi) d\beta,$$

where $W^{\text{eff}}_{j\tau}(\chi, \varphi) = W_{j\tau}(\chi, \varphi)/\varphi$ is the effective wage of the skill type χ at firm j. Given any density $f_{\beta|\chi,\tau}$ and firm amenities $\{a_{k\chi}\}_k$, the firm-specific labor supply elasticity, defined as $\boldsymbol{\varepsilon}_{j\tau}(\chi, \varphi) = \partial \log L_{j\tau}(\chi, \varphi)/\partial \log W_{j\tau}(\chi, \varphi)$, equals a χ-specific function of $\{W^{\text{eff}}_{k\tau}(\chi, \varphi)\}_k$:

$$\boldsymbol{\varepsilon}_{j\tau}(\chi, \varphi) = \int \beta \times \frac{\left[\exp\left(\beta \log W^{\text{eff}}_{j\tau}(\chi)\right) / I^{\text{eff}}_{\tau}(\beta, \chi, \varphi)\right] f_{\beta|\chi,\tau}(\beta|\chi, \tau)}{\int \left[\exp\left(\beta' \log W^{\text{eff}}_{j\tau}(\chi)\right) / I^{\text{eff}}_{\tau}(\beta', \chi, \varphi)\right] f_{\beta|\chi,\tau}(\beta'|\chi, \tau) d\beta'} d\beta,$$

where $I^{\text{eff}}_{\tau}(\beta, \chi, \varphi) = \sum_{k=1}^{J} \exp(\beta \log W^{\text{eff}}_{k\tau}(\chi, \varphi) + a_{k\chi})$. The equilibrium wage equations are:

$$W^{\text{eff}}_{j\tau}(\chi, \varphi) = \left(\frac{\boldsymbol{\varepsilon}_{\chi}\left(W^{\text{eff}}_{j\tau}(\chi, \varphi) \mid \{W^{\text{eff}}_{k\tau}(\chi, \varphi)\}_{k \neq j}\right)}{1 + \boldsymbol{\varepsilon}_{\chi}\left(W^{\text{eff}}_{j\tau}(\chi, \varphi) \mid \{W^{\text{eff}}_{k\tau}(\chi, \varphi)\}_{k \neq j}\right)}\right) \times T_j (1 - \alpha_j) \theta_{j\chi} \left(L^{\text{eff}}_j(\chi)\right)^{\rho_j - 1} \left(\sum_{\chi' \in \mathcal{X}} \theta_{j\chi'} \left(L^{\text{eff}}_j(\chi')\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j} - 1},$$

for any firm j and vector of skills $(\chi, \varphi) \in \mathcal{X} \times \mathbb{R}_{++}$. By construction, the right-hand-side of these equations is a (j, χ)-specific function of the effective wages $\{W^{\text{eff}}_{k\tau}(\chi, \varphi)\}_k$. As shown by Lemma 2, there is almost always a unique, profit-maximizing solution to these equations. Thus, with probability one, the effective wages satisfy $W^{\text{eff}}_{j\tau}(\chi, \varphi) = W^{\text{eff}}_{j\tau}(\chi)$ in equilibrium.

$\square$

C.2. Identification of Labor Supply Elasticities

My identification strategy relies on a *common trends* assumption, which asserts that the difference in untreated potential outcomes over time is the same, on average, between treated and untreated firms in the economy. To assess when this assumption is valid, it is necessary to understand how untreated potential outcomes evolve. In my setting, these outcomes evolve due to labor supply shifts resulting from the TFP shocks at treated firms. These shifts occur through a common channel: workers' wage indices $\{I(\beta, X)\}_{\beta,X}$ and the joint distribution $F_{\beta,X}$. Therefore, for the common trends assumption to hold, I require that treated and untreated firms are affected in the same way, on average by any change in $\{I(\beta, X)\}_{\beta,X}$ and $F_{\beta,X}$.

Proof of Proposition 2.

To begin, I consider the evolution of labor supply and effective wages for any untreated firm j between time periods τ_0 and τ_1. This firm does not experience a TFP shock. Yet, its wages and labor shift due to changes in $\{I(\beta, X)\}_{\beta,X}$ and $F_{\beta,X}$. The proof follows two steps.

Step 1. Derive a sufficient statistic for $w^{\text{eff}}_{j\tau_1,0}(\chi) - w^{\text{eff}}_{j\tau_0,0}(\chi)$ *and* $\ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi)$.

Before proceeding, I introduce some new notation. Let $f_{\beta|\chi,\tau_0}(\beta_i|\chi)$ and $f_{\beta|\chi,\tau_1}(\beta_i|\chi)$ denote the conditional densities of workers' preference parameters β_i given their skill type χ at time periods τ_0 and τ_1, respectively. I also define $f_{\chi|\tau_0}(\chi)$ and $f_{\chi|\tau_1}(\chi)$ as the corresponding densities of worker skill types. Next, for $\Gamma_j = \left(\rho_j, \alpha_j, \{\theta_{j\chi}\}_\chi, \{a_{j\chi} - a_{j\chi'}\}_{\chi,\chi'}\right)'$, I define:

$$g_{\chi,\tau}(w|\Gamma_j) = \log(1-\alpha_j) + \log\theta_{j\chi} - (1-\rho_j)h_{\chi,\tau}(w|\Gamma_j) + \frac{1-\alpha_j-\rho_j}{\rho_j}\log\sum_{\chi'\in X}\theta_{j\chi'}\left(h_{\chi',\tau}(w|\Gamma_j)\right)^{\rho_j}$$

$$h_{\chi,\tau}(w|\Gamma_j) = \int\int\varphi\left(\frac{\exp\left(\beta w + a_{j\chi} - a_{j\chi^*}\right)}{I^{\text{eff}}_\tau(\beta,\chi)}\right) f_{\beta|\chi,\tau}(\beta|\chi) f_{\varphi|\chi,\tau}(\varphi) f_{\chi|\tau}(\chi)\, d\beta\, d\varphi.$$

for $\tau \in \{\tau_0, \tau_1\}$, where χ^* denotes some "reference skill type" at firm j. I assume that each firm j takes the effective wage index $I^{\text{eff}}_\tau(\beta,\chi)$ as given. I can interpret $g_{\chi,\tau}(w|\Gamma_j)$ as the equilibrium wage equation for a price-taking firm (with zero markdowns) where $\log(T_j) = 0$ and $a_{j\chi} = a_{j\chi} - a_{j\chi^*}$. As shown in Appendix B.2, there is a unique fixed point solution to this system since the Jacobian matrix is negative definite. So, I can write: $g_{\chi,\tau}(w^*|\Gamma_j) = g^*_{\chi,\tau}(\Gamma_j)$.

By construction, the effective labor of skill type χ for an untreated firm j in period τ_1 equals $L^{\text{eff}}_{j\tau_1,0}(\chi) = \exp(a_{j\chi^*}) \times h_{\chi,\tau}(w^{\text{eff}}_{j\tau_1,0}(\chi)|\Gamma_j)$. So, the potential outcome $w^{\text{eff}}_{j,\tau_1,0}(\chi)$ is:

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

Since 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 for an untreated firm j during time period τ_1 evaluated at the pre-period effective wage $w^{\text{eff}}_{j\tau_0,0}(\chi)$. By this property, I can write:

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

Additionally, I can write the change in untreated labor outcomes $\ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi)$ to be:

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

Step 2. Demonstrate that the common trends assumption holds under Assumption I.

To begin, I write down the difference-in-differences estimand in the following way:

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

In terms of the potential outcomes in the model, this estimand may be re-written to be:

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

Using 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 can write:

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

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

In both these equations, the second equality directly follows from Assumption I. Next, I write:

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

Letting $\Delta\ell_{j,\tau_1}(\chi) = \ell_{j\tau_1,1}(\chi) - \ell_{j\tau_1,0}(\chi)$ and $\Delta w^{\text{eff}}_{j\tau_1}(\chi) = w^{\text{eff}}_{j\tau_1,1}(\chi) - w^{\text{eff}}_{j\tau_1,0}(\chi)$, I write:

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

where the second equality uses the fact that $\boldsymbol{\varepsilon}_{\tau_1,0}(\chi, \cdot)$ is a continuous function of the (log) wage. Since the TFP shocks are infinitesimal, the difference $\Delta w^{\text{eff}}_{j\tau_1}(\chi)$ is also infinitesimal for any firm j. By continuity, the difference $\boldsymbol{\varepsilon}_{j\tau_1}(\chi) - \boldsymbol{\varepsilon}_{\tau_1}(\chi, w^{\text{eff}}_{j\tau_0}(\chi))$ is also infinitesimal. Thus, from the equation above, I conclude that $\boldsymbol{\varepsilon}_{\tau_1}(\chi, w) = \mathrm{DiD}_{\tau_0,\tau_1}(w)$ for any wage $w \in \mathbb{R}$.

□

C.3. Identification of Technology

Proof of Proposition 3.

Consider two skill types χ and χ'. For any time period $\tilde{\tau} \in \{\tau, \tau'\}$, equation (10) implies:

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

Because ρ_j and $\{\theta_{j\chi}\}_\chi$ are fixed 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.

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

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

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 at the firm. I write:

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

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

□

C.4. *Identification of Non-Wage Amenities*

Proof of Proposition 5.

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

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

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

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

Note that Proposition 2 establishes that $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$ is point-identified from $\mathrm{DiD}_{\tau_0,\tau_1}\big(w^{\mathrm{eff}}_{j\tau_0}(\chi)|\chi\big)$.

□

C.5. *Identification of Worker Preferences*

Proof of Proposition 6.

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:

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

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

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

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 formula:

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

□

C.6. *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:

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

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

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

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

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:

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

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

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

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