

Optimal Taxation and Market Power

Jan Eeckhout, Chunyang Fu, Wenjian Li, and Xi Weng

I. Environment

1.1. The Cournot Competitive Tax Equilibrium

When the first-order conditions are both necessary and sufficient for the individuals' and the final good producer's problem, the equilibrium allocations are governed by conditions (12) to (18) alongside the individuals' budget constraints. Given the technology considered in this paper and assuming $\phi_o(l_o) = \frac{l_o^{1+\frac{1}{\varepsilon_o}}}{1+\frac{1}{\varepsilon_o}}$, the following conditions hold in the symmetric equilibrium:

1. First-order conditions

$$(A1) \quad P(\theta_e) = [N_e f_e(\theta_e)]^{-\frac{1}{\sigma}} \zeta(\theta_e) A^{\frac{\sigma-1}{\sigma}} Q_{ij}(\theta_e)^{-\frac{1}{\sigma}} Q^{\frac{1}{\sigma}},$$

and

$$(A2) \quad \left[1 + \frac{\partial \ln P_{ij}(Q_{ij}(\theta_e), Q_{-ij}(\theta_e), \theta_e)}{\partial \ln Q_{ij}(\theta_e)} \right] \frac{\zeta P(\theta_e) Q_{ij}(\theta_e)}{L_w(\theta_e)} - \frac{W}{1-t_s} = 0$$

and

$$(A3) \quad W x_w(\theta_w) [1 - T'_w(W x_w(\theta_w) l_w(\theta_w))] = l_w(\theta_w)^{\frac{1}{\varepsilon_w}},$$

and

$$\left[1 + \frac{\partial \ln P_{ij}}{\partial \ln Q_{ij}(\theta_e)} \right] P(\theta_e) Q_{ij}(\theta_e) (1-t_s) [1 - T'_e(y_e(\theta_e))] = l_e(\theta_e)^{1+\frac{1}{\varepsilon_e}}, \theta_e \in \Theta_o.$$

Combination of (A2) and (19) (i.e., $\mu(\theta_e) = \frac{P(\theta_e)}{W \left[\frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)} (1-t_s) \right]} = \frac{\zeta P(\theta_e) Q_{ij}(\theta_e) (1-t_s)}{W L_w(\theta_e)}$) delivers (20). Substituting

$1 + \frac{\partial \ln P(Q_{ij}(\theta_e), \theta_e)}{\partial \ln Q_{ij}(\theta_e)}$ by (20), we have:

$$(A4) \quad W L_w(\theta_e) = \frac{\zeta (1-t_s)}{\mu(\theta_e)} P(\theta_e) Q_{ij}(\theta_e),$$

and

$$(A5) \quad \frac{P(\theta_e) Q_{ij}(\theta_e) (1-t_s)}{\mu(\theta_e)} [1 - T'_e(y_e(\theta_e))] = l_e(\theta_e)^{1+\frac{1}{\varepsilon_e}}, \theta_e \in \Theta_e.$$

2. Inverse demand function

$$(A6) \quad \begin{aligned} & P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) \\ &= \chi(\theta_e) A^{\frac{\sigma-1}{\sigma}} Q_{ij}^{-\frac{1}{\eta(\theta_e)}} I^{-\left[\frac{1}{\eta(\theta_e)} - \frac{1}{\sigma}\right] \frac{\eta(\theta_e)}{\eta(\theta_e)-1}} \\ & \quad \times \left[(I-1) Q_{-ij}(\theta_e)^{\frac{\eta(\theta_e)-1}{\eta(\theta_e)}} + Q_{ij}^{\frac{\eta(\theta_e)-1}{\eta(\theta_e)}} \right]^{\left[\frac{1}{\eta(\theta_e)} - \frac{1}{\sigma}\right] \frac{\eta(\theta_e)}{\eta(\theta_e)-1}} \left[\frac{Q}{N} \right]^{\frac{1}{\sigma}}, \end{aligned}$$

where $Q_{ij}(\theta_e)$ is treated as given by the entrepreneurs.

3. Labor market clear condition

$$(A7) \quad \int_{\theta_w} x_w(\theta_w) l_w(\theta_w) f_w(\theta_w) d\theta_w = W^{\varepsilon_w} \int_{\theta_w} \zeta(\theta_w)^{\varepsilon_w+1} [1 - \tau_w(\theta_w)]^{\varepsilon_w} f_w(\theta_w) d\theta_w$$

4. Meanwhile, in the equilibrium, we have:

$$(A8) \quad Q = \int_{\theta_e} N_e f_e(\theta_e) P(\theta_e) Q_{ij}(\theta_e) d\theta_e.$$

The above parts 1 to 4 solve the equilibrium allocation $\{L_w(\theta_e), l_e(\theta_e), l_w(\theta_w)\}$, price system $\{P(\theta_e), W\}$, and total output Q . Then one can derive other allocations with individuals' budget constraints.

1.2. Elasticities

(I) ELASTICITIES OF PRICE AND WAGE.

We define the elasticity of firm-level outputs with respect to the entrepreneurial effort $l_e(\theta_e)$ and labor inputs $L_w(\theta_e)$ respectively as:

$$\varepsilon_{l_e}^{Q_{ij}}(\theta_e) \equiv \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln l_e(\theta_e)} \quad \text{and} \quad \varepsilon_{L_w}^{Q_{ij}}(\theta_e) \equiv \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln L_w(\theta_e)}.$$

In the nested-CES economy, $\varepsilon_{l_e}^{Q_{ij}}(\theta_e) = 1$ and $\varepsilon_{L_w}^{Q_{ij}}(\theta_e) = \zeta$ are constants.

We define the sales elasticity and price elasticity, respectively as

$$\varepsilon_{Q_{ij}}^S(\theta_e) \equiv \frac{\partial \ln [P_{ij}(Q_{ij}(\theta_e), Q_{ij}(\theta_e), \theta_e) Q_{ij}(\theta_e)]}{\partial \ln Q_{ij}(\theta_e)} \quad \text{and} \quad \varepsilon_{Q_{ij}}^P(\theta_e) \equiv \frac{\partial \ln P(\theta_e)}{\partial \ln Q_{ij}(\theta_e)}.$$

Define the own-inverse demand elasticity and cross-inverse demand elasticity as:

$$\varepsilon_{Q_{ij}}^{P,own}(\theta_e) \equiv \frac{\partial \ln P_{ij}}{\partial \ln Q_{ij}}|_{Q_{ij}=Q_{ij}(\theta_e)} \quad \text{and} \quad \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e) \equiv \frac{\partial \ln P_{ij}}{\partial \ln Q_{ij}(\theta_e)}|_{Q_{ij}=Q_{ij}(\theta_e)}.$$

By definitions, $\varepsilon_{Q_{ij}}^P(\theta_e) = \varepsilon_{Q_{ij}}^{P,own}(\theta_e) + \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e)$ and $\varepsilon_{Q_{ij}}^S(\theta_e) = 1 + \varepsilon_{Q_{ij}}^P(\theta_e)$. Moreover, notice that $\mu(\theta_e) = \frac{1}{1 + \varepsilon_{Q_{ij}}^{P,own}(\theta_e)}$. We have

$$(A9) \quad \varepsilon_{Q_{ij}}^S(\theta_e) = \frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e), \quad \forall \theta_e \in \Theta_e.$$

Following [Sachs, Tsyvinski and Werquin \(2020\)](#), we denote by $\varepsilon_{L_w}^\omega(\theta'_e, \theta_e)$ and $\varepsilon_{l_e}^\omega(\theta'_e, \theta_e)$ the cross elasticities of wage with respect to $L_w(\theta_e)$ and $l_e(\theta_e)$ for any $(\theta_e, \theta'_e) \in \Theta_e^2$:

$$\varepsilon_{L_w}^\omega(\theta'_e, \theta_e) = \begin{cases} \frac{\partial \ln \omega(\theta'_e)}{\partial \ln L_w(\theta_e)}, & \theta'_e \neq \theta_e, \\ \lim_{\theta'_e \rightarrow \theta_e} \frac{\partial \ln \omega(\theta'_e)}{\partial \ln L_w(\theta_e)}, & \theta'_e = \theta_e, \end{cases}$$

and

$$\varepsilon_{l_e}^\omega(\theta'_e, \theta_e) = \begin{cases} \frac{\partial \ln \omega(\theta'_e)}{\partial \ln l_e(\theta_e)}, & \theta'_e \neq \theta_e, \\ \lim_{\theta'_e \rightarrow \theta_e} \frac{\partial \ln \omega(\theta'_e)}{\partial \ln l_e(\theta_e)}, & \theta'_e = \theta_e. \end{cases}$$

We denote $\varepsilon_{L_w}^\omega(\theta_e)$ and $\varepsilon_{l_e}^\omega(\theta_e)$ as the own elasticities of wages with respect to $L_w(\theta_e)$ and $l_e(\theta_e)$. These own

elasticities of wages are defined by the following relationships:

$$\frac{\partial \ln \omega(\theta_e)}{\partial \ln L_w(\theta_e)} = \varepsilon_{L_w}^\omega(\theta_e, \theta_e) + \varepsilon_{L_w}^\omega(\theta_e) \delta(\theta'_e - \theta_e),$$

and

$$\frac{\partial \ln \omega(\theta_e)}{\partial \ln l_e(\theta_e)} = \varepsilon_{l_e}^\omega(\theta_e, \theta_e) + \varepsilon_{l_e}^\omega(\theta_e) \delta(\theta'_e - \theta_e),$$

where δ denotes the Dirac delta function, $(\theta_e, \theta'_e) \in \Theta_e^2$. See Supplemental Appendix I.1.2 for details about the above elasticities.

Recall that in the nested-CES economy, we have:

$$\begin{aligned} P_{ij}(\theta_e) &= N^{-\frac{1}{\sigma}} A^{\frac{\sigma-1}{\sigma}} \chi(\theta_e) Q_{ij}(\theta_e)^{-\frac{1}{\sigma}} Q^{\frac{1}{\sigma}}, \\ Q_{ij}(\theta_e) &= x_e(\theta_e) l_e(\theta_e) L_w(\theta_e)^\zeta, \\ \omega(\theta_e) &= \frac{\chi(\theta_e) N^{-\frac{1}{\sigma}} A^{\frac{\sigma-1}{\sigma}} Q_{ij}(\theta_e)^{-\frac{1}{\sigma}} Q^{\frac{1}{\sigma}}}{\mu(\theta_e)} \zeta x_e(\theta_e) l_e(\theta_e) L_w(\theta_e)^{\zeta-1}, \quad \forall \theta_e \in \Theta_e. \end{aligned}$$

It's easy to see that $\varepsilon_{L_w}^\omega(\theta'_e, \theta_e)$ and $\varepsilon_{l_e}^\omega(\theta'_e, \theta_e)$ are independent of θ'_e . By definition,

$$(A10) \quad \varepsilon_{L_w}^\omega(\theta_e) = \zeta \left(1 - \frac{1}{\sigma}\right) - 1 < 0, \quad \text{and} \quad \varepsilon_{l_e}^\omega(\theta) = 1 - \frac{1}{\sigma} > 0.$$

Note that both $\varepsilon_{L_w}^\omega(\theta_e)$ and $\varepsilon_{l_e}^\omega(\theta)$ are constants.

By solving the final good producer's problem, we immediately derive the price equation (A1) and the inverse demand function (A6). Utilizing the definitions of price elasticities, we obtain the following results for this economy:

$$\begin{aligned} \varepsilon_{Q_{ij}}^P(\theta_e) &= -\frac{1}{\sigma}, \quad \varepsilon_{Q_{ij}}^S(\theta_e) = \frac{\sigma-1}{\sigma}, \\ \varepsilon_{Q_{ij}}^{P,own}(\theta_e) &= -\left[\frac{1}{\eta(\theta_e)} \frac{I-1}{I} + \frac{1}{\sigma} \frac{1}{I} \right], \\ \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e) &= \left[\frac{1}{\eta(\theta_e)} - \frac{1}{\sigma} \right] \frac{I-1}{I}, \quad \forall \theta_e \in \Theta_e. \end{aligned}$$

Notice that $\mu(\theta_e) = \frac{1}{1 + \varepsilon_{Q_{ij}}^{P,own}(\theta_e)}$, we have:

$$(A11) \quad \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e) = -\frac{1}{\mu(\theta_e)} + \frac{\sigma-1}{\sigma}, \quad \forall \theta_e \in \Theta_e.$$

Under our production technology, we have

$$\begin{aligned} (A12) \quad & \frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} \\ &= \frac{d \ln P(\theta_e)}{d \theta_e} - \varepsilon_{Q_{ij}}^{P,own}(\theta_e) \frac{d \ln Q_{ij}(\theta_e)}{d \theta_e} \\ &= \frac{\chi'(\theta_e)}{\chi(\theta_e)} - \frac{1}{\sigma} \frac{Q'_{ij}(\theta_e)}{Q_{ij}(\theta_e)} + \left[\frac{I-1}{I} \frac{1}{\eta(\theta_e)} + \frac{1}{I} \frac{1}{\sigma} \right] \frac{Q'_{ij}(\theta_e)}{Q_{ij}(\theta_e)} \\ &= \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e) \frac{d \ln Q_{ij}(\theta_e)}{d \theta_e}, \quad \forall \theta_e \in \Theta_e. \end{aligned}$$

Specially, when $I = 1$, we have $\frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} = \frac{\chi'(\theta_e)}{\chi(\theta_e)}$.

(II) ELASTICITY OF PROFIT.

Consider a small increase (i.e., $d\tau$) in the marginal tax rate faced by a θ_e -type entrepreneur. In this case, the tax reform has no first-order effects on the aggregate values and the actions of other types. Thus, aggregate variables, including the outputs of final goods and the price of labor factors, remain unchanged. As in [Scheuer and Werning \(2017\)](#), the elasticity derived here is a micro elasticity.

The optimal choices of the θ_e -type entrepreneur (i.e., l_e and L_w) satisfy the following first-order conditions in the equilibrium:

$$(A13) \quad WL_w = (1 - t_s) \frac{P_{ij} Q_{ij}}{\mu(\theta_e)} \frac{\partial \ln Q_{ij}}{\partial \ln L_w},$$

and

$$(A14) \quad \phi'_e(l_e) = [1 - T'_e(P_{ij} Q_{ij} (1 - t_s) - WL_w) - d\tau] \frac{P_{ij} Q_{ij} (1 - t_s)}{l_e \mu(\theta_e)},$$

where P_{ij} and Q_{ij} refer to $P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)$ and $Q_{ij}(x_e(\theta_e)l_e, L_w)$, respectively. Note that $Q_{ij}(\theta_e)$ in $P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)$ is treated as given by the agent. Equation (A13) is derived by (16) and (20), and equation (A14) is derived by (18) and (20).

The cases we considered has constant $\frac{\partial \ln Q_{ij}}{\partial \ln L_w}$ and exogenous $\mu(\theta_e)$. Set $\xi = \frac{\partial \ln Q_{ij}}{\partial \ln L_w}$. Combination of (A13) and (A14) gives:

$$(A15) \quad WL_w = (1 - t_s) \frac{P_{ij} Q_{ij}}{\mu(\theta_e)} \xi,$$

and

$$(A16) \quad \phi'_e(l_e) = \left[1 - T'_e \left(\left(\frac{\mu(\theta_e)}{\xi} - 1 \right) WL_w \right) - d\tau \right] \frac{WL_w}{\xi} \frac{1}{l_e}.$$

The θ_e -type entrepreneur's reaction to the tax reform can be described by differential equations of the first-order conditions. Total differential of (A16) yields:

$$(A17) \quad \frac{1 + \varepsilon_e}{\varepsilon_e} \frac{dl_e}{l_e} = \frac{dL_w}{L_w} \left[1 - \frac{y_e T''_e(y_e)}{1 - T'_e(y_e)} \right] - \frac{d\tau}{1 - T'_e(y_e)}.$$

Total differential of (A13) gives:

$$\frac{dL_w}{L_w} = \frac{1}{\mu(\theta_e)} \left[\frac{dl_e}{l_e} + \xi \frac{dL_w}{L_w} \right].$$

Note that $Q_{-ij}(\theta_e)$ also changes with the tax reform, which is captured by $\varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)$.

A combination of the above two equations gives

$$(A18) \quad -\frac{\frac{dL_w}{L_w}}{\frac{d\tau}{1 - T'_e(y_e)}} = \frac{1}{\frac{1 + \varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \xi] - \left[1 - \frac{y_e T''_e(y_e)}{1 - T'_e(y_e)} \right]}.$$

Notice that $y_e = PQ_{ij}(1 - t_s) - WL_w$ and $WL_w = (1 - t_s) PQ_{ij} \frac{\xi}{\mu(\theta_e)}$. We have $y_e = WL_w \left(\frac{\mu(\theta_e)}{\xi} - 1 \right)$ and

$$(A19) \quad \frac{dy_e}{y_e} = \frac{dL_w}{L_w}.$$

Define $\varepsilon_{1-\tau_e}^{y_e}(\theta_e) \equiv -\frac{dy_e(\theta_e)}{y_e(\theta_e)} / \frac{d\tau}{1-T'_e(y_e(\theta_e))}$ as the non-linear profit elasticity. We have:

$$(A20) \quad \varepsilon_{1-\tau_e}^{y_e}(\theta_e) = \frac{1}{\frac{1+\varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \tilde{\zeta}] - \left[1 - \frac{y_e(\theta_e)T'_e(y_e(\theta_e))}{1-T'_e(y_e(\theta_e))}\right]}.$$

(III) GENERAL ELASTICITY OF PROFIT.

Consider again a small increase (i.e., $d\tau$) in the marginal tax rate faced by the θ_e -type entrepreneur. While this tax reform has no first-order effects on aggregate values or the actions of other types, it now affects aggregate variables, including the outputs of final goods and the price of labor factors, through changes in $Q_{-ij}(\theta_e)$. Define $\tilde{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e) \equiv -\frac{dy_e(\theta_e)}{y_e(\theta_e)} / \frac{d\tau}{1-T'_e(y_e(\theta_e))}$ as the elasticity of profit in this context. We refer to $\tilde{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)$ the *general elasticity of profit* and $\varepsilon_{1-\tau_e}^{y_e}(\theta_e)$ the *partial elasticity of profit* to distinguish them from each other.

The conditions in equations (A15) and (A16) still hold. However, when taking total differential, the change in $Q_{-ij}(\theta_e)$ must now be considered. Note that $Q_{-ij}(\theta_e)$ equals $Q_{ij}(\theta_e)$. Taking the total differential of equation (A15) yields:

$$\begin{aligned} d \ln L_w &= \left[\frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d \ln Q_{ij}} + 1 \right] d \ln Q_{ij} \\ &= \left[1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e) + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \right] [d \ln l_e + \tilde{\zeta} d \ln L_w] \\ &= \left[\frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \right] [d \ln l_e + \tilde{\zeta} d \ln L_w] \\ &= \varepsilon_{Q_{ij}}^S(\theta_e) [d \ln l_e + \tilde{\zeta} d \ln L_w] \end{aligned}$$

i.e.,

$$\frac{dl_e}{l_e} = \frac{dL_w}{L_w} \frac{1 - \varepsilon_{Q_{ij}}^S(\theta_e) \tilde{\zeta}}{\varepsilon_{Q_{ij}}^S(\theta_e)},$$

where $\varepsilon_{Q_{ij}}^S(\theta_e) = \frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) = \frac{\sigma-1}{\sigma}$.

$$\frac{dl_e}{l_e} = \frac{dL_w}{L_w} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \tilde{\zeta} \right],$$

On the other hand, we still have (A17). Substitute $\frac{dl_e}{l_e}$ in (A17) by the above equation:

$$(A21) \quad \tilde{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e) = \frac{1}{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln l_e(\theta_e)} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e) \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln l_e(\theta_e)}} - \tilde{\zeta} \right] - \left[1 - \frac{y_e(\theta_e)T'_e(y_e(\theta_e))}{1-T'_e(y_e(\theta_e))} \right]}.$$

The general profit elasticity captures the extent to which a marginal profit tax reform affects firm-level profits by altering both the firm's own decisions and its competitors' outputs. In contrast, the partial elasticity reflects only the firm's direct response to the tax reform.

(IV) ELASTICITY OF PROFIT TO THE SKILL.

In this section, we provide the elasticity of profit to the skill. Remember that the entrepreneur's FOCs imply

$$\phi'_e(l_e(\theta_e)) = \left[1 - T'_e \left(\left(\frac{\mu(\theta_e)}{\tilde{\zeta}} - 1 \right) W L_w(\theta_e) \right) \right] \frac{W L_w(\theta_e)}{\tilde{\zeta}} \frac{1}{l_e(\theta_e)},$$

and

$$WL_w(\theta_e) = (1 - t_s) P(\theta_e) Q_{ij}(\theta_e) \frac{\xi}{\mu(\theta_e)}.$$

Take the derivative of both sides of the above equations with respect to θ :

$$\left(1 + \frac{1}{\varepsilon_e}\right) \frac{l'_e(\theta_e)}{l_e(\theta_e)} = \frac{L'_w(\theta_e)}{L_w(\theta_e)} \left(1 - \frac{T'_e(y_e(\theta_e))y_e(\theta_e)}{1 - T'_e(y_e(\theta))}\right) - \frac{T''_e(y_e(\theta_e))y_e(\theta_e)}{1 - T'_e(y_e(\theta_e))} \frac{\frac{\mu'(\theta_e)}{\xi}}{\frac{\mu(\theta_e)}{\xi} - 1}$$

and

$$(A22) \quad \frac{L'_w(\theta_e)}{L_w(\theta_e)} \left(1 - \xi \varepsilon_{Q_{ij}}^S(\theta_e)\right) = \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{l'_e(\theta_e)}{l_e(\theta_e)}\right] \varepsilon_{Q_{ij}}^S(\theta_e) + \frac{\chi'(\theta_e)}{\chi(\theta_e)} - \frac{\mu'(\theta_e)}{\mu(\theta_e)},$$

where we use

$$(A23) \quad \frac{Q'_{ij}(\theta_e)}{Q_{ij}(\theta_e)} = \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \xi \frac{L'_w(\theta_e)}{L_w(\theta_e)}.$$

Combine the above the equations:

$$\frac{L'_w(\theta_e)}{L_w(\theta_e)} = \frac{\frac{1+\varepsilon_e}{\varepsilon_{Q_{ij}}^S(\theta_e)} \left[\frac{X'_e(\theta_e)}{X_e(\theta_e)} - \frac{\mu'(\theta_e)}{\mu(\theta_e)}\right] - \frac{T''_e(y_e(\theta))y_e(\theta)}{1 - T'_e(y_e(\theta))} \frac{\frac{\mu'(\theta_e)}{\xi}}{\frac{\mu(\theta_e)}{\xi} - 1}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \xi\right] - \left[1 - \frac{T''_e(y_e(\theta))y_e(\theta)}{1 - T'_e(y_e(\theta))}\right]}.$$

Last, notice that $y_e(\theta_e) = WL_w(\theta_e) \left(\frac{\mu(\theta_e)}{\xi} - 1\right)$. We have:

$$(A24) \quad \frac{y'_e(\theta_e)}{y_e(\theta_e)} = \frac{\mu'(\theta_e)}{\mu(\theta_e) - \xi} + \frac{L'_w(\theta)}{L_w(\theta)}, \forall \theta_e \in \Theta_e.$$

Combine the above two equations:

$$\frac{y'_e(\theta_e)}{y_e(\theta_e)} = \frac{\frac{1+\varepsilon_e}{\varepsilon_{Q_{ij}}^S(\theta_e)} \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)}\right] - \left[1 - \left(\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \xi\right) \frac{1+\varepsilon_e}{\varepsilon_e}\right] \frac{\frac{\mu'(\theta_e)}{\xi}}{\frac{\mu(\theta_e)}{\xi} - 1}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \xi\right] - \left[1 - \frac{T''_e(y_e(\theta))y_e(\theta)}{1 - T'_e(y_e(\theta))}\right]}.$$

Use $\tilde{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e) = \frac{1}{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{1-\frac{\sigma-1}{\sigma}\xi}{\frac{\sigma-1}{\sigma}} - \left[1 - \frac{\pi T'_e(y_e)}{1 - T'_e(y_e)}\right]}$:

$$\frac{y'_e(\theta_e)}{y_e(\theta_e)} = \tilde{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)}\right] - \left[1 - \left(\frac{1}{\frac{\sigma-1}{\sigma}} - \xi\right) \frac{1+\varepsilon_e}{\varepsilon_e}\right] \frac{\frac{\mu'(\theta_e)}{\xi}}{\frac{\mu(\theta_e)}{\xi} - 1} \right].$$

1.3. Preliminary Analysis of the Equilibrium

(I) ALLOCATIONS AND PRICES.

To simplify the expression, in the following analysis, we consider $t_s = 0$. Besides, we use $X_e(\theta_e) = A^{\frac{\sigma-1}{\sigma}} x_e(\theta_e)^{\frac{\sigma-1}{\sigma}} \chi(\theta_e)$ and $f_w(\theta_w) = f_e(\theta_e) = 1$ to shorten the expressions. Combination of (A4), (A5), and (A1)

gives:

$$(A25) \quad l_e(\theta_e) = \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma \varepsilon_e}{\varepsilon_e + \sigma}} L_w(\theta_e)^{\frac{\xi(\sigma-1)\varepsilon_e}{\varepsilon_e + \sigma}} [1 - \tau_e(\theta_e)]^{\frac{\varepsilon_e \sigma}{\sigma + \varepsilon_e}}.$$

Substituting $P(\theta_e)$ and $Q_{ij}(\theta_e)$ in (A4) with (A1) and $Q_{ij}(\theta_e) = x_e(\theta_e) l_e(\theta_e) L_w(\theta_e)^\xi$, respectively, we have:

$$(A26) \quad \begin{aligned} L_w(\theta_e) &= \frac{\xi}{W \mu(\theta_e)} \chi(\theta_e) \left[x_e(\theta_e) l_e(\theta_e) L_w(\theta_e)^\xi \right]^{\frac{\sigma-1}{\sigma}} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} [1 - \tau_e(\theta_e)]^{\frac{\varepsilon_e(\sigma-1)}{\sigma + \varepsilon_e}} \\ &= \frac{\xi}{W} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e + \sigma}} L_w(\theta_e)^{\frac{\xi(\sigma-1)(\varepsilon_e+1)}{\sigma + \varepsilon_e}} [1 - \tau_e(\theta_e)]^{\frac{\varepsilon_e(\sigma-1)}{\sigma + \varepsilon_e}}, \end{aligned}$$

where we substitute $l_e(\theta_e)$ with (A25) in the second equation.

Rearranging the above equation gives:

$$(A27) \quad L_w(\theta_e) = \left(\frac{\xi}{W} \right)^{\frac{\sigma + \varepsilon_e}{\sigma + \varepsilon_e - \xi(\sigma-1)(\varepsilon_e+1)}} \left\{ \frac{X_e(\theta_e) [1 - \tau_e(\theta_e)]^{\frac{(\sigma-1)\varepsilon_e}{\sigma(\varepsilon_e+1)}}}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right\}^{\frac{\sigma(\varepsilon_e+1)}{\sigma + \varepsilon_e - \xi(\sigma-1)(\varepsilon_e+1)}}.$$

Substituting the above equation into (A25), we have:

$$(A28) \quad P(\theta_e) Q_{ij}(\theta_e) = \mu(\theta_e) \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)(\varepsilon_e+1)}{\sigma + \varepsilon_e - \xi(\sigma-1)(\varepsilon_e+1)}} \left\{ \frac{X_e(\theta_e) [1 - \tau_e(\theta_e)]^{\frac{\varepsilon_e(\sigma-1)}{\sigma(\varepsilon_e+1)}}}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right\}^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e + \sigma - \xi(\sigma-1)(\varepsilon_e+1)}}$$

and

$$(A29) \quad l_e(\theta_e) = \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)\varepsilon_e}{\sigma + \varepsilon_e - \xi(\sigma-1)(\varepsilon_e+1)}} \left\{ \frac{X_e(\theta_e) [1 - \tau_e(\theta_e)]^{\frac{\sigma - \xi(\sigma-1)}{\sigma}}}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right\}^{\frac{\sigma \varepsilon_e}{\varepsilon_e + \sigma - \xi(\sigma-1)(\varepsilon_e+1)}}.$$

Equation (A3) gives:

$$(A30) \quad l_w(\theta_w) = \{W x_w(\theta_w) [1 - \tau_w(\theta_w)]\}^{\varepsilon_w}.$$

The three equations above together with (A7) and (A8) solve the symmetric equilibrium prices $\{P(\theta_e), W\}$, allocations $\{L_w(\theta_e), l_e(\theta_e), l_w(\theta_w)\}$, and total output Q . Lastly, one can derive other allocations with individuals' budget constraints. See below for details.

For later use, we define

$$(A31) \quad \begin{aligned} A_1 &= \int_{\theta_e} N \mu(\theta_e) \left[\frac{X_e(\theta_e) [1 - \tau_e(\theta_e)]^{\frac{(\sigma-1)\varepsilon_e}{\sigma(\varepsilon_e+1)}}}{\mu(\theta_e) N_e^{\frac{1}{\sigma}}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e + \sigma - \xi(\sigma-1)(\varepsilon_e+1)}} d\theta_e, \\ A_2 &= \int N \left[\frac{X_e(\theta_e) [1 - \tau_e(\theta_e)]^{\frac{(\sigma-1)\varepsilon_e}{\sigma(\varepsilon_e+1)}}}{\mu(\theta_e) N_e^{\frac{1}{\sigma}}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e + \sigma - \xi(\sigma-1)(\varepsilon_e+1)}} d\theta_e, \\ A_3 &= N_w \xi^{\varepsilon_w} \int_{\theta_w} x(\theta_w)^{\varepsilon_w+1} [1 - \tau_w(\theta_w)]^{\varepsilon_w} d\theta_w. \end{aligned}$$

Substituting $L_w(\theta_e)$ in (A4) with (A27), we have:

$$(A32) \quad P(\theta_e)Q_{ij}(\theta_e) = \mu(\theta_e) \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \\ \times [1 - \tau_e(\theta_e)]^{\frac{(\sigma-1)\varepsilon_e}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}}.$$

Substituting $P(\theta_e)Q_{ij}(\theta_e)$ in (A8) with (A32), we have

$$Q = \int_{\theta_e} N_e \mu(\theta_e) \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} [1 - \tau_e(\theta_e)]^{\frac{(\sigma-1)\varepsilon_e}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} d\theta_e,$$

which gives the following equation by the definition of A_1 :

$$(A33) \quad Q = \left(\frac{\xi}{W} \right)^{\frac{\xi(\varepsilon_e+1)}{1-\xi(\varepsilon_e+1)}} A_1^{\frac{\sigma+\varepsilon_e-\xi(\sigma-1)(\varepsilon_e+1)}{1-\xi(\varepsilon_e+1)} \frac{1}{\sigma-1}}.$$

Similarly, substituting $L_w(\theta_e)$ in (A7) with (A27), we have the aggregate labor demand

$$(A34) \quad L^D \equiv \left(\frac{\xi}{W} \right)^{\frac{1}{1-\xi(\varepsilon_e+1)}} (A_1)^{\frac{\varepsilon_e+1}{(\sigma-1)[1-\xi(\varepsilon_e+1)]}} A_2.$$

On the other hand, according to (A7) and (A30), we have the aggregate labor supply

$$(A35) \quad L^S \equiv N_w [W]^{\varepsilon_w} \int_{\theta_w} x(\theta_w)^{\varepsilon_w+1} f_w(\theta_w) d\theta_w = \left[\frac{W}{\xi} \right]^{\varepsilon_w} A_3.$$

Combining (A34) and (A35) gives

$$\left[\frac{W}{\xi} \right]^{\varepsilon_w + \frac{1}{1-\xi(\varepsilon_e+1)}} = (A_1)^{\frac{\varepsilon_e+1}{(\sigma-1)[1-\xi(\varepsilon_e+1)]}} \frac{A_2}{A_3},$$

that is

$$(A36) \quad W = \xi \left[(A_1)^{\frac{\varepsilon_e+1}{(\sigma-1)[1-\xi(\varepsilon_e+1)]}} \frac{A_2}{A_3} \right]^{\frac{1}{\varepsilon_w + \frac{1}{1-\xi(\varepsilon_e+1)}}}.$$

Lastly, substituting W in (A33) with (A36), we have

$$(A37) \quad Q = \left[\frac{A_3}{A_2 A_1^{\frac{\varepsilon_e+1}{(\sigma-1)[1-\xi(\varepsilon_e+1)]}}} \right]^{\frac{1}{\varepsilon_w + \frac{1}{1-\xi(\varepsilon_e+1)}} \frac{\xi(\varepsilon_e+1)}{1-\xi(\varepsilon_e+1)}} A_1^{\frac{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}{(\sigma-1)[1-\xi(\varepsilon_e+1)]}}.$$

Then we can derive $l_w(\theta_w)$, $L_w(\theta_e)$, and $l_e(\theta_e)$ by substituting Q and W into (A27), (A29), and (A30). Moreover, by definition, we have

$$(A38) \quad Q_{ij}(\theta_e) = x_e(\theta_e) \left(\frac{\xi}{W} \right)^{\frac{\xi\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)\xi+\sigma\varepsilon_e}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} [1 - \tau_e(\theta_e)]^{\frac{\sigma\varepsilon_e}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}},$$

and

$$(A39) \quad P(\theta_e) = \frac{\mu(\theta_e)}{x_e(\theta_e)} \left(\frac{\xi}{W} \right)^{\frac{-\xi(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma-\sigma(\varepsilon_e+1)\xi}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} [1 - \tau_e(\theta_e)]^{\frac{\sigma-\xi(\sigma-1)(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}}.$$

According to the above results, we have:

$$(A40) \quad \frac{d \ln l_e(\theta_e)}{d\theta_e} = \frac{\frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + \frac{\sigma-\xi(\sigma-1)}{\sigma} \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \frac{\sigma-1}{\sigma} \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \xi \right)},$$

$$(A41) \quad \frac{d \ln L_w(\theta_e)}{d\theta_e} = \frac{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + \frac{\sigma-1}{\sigma} \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \frac{\sigma-1}{\sigma} \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \xi \right)},$$

$$\frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} = \frac{d \ln x_e(\theta_e)}{d\theta_e} + \frac{\left(\xi + \frac{\varepsilon_e}{1+\varepsilon_e} \right) \frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1}{\sigma} + \frac{\sigma-1}{\sigma} \left(\frac{1}{1+\varepsilon_e} - \xi \right)},$$

$$\frac{d \ln [Q_{ij}(\theta_e) P(\theta_e)]}{d\theta_e} = \frac{\frac{1-\sigma}{\sigma} \left(\xi + \frac{\varepsilon_e}{1+\varepsilon_e} \right) \frac{d \ln \mu(\theta_e)}{d\theta_e} + \frac{d \ln X_e(\theta_e)}{d\theta_e}}{\frac{1}{\sigma} + \frac{\sigma-1}{\sigma} \frac{1}{1+\varepsilon_e} - \xi} + \frac{(\sigma-1) \frac{\varepsilon_e}{1+\varepsilon_e} \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e}}{1 + (\sigma-1) \frac{1}{1+\varepsilon_e} - \xi}.$$

Moreover, we have

$$\begin{aligned} \frac{d \ln y_e(\theta_e)}{d\theta_e} &= \frac{d \ln \frac{\mu(\theta_e)-\xi}{\xi} W L_w(\theta_e)}{d\theta_e} = \frac{d \ln [\mu(\theta_e) - \xi]}{d\theta_e} + \frac{d \ln L_w(\theta_e)}{d\theta_e} \\ &= \frac{d \ln [\mu(\theta_e) - \xi]}{d\theta_e} + \frac{\frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + \frac{\sigma-1}{\sigma} \frac{\varepsilon_e}{1+\varepsilon_e} \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e}}{1 - \frac{\sigma-1}{\sigma} \left(\frac{\varepsilon_e}{1+\varepsilon_e} + \xi \right)} \end{aligned}$$

and

$$\begin{aligned} &\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} \\ &= \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \mu(\theta_e) \left\{ \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[\frac{\sigma-1}{\sigma} - \frac{1}{\mu(\theta_e)} \right] \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} \right\} \\ &= \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \mu(\theta_e) \left\{ \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[\frac{\sigma-1}{\sigma} - \frac{1}{\mu(\theta_e)} \right] \left[\frac{d \ln x_e(\theta_e)}{d\theta_e} + \frac{d \ln l_e(\theta_e)}{d\theta_e} + \xi \frac{d \ln L_w(\theta_e)}{d\theta_e} \right] \right\} \\ &= \mu(\theta_e) \frac{X'(\theta_e)}{X(\theta_e)} + \left[\frac{\sigma-1}{\sigma} \mu(\theta_e) - 1 \right] \left[\frac{d \ln l_e(\theta_e)}{d\theta_e} + \xi \frac{d \ln L_w(\theta_e)}{d\theta_e} \right] \\ &= \mu(\theta_e) \frac{X'(\theta_e)}{X(\theta_e)} + \left[\frac{\sigma-1}{\sigma} \mu(\theta_e) - 1 \right] \left[\frac{\left(1 + \xi \frac{1+\varepsilon_e}{\varepsilon_e} \right) \frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \frac{\sigma-1}{\sigma} \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \xi \right)} \right] \\ &= [\mu(\theta_e) - \xi] \frac{d \ln [\mu(\theta_e) - \xi]}{d\theta_e} + \frac{\left[\left[\mu(\theta_e) - \left(\frac{\varepsilon_e}{1+\varepsilon_e} + \xi \right) \right] \frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + \left[\frac{\sigma-1}{\sigma} \mu(\theta_e) - 1 \right] \frac{\varepsilon_e}{1+\varepsilon_e} \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e} \right]}{1 - \frac{\sigma-1}{\sigma} \left(\frac{\varepsilon_e}{1+\varepsilon_e} + \xi \right)}, \end{aligned}$$

where we use (A40) and (A41) to substitute $\frac{d \ln l_e(\theta_e)}{d\theta_e}$ and $\frac{d \ln L_w(\theta_e)}{d\theta_e}$ in the fourth equation.

(II) SKILL GAP.

To derive the optimal profit tax formula in terms of parameters, we need $\frac{d \ln L_w(\theta_e)}{d\theta_e}$, $\frac{d \ln l_e(\theta_e)}{d\theta_e}$ and $\frac{d \ln y_e(\theta_e)}{d\theta_e}$ in terms of θ_e and profit tax rate. The entrepreneur's choice of labor inputs satisfies (A4), i.e., $WL_w(\theta_e) = \frac{\xi}{\mu(\theta_e)} P(\theta_e) Q_{ij}(\theta_e)$. Substitute $P(\theta_e)$ in (A4) by the inverse demand function (A6), and take the total differential on both sides of the derived equation:

$$(A42) \quad \frac{d \ln L_w(\theta_e)}{d\theta_e} = \frac{d \ln \chi(\theta_e)/\mu(\theta_e)}{d\theta_e} + \left[\frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \right] \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{d \ln l_e(\theta_e)}{d\theta_e} + \xi \frac{d \ln L_w(\theta_e)}{d\theta_e} \right].$$

Rearrange the above equation:

$$(A43) \quad \begin{aligned} \frac{d \ln L_w(\theta_e)}{d\theta_e} &= \frac{\frac{d \ln \chi(\theta_e)/\mu(\theta_e)}{d\theta_e} + \left[\frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \right] \frac{x'_e(\theta_e)}{x_e(\theta_e)}}{1 - \left[\frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \right] \xi} \\ &\quad + \frac{\frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)}{1 - \left[\frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \right] \xi} \frac{d \ln l_e(\theta_e)}{d\theta_e} \\ &= \frac{1}{1 - \xi \varepsilon_{Q_{ij}}^S(\theta_e)} \frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + \frac{\varepsilon_{Q_{ij}}^S(\theta_e)}{1 - \xi \varepsilon_{Q_{ij}}^S(\theta_e)} \frac{l'_e(\theta_e)}{l_e(\theta_e)}, \end{aligned}$$

where $\varepsilon_{Q_{ij}}^S(\theta_e) = \frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) = \frac{\sigma-1}{\sigma}$.

The entrepreneurial effort $l_e(\theta_e)$ satisfies the first-order condition (A5), i.e., $\frac{P(\theta_e) Q_{ij}(\theta_e)}{\mu(\theta_e)} [1 - \tau_e(\theta_e)] = l_e(\theta_e)^{1+\frac{1}{\varepsilon_e}}$, where $\tau_e(\theta_e) = T'(y_e(\theta_e))$. Notice that (A5) and (A4). We have $\frac{WL_w(\theta_e)}{\xi} [1 - \tau_e(\theta_e)] = l_e(\theta_e)^{1+\frac{1}{\varepsilon_e}}$. Therefore,

$$(A44) \quad \left(1 + \frac{1}{\varepsilon_e} \right) \frac{d \ln l_e(\theta_e)}{d\theta_e} = \frac{d \ln L_w(\theta_e)}{d\theta_e} + \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}.$$

Combination of (A43) and (A44) delivers:

$$(A45) \quad \frac{d \ln l_e(\theta_e)}{d\theta_e} = \frac{\frac{d \ln X_e(\theta_e)/\mu(\theta_e)}{d\theta_e} + [1 - \xi \frac{\sigma-1}{\sigma}] \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} (1 - \xi \frac{\sigma-1}{\sigma}) - \frac{\sigma-1}{\sigma}}$$

Combination of firm's first-order condition (A4) and $y_e(\theta_e) = P_{ij}(\theta_e) Q_{ij}(\theta_e) (1 - t_s) - WL_{w,ij}(\theta_e)$ gives:

$$y_e(\theta_e) = \frac{\mu(\theta_e) - \xi}{\xi} WL_w(\theta_e).$$

Take the total differential on both sides of the above equation:

$$(A46) \quad \begin{aligned} \frac{d \ln y_e(\theta_e)}{d\theta_e} &= \frac{d \ln [\mu(\theta_e) - \xi]}{d\theta_e} + \frac{d \ln L_w(\theta_e)}{d\theta_e} \\ &= \frac{d \ln [\mu(\theta_e) - \xi]}{d\theta_e} + \left(1 + \frac{1}{\varepsilon_e} \right) \frac{d \ln l_e(\theta_e)}{d\theta_e} - \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e} \\ &= \frac{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} \frac{d \ln [X_e(\theta_e)/\mu(\theta_e)]}{d\theta_e} + \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \xi \right) - 1} + \frac{d \ln [\mu(\theta_e) - \xi]}{d\theta_e}, \end{aligned}$$

where the second and third equations are derived by (A44) and (A45). Therefore,

$$\frac{d \ln y_e^o(\theta_e)}{d\theta_e} = \frac{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{1}{\varepsilon_{Qij}^S(\theta_e)} \frac{d \ln [X_e(\theta_e)/\mu(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{1}{\varepsilon_{Qij}^S(\theta_e)} - \zeta \right) - 1} + \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e},$$

and

$$(A47) \quad \begin{aligned} \frac{\gamma_e'(\theta_e)}{\gamma_e(\theta_e)} &= \frac{d \ln y_e^o(\theta_e)}{d\theta_e} \frac{1}{\varepsilon_{1-\tau_e}^{y_e^o}(\theta_e)} \frac{\varepsilon_e}{1+\varepsilon_e} \\ &= \left[\frac{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{1}{\varepsilon_{Qij}^S(\theta_e)} \frac{d \ln [X_e(\theta_e)/\mu(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{1}{\varepsilon_{Qij}^S(\theta_e)} - \zeta \right) - 1} + \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \right] \frac{1}{\varepsilon_{1-\tau_e}^{y_e^o}(\theta_e)} \frac{\varepsilon_e}{1+\varepsilon_e} \end{aligned}$$

and

$$(A48) \quad \begin{aligned} H(\theta_e) &= \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{1}{\varepsilon_{Qij}^S(\theta_e)} \frac{d \ln [X_e(\theta_e)/\mu(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{1}{\varepsilon_{Qij}^S(\theta_e)} - \zeta \right) - 1} + \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \right] \\ &= \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{\gamma_e'(\theta_e)}{\gamma_e(\theta_e)} \varepsilon_{1-\tau_e}^{y_e^o}(\theta_e) \frac{1 + \varepsilon_e}{\varepsilon_e}. \end{aligned}$$

Under the nested-CES economy, equation (A47) is equivalent to (30). Moreover, equation (A47) suggests that:

$$(A49) \quad \varepsilon_{1-\tau_e}^{y_e^o}(\theta_e) = \frac{\frac{d \ln y_e^o(\theta_e)}{d\theta_e}}{\frac{d \ln \gamma_e(\theta_e)}{d\theta_e}} \frac{\varepsilon_e}{1 + \varepsilon_e}.$$

$H(\theta_e)$ is the hazard ratio of profit $\left(\frac{1-F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln y_e^o(\theta_e)}{d\theta_e} \right)$ when the profit tax rate is constant, i.e., $\tau_e'(\theta_e) = 0$. Combination of (A48) and (A46) implies:

$$(A50) \quad H(\theta_e) = \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{d \ln y_e(\theta_e)}{d\theta_e} - \frac{\frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{1}{\varepsilon_{Qij}^S(\theta_e)} - \zeta \right) - 1} \right].$$

We now derive a more explicit expression of $\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x_e'(\theta_e)}{x_e(\theta_e)}$. Remember that $\frac{d \ln P_{ij}}{d\theta_e} \big|_{Q_{ij}=Q_{ij}(\theta_e)} = \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q-ij}^{P, cross}(\theta_e) \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e}$ (see (A12)), where:

$$(A51) \quad \begin{aligned} \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} &= \frac{x_e'(\theta_e)}{x_e(\theta_e)} + \frac{d \ln l_e(\theta_e)}{d\theta_e} + \zeta \frac{d \ln L_w(\theta_e)}{d\theta_e} \\ &= \frac{x_e'(\theta_e)}{x_e(\theta_e)} + \left(1 + \zeta \frac{1 + \varepsilon_e}{\varepsilon_e} \right) \frac{d \ln l_e(\theta_e)}{d\theta_e} - \zeta \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e} \\ &= \frac{x_e'(\theta_e)}{x_e(\theta_e)} + \frac{1}{1 - \zeta \varepsilon_{Qij}^S(\theta_e)} \frac{l_e'(\theta_e)}{l_e(\theta_e)} + \frac{\zeta}{1 - \zeta \varepsilon_{Qij}^S(\theta_e)} \frac{d}{d\theta_e} \ln \frac{X_e(\theta_e)}{\mu(\theta_e)}. \end{aligned}$$

The second and third equations are derived by (A44) and (A43). Substitute $\frac{l'_e(\theta_e)}{l_e(\theta_e)}$ in (A51) by (A45):

$$(A52) \quad \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} = \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{\left(1 + \zeta \frac{1+\varepsilon_e}{\varepsilon_e}\right) \frac{d}{d\theta_e} \ln \frac{X_e(\theta_e)}{\mu(\theta_e)} + \frac{d \ln[1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \zeta\right)}.$$

We have:

$$\begin{aligned} & \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} = \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} \\ &= \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \left[1 + \zeta \frac{1+\varepsilon_e}{\varepsilon_e}\right] \frac{l'_e(\theta_e)}{l_e(\theta_e)} - \zeta \frac{d \ln[1-\tau_e(\theta_e)]}{d\theta_e} \right] \\ &= \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{1}{1 - \zeta \varepsilon_{Q_{ij}}^S(\theta_e)} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{\zeta}{1 - \zeta \varepsilon_{Q_{ij}}^S(\theta_e)} \frac{d \ln \frac{X_e(\theta_e)}{\mu(\theta_e)}}{d\theta_e} \right]. \end{aligned}$$

Substitute $\frac{d \ln Q_{ij}(\theta_e)}{d\theta_e}$ by (A52):

$$\begin{aligned} & \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\ &= \mu(\theta_e) \left[\frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} \right] + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\ &= \mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) + 1 \right] \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \\ & \quad \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \frac{\left(1 + \zeta \frac{1+\varepsilon_e}{\varepsilon_e}\right) \frac{d}{d\theta_e} \ln \frac{X_e(\theta_e)}{\mu(\theta_e)} + \frac{d \ln[1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \zeta\right)}, \end{aligned}$$

where

$$\begin{aligned} & \mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) + 1 \right] \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\ &= \mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \mu(\theta_e) \varepsilon_{Q_{ij}}^S(\theta_e) \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\ &= \mu(\theta_e) \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} \right] + \mu(\theta_e) \frac{d \ln \mu(\theta_e)}{d\theta_e} \\ &= \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) + 1}{\varepsilon_{Q_{ij}}^S(\theta_e)} \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} \right] + \mu(\theta_e) \frac{d \ln \mu(\theta_e)}{d\theta_e}. \end{aligned}$$

The first and third equations of the above equations are derived by $\varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) = -\frac{1}{\mu(\theta_e)} + \frac{\sigma-1}{\sigma}$, and the

second equation is derived by $X_e(\theta_e) = A^{\frac{\sigma-1}{\sigma}} \chi(\theta_e) x_e(\theta_e)^{\frac{\sigma-1}{\sigma}}$. In conclusion,

$$\begin{aligned}
(A53) \quad & \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\
&= \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) + 1}{\varepsilon_{Q_{ij}}^S(\theta_e)} \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} \right] + \mu(\theta_e) \frac{d \ln \mu(\theta_e)}{d\theta_e} \\
&\quad + \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \frac{\left(1 + \zeta \frac{1+\varepsilon_e}{\varepsilon_e}\right) \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} \right] + \frac{d \ln[1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \zeta\right)} \\
&= \mu(\theta_e) \frac{d \ln X_e(\theta_e)}{d\theta_e} + \left[\mu(\theta_e) \varepsilon_{Q_{ij}}^S(\theta_e) - 1 \right] \frac{\left(1 + \zeta \frac{1+\varepsilon_e}{\varepsilon_e}\right) \frac{d}{d\theta_e} \ln \frac{X_e(\theta_e)}{\mu(\theta_e)} + \frac{d \ln[1-\tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \zeta\right)}.
\end{aligned}$$

When $\tau'_e(\theta_e) = 0$,

$$\begin{aligned}
& \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\
&= \mu(\theta_e) \frac{d \ln X_e(\theta_e)}{d\theta_e} + \left[\mu(\theta_e) \varepsilon_{Q_{ij}}^S(\theta_e) - 1 \right] \frac{\left(1 + \zeta \frac{1+\varepsilon_e}{\varepsilon_e}\right) \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} \right]}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \zeta\right)}.
\end{aligned}$$

By (A6), we have:

$$\frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} = \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[\frac{\sigma-1}{\sigma} - \frac{1}{\mu(\theta_e)} \right] \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e},$$

where, by definition,

$$\begin{aligned}
\frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} &= \frac{d \ln x_e(\theta_e)}{d\theta_e} + \frac{d \ln l_e(\theta_e)}{d\theta_e} + \zeta \frac{d \ln L_w(\theta_e)}{d\theta_e} \\
&= \frac{\sigma}{\sigma-1} \left[\frac{d \ln \mu(\theta_e)}{d\theta_e} - \frac{d \ln x_e(\theta_e)}{d\theta_e} + \frac{\sigma(\varepsilon_e+1)}{\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)} \left[\frac{d \ln X_e(\theta_e)}{d\theta_e} - \frac{d \ln \mu(\theta_e)}{d\theta_e} \right] \right] \\
&\quad + \frac{\sigma \varepsilon_e \frac{d \ln[1-\tau_e(\theta_e)]}{d\theta_e}}{\sigma + \varepsilon_e - \zeta(\sigma-1)(\varepsilon_e+1)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
(A54) \quad & \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\
&= \frac{\sigma(\varepsilon_e+1) [\mu(\theta_e) - \zeta] - \sigma \varepsilon_e}{\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln X(\theta_e)}{d\theta_e} \\
&\quad - \frac{(\sigma-1) [\varepsilon_e + \zeta(\varepsilon_e+1)] [\mu(\theta_e) - \frac{\sigma}{\sigma-1}]}{\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln \mu(\theta_e)}{d\theta_e} \\
&\quad + \frac{\varepsilon_e [(\sigma-1)\mu(\theta_e) - \sigma]}{\sigma + \varepsilon_e - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln [1-\tau_e(\theta_e)]}{d\theta_e}.
\end{aligned}$$

Substitute $\left[\mu(\theta_e) \frac{\partial \ln P(Q_{ij}(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]$ in $V'_e(\theta_e) = l_e(\theta_e)^{\frac{1+\varepsilon_e}{\varepsilon_e}} \left[\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]$:

$$V'(\theta_e) = l_e(\theta_e)^{\frac{1+\varepsilon_e}{\varepsilon_e}} \left[\begin{aligned} & \frac{\sigma(\varepsilon_e+1)\mu(\theta_e) - \sigma\varepsilon_e - \zeta\sigma(\varepsilon_e+1)}{\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln X(\theta_e)}{d\theta_e} \\ & - \frac{(\sigma-1)[\varepsilon_e + \zeta(\varepsilon_e+1)] [\mu(\theta_e) - \frac{\sigma}{\sigma-1}]}{\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln \mu(\theta_e)}{d\theta_e} \\ & + \frac{\varepsilon_e[(\sigma-1)\mu(\theta_e) - \sigma]}{\sigma + \varepsilon_e - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e} \end{aligned} \right], \forall \theta_e \in \Theta_e.$$

(III) GROSS UTILITY.

Note that $\frac{1+\varepsilon_e}{\varepsilon_e} - \frac{\sigma-1}{\sigma} \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \zeta\right)$, $\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)$ and $\frac{\sigma}{\sigma-1} - \left(\frac{\varepsilon_e}{1+\varepsilon_e} + \zeta\right)$ are positive under condition (22). Under such a condition, whether $l_e(\theta_e)$, $L_w(\theta_e)$, $Q_{ij}(\theta_e)$, $Q_{ij}(\theta_e)P(\theta_e)$ and $V(\theta_e)$ increases with θ_e is determined by the relative change of $X_e(\theta_e)$ to $\mu(\theta_e)$. In particular, we have

$$\begin{aligned} \frac{d \ln V(\theta_e)}{d\theta_e} &= \left[\begin{aligned} & \frac{\mu(\theta_e)}{\mu(\theta_e) - \left(\zeta + \frac{\varepsilon_e}{1+\varepsilon_e}\right)} \\ & - \frac{\frac{\sigma}{\sigma-1}}{\frac{\sigma}{\sigma-1} - \left(\zeta + \frac{\varepsilon_e}{1+\varepsilon_e}\right)} \end{aligned} \right] \frac{\mu'(\theta_e)}{\mu(\theta_e)} + \frac{\frac{\sigma}{\sigma-1} \frac{d \ln X_e(\theta_e)}{d\theta_e}}{\frac{\sigma}{\sigma-1} - \left(\frac{\varepsilon_e}{1+\varepsilon_e} + \zeta\right)} \\ &\geq \frac{\frac{\sigma}{\sigma-1} \frac{d \ln X_e(\theta_e)}{d\theta_e}}{\frac{\sigma}{\sigma-1} - \left(\frac{\varepsilon_e}{1+\varepsilon_e} + \zeta\right)}, \end{aligned}$$

where the second inequality is derived by $\mu(\theta_e) \leq \frac{\sigma}{\sigma-1}$. Thus, $\frac{d \ln V(\theta_e)}{d\theta_e}$ increases with $\frac{d \ln X_e(\theta_e)}{d\theta_e}$ and introducing market power inequality rises $\frac{d \ln V(\theta_e)}{d\theta_e}$.

We also have:

$$(A55) \quad V'_e(\theta_e) = l_e(\theta_e) \phi'(l_e(\theta_e)) \left[\mu(\theta_e) \frac{\partial \ln P(Q_{ij}(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right],$$

with

$$\begin{aligned} (A56) \quad & \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\ &= \mu(\theta_e) \left\{ \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[\frac{\sigma-1}{\sigma} - \frac{1}{\mu(\theta_e)} \right] \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} \right\} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\ &= \frac{\sigma(\varepsilon_e+1)[\mu(\theta_e) - \zeta] - \sigma\varepsilon_e}{\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln X_e(\theta_e)}{d\theta_e} \\ & \quad + \frac{(\sigma-1)[\varepsilon_e + \zeta(\varepsilon_e+1)] \left[\frac{\sigma}{\sigma-1} - \mu(\theta_e) \right]}{\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)} \frac{d \ln \mu(\theta_e)}{d\theta_e}. \end{aligned}$$

Notice that $\varepsilon_e + \sigma - \zeta(\sigma-1)(\varepsilon_e+1)$ is positive under condition (22). The entrepreneurial skill gap increases with $\frac{d \ln X_e(\theta_e)}{d\theta_e}$. Moreover, since $\mu(\theta_e) \leq \frac{\sigma}{\sigma-1}$, the second term on the right side of (A56) is positive under condition (22), which suggests that the θ_e -type entrepreneur's skill gap increases in $\frac{d \ln \mu(\theta_e)}{d\theta_e}$.

(IV) TECHNOLOGY AND EQUILIBRIUM.

$x_e(\cdot)$ and $\chi(\cdot)$ have different economic meanings. They can refer to quantity-augmenting and quality-augmenting (Rosen (1981)), ability and talent (Sattinger (1975b)), and effort-augmenting and total-productivity-augmenting (non-effort-augmenting) elements (Ales, Bellofatto and Wang (2017)), all of which catch the difference between an entrepreneur and a worker.

The expressions for allocations and prices in Supplemental Appendix I.1.1 show that $Q_{ij}(\theta_e)$ and $P(\theta_e)$ are generally dependent on the specific values of $x_e(\theta_e)$ and $\chi(\theta_e)$ instead of only depending on the value of $X_e(\theta_e) = A^{\frac{\sigma-1}{\sigma}} x_e(\theta_e)^{\frac{\sigma-1}{\sigma}} \chi(\theta_e)$.

Consider the economy without taxes. According to (A38) and (A39), we have:

$$Q_{ij}(\theta_e) = x_e(\theta_e) \left(\frac{\xi}{W} \right)^{\frac{\xi\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)\xi+\sigma\varepsilon_e}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}},$$

and

$$P(\theta_e) = \frac{\mu(\theta_e)}{x_e(\theta_e)} \left(\frac{\xi}{W} \right)^{\frac{-\xi(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma-\sigma(\varepsilon_e+1)\xi}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}}, \forall \theta_e \in \Theta.$$

On the other hand, given $X_e(\theta_e)$, we see that $P(\theta_e)Q_{ij}(\theta_e)$, $L_w(\theta_e)$, $l_e(\theta_e)$, and $V_e(\theta_e)$ are independent of the specific values of $\chi(\theta_e)$ and $x_e(\theta_e)$. According to (A27) to (A29), we have the following results for any $\theta_e \in \Theta_e$:

$$\begin{aligned} L_w(\theta_e) &= \left(\frac{\xi}{W} \right)^{\frac{\sigma+\varepsilon_e}{\sigma+\varepsilon_e-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\sigma+\varepsilon_e-\xi(\sigma-1)(\varepsilon_e+1)}}, \\ P(\theta_e)Q_{ij}(\theta_e) &= \mu(\theta_e) \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}}, \\ l_e(\theta_e) &= \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)\varepsilon_e}{\sigma+\varepsilon_e-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma\varepsilon_e}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}}, \end{aligned}$$

In addition, by $c_e(\theta_e) = P(\theta_e)Q_{ij}(\theta_e) - WL_{w,ij}(\theta_e)$ and $V(\theta_e) = c_e(\theta_e) - \phi(l_e(\theta_e))$, we have:

$$\begin{aligned} (A57) \quad V(\theta_e) &= \left[\mu(\theta_e) - \xi - \frac{\varepsilon_e}{1+\varepsilon_e} \right] \phi'(l_e(\theta_e)) l_e(\theta_e) \\ &= \left[\mu(\theta_e) - \xi - \frac{\varepsilon_e}{1+\varepsilon_e} \right] \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)(\varepsilon_e+1)}{\sigma+\varepsilon_e-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \end{aligned}$$

and

$$(A58) \quad c_e(\theta_e) = [\mu(\theta_e) - \xi] \left(\frac{\xi}{W} \right)^{\frac{\xi(\sigma-1)(\varepsilon_e+1)}{\sigma+\varepsilon_e-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{X_e(\theta_e)}{\mu(\theta_e)} \left(\frac{Q}{N_e} \right)^{\frac{1}{\sigma}} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}}.$$

Similarly, one can see that W , $l_w(\theta_w)$, and $V_w(\theta_w)$ are also only dependent on $X_e(\theta_e)$.

Lastly, we find that given $\frac{d \ln X_e(\theta_e)}{d \theta_e}$, $\frac{V'(\theta_e)}{V(\theta_e)}$ is independent of the specific values of $\chi(\theta_e)$ and $x_e(\theta_e)$. Combining (A55) and (A56) gives:

$$V'(\theta_e) = [l_e(\theta_e)]^{\frac{1+\varepsilon_e}{\varepsilon_e}} \left[\begin{aligned} &\frac{\sigma(\varepsilon_e+1)\mu(\theta_e)-\sigma\varepsilon_e-\xi\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)} \frac{d \ln X_e(\theta_e)}{d \theta_e} \\ &- \frac{(\sigma-1)[\varepsilon_e+\xi(\varepsilon_e+1)]\left[\mu(\theta_e)-\frac{\sigma}{\sigma-1}\right]}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)} \frac{d \ln \mu(\theta_e)}{d \theta_e} \end{aligned} \right], \forall \theta_e \in \Theta_e.$$

Combining (A57) and (A55) gives

$$\begin{aligned} (A59) \quad \frac{V'(\theta_e)}{V(\theta_e)} &= \frac{\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)}}{\mu(\theta_e) - \xi - \frac{\varepsilon_e}{1+\varepsilon_e}} \\ &= \frac{\frac{\sigma(\varepsilon_e+1)[\mu(\theta_e)-\xi]-\sigma\varepsilon_e}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)} \frac{d \ln X_e(\theta_e)}{d \theta_e} + \frac{(\sigma-1)[\varepsilon_e+\xi(\varepsilon_e+1)]\left[\frac{\sigma}{\sigma-1}-\mu(\theta_e)\right]}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)} \frac{d \ln \mu(\theta_e)}{d \theta_e}}{\mu(\theta_e) - \xi - \frac{\varepsilon_e}{1+\varepsilon_e}}, \end{aligned}$$

where the second equation is derived by (A56) and (A55), $\theta_e \in \Theta_e$. Specially, when markup is constant, we have

$$\frac{V'(\theta_e)}{V(\theta_e)} = \frac{\sigma(\varepsilon_e + 1)}{\varepsilon_e + \sigma - \zeta(\sigma - 1)(\varepsilon_e + 1)} \frac{d \ln X_e(\theta_e)}{d \theta_e}, \forall \theta_e \in \Theta_e.$$

II. Solution

2.1. Proof of Proposition 1

Part 1 of Proposition 1 can be derived by (16) and (20). We now prove part 2. By (A37), (A36), and (A35), we have

$$(A60) \quad \nu(I) \triangleq \frac{WL}{\xi Q} = \frac{A_2}{A_1},$$

where L is the aggregate labor input. Substituting A_1 and A_2 by (A31), we have

$$(A61) \quad \nu(I) = \frac{\int f_e(\theta_e) \left[\frac{x_e(\theta_e)^{\frac{\sigma-1}{\sigma}} \chi(\theta_e)}{\mu(\theta_e)} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\zeta(\sigma-1)(\varepsilon_e+1)}} d\theta_e}{\int_{\theta_e} f_e(\theta_e) \mu(\theta_e) \left[\frac{x_e(\theta_e)^{\frac{\sigma-1}{\sigma}} \chi(\theta_e)}{\mu(\theta_e)} \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\zeta(\sigma-1)(\varepsilon_e+1)}} d\theta_e}.$$

For the convenience of analysis, define

$$\begin{aligned} m(\theta_e) &\equiv f_e(\theta_e) \left[x_e(\theta_e)^{\frac{\sigma-1}{\sigma}} \chi(\theta_e) \right]^{\frac{\sigma(\varepsilon_e+1)}{\varepsilon_e+\sigma-\zeta(\sigma-1)(\varepsilon_e+1)}}, \\ g(\theta_e, I) &\equiv \left[\frac{1}{\mu(\theta_e)} \right]^{(\sigma-1) \frac{\varepsilon_e+\zeta(\varepsilon_e+1)}{\varepsilon_e+\sigma-\zeta(\sigma-1)(\varepsilon_e+1)}}, \\ f_s(\theta_e, I) &\equiv \frac{g(\theta_e, I) m(\theta_e)}{\int_{\theta_e} g(\theta_e, I) m(\theta_e) d\theta_e}, \forall \theta_e \in \Theta_e. \end{aligned}$$

Then we have:

$$(A62) \quad \nu(I) = \int_{\theta_e} \frac{f_s(\theta_e, I)}{\mu(\theta_e)} d\theta_e.$$

In addition,

$$\begin{aligned} \frac{\frac{d\nu(I)}{d \ln I}}{\frac{(\sigma-1)(\varepsilon_e+1)}{\varepsilon_e+\sigma-\zeta(\sigma-1)(\varepsilon_e+1)}} &= \int_{\theta_e} f_s(\theta_e, I) \left[\left(1 + \frac{1}{\sigma-1} \right) \frac{1}{\mu(\theta_e)} - \nu(I) \left(1 - \frac{1}{1+\varepsilon_e} + \zeta \right) \right] \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} d\theta \\ &= \left(\frac{1}{\sigma-1} + \frac{1}{1+\varepsilon_e} - \zeta \right) \int_{\theta_e} f_s(\theta_e, I) \frac{1}{\mu(\theta_e)} \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} \\ &\quad + \left(\frac{\varepsilon_e}{1+\varepsilon_e} + \zeta \right) \int_{\theta_e} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} d\theta. \end{aligned}$$

Since $\varepsilon_e + \sigma - \zeta(\sigma - 1)(\varepsilon_e + 1) > 0$ the sign of $\frac{d\nu(I)}{d \ln I}$ is same with the right equations, and $\left(\frac{1}{\sigma-1} + \frac{1}{1+\varepsilon_e} - \zeta \right) >$

0 and $\frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} > 0$, we have:

$$\left(\frac{1}{\sigma - 1} + \frac{1}{1 + \varepsilon_e} - \xi \right) \int_{\theta_e} f_s(\theta_e, I) \frac{1}{\mu(\theta_e)} \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} d\theta_e > 0.$$

Notice that $\frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} = \left[1 - \frac{\sigma - 1}{\sigma} \mu(\theta_e) \right] \frac{I}{I - 1}$ decrease in $\mu(\theta_e)$. We now try to prove that:

$$\int_{\theta_e} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} d\theta_e \geq 0.$$

To do this, note that by (A62), we use

$$\int_{\theta_e} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] d\theta_e = 0,$$

where $f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right]$ is positive if and only if $\frac{1}{\mu(\theta_e)} - \nu(I)$ is positive. Set $\Omega \equiv \left\{ \theta_e | \mu(\theta_e) < \frac{1}{\nu(I)} \right\}$. $f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] > 0$ if and only if $\theta_e \in \Omega$.

Notice that:

$$\int_{\theta_e \in \Omega_*} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] d\theta_e + \int_{\theta_e \notin \Omega_*} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] d\theta_e = 0,$$

and

$$\int_{\theta_e \in \Omega_*} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] d\theta_e > 0$$

and $\frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} < 0$. One can see that for any $\theta_e \in \Omega$,

$$\frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} \geq \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} \Big|_{\mu(\theta_e) = \nu(I)} = \left[1 - \frac{\sigma - 1}{\sigma} \nu(I) \right] \frac{I}{I - 1}.$$

Therefore, we have:

$$\begin{aligned} \text{(A63)} \quad & \int_{\theta_e \in \Omega_*} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} d\theta_e \\ & \geq \left[1 - \frac{\sigma - 1}{\sigma} \nu(I) \right] \frac{I}{I - 1} \int_{\theta_e \in \Omega_*} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] d\theta_e. \end{aligned}$$

On the other hand, for any $\theta_e \notin \Omega$:

$$\frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} \leq \left[1 - \frac{\sigma - 1}{\sigma} \nu(I) \right] \frac{I}{I - 1}, f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] \leq 0.$$

Therefore, we have:

$$\begin{aligned} \text{(A64)} \quad & \int_{\theta_e \notin \Omega_*} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} d\theta_e \\ & \geq \left[1 - \frac{\sigma - 1}{\sigma} \nu(I) \right] \frac{I}{I - 1} \int_{\theta_e \notin \Omega_*} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] d\theta_e. \end{aligned}$$

Combination of (A63) and (A64) gives:

$$\int_{\theta_e} f_s(\theta_e, I) \left[\frac{1}{\mu(\theta_e)} - \nu(I) \right] \frac{d \ln \left[\frac{1}{\mu(\theta_e)} \right]}{d \ln I} d\theta_e \geq 0,$$

which suggests $\frac{d\nu(I)}{d \ln I} \geq 0$.

The proof for the first point of part three is below. Remember that markups are constant. Equation (A36) and (A37) give:

$$\frac{W}{\xi} = \left[(A_1)^{\frac{\varepsilon_e+1}{(\sigma-1)[1-\xi(\varepsilon_e+1)]}} \frac{A_2}{A_3} \right]^{\frac{1-\xi(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}} \propto \left(\frac{1}{\mu} \right)^{\frac{(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}$$

and

$$Q \propto \left(\frac{1}{\mu} \right)^{-\frac{\xi(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1} \frac{(\varepsilon_e+1)}{1-\xi(\varepsilon_e+1)}} \mu \left(\frac{1}{\mu} \right)^{\frac{(\varepsilon_e+1)}{1-\xi(\varepsilon_e+1)}} = \left(\frac{1}{\mu} \right)^{\frac{(\varepsilon_w+1)\varepsilon_e+\varepsilon_w\xi(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}.$$

Substituting W and Q in (A27) with the above equations, we have

$$L_w(\theta_e) \propto \left[\frac{1}{\mu} \right]^{\frac{(\sigma-1)(\varepsilon_e+1)}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} \left[\frac{1}{\mu} \right]^{\frac{\varepsilon_w(\varepsilon_e+1)-\sigma+1}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1} \frac{\varepsilon_e+1}{\varepsilon_e+\sigma-\xi(\sigma-1)(\varepsilon_e+1)}} = \left[\frac{1}{\mu} \right]^{\frac{(\varepsilon_e+1)\varepsilon_w}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \forall \theta_e \in \Theta_e.$$

Similarly, we have for all $\theta_w \in \Theta_w$ and $\theta_e \in \Theta_e$:

$$\begin{aligned} S(\theta_e) &\propto \left[\frac{1}{\mu} \right]^{\frac{\varepsilon_e(\varepsilon_w+1)+\varepsilon_w\xi(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \\ l_e(\theta_e) &\propto \left[\frac{1}{\mu} \right]^{\frac{(\varepsilon_w+1)\varepsilon_e}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \\ Q_{ij}(\theta_e) &\propto \left[\frac{1}{\mu} \right]^{\frac{(\varepsilon_e+1)\varepsilon_w\xi+(\varepsilon_w+1)\varepsilon_e}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \\ P(\theta_e) &\propto \left[\frac{1}{\mu} \right]^0, \\ l_w(\theta_w) &\propto \left(\frac{1}{\mu} \right)^{\frac{\varepsilon_w(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \\ c_w(\theta_w) &\propto \left(\frac{1}{\mu} \right)^{\frac{(\varepsilon_w+1)(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \\ V_w(\theta_w) &\propto \left(\frac{1}{\mu} \right)^{\frac{(\varepsilon_w+1)(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}. \end{aligned}$$

It's easy to see that under the conditions (23) and (22), $L_w(\theta_e)$, $S(\theta_e)$, $l_e(\theta_e)$, $Q_{ij}(\theta_e)$, and $P(\theta_e)$ go down with the decrease of I . Since the markup is uniform, firm-level labor shares must go down too.

We now prove the part three. By (A58) and (A57) we have:

$$\begin{aligned} c_e(\theta_e) &\propto [\mu - \xi] \left[\frac{1}{\mu} \right]^{\frac{(\varepsilon_w+1)(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \\ V_e(\theta_e) &\propto \left[\mu - \xi - \frac{\varepsilon_e}{1+\varepsilon_e} \right] \left[\frac{1}{\mu} \right]^{\frac{(\varepsilon_w+1)(\varepsilon_e+1)}{\varepsilon_w[1-\xi(\varepsilon_e+1)]+1}}, \forall \theta_e \in \Theta_e. \end{aligned}$$

Notice that

$$\frac{d \ln c(\theta_e)}{d \ln \mu} \geq 0 \Leftrightarrow \frac{\mu - \zeta}{\mu} \leq \frac{\varepsilon_w + 1 - \varepsilon_w \zeta (\varepsilon_e + 1)}{(\varepsilon_w + 1)(\varepsilon_e + 1)}.$$

One can see that the condition for $\frac{d \ln c_{ij}(\theta_e)}{d \ln \mu} \geq 0$ is

$$\mu \leq \frac{\zeta}{\frac{\varepsilon_e}{1+\varepsilon_e} + \frac{\varepsilon_w}{\varepsilon_w+1} \zeta}$$

On the other hand,

$$\frac{d V_e(\theta_e)}{d \ln \mu} \propto \left[\zeta + \frac{\varepsilon_e}{1+\varepsilon_e} \right] - \left[\frac{\varepsilon_w}{1+\varepsilon_w} \zeta + \frac{\varepsilon_e}{1+\varepsilon_e} \right] \mu.$$

Therefore, $\mu \leq \frac{\zeta + \frac{\varepsilon_e}{1+\varepsilon_e}}{\frac{\varepsilon_e}{1+\varepsilon_e} + \frac{\varepsilon_w}{1+\varepsilon_w} \zeta}$ is a condition for $\frac{d V_e(\theta_e)}{d \ln \mu} \geq 0$. ■

2.2. Proof of Lemma 1

To simplify notation, in the following analysis, we set $P(Q_{ij}, \theta_e) = P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)$ for any $\theta_e \in \Theta_e$ and $Q_{ij} \in \mathbb{R}_+$.

We first prove part (i) of Lemma 1 (i.e., given (A65), (A68) is satisfied if and only if (27) is satisfied). According to the definition of $V_e(\theta)$, we have

$$(A65) \quad V_e(\theta_e) = c_e(\theta_e) - \phi_e(l_e(\theta_e)), \quad \forall \theta_e \in \Theta_e,$$

Notice that

$$V_e(\theta'_e|\theta_e) = c_e(\theta'_e) - \phi_e(l_e(\theta'_e|\theta_e)),$$

where $l_e(\theta'_e|\theta_e)$ is the effort needed to finish the θ'_e task:

$$(A66) \quad \begin{aligned} l_e(\theta'_e|\theta_e) &= \arg \min_{l_e, L_w} l_e \\ \text{s.t. } &P(Q_{ij}(x_e(\theta_e)l_e, L_w), Q_{-ij}(\theta_e), \theta_e) Q_{ij}(x_e(\theta_e)l_e, L_w)(1-t_s) - WL_w = y_e(\theta'_e) \end{aligned}$$

Obviously, $P(Q_{ij}(x_e(\theta_e)l_e, L_w), Q_{-ij}(\theta_e), \theta_e) \cdot Q_{ij}(x_e(\theta_e)l_e, L_w)(1-t_s) - WL_w$ increases in l_e . Denote by $L_w(\theta_e|l_e)$ the solution to

$$\begin{aligned} &L_w(\theta_e|l_e) \\ &= \arg \max_{L_w} \{P(Q_{ij}(x_e(\theta_e)l_e, L_w), Q_{-ij}(\theta_e), \theta_e) Q_{ij}(x_e(\theta_e)l_e, L_w)(1-t_s) - WL_w\} \end{aligned}$$

for $l_e > 0$. In Supplemental Appendix II.2.5, we show that for any $l_e > 0$, the first-order condition for solving $L_w(\theta_e|l_e)$ is not only necessary but also sufficient and there is a unique solution. Meanwhile, for $L_w > 0$,

$$P(Q_{ij}(x_e(\theta_e)l_e, L_w), Q_{-ij}(\theta_e), \theta_e) Q_{ij}(x_e(\theta_e)l_e, L_w)(1-t_s) - WL_w$$

strictly increases in l_e . Therefore, there must exist a unique solution to problem (A66). Denote by $L_w(\theta'_e|\theta_e)$ the optimal labor input given that θ_e entrepreneur reports θ'_e . When $y_e(\theta'_e) = 0$, $L_w(\theta'_e|\theta_e) = l_e(\theta'_e|\theta_e) = 0$. Otherwise, $L_w(\theta'_e|\theta_e) > 0$ and $l_e(\theta'_e|\theta_e) > 0$ are determined by the first-order conditions. In particular, $L_w(\theta'_e|\theta_e)$ and $l_e(\theta'_e|\theta_e)$ satisfy:

$$\begin{aligned} WL_w(\theta'_e|\theta_e) + y_e(\theta'_e) &= P(Q_{ij}(x_e(\theta_e)l_e(\theta'_e|\theta_e), Q_{-ij}(\theta_e), L_w(\theta'_e|\theta_e)), \theta_e) \\ &\quad \times Q_{ij}(x_e(\theta_e)l_e(\theta'_e|\theta_e), L_w(\theta'_e|\theta_e))(1-t_s). \end{aligned}$$

The above equation and problem (A66) implies:

$$\begin{aligned}
(A67) \quad \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta_e} &= - \frac{\frac{\partial [P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) Q_{ij}]}{\partial Q_{ij}} \frac{\partial Q_{ij}}{\partial \theta_e} + \frac{dP_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} Q_{ij}}{\frac{\partial [P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) Q_{ij}]}{\partial Q_{ij}} \frac{\partial Q_{ij}}{\partial l_e}} \\
&= - \frac{\frac{\partial Q_{ij}}{\partial \theta_e} + \frac{\frac{dP_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} Q_{ij}}{P(Q_{ij}, \theta_e) \left[1 + \frac{\partial \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{\partial \ln Q_{ij}} \right]}}{\frac{\partial Q_{ij}}{\partial l_e}} \\
&= - \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{\frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e}}{1 + \frac{\partial \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{\partial \ln Q_{ij}}} \right] l_e(\theta'_e|\theta_e) < 0.
\end{aligned}$$

The first-order incentive condition ($\frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e} \big|_{\theta'_e=\theta_e} = 0$) can be expressed as

$$(A68) \quad 0 = \left[c'_e(\theta'_e) - \phi'_e(l_e(\theta'_e|\theta_e)) \frac{\partial l_e(\theta'|\theta)}{\partial \theta'} \right] \big|_{\theta'=\theta}, \quad \forall \theta_e \in \Theta_e.$$

First, note that by

$$V_e(\theta_e) = \max_{\theta'_e} V_e(\theta'_e|\theta_e),$$

we have

$$(A69) \quad V'_e(\theta_e) = \frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e^*(\theta_e)} \frac{d\theta_e^*(\theta_e)}{d\theta_e} + \frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e}$$

where we use $\theta_e^*(\theta_e)$ to denote the optimal choice of θ_e entrepreneur.

Second, by the definition of $V_e(\theta'_e|\theta_e)$, we have

$$(A70) \quad \frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e} = -\phi'_e(l_e(\theta_e^*(\theta_e)|\theta_e)) \frac{\partial l_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e},$$

where by (A67), we have

$$\frac{\partial l_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e} = -\frac{x'_e(\theta_e)}{x_e(\theta_e)} l_e(\theta_e^*(\theta_e)|\theta_e) - \frac{\partial \ln P(Q_{ij}(x_e(\theta_e), l_e(\theta_e^*(\theta_e)|\theta_e), L_w(\theta_e^*(\theta_e)|\theta_e)), \theta_e)}{\partial \theta_e} \frac{l_e(\theta_e^*(\theta_e)|\theta_e)}{1 + \varepsilon_{Q_{ij}}^{P,own}(\theta_e)}.$$

Combining (A69), (A70), and (A71) gives

$$V'_e(\theta_e) = \frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e^*(\theta_e)} \frac{d\theta_e^*(\theta_e)}{d\theta_e} - \phi'_e(l_e(\theta_e^*(\theta_e)|\theta_e)) \times \left[\frac{-\frac{x'_e(\theta_e)}{x_e(\theta_e)} l_e(\theta_e^*(\theta_e)|\theta_e) - \frac{\partial \ln P(Q_{ij}(x_e(\theta_e), l_e(\theta_e^*(\theta_e)|\theta_e), L_w(\theta_e^*(\theta_e)|\theta_e)), \theta_e)}{\partial \theta_e} \frac{l_e(\theta_e^*(\theta_e)|\theta_e)}{1 + \varepsilon_{Q_{ij}}^{P,own}(\theta_e)}}{\frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e^*(\theta_e)} \frac{d\theta_e^*(\theta_e)}{d\theta_e}} \right]$$

which implies that for any $\theta_e \in \Theta$,

$$V'_e(\theta_e) = \phi'_e(l_e(\theta_e^*(\theta_e)|\theta_e)) l_e(\theta_e^*(\theta_e)|\theta_e) \times \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \mu(\theta_e) \frac{d \ln P_{ij}}{d\theta_e} \big|_{Q_{ij}=Q_{ij}(x_e(\theta_e), l_e(\theta_e^*(\theta_e)|\theta_e), L_w(\theta_e^*(\theta_e)|\theta_e))} \right]$$

if and only if $\frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e^*(\theta_e)} \frac{d\theta_e^*(\theta_e)}{d\theta_e} = 0$.

Notice that mass points are ruled out by Assumption 1: $\frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e^*(\theta_e)} \frac{d\theta_e^*(\theta_e)}{d\theta_e} = 0$ if and only if $\frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e^*(\theta_e)} = 0$.

Therefore, (27) or the above equation implies the first-order necessary condition $\frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e}|_{\theta'_e=\theta_e} = 0$ when $\theta_e^*(\theta_e) = \theta_e$; and the first-order necessary condition $\frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e}|_{\theta'_e=\theta_e} = 0$, i.e., $\frac{\partial V_e(\theta_e^*(\theta_e)|\theta_e)}{\partial \theta_e^*(\theta_e)}|_{\theta_e^*(\theta_e)=\theta_e} = 0$, implies (27) when the agent reports the true types. In conclusion, (27) is a necessary condition for the truth-telling strategy to be the optimum choice of agents and it implies $\frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e}|_{\theta'_e=\theta_e} = 0$. ■

2.3. Optimal Taxation

Lagrangian and First-order Conditions. We now take Lagrange multipliers to solve the planner's optimization problem.⁵² The Lagrangian function for the planner's problem is

$$\begin{aligned} & \mathcal{L}(L_w, l_w, l_e, V_w, V_e, \delta; \lambda, \lambda', \psi_w, \psi_e, \kappa, \varphi) \\ = & \sum_{o \in \{w, e\}} N_o \int_{\theta_o} G(V_o(\theta_o)) \tilde{f}_o(\theta_o) d\theta_o \\ & + \lambda \left[Q - \sum_{o \in \{w, e\}} N_o \int_{\theta_o} [V_o(\theta_o) + \phi_o(l_o(\theta_o))] f_o(\theta_o) d\theta_o - R \right] \\ & + \lambda' \left[N_w \int_{\theta_w} x_w(\theta_w) l_w(\theta_w) f_w(\theta_w) d\theta_w - N_e \int_{\theta_e} L_w(\theta_e) f_e(\theta_e) d\theta_e \right] \\ & + \int_{\theta_e} \varphi(\theta_e) \frac{d \ln \omega(\theta_e)}{d\theta_e} d\theta_e + \int_{\theta_e} \kappa(\theta_e) \left[\delta(\theta_e) - \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} \right] d\theta_e \\ & + \int_{\theta_w} \psi_w(\theta_w) \left[l_w(\theta_w) \phi'_w(l_w(\theta_w)) \frac{x'_w(\theta_e)}{x_w(\theta_e)} - V'_w(\theta_w) \right] d\theta_w \\ & + \int_{\theta_e} \psi_e(\theta_e) \left[\phi'_e(l_e(\theta_e)) l_e(\theta_e) \left[\frac{\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \mu(\theta_e) \times}{\left[\frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q-ij}^{P, cross}(\theta) \delta(\theta_e) \right]} \right] - V'_e(\theta_e) \right] d\theta_e, \end{aligned}$$

where

$$\frac{\chi'(\theta_e)}{\chi(\theta_e)} + \varepsilon_{Q-ij}^{P, cross}(\theta_e) \delta(\theta_e) = \frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)},$$

and

$$\omega(\theta_e) = \frac{\partial [P_{ij}(Q_{ij}(\theta_e), Q_{-ij}(\theta_e), \theta_e) Q_{ij}(\theta_e)]}{\partial L_w(\theta_e)}.$$

Note that we have introduced $\delta(\theta_e) = \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e}$ as a control value and that $\ln Q_{ij}(\theta_e)$ can be treated as a state variable. Constraint $\frac{d \ln \omega(\theta_e)}{d\theta_e} = 0$ is used to guarantee that $\omega(\theta_e) = \frac{\partial [P_{ij}(Q_{ij}(\theta_e), Q_{-ij}(\theta_e), \theta_e) Q_{ij}(\theta_e)]}{\partial L_w(\theta_e)}$ (equivalently $\omega(\theta_e) = \frac{P(\theta_e)}{\mu(\theta_e)} \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)}$) is constant, which is a result of uniform sales taxes on the goods produced by firms.

Taking partial integrals yields the following

$$\begin{aligned} - \int_{\theta_e} \kappa(\theta_e) \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} d\theta_e &= \ln Q_{ij}(\underline{\theta_e}) \kappa(\underline{\theta_e}) - \ln Q_{ij}(\bar{\theta_e}) \kappa(\bar{\theta_e}) \\ &+ \int_{\theta_e} \kappa'(\theta_e) \ln Q_{ij}(\theta_e) d\theta_e, \end{aligned}$$

and

$$\int_{\theta_e} \varphi(\theta_e) \frac{d \ln \omega(\theta_e)}{d\theta_e} d\theta_e = \varphi(\bar{\theta_e}) \ln \omega(\bar{\theta_e}) - \varphi(\underline{\theta_e}) \ln \omega(\underline{\theta_e}) - \int_{\theta_e} \varphi'(\theta_e) \ln \omega(\theta_e) d\theta_e,$$

⁵²See Luenberger (1997) for details about the Lagrangian techniques, and Mirrlees (1976), Golosov, Tsyvinski and Werquin (2016), Findeisen and Sachs (2017) for its application in the field of public economics.

and

$$-\int_{\theta_o} \psi_o(\theta_e) V_o'(\theta_o) d\theta_o = V_o(\underline{\theta}_o) \psi_o(\underline{\theta}_o) - V_o(\bar{\theta}_o) \psi_o(\bar{\theta}_o) + \int_{\theta_o} \psi_o'(\theta_o) V_o(\theta_o) d\theta_o.$$

The derivatives with respect to the endpoint conditions yield boundary conditions:

$$(A71) \quad \kappa(\underline{\theta}_e) = \kappa(\bar{\theta}_e) = \varphi(\bar{\theta}_e) = \varphi(\underline{\theta}_e) = \psi_o(\underline{\theta}_o) = \psi_o(\bar{\theta}_o) = 0, \quad o \in \{w, e\}.$$

Thus,

$$(A72) \quad \int_{\theta_e} \varphi'(\theta_e) d\theta_e = 0,$$

Substituting the above conditions into the Lagrangian function yields the following first-order conditions for any $\theta_o \in \Theta_o$:

$$(A73) \quad \frac{\partial \mathcal{L}}{\partial V_o(\theta_o)} = G'(V_o(\theta_o)) N_o \tilde{f}_o(\theta_o) + \psi_o'(\theta_o) - \lambda N_o f_o(\theta_o) = 0,$$

$$(A74) \quad \frac{\partial \mathcal{L}}{\partial \delta(\theta_e)} = \kappa(\theta_e) + \psi_e(\theta_e) \phi_e'(l_e(\theta_e)) l_e(\theta_e) \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) = 0,$$

$$(A75) \quad \frac{\partial \mathcal{L}}{\partial l_w(\theta_w)} = \left\{ \begin{aligned} &[-\lambda \phi_w'(l_w(\theta_w)) + \lambda' x_w(\theta_w)] N_w f_w(\theta_w) \\ &+ \psi_w(\theta_w) \frac{\phi_w'(l_w(\theta_w))}{x_w(\theta_w)} \frac{1 + \varepsilon_w}{\varepsilon_w} \end{aligned} \right\} = 0,$$

$$(A76) \quad \frac{\partial \mathcal{L}}{\partial L_w(\theta_e)} = \left[\lambda P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)} - \lambda' \right] N_e f_e(\theta_e) + \left[\frac{\frac{\kappa'(\theta_e)}{L_w(\theta_e)} \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln L_w(\theta_e)} - \frac{\int_{\Theta_e} \varphi'(\theta_e') \frac{\partial \ln \omega(\theta_e')}{\partial \ln L_w(\theta_e)} d\theta_e'}{L_w(\theta_e)} \right] = 0,$$

and

$$(A77) \quad \begin{aligned} \frac{\partial \mathcal{L}}{\partial l_e(\theta_e)} &= \psi_e(\theta_e) \phi_e'(l_e(\theta_e)) \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x_e'(\theta_e)}{x_e(\theta_e)} \right] \\ &+ \lambda \left[P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)} - \phi_e'(l_e(\theta_e)) \right] N_e f_e(\theta_e) \\ &+ \frac{\kappa'(\theta_e)}{l_e(\theta_e)} \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln l_e(\theta_e)} - \frac{\int_{\Theta} \varphi'(\theta_e') \frac{\partial \ln \omega(\theta_e')}{\partial \ln l_e(\theta_e)} d\theta_e'}{l_e(\theta_e)} = 0. \end{aligned}$$

Social Welfare Weight. Unless otherwise specified, the following equations in this subsection are derived for any $\theta_o \in \Theta_o$. According to $\frac{\partial \mathcal{L}}{\partial V_o(x)}$ and $\psi_o(\underline{\theta}_o) = \psi_o(\bar{\theta}_o) = 0$, we have:

$$(A78) \quad \lambda = \int_{\theta_o} G'(V_o(\theta_o)) \tilde{f}_o(\theta_o) d\theta_o.$$

Set

$$(A79) \quad g_o(\theta_o) = \frac{G'(V_o(\theta_o)) \tilde{f}_o(\theta_o)}{\lambda f_o(\theta_o)}$$

as the monetary marginal social welfare weight for θ_o agent of o occupation. Set:

$$(A80) \quad \bar{g}_o(\theta_o) = \frac{\int_{\theta_o}^{\bar{\theta}_o} g(x) f_o(x) dx}{1 - F_o(\theta_o)}$$

as the weighted monetary social welfare weight for agents whose abilities are higher than θ_e .

Substituting $g_o(\theta_o)$ into $\frac{\partial \mathcal{L}}{\partial V_o(\theta_o)}$ gives

$$(A81) \quad \frac{\psi'_o(\theta_o)}{\lambda N_o f_o(\theta_o)} = 1 - g_o(\theta_o)$$

Taking integration and using the boundary conditions gives

$$(A82) \quad -\frac{\psi_o(\theta_o)}{\lambda N_o} = \int_{\theta_o}^{\bar{\theta}_o} [1 - g_o(x)] f_o(x) dx = [1 - \bar{g}_o(\theta_o)] [1 - F_o(\theta_o)].$$

In addition, based on $\frac{\partial \mathcal{L}}{\partial \delta(\theta_e)}$, we have:

$$(A83) \quad \begin{aligned} \kappa(\theta_e) &= -\psi_e(\theta_e) \phi'_e(l_e(\theta_e)) l_e(\theta_e) \mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \\ &= -\psi_e(\theta_e) P(\theta_e) Q_{ij}(\theta_e) [1 - \tau_e(\theta_e)] (1 - \tau_s) \varepsilon_{Q-ij}^{P, cross}(\theta_e), \end{aligned}$$

where the second equation is derived by

$$(A84) \quad \phi'_e(l_e(\theta_e)) l_e(\theta_e) = \frac{P(\theta_e) Q_{ij}(\theta_e)}{\mu(\theta_e)} [1 - \tau_e(\theta_e)] (1 - \tau_s).$$

In addition, we have:

$$(A85) \quad \begin{aligned} \kappa'(\theta_e) &= -\frac{d \left[\psi_e(\theta_e) \phi'_e(l_e(\theta_e)) l_e(\theta_e) \mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \right]}{d\theta_e} \\ &= -\phi'_e(l_e(\theta_e)) l_e(\theta_e) \mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \psi_e(\theta_e) \left[\frac{1 + \varepsilon_e \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{\psi'_e(\theta_e)}{\psi_e(\theta_e)}}{\varepsilon_e} + \frac{d \ln \left[\mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \right]}{d\theta_e} \right] \\ &= -P(\theta_e) Q_{ij}(\theta_e) [1 - \tau_e(\theta_e)] (1 - \tau_s) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \\ &\quad \times \psi_e(\theta_e) \left[\frac{1 + \varepsilon_e \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{\psi'_e(\theta_e)}{\psi_e(\theta_e)}}{\varepsilon_e} + \frac{d \ln \left[\mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \right]}{d\theta_e} \right]. \end{aligned}$$

Substituting $\psi_e(\theta_e)$ and $\psi'_e(\theta_e)$ in the above equation and (A83) with (A81) and (A82), we have:

$$\begin{aligned} \frac{\kappa(\theta_e)}{\lambda N_e} &= [1 - \bar{g}_e(\theta_e)] [1 - F_e(\theta_e)] \phi'_e(l_e(\theta_e)) l_e(\theta_e) \mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \\ &= [1 - \bar{g}_e(\theta_e)] [1 - F_e(\theta_e)] P(\theta_e) Q_{ij}(\theta_e) [1 - \tau_e(\theta_e)] (1 - \tau_s) \varepsilon_{Q-ij}^{P, cross}(\theta_e), \end{aligned}$$

and

$$(A86) \quad \frac{\kappa'(\theta_e)}{\lambda N_e f_e(\theta_e)} = -P(\theta_e) Q_{ij}(\theta_e) [1 - \tau_e(\theta_e)] (1 - \tau_s) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \times \left[\left[1 - g_e(\theta_e) \right] - \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \times \left[\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{d \ln [\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)]}{d\theta_e} \right] \right].$$

2.4. Proof of Theorem 1

Unless otherwise specified, the following equations in this subsection are derived for any $\theta_o \in \Theta_o$ and $\tau_s = 0$.

(i) $\frac{\partial \mathcal{E}}{\partial L_w(\theta_e)} = 0$ implies:

$$\begin{aligned} P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)} &= \frac{\lambda'}{\lambda} - \frac{\kappa'(\theta_e)}{\lambda L_w(\theta_e) N_e f_e(\theta_e)} \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln L_w(\theta_e)} + \frac{\int_{\theta_e} \varphi'(\theta'_e) \frac{\partial \ln \omega(\theta'_e)}{\partial \ln L_w(\theta_e)} d\theta'_e}{\lambda L_w(\theta_e) N_e f_e(\theta_e)} \\ &= \frac{\lambda'}{\lambda} - \frac{\kappa'(\theta_e) \xi}{\lambda L_w(\theta_e) N_e f_e(\theta_e)} + \frac{\varphi'(\theta_e) \varepsilon_{L_w}^\omega(\theta_e)}{\lambda L_w(\theta_e) N_e f_e(\theta_e)}, \end{aligned}$$

where $\int_{\theta_e} \varphi'(\theta'_e) \frac{\partial \ln \omega(\theta'_e)}{\partial \ln L_w(\theta_e)} d\theta'_e = \varphi'(\theta_e) \varepsilon_{L_w}^\omega(\theta_e)$ since $\varepsilon_{L_w}^\omega(\theta'_e, \theta_e)$ is independent of θ'_e and $\int_{\theta_e} \varphi'(\theta'_e) d\theta'_e = 0$. Substituting $P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)}$ by $W\mu(\theta_e)$ gives:

$$(A87) \quad W\mu(\theta_e) = \frac{\lambda'}{\lambda} - \frac{\kappa'(\theta_e) \xi}{\lambda L_w(\theta_e) N_e f_e(\theta_e)} + \frac{\varphi'(\theta_e) \varepsilon_{L_w}^\omega(\theta_e)}{\lambda L_w(\theta_e) N_e f_e(\theta_e)}.$$

Dividing both sides of the above equation by $\frac{\varepsilon_{L_w}^\omega(\theta_e)}{L_w(\theta_e) N_e f_e(\theta_e)}$ and integrating across θ_e gives:

$$W \int_{\theta_e} \mu(\theta_e) \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e = \frac{\lambda'}{\lambda} \int_{\theta_e} \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e - \int_{\theta_e} \frac{\kappa'(\theta_e)}{\lambda \varepsilon_{L_w}^\omega(\theta_e)} \xi d\theta_e,$$

where we use $\int_{\theta_e} \varphi'(\theta'_e) d\theta'_e = 0$ again. Reformation of the above equation gives:

$$\begin{aligned} 1 &= \frac{\frac{\lambda'}{\lambda} \int_{\theta_e} \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e}{W \int_{\theta_e} \mu(\theta_e) \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e} - \frac{\int_{\theta_e} \frac{\kappa'(\theta_e)}{\lambda \varepsilon_{L_w}^\omega(\theta_e)} \xi d\theta_e}{W \int_{\theta_e} \mu(\theta_e) \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e} \\ &= \frac{\frac{\lambda'}{\lambda} \int_{\theta_e} \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e}{W \int_{\theta_e} \mu(\theta_e) \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e} + \int_{\theta_e} \frac{\kappa(\theta_e)}{\lambda} \frac{d \frac{\xi}{\varepsilon_{L_w}^\omega(\theta_e)} / d\theta_e}{W \int_{\theta_e} \mu(\theta_e) \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} d\theta_e} d\theta_e, \end{aligned}$$

where the second equation is derived by $\kappa(\underline{\theta}_e) = \kappa(\bar{\theta}_e) = 0$ and integration by parts. Note that under our production function $\varepsilon_{L_w}^\omega(\theta_e)$ is independent of θ_e (see, e.g., (A10)). Thus, $d \frac{\xi}{\varepsilon_{L_w}^\omega(\theta_e)} / d\theta_e = 0$ and the above equation implies:

$$(A88) \quad 1 = \frac{\lambda'}{\lambda W \mu}.$$

According to (A75), we have:

$$\frac{1}{\frac{\phi'_w(l_w(\theta_w))}{x_w(\theta_w)}} = \frac{\lambda}{\lambda'} \left[1 - \frac{x'_w(\theta_w)}{x_w(\theta_w)} \frac{\psi_w(\theta_w)}{\lambda N_w f_w(\theta_w)} \frac{1 + \varepsilon_w}{\varepsilon_w} \right].$$

Substitute $\frac{\phi'_w(l_w(\theta_w))}{x_w(\theta_w)}$ by $[1 - \tau_w(\theta_w)] W$:

$$(A89) \quad \frac{1}{1 - \tau_w(\theta_w)} = \frac{W\lambda}{\lambda'} \left[1 - \frac{x'_w(\theta_w)}{x_w(\theta_w)} \frac{\psi_w(\theta_w)}{\lambda N_w f_w(\theta_w)} \frac{1 + \varepsilon_w}{\varepsilon_w} \right].$$

Use (A82), and (A88) to substitute $\frac{\psi_w(\theta_w)}{\lambda N_w f_w(\theta_w)}$ and $\frac{\lambda}{\lambda'}$ in (A89):

$$(A90) \quad \frac{1}{1 - \tau_w(\theta_w)} = \frac{1}{\mu} \left[1 + [1 - \bar{g}_w(\theta_w)] \frac{1 - F_w(\theta_w)}{f_w(\theta_w)} \frac{x'_w(\theta_w)}{x_w(\theta_w)} \frac{1 + \varepsilon_w}{\varepsilon_w} \right].$$

(ii) In the following analysis, we first derive an optimal profit tax formula in part (a). Then we simplify the expression in parts (b) and (c).

(a) Divide both sides of (A77) by $\lambda N_e f_e(\theta_e) P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}$:

$$\begin{aligned} & 1 - \frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}} \\ &= - \frac{\psi_e(\theta_e)}{\lambda N_e f_e(\theta_e)} \frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\ & \quad - \frac{\kappa'(\theta_e)}{\lambda l_e(\theta_e) P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)} N_e f_e(\theta_e)} + \frac{\varphi'(\theta_e) \varepsilon_{l_e}^\omega(\theta_e)}{\lambda l_e(\theta_e) P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)} N_e f_e(\theta_e)}, \end{aligned}$$

where we use $\frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln l_e(\theta_e)} = 1$ and $\int_{\Theta} \varphi'(\theta'_e) \frac{\partial \ln \omega(\theta'_e)}{\partial \ln l_e(\theta'_e)} d\theta_e = \varphi'(\theta_e) \varepsilon_{l_e}^\omega(\theta_e)$ to simplify the expression. Moreover, from the definitions of the elasticities,

$$\int_{\Theta} \varphi'(\theta'_e) \frac{\partial \ln \omega(\theta'_e)}{\partial \ln l_e(\theta'_e)} d\theta_e = \varphi'(\theta_e) \varepsilon_{l_e}^\omega(\theta_e)$$

since $\varepsilon_{l_e}^\omega(\theta'_e, \theta_e)$ is independent of θ'_e and $\int_{\theta_e} \varphi'(\theta'_e) d\theta' = 0$.

For the convenience of derivation, we define:

$$1 - \tilde{\tau}_e(\theta_e) \equiv \frac{[1 - \tau_e(\theta_e)] (1 - \tau_s)}{\mu(\theta_e)} = \frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}}.$$

Then one has

$$\begin{aligned} \tilde{\tau}_e(\theta_e) &= - \frac{\psi_e(\theta_e)}{\lambda N_e f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] [1 - \tilde{\tau}_e(\theta_e)] \\ & \quad - \frac{\kappa'(\theta_e)}{\lambda P(\theta_e) Q_{ij}(\theta_e) N_e f_e(\theta_e)} + \frac{\varphi'(\theta_e) \varepsilon_{l_e}^\omega(\theta_e)}{\lambda P(\theta_e) Q_{ij}(\theta_e) N_e f_e(\theta_e)}, \end{aligned}$$

where we use $\frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln l_e(\theta_e)} = 1$ to simplify the expression. In the same vein, we have

$$(A91) \quad \frac{\tilde{\tau}_e(\theta_e)}{1 - \tilde{\tau}_e(\theta_e)} = -\frac{\psi_e(\theta_e)}{\lambda N_e f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] - \frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{1}{P(\theta_e) Q_{ij}(\theta_e)} \left[\frac{\kappa'(\theta_e)}{\lambda N_e f_e(\theta_e)} - \frac{\varphi'(\theta_e)}{\lambda N_e f_e(\theta_e)} \varepsilon_{l_e}^\omega(\theta_e) \right]$$

Combining (A91) and (A82) gives:

$$\frac{\tilde{\tau}_e(\theta_e)}{1 - \tilde{\tau}_e(\theta_e)} = [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] - \frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{1}{P(\theta_e) Q_{ij}(\theta_e)} \frac{\kappa'(\theta_e)}{\lambda N_e f_e(\theta_e)} + \frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{1}{P(\theta_e) Q_{ij}(\theta_e)} \frac{\varphi'(\theta_e) \varepsilon_{l_e}^\omega(\theta_e)}{\lambda N_e f_e(\theta_e)}.$$

Using (A87) to substitute $\frac{\varphi'(\theta_e) \varepsilon_{l_e}^\omega(\theta_e)}{\lambda N_e f_e(\theta_e)}$ in the above equation,⁵³ we have:

$$(A92) \quad \frac{\tilde{\tau}_e(\theta_e)}{1 - \tilde{\tau}_e(\theta_e)} = [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] - \frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{1}{P(\theta_e) Q_{ij}(\theta_e)} \frac{\kappa'(\theta_e)}{\lambda N_e f_e(\theta_e)} \left[1 - \xi \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} \right] - \frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{L_w(\theta_e)}{P(\theta_e) Q_{ij}(\theta_e)} \frac{\lambda'}{\lambda} \left[1 - \frac{\lambda}{\lambda'} \frac{W \mu(\theta_e)}{1 - \tau_s} \right] \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)}.$$

We now transform the three terms on the right side of the above equations one by one. First, substituting $\kappa'(\theta_e)$ with (A86), we have the following equation:⁵⁴

$$-\frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{1}{P(\theta_e) Q_{ij}(\theta_e)} \frac{\kappa'(\theta_e)}{\lambda N_e f_e(\theta_e)} = \frac{1 - \tau_e(\theta_e)}{1 - \tilde{\tau}_e(\theta_e)} \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \left[[1 - g_e(\theta_e)] - \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \right] \times \left[\frac{\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{d \ln [\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)]}{d \theta_e}}{\varepsilon_{L_w}^\omega(\theta_e)} \right].$$

Second, notice that $\frac{L_w(\theta_e) W}{P(\theta_e) Q_{ij}(\theta_e)} = \frac{\xi}{\mu(\theta_e)}$ and $\frac{\lambda'}{\lambda W} = \mu$ (see, e.g., (A88)). The last term of (A92) equals:

$$(A93) \quad -\frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{L_w(\theta_e)}{P(\theta_e) Q_{ij}(\theta_e)} \frac{\lambda'}{\lambda} \left[1 - \frac{\lambda}{\lambda'} \frac{W \mu(\theta_e)}{1 - \tau_s} \right] \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} = -\frac{1 - \tau_s}{1 - \tilde{\tau}_e(\theta_e)} \frac{\xi}{\mu(\theta_e)} \mu \left[1 - \frac{\mu(\theta_e)}{\mu} \right] \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} = -\frac{\xi}{1 - \tilde{\tau}_e(\theta_e)} \left[\frac{\mu}{\mu(\theta_e)} - 1 \right] \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)}.$$

Substituting the second and third terms of the right side of (A92) by (A93) and (A93) gives:

⁵³Equation (A87) suggests that $\frac{\varphi'(\theta_e) \varepsilon_{l_e}^\omega(\theta_e)}{\lambda N_e f_e(\theta_e)} = \left[\left[\frac{W \mu(\theta)}{1 - \tau_s} - \frac{\lambda'}{\lambda} \right] L_w(\theta_e) + \frac{\kappa'(\theta_e) \xi}{\lambda N_e f_e(\theta_e)} \right] \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)}$.

⁵⁴Note that we consider the case with $\tau_s = 0$.

$$\begin{aligned}
(A94) \quad & \frac{1}{1 - \tau_e(\theta_e)} \\
&= \frac{1 + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]}{\mu(\theta_e)} \\
&+ \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \left[\frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \left[\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{d \ln [\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)]}{d\theta_e} \right] \right] \left[1 - \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} \right] \\
&+ \frac{1}{1 - \tau_e(\theta_e)} \left[1 - \frac{\mu}{\mu(\theta_e)} \right] \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)},
\end{aligned}$$

where we have substituted $1 - \tilde{\tau}_e(\theta_e)$ by $\frac{1 - \tau_e(\theta_e)}{\mu(\theta_e)}$. Using $RE(\theta_e) \equiv \frac{\mu}{\mu(\theta_e)} - 1$ and

$$(A95) \quad \widetilde{IRE}(\theta_e) \equiv \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \left[\frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \left[\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{d \ln [\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)]}{d\theta_e} \right] \right],$$

we have

$$\begin{aligned}
\frac{1}{1 - \tau_e(\theta_e)} &= \frac{1 + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]}{\mu(\theta_e)} \\
&+ \widetilde{IRE}(\theta_e) \left[1 - \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} \right] - \frac{1}{1 - \tau_e(\theta_e)} \zeta RE(\theta_e) \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)},
\end{aligned}$$

which is equivalent to

$$(A96) \quad \frac{1}{1 - \tau_e(\theta_e)} = \frac{\left[\frac{1 + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]}{\mu(\theta_e)} + \widetilde{IRE}(\theta_e) \left[1 - \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} \right] \right]}{1 + RE(\theta_e) \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)}}.$$

(b) We now try to express the right side of the above equation in terms of parameters. Using (A53), we

have:

$$\begin{aligned}
& \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\
&= \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left\{ \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) + 1}{\varepsilon_{Q_{ij}}^S(\theta_e)} \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} \right] + \mu(\theta_e) \frac{d \ln \mu(\theta_e)}{d\theta_e} \right. \\
&\quad \left. + \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \frac{(1 + \zeta \frac{1 + \varepsilon_e}{\varepsilon_e}) \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} \right] + \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)} \right\} \\
&= \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \frac{1 + \varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \zeta] \\
&\quad + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{d}{d\theta_e} \ln \frac{X_e(\theta_e)}{\mu(\theta_e)}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)} \left[\frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) + 1}{\varepsilon_{Q_{ij}}^S(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \right. \\
&\quad \left. - \left(1 + \zeta \frac{1 + \varepsilon_e}{\varepsilon_e} \right) \right] \\
&\quad + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)},
\end{aligned}$$

where the third equation is derived by $\mu(\theta_e) \frac{d \ln \mu(\theta_e)}{d\theta_e} = [\mu(\theta_e) - \zeta] \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e}$ and combine terms multiplied by $\frac{d}{d\theta_e} \ln \frac{X_e(\theta_e)}{\mu(\theta_e)}$. Notice that:

$$\frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{d}{d\theta_e} \ln \frac{X_e(\theta_e)}{\mu(\theta_e)}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \frac{\sigma - 1}{\sigma} \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)} = H(\theta_e) - \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e},$$

where $H(\theta_e)$ is given by (A48). We have

$$\begin{aligned}
& \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\
&= \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) + \frac{\varepsilon_e}{1 + \varepsilon_e} - \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross} + 1}{\varepsilon_{Q_{ij}}^S(\theta_e)} \right] \\
&\quad + H(\theta_e) \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross} + 1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \frac{\varepsilon_e}{1 + \varepsilon_e} - \zeta \right] + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross} \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)},
\end{aligned}$$

where, according to $\varepsilon_{Q_{-ij}}^{P, cross} = -\frac{1}{\mu(\theta_e)} + \varepsilon_{Q_{ij}}^S(\theta_e)$, one has $\frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross} + 1}{\varepsilon_{Q_{ij}}^S(\theta_e)} = \mu(\theta_e)$. Therefore,

$$\begin{aligned}
(A97) \quad & \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\
&= \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} + H(\theta_e) \left[\frac{1 + \varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \zeta] - 1 \right] \\
&\quad + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross} \frac{d \ln [1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \varepsilon_{Q_{ij}}^S(\theta_e) \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)}.
\end{aligned}$$

Substituting $\frac{l'_e(\theta_e)}{l_e(\theta_e)}$ in $\widetilde{IRE}(\theta_e)$ by (A45) and utilizing (A48), i.e.,

$$H(\theta_e) = \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{d \ln[X_e(\theta_e)/\mu(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[1 - \xi \varepsilon_{Qij}^S(\theta_e) \right] - \varepsilon_{Qij}^S(\theta_e)} + \frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e} \right],$$

we have:

$$\widetilde{IRE}(\theta_e) = \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left\{ \left[\begin{aligned} & [1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] \times \\ & \frac{\frac{1+\varepsilon_e}{\varepsilon_e} \left[1 - \xi \varepsilon_{Qij}^S(\theta_e) \right] \frac{d \ln[1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Qij}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \xi \right)} \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} + H(\theta_e) - \\ & \left[1 - \frac{\frac{d \ln[\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e}}{\frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e}} \right] \frac{[1 - F_e(\theta_e)]}{f_e(\theta_e)} \frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e} \end{aligned} \right] \right\}.$$

Notice that $\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) + 1 = \mu(\theta_e) \varepsilon_{Qij}^S(\theta_e)$. We have $\frac{d \ln[\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e} / \frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e} = \frac{\varepsilon_{Qij}^S(\theta_e) [\mu(\theta_e) - \xi]}{\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)}$ and

$$(A98) \quad \widetilde{IRE}(\theta_e) = \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left\{ \left[\begin{aligned} & [1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] \times \\ & H(\theta_e) + \frac{\left[1 - \xi \varepsilon_{Qij}^S(\theta_e) \right] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e}}{\mu(\theta_e) \varepsilon_{Q-ij}(\theta_e)} \\ & + \frac{\left[1 - \xi \varepsilon_{Qij}^S(\theta_e) \right] \frac{1+\varepsilon_e}{\varepsilon_e} \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln[1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Qij}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \xi \right)} \end{aligned} \right] \right\}.$$

Last, substituting $\frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\partial \ln P(Q_{ij}(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]$ and $\widetilde{IRE}(\theta_e)$ in (A96) by (A97) and (A98), respectively, we have, for any $\theta_e \in \Theta_e$:

$$(A99) \quad \begin{aligned} & \frac{1 + RE(\theta_e) \xi \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)}}{1 - \tau_e(\theta_e)} \\ &= \frac{1 + [1 - \bar{g}_e(\theta_e)] \left[\frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e} + \frac{\frac{1+\varepsilon_e}{\varepsilon_e} \mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \frac{d \ln[1 - \tau_e(\theta_e)]}{d\theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Qij}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \xi \right)} \right] \right]}{\mu(\theta_e)} \\ &+ \left[1 - \xi \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)} \right] \varepsilon_{Q-ij}^{P,cross}(\theta_e) [1 - g_e(\theta_e)] - \left[1 - \xi \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)} \right] \varepsilon_{Q-ij}^{P,cross}(\theta_e) \\ &\times [1 - \bar{g}_e(\theta_e)] \left\{ \begin{aligned} & H(\theta_e) + \frac{\left[1 - \xi \varepsilon_{Qij}^S(\theta_e) \right] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e}}{\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)} \\ & + \frac{\left[1 - \xi \varepsilon_{Qij}^S(\theta_e) \right] \frac{1+\varepsilon_e}{\varepsilon_e} \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln[1 - \tau_e(\theta_e)]}{d\theta_e}}{\left[\frac{1+\varepsilon_e}{\varepsilon_e} - \varepsilon_{Qij}^S(\theta_e) \left(1 + \frac{1+\varepsilon_e}{\varepsilon_e} \xi \right) \right]} \end{aligned} \right\} \end{aligned}$$

Notice that $1 - \xi \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)} = \frac{1}{1 - \xi \varepsilon_{Qij}^S(\theta_e)}$. One can see the sum of terms multiplied by $\frac{d \ln[1 - \tau_e(\theta_e)]}{d\theta_e}$ of the above equation equals zero. Moreover, the sum of terms multiplied by $\frac{d \ln[\mu(\theta_e) - \xi]}{d\theta_e}$ also equals zero. Last, using the

definition of $IRE(\theta_e)$ (see, e.g., (34)), we have (A100):

$$(A100) \quad \frac{1 + RE(\theta_e) \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)}}{1 - \tau_e(\theta_e)} = \frac{1 + [1 - \bar{g}_e(\theta_e)] \left\{ H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \zeta] - 1 \right] \right\}}{\mu(\theta_e)} \\ + \left[1 - \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} \right] [1 - g_e(\theta_e)] \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \\ - \left[1 - \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} \right] \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) [1 - \bar{g}_e(\theta_e)] H(\theta_e),$$

Notice that $1 - \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{L_w}^\omega(\theta_e)} = \frac{1}{1 - \zeta \frac{\sigma-1}{\sigma}}$. Equation (A100) is equivalent to (32) in the nested-CES economy. ■

2.5. Validity of FOA

The solution to the relaxed planner's problem might not be a solution to the original planner's problem, because the first-order conditions are only necessary conditions. In a perfectly competitive economy, the validity of the FOA is guaranteed by the standard Spence-Mirrlees condition on preference and a monotonicity condition on the allocation. However, under the Cournot competition, the conditions are more involved. Lemma A1 provides an insight for when the first-order approach is valid. For the quantitative analysis in Section IV, we verify incentive compatibility as well as any omitted non-negativity constraints numerically.

LEMMA A1: Under Assumption 1, the first-order incentive condition $\frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e} \big|_{\theta'_e=\theta_e} = 0$ is not only a necessary but also a sufficient condition for the agent's problem if for any $(\theta_e, \theta'_e) \in \Theta^2$, $\frac{1+\varepsilon_e}{\varepsilon_e} \frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e} \frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta'_e} + \frac{\partial^2 \ln l_e(\theta'_e|\theta_e)}{\partial \theta'_e \partial \theta_e} < 0$.

PROOF:

(i) We first show that there is a unique solution to problem (26). Remember that we have set $Q_{ij}(\theta'_e|\theta_e) = Q_{ij}(x_e(\theta_e) l_e(\theta'_e|\theta_e), L_w(\theta'_e|\theta_e))$ and $P_{ij}(\theta'_e|\theta_e) = P_{ij}(Q_{ij}(\theta'_e|\theta_e), Q_{-ij}(\theta_e), \theta_e)$. The first-order condition of problem (26),

$$(A101) \quad \frac{\partial [P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e) (1 - t_s) - W L_w(\theta'_e|\theta_e)]}{\partial L_w(\theta'_e|\theta_e)} = 0,$$

is equivalent to

$$\left[1 + \frac{\partial \ln P_{ij}(\theta'_e|\theta_e)}{\partial \ln Q_{ij}(\theta'_e|\theta_e)} \right] \frac{P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)}{L_w(\theta'_e|\theta_e)} \zeta (1 - t_s) - W = 0.$$

The second-order condition of problem (26),

$$0 > \frac{\partial \left[1 + \frac{\partial \ln P_{ij}(\theta'_e|\theta_e)}{\partial \ln Q_{ij}(\theta'_e|\theta_e)} \right]}{\partial L_w(\theta'_e|\theta_e)} \frac{P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)}{L_w(\theta'_e|\theta_e)} \zeta (1 - t_s) - \\ \left[1 + \frac{\partial \ln P_{ij}(\theta'_e|\theta_e)}{\partial \ln Q_{ij}(\theta'_e|\theta_e)} \right] \frac{P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)}{L_w(\theta'_e|\theta_e)^2} \zeta (1 - t_s) + \\ \left[1 + \frac{\partial \ln P_{ij}(\theta'_e|\theta_e)}{\partial \ln Q_{ij}(\theta'_e|\theta_e)} \right] \frac{1}{L_w(\theta'_e|\theta_e)} \frac{\partial [P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)]}{\partial L_w(\theta'_e|\theta_e)} \zeta (1 - t_s),$$

holds because

$$\frac{\partial \left[1 + \frac{\partial \ln P_{ij}(\theta'_e|\theta_e)}{\partial \ln Q_{ij}(\theta'_e|\theta_e)} \right]}{\partial L_w(\theta'_e|\theta_e)} \frac{P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)}{L_w(\theta'_e|\theta_e)} \zeta (1 - t_s) < 0$$

and

$$\begin{aligned}
& \frac{1}{L_w(\theta'_e|\theta_e)} \frac{\partial [P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)]}{\partial L_w(\theta'_e|\theta_e)} - \frac{P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)}{L_w(\theta'_e|\theta_e)^2} \\
&= \frac{P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)}{L_w(\theta'_e|\theta_e)^2} \left[\frac{\partial \ln [P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)]}{\partial \ln L_w(\theta'_e|\theta_e)} - 1 \right] \\
&= \frac{P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)}{L_w(\theta'_e|\theta_e)^2} \left[\frac{\partial \ln [P_{ij}(\theta'_e|\theta_e) Q_{ij}(\theta'_e|\theta_e)]}{\partial \ln Q_{ij}(\theta'_e|\theta_e)} \xi - 1 \right] < 0
\end{aligned}$$

hold for any $L_w(\theta'_e|\theta_e)$ and $l_e(\theta'_e|\theta_e)$. Therefore, there must be a unique solution to the problem (26).

(ii) According to the definition of $V_e(\theta'_e|\theta_e)$, we have

$$(A102) \quad \frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e} = c'_e(\theta'_e) - \phi'_e(l_e(\theta'_e|\theta_e)) \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta'_e}.$$

The first-order necessary condition $\frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e}|_{\theta'_e=\theta_e} = 0$ implies $\lim_{\theta'_e \rightarrow \theta_e} \frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e} = 0$, i.e.,

$$(A103) \quad 0 = \left[c'_e(\theta'_e) - \phi'_e(l_e(\theta'_e|\theta_e)) \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta'_e} \right] |_{\theta'_e=\theta_e},$$

Adding (A102) into (A103), we have

$$\frac{\partial V_e(\theta'_e|\theta_e)}{\partial \theta'_e} = \left[\phi'_e(l_e(\theta'_e|\theta_e)) \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta'_e} \right] |_{\theta'_e=\theta_e} - \phi'_e(l_e(\theta'_e|\theta_e)) \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta'_e}$$

Using the mean value theorem, the sign of the right-hand side is given by

$$\frac{d \left[\phi'_e(l_e(\theta'_e|\theta_e^*)) \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} \right]}{d\theta^*} (\theta'_e - \theta_e)$$

for some θ_e^* that lies between θ'_e and θ_e . If one has $\frac{d \left[\phi'_e(l_e(\theta'_e|\theta_e^*)) \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} \right]}{d\theta^*} < 0$ for any $(\theta_e^*, \theta'_e) \in \Theta^2$, the function $V_e(\theta'_e|\theta_e)$ will increase with θ'_e until $\theta'_e = \theta_e$ and then decreases with θ'_e . In conclusion, there is a unique the local maximum point that is also the global maximizer of $V_e(\theta'_e|\theta_e)$. Then under Assumption 1, the first-order incentive condition is not only necessary but also sufficient for the agent's problem.

Notice that

$$\frac{\partial^2 \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^* \partial \theta'_e} = \frac{1}{l_e(\theta'_e|\theta_e^*)} \frac{\partial^2 l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^* \partial \theta'_e} - \frac{1}{l_e(\theta'_e|\theta_e^*)^2} \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^*} \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e}$$

and

$$\begin{aligned}
\frac{d \left[\phi'_e(l_e(\theta'_e|\theta_e^*)) \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} \right]}{d\theta^*} &= \phi_e''(l_e(\theta'_e|\theta_e^*)) \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^*} \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} + \phi'_e(l_e(\theta'_e|\theta_e^*)) \frac{\partial^2 l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^* \partial \theta'_e} \\
&= \phi'_e(l_e(\theta'_e|\theta_e^*)) l_e(\theta'_e|\theta_e^*) \left[\frac{\phi_e''(l_e(\theta'_e|\theta_e^*)) l_e(\theta'_e|\theta_e^*)}{\phi'_e(l_e(\theta'_e|\theta_e^*))} \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^*} \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} \right. \\
&\quad \left. + \frac{1}{l_e(\theta'_e|\theta_e^*)} \frac{\partial^2 l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^* \partial \theta'_e} \right].
\end{aligned}$$

We have

$$\frac{d \left[\phi'_e(l_e(\theta'_e|\theta_e^*)) \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} \right]}{d\theta^*} = \phi'_e(l_e(\theta'_e|\theta_e^*)) l_e(\theta'_e|\theta_e^*) \left[\frac{\left[\frac{\phi_e''(l_e(\theta'_e|\theta_e^*)) l_e(\theta'_e|\theta_e^*)}{\phi'_e(l_e(\theta'_e|\theta_e^*))} + 1 \right] \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^*} \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e}}{+ \frac{\partial^2 \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^* \partial \theta'_e}} \right].$$

It can be seen that $\frac{d \left[\phi'_e(l_e(\theta'_e|\theta_e^*)) \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} \right]}{d\theta^*} < 0$ just means that the second-order necessary condition is satisfied. To see this, notice that the first-order necessary condition implies $\left[\frac{\partial^2 V_e(\theta'_e|\theta_e)}{\partial (\theta'_e)^2} d\theta'_e + \frac{\partial^2 V_e(\theta'_e|\theta_e)}{\partial \theta'_e \partial \theta_e} d\theta_e \right] \Big|_{\theta'_e=\theta_e} = 0$. The second-order necessary condition $\frac{\partial^2 V_e(\theta'_e|\theta_e)}{\partial (\theta'_e)^2} \leq 0$ is equivalent to $\frac{\partial^2 V_e(\theta'_e|\theta_e)}{\partial \theta'_e \partial \theta_e} \geq 0$ under the first-order necessary condition, i.e.,

$$- \left[\frac{\phi''_e(l_e(\theta'_e|\theta_e)) l_e(\theta'_e|\theta_e)}{\phi'_e(l_e(\theta'_e|\theta_e))} \frac{1}{l_e(\theta'_e|\theta_e)} \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta_e} \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta'_e} + \frac{\partial^2 l_e(\theta'_e|\theta_e)}{\partial \theta'_e \partial \theta_e} \right] \phi'_e(l_e(\theta'_e|\theta_e)) \geq 0.$$

Notice that $\frac{\partial^2 l_e(\theta'_e|\theta_e)}{\partial \theta_e \partial \theta'_e} = \frac{\partial^2 \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e \partial \theta'_e} l_e(\theta'_e|\theta_e) + \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta_e} \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta'_e} \frac{1}{l_e(\theta'_e|\theta_e)}$. The above inequality is equivalent to

$$- \left[\left(1 + \frac{1}{\varepsilon_e} \right) \frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e} \frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta'_e} + \frac{\partial^2 \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e \partial \theta'_e} \right] l_e(\theta'_e|\theta_e) \phi'_e(l_e(\theta'_e|\theta_e)) \geq 0.$$

Now that

$$\text{sign} \left(\frac{d \left[\phi'_e(l_e(\theta'_e|\theta_e^*)) \frac{\partial l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} \right]}{d\theta^*} \right) = \text{sign} \left(\left(1 + \frac{1}{\varepsilon_e} \right) \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^*} \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} + \frac{\partial^2 \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^* \partial \theta'_e} \right),$$

one can see that $\left(1 + \frac{1}{\varepsilon_e} \right) \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^*} \frac{\partial \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta'_e} + \frac{\partial^2 \ln l_e(\theta'_e|\theta_e^*)}{\partial \theta_e^* \partial \theta'_e} < 0$ for any $(\theta_e^*, \theta'_e) \in \Theta^2$ is a sufficient condition for the validity of FOA. ■

Notice that $\frac{1+\varepsilon_e}{\varepsilon_e} > 0$, a stronger condition for the validity of the first-order approach is $\frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e} < 0$, $\frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta'_e} > 0$ and $\frac{\partial^2 \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e \partial \theta'_e} \leq 0$. $\frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e} < 0$ means the higher the ability of the entrepreneur, the lower the effort needed to finish the task. $\frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta'_e} > 0$ means the higher the reported type, the more effort needed to finish the task. Last, $\frac{\partial^2 \ln l_e(\theta'_e|\theta_e)}{\partial \theta_e \partial \theta'_e} \leq 0$ means $\frac{\partial \ln l_e(\theta'_e|\theta_e)}{\partial \theta'_e}$ non-increase in θ_e . All of the above conditions hold in the monopoly competitive case with $X'_e(\cdot) > 0$. Incentive compatibility in more general case needs to be discussed with specific examples.

Here we establish implementability of the constrained optimal allocation.^{55,56}

LEMMA A2: Suppose that the FOCs of the agents and the final good producers are both necessary and sufficient. Symmetric Cournot competitive tax equilibrium $\{\mathcal{A}, \mathcal{T}, \mathcal{P}\}$ satisfies the following conditions jointly:

- (i) Incentive conditions (24) and (27) are satisfied;
- (ii) Prices and wage satisfy (15) and (16);
- (iii) Market clearing conditions (12) to (14) are satisfied.

Conversely, suppose the allocation $\mathcal{A}$ and price $\mathcal{P}$ satisfy the above parts (i) to (iii). Then there exists a tax system $\mathcal{T}$ with $t_s = 0$ such that the allocation $\mathcal{A}$ can be implemented at the prices $\mathcal{P}$ by the tax system $\mathcal{T}$.

PROOF:

We first show that a symmetric Cournot competitive tax equilibrium must satisfy parts 1 to 3. First, by the definition of SCCTE, (15) to (16) and (12) to (14) must be satisfied. Second, by the definition of SCCTE, agents maximize their utilities, which, by envelop theory, means (24) and (27) must be satisfied (see, e.g., Lemma 1).

⁵⁵To see why the sales tax is redundant, suppose that $\{T_w(y), T_e(\pi), t_s\}$ is the optimal tax that implements the second-best allocation and that there exists another optimal tax system $\{T_w^\#(y), T_e^\#(\pi), t_s^\#\}$ that can implement the second-best allocation with $t_s^\# = 0$. Then the tax system can be constructed such that $1 - T_o^\#(x) = [1 - T_o^I(x)](1 - t_s)$, $x \in \mathbb{R}_+$.

⁵⁶In our model, we allow profit and labor income tax to be different, which governs the wage rate, so that there is no need to use the sales tax to manipulate W to achieve indirect redistribution between the entrepreneur and worker. The sales tax is not redundant if income taxes on labor income and profit are restricted to be uniform (see, e.g., Scheuer (2014)).

Next, suppose that we are given an allocation $\mathcal{A}$ and price $\mathcal{P}$ to satisfy the properties in parts 1 to 3. We now construct the tax system $\mathcal{T}$ (with $t_s = 0$), which together with the given allocation $\mathcal{A}$ and price $\mathcal{P}$ constructs a SCCTE. We first construct a policy system with the given allocation $\mathcal{A}$ and price $\mathcal{P}$. We then show that this constructed policy system together with $\mathcal{A}$ and $\mathcal{P}$ constructs a SCCTE.

First, we construct the policy system. By the definition of tax wedges and $t_s = 0$, the marginal tax rates are constructed as follows:

$$T'_w(y_w(\theta_w)) = 1 - \frac{\phi'_w(l_w(\theta_w))}{\frac{P(\theta_e)}{\mu(\theta_e)} \frac{\partial Q_{ij}(\theta_e)}{\partial l_w(\theta_e)} x_w(\theta_w)}$$

and

$$T'_e(y_e(\theta_e)) = 1 - \frac{\phi'_e(l_e(\theta_e))}{\frac{P(\theta_e)}{\mu(\theta_e)} \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}}.$$

We use agents' budget constraints to fix the labor income taxes. We first construct $T_w(\cdot)$. To do this, we substitute

$$T_w(y_w(\theta_w)) = T_w(y_w(\underline{\theta}_w)) + \int_{y_w(\underline{\theta}_w)}^{y_w(\theta_w)} T'_w(y_w) dy$$

into

$$y_w(\underline{\theta}_w) - T_w(y_w(\underline{\theta}_w)) - c_w(\underline{\theta}_w) = 0.$$

We now show that the constructed labor income tax is consistent with the agent's problems. By construction, we have shown that the constructed $T_w(\cdot)$ is consistent with the $\underline{\theta}_w$ -type agent's action. We should show the constructed labor income tax is also consistent with other agents' choices. In particular, since by construction the first-order conditions are already satisfied, we should show that the budget constraints are also satisfied. This is equivalent to say that for any $\theta_w \in \Theta_w$,

$$y'_w(\theta_w) [1 - T'_w(y_w(\theta_w))] - c'_w(\theta_w) = 0.$$

Substituting $1 - T'_w(y_w(\theta_w))$ with the FOC (17), the above equation is equivalent to

$$c'_w(\theta_w) - \frac{\phi'_w(l_w(\theta_w))}{W x_w(\theta_w)} y'_w(\theta_w) = 0.$$

The above equations holds because for any $\theta_w \in \Theta_w$ the allocation satisfies:

$$\begin{aligned} (A104) \quad V'_w(\theta_w) &= \frac{\phi'_w(l_w(\theta_w)) l_w(\theta_w) x'_w(\theta_w)}{x_w(\theta_w)} \\ &= c'_w(\theta_w) - \frac{y'_w(\theta_w)}{W x_w(\theta_w)} \phi'_w(l_w(\theta_w)) + \frac{\phi'_w(l_w(\theta_w)) l_w(\theta_w) x'_w(\theta_w)}{x_w(\theta_w)}. \end{aligned}$$

The first equation of (A104) is the incentive condition, and the second equation is derived through the definition of $V_w(\theta_w)$. In conclusion, given the allocation, we can construct a unique labor income tax that is consistent with the allocation in the equilibrium.

The construction of $T_e(\cdot)$ is similar to the construction of $T_w(\cdot)$. Note that $T_w(y_w(\underline{\theta}_w))$ can be different from $T_e(y_e(\underline{\theta}_e))$. We substitute

$$T_e(y_e(\theta_e)) = T_e(y_e(\underline{\theta}_e)) + \int_{y_e(\underline{\theta}_e)}^{y_e(\theta_e)} T'_e(y_e) d\pi$$

into

$$y_e(\theta_e) - T_e(y_e(\theta_e)) - c_e(\theta_e) = 0$$

and show that there exists $T_e(\cdot)$ such that given allocation $\mathcal{A}$, price $\mathcal{P}$ and the constructed marginal profit income tax rates, the above equation is satisfied for any $\theta_e \in \Theta_e$. To be consistent with the $\underline{\theta}_e$ -type agent's

budget constraint, $T_e(y_e(\theta_e))$ must satisfy

$$y_e(\theta_e) - T_e(y_e(\theta_e)) - c_e(\theta_e) = 0.$$

We should show this constructed profit tax on $y_e(\theta_e)$, i.e., $T_e(y_e(\theta_e))$, is also consistent with other agents' budget constraints. This is equivalent to say that

$$y_e(\theta_e) [1 - T'_e(y_e(\theta_e))] - c'_e(\theta_e) = 0.$$

Substituting $1 - T'_e(y_e(\theta_e))$ with the FOC (18), the above equation is equivalent to

$$c'_e(\theta_e) - \frac{\phi'_e(l_e(\theta_e))}{\frac{P(\theta_e)}{\mu(\theta_e)} \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}} y_e(\theta_e) = 0,$$

which is further equivalent to

$$(A105) \quad c'_e(\theta_e) - \mu(\theta_e) \frac{\phi'_e(l_e(\theta_e)) l_e(\theta_e)}{P(\theta_e) Q_{ij}(\theta_e)} y_e(\theta_e) = 0.$$

The above equations are true since we have

$$(A106) \quad \begin{aligned} V'_e(\theta_e) &= \phi'_e(l_e(\theta_e)) l_e(\theta_e) \left[\mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{\zeta'_e(\theta_e)}{\zeta_e(\theta_e)} \right] \\ &= c'_e(\theta_e) - \phi'_e(l_e(\theta_e)) l_e(\theta_e) \frac{l'_e(\theta_e)}{l_e(\theta_e)}, \end{aligned}$$

and

$$(A107) \quad y_e(\theta_e) = Q_{ij}(\theta_e) P(\theta_e) \left[\frac{\frac{d \ln P_{ij}(Q_{ij}, Q_{ij}(\theta_e), \theta_e)}{d \theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \left[1 + \frac{\partial \ln P_{ij}(Q_{ij}(\theta_e), Q_{-ij}, \theta_e)}{\partial \ln Q_{ij}(\theta_e)} \Big|_{Q_{-ij}=Q_{ij}(\theta_e)} \right] \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{l'_e(\theta_e)}{l_e(\theta_e)} \right] \right].$$

Substituting $\frac{l'_e(\theta_e)}{l_e(\theta_e)}$ in (A106) by (A107) and using $1 + \frac{\partial \ln P_{ij}(Q_{ij}(\theta_e), Q_{-ij}, \theta_e)}{\partial \ln Q_{ij}(\theta_e)} \Big|_{Q_{-ij}=Q_{ij}(\theta_e)} = \frac{1}{\mu(\theta_e)}$ delivers (A105) immediately. The first equation of (A106) is the incentive condition, and the second equation is derived through the definition of $V_e(\theta_e)$. (A107) is derived from the definition of $y_e(\theta_e)$ (i.e., $y_e(\theta_e) = P(\theta_e) Q_{ij}(\theta_e) - WL_w(\theta_e)$) and the fact that the derivative of $y_e(\theta_e)$ with respect to $L_w(\theta_e)$ is zero.

In conclusion, given the allocation, we can construct a unique combination of labor income tax and profit tax that is consistent with the allocation in the equilibrium.

We now show that the allocation $\mathcal{A}$ and price $\mathcal{P}$ satisfying parts 1 to 3 and the constructed tax system $\mathcal{T}$ construct an SCCTE. First, the constructed allocation and taxation satisfy the first-order conditions. Since we have assumed that the first-order conditions are both necessary and sufficient, the allocation solves agents' problems. Second, the price $\mathcal{P}$ satisfies (15) and (16). Third, the market clear conditions (12) to (14) are satisfied. Lastly, agents' budget constraints (8) and (11) are embedded in the definitions of gross utilities and the construction of income taxes. In conclusion, the constructed tax system $\mathcal{T}$ together with the given allocation $\mathcal{A}$ and price $\mathcal{P}$ constructs an SCCTE. ■

III. Main Results

3.1. An Explicit Formula for the Tax Rate

Theorem 1 fully describes the optimal tax wedges, but the tax rate for the entrepreneurs $\tau_e(\theta_e)$ cannot be written explicitly because the weights $\omega(\theta_e)$ in equation (35) and hence the average markup μ is a function of the tax rate $\tau_e(\theta_e)$. In what follows, we can write the tax rate explicitly under a particular parameter

configuration. Then we can solve explicitly for the weights $\omega(\theta_e)$ and therefore we can write the average markup μ explicitly in the following Corollary:

COROLLARY A1: When $\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \zeta \right) = 2$, we have

$$\mu = \frac{\left(\frac{\sigma}{\sigma-1} + \zeta \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e}{2\zeta} - \frac{\sqrt{\Delta}}{2\zeta},$$

where

$$\bar{\gamma}(\theta_e) = \frac{\gamma(\theta_e)}{\int_{\theta_e} \gamma(\theta_e) d\theta_e},$$

$$\gamma(\theta_e) = \frac{\left(\frac{X_e(\theta_e)}{\mu(\theta_e)} \right)^{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{\sigma}{\sigma-1}} f_e(\theta_e)}{1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[\frac{1}{\varepsilon_e} \mu(\theta_e) - \frac{1+\varepsilon_e}{\varepsilon_e} \zeta + \frac{1}{\sigma-1} \right] + [\mu(\theta_e) - \frac{\sigma}{\sigma-1}] [1 - g_e(\theta_e)]},$$

$$\Delta = \left(\frac{\sigma}{\sigma-1} - \zeta \right)^2 \left[\int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e \right]^2 - 4 \frac{\sigma}{\sigma-1} \zeta \left[\int_{\theta_e} \mu(\theta_e)^2 \bar{\gamma}(\theta_e) d\theta_e - \left(\int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e \right)^2 \right].$$

PROOF:

Substituting elasticities, $RE(\theta_e)$ and $IRE(\theta_e)$ in (32) by (33) and (34), we have:

$$1 - \tau_e(\theta_e) = \frac{\frac{\frac{\sigma}{\sigma-1}}{\frac{\sigma}{\sigma-1} - \zeta} - \frac{\zeta}{\frac{\sigma}{\sigma-1} - \zeta} \frac{\mu}{\mu(\theta_e)}}{1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[\frac{1}{\varepsilon_e} \mu(\theta_e) - \frac{1+\varepsilon_e}{\varepsilon_e} \zeta + \frac{1}{\sigma-1} \right] + [\mu(\theta_e) - \frac{\sigma}{\sigma-1}] [1 - g_e(\theta_e)] \mu(\theta_e)}.$$

Notice that $1 < \sigma < \eta(\theta_e)$, $\frac{\sigma}{\sigma-1} \geq \mu(\theta_e) = \frac{1}{1 - \left[\frac{1}{\eta(\theta_e)} \frac{l-1}{l} + \frac{1}{\sigma} \frac{1}{l} \right]} > 1$, $0 \leq \zeta < 1$. We have

$$1 - \frac{\zeta}{\frac{\sigma}{\sigma-1} - \zeta} RE(\theta_e) = \frac{1}{\frac{\sigma}{\sigma-1} - \zeta} \left[\frac{\sigma}{\sigma-1} \frac{\mu(\theta_e)}{\mu} - \zeta \right] > 0.$$

Then, we have

$$1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[\frac{1}{\varepsilon_e} \mu(\theta_e) - \frac{1+\varepsilon_e}{\varepsilon_e} \zeta + \frac{1}{\sigma-1} \right] + \left[\mu(\theta_e) - \frac{\sigma}{\sigma-1} \right] [1 - g_e(\theta_e)] \geq 0,$$

since $1 - \tau_e(\theta_e) \geq 0$. According to the above inequality, we have $\gamma(\theta_e) \geq 0$ and $\bar{\gamma}(\theta_e) \geq 0$. When

$\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \zeta \right) = 2$, we have

$$\begin{aligned} \mu &= \frac{\int_{\theta_e} \mu(\theta_e) [1 - \tau_e(\theta_e)] \left(\frac{X_e(\theta_e)}{\mu(\theta_e)} \right)^{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{\sigma}{\sigma-1}} f_e(\theta_e) d\theta_e}{\int_{\theta_e} [1 - \tau_e(\theta_e)] \left(\frac{X_e(\theta_e)}{\mu(\theta_e)} \right)^{\frac{1+\varepsilon_e}{\varepsilon_e} \frac{\sigma}{\sigma-1}} f_e(\theta_e) d\theta_e} \\ &= \frac{\int_{\theta_e} \mu(\theta_e) \gamma(\theta_e) \left[\frac{\sigma}{\sigma-1} \mu(\theta_e) - \zeta \mu \right] d\theta_e}{\int_{\theta_e} \gamma(\theta_e) \left[\frac{\sigma}{\sigma-1} \mu(\theta_e) - \zeta \mu \right] d\theta_e}, \end{aligned}$$

where the second equation is derived by the relationship between $\gamma(\theta_e)$ and $[1 - \tau_e(\theta_e)]$. Accordingly, we

have

$$\begin{aligned} & \frac{\sigma}{\sigma-1} \mu \int_{\theta_e} \gamma(\theta_e) \mu(\theta_e) d\theta_e - \xi \mu^2 \int_{\theta_e} \gamma(\theta_e) d\theta_e \\ &= \frac{\sigma}{\sigma-1} \int_{\theta_e} \mu(\theta_e)^2 \gamma(\theta_e) d\theta_e - \xi \mu \int_{\theta_e} \mu(\theta_e) \gamma(\theta_e) d\theta_e, \end{aligned}$$

or equivalently

$$(A108) \quad 0 = \xi \mu^2 - \left[\left(\frac{\sigma}{\sigma-1} + \xi \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e \right] \mu + \frac{\sigma}{\sigma-1} \int_{\theta_e} \mu(\theta_e)^2 \bar{\gamma}(\theta_e) d\theta_e,$$

which is a quadratic equation of μ .

We define

$$\Delta = \left(\frac{\sigma}{\sigma-1} - \xi \right)^2 \left[\int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e \right]^2 - 4 \frac{\sigma}{\sigma-1} \xi \left[\int_{\theta_e} \mu^2(\theta_e) \bar{\gamma}(\theta_e) d\theta_e - \left(\int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e \right)^2 \right]$$

as the discriminant of (A108). Set

$$\mathbf{E}_{\bar{\gamma}}(\mu) = \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e, \mathbf{Var}_{\bar{\gamma}}(\mu) = \left[\int_{\theta_e} \mu^2(\theta_e) \bar{\gamma}(\theta_e) d\theta_e - \left(\int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e \right)^2 \right]$$

and

$$\mu^*(\theta_e) = (\sigma-1) [\mu(\theta_e) - 1].$$

We have $\mu^*(\theta_e) \in (0, 1]$, because $\mu(\theta_e) \in (1, \frac{\sigma}{\sigma-1}]$. Set

$$\begin{aligned} \mathbf{E}_{\bar{\gamma}}(\mu^*) &= \int_{\theta_e} \bar{\gamma}(\theta_e) \mu^*(\theta_e) d\theta_e, \mathbf{E}_{\bar{\gamma}}(\mu^{*2}) = \int_{\theta_e} \bar{\gamma}(\theta_e) \mu^{*2}(\theta_e) d\theta_e, \\ \mathbf{Var}_{\bar{\gamma}}(\mu^*) &= \left[\int_{\theta_e} \mu^{*2}(\theta_e) \bar{\gamma}(\theta_e) d\theta_e - \left(\int_{\theta_e} \bar{\gamma}(\theta_e) \mu^*(\theta_e) d\theta_e \right)^2 \right]. \end{aligned}$$

Then, we have

$$\Delta = \left(\frac{\sigma}{\sigma-1} - \xi \right)^2 \left[1 + \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right]^2 - 4 \frac{\sigma}{\sigma-1} \frac{\xi}{(\sigma-1)^2} \mathbf{Var}_{\bar{\gamma}}(\mu).$$

One necessary condition for the exist of solution to (A108) is $\Delta \geq 0$. Notice that $\mathbf{E}_{\bar{\gamma}}(\mu^{*2}) \leq \mathbf{E}_{\bar{\gamma}}(\mu^*)$. We have

$$\begin{aligned} (A109) \quad \Delta &= \left(\frac{\sigma}{\sigma-1} - \xi \right)^2 + \frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1} - \xi \right)^2 \mathbf{E}_{\bar{\gamma}}(\mu^*) + \frac{\left(\frac{\sigma}{\sigma-1} + \xi \right)^2}{(\sigma-1)^2} [\mathbf{E}_{\bar{\gamma}}(\mu^*)]^2 \\ &\quad - \frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1} \right)^2 \frac{2\xi}{\sigma} \mathbf{E}_{\bar{\gamma}}(\mu^{*2}) \\ &\geq \Delta_H(\mathbf{E}_{\bar{\gamma}}(\mu^*)), \end{aligned}$$

where we set

$$\begin{aligned}\Delta_H(\mathbf{E}_{\bar{\gamma}}(\mu^*)) &= \left(\frac{\sigma}{\sigma-1} - \xi\right)^2 + \frac{\left(\frac{\sigma}{\sigma-1} + \xi\right)^2}{(\sigma-1)^2} [\mathbf{E}_{\bar{\gamma}}(\mu^*)]^2 \\ &\quad - \frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1}\right)^2 \left[\frac{2\xi}{\sigma} - \left(1 - \frac{\sigma-1}{\sigma}\xi\right)^2\right] \mathbf{E}_{\bar{\gamma}}(\mu^*)\end{aligned}$$

as a quadratic function of $\mathbf{E}_{\bar{\gamma}}(\mu^*)$.

The minimum value of Δ_H is derived at $\mu^* \equiv \frac{(\sigma-1)\left[\frac{2\sigma-1}{\sigma^2}\xi^2 - (1-\xi)^2\right]}{(1+\frac{\sigma-1}{\sigma}\xi)^2} < 1$. However, $\mathbf{E}_{\bar{\gamma}}(\mu^*) \in (0, 1]$, and thus μ^* may not belong to the domain of $\mathbf{E}_{\bar{\gamma}}(\mu^*)$.

If $\mu^* \leq 0$, to prove that $\Delta \geq 0$, we only need to prove that $\Delta_H \geq 0$ when $\mathbf{E}_{\bar{\gamma}}(\mu^*) = 1$ and $\mathbf{E}_{\bar{\gamma}}(\mu^*) = 0$. This is true. According to (A109), when $\mathbf{E}_{\bar{\gamma}}(\mu^*) = 1$, we have

$$\begin{aligned}\Delta_H &= \left(\frac{\sigma}{\sigma-1} - \xi\right)^2 - \frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1}\right)^2 \left[\frac{2\xi}{\sigma} - \left(1 - \frac{\sigma-1}{\sigma}\xi\right)^2\right] + \frac{\left(\frac{\sigma}{\sigma-1} + \xi\right)^2}{(\sigma-1)^2} \\ &= \left(\frac{\sigma}{\sigma-1}\right)^2 \left(\frac{\sigma}{\sigma-1} - \xi\right)^2 - \left(\frac{\sigma}{\sigma-1}\right)^2 \frac{2}{\sigma-1} \frac{2\xi}{\sigma} + \left(\frac{\sigma}{\sigma-1}\right)^2 \frac{2}{\sigma-1} \frac{2\xi}{\sigma} \\ &= \left(\frac{\sigma}{\sigma-1}\right)^2 \left(\frac{\sigma}{\sigma-1} - \xi\right)^2 > 0.\end{aligned}$$

When $\mathbf{E}_{\bar{\gamma}}(\mu^*) = 0$, we have

$$\Delta_H = \left(\frac{\sigma}{\sigma-1} - \xi\right)^2 > 0.$$

If $\mu^* > 0$ (note that μ^* must be lower than one), to prove $\Delta \geq 0$, we only need to prove that $\Delta_H(\mu^*) \geq 0$. To see this, first note that when $\mu^* \in (0, 1)$, we have

$$\frac{2\sigma-1}{\sigma^2} > \left(\frac{1}{\xi} - 1\right)^2$$

and

$$\frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1}\right)^2 \left[\frac{2\xi}{\sigma} - \left(1 - \frac{\sigma-1}{\sigma}\xi\right)^2\right] > 0.$$

We set

$$\tilde{\Delta} = \left\{ \frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1}\right)^2 \left[\frac{2\xi}{\sigma} - \left(1 - \frac{\sigma-1}{\sigma}\xi\right)^2\right] \right\}^2 - 4 \left(\frac{\sigma}{\sigma-1} - \xi\right)^2 \frac{\left(\frac{\sigma}{\sigma-1} + \xi\right)^2}{(\sigma-1)^2}$$

as the discriminant of $\Delta_H(\mathbf{E}_{\bar{\gamma}}(\mu^*)) > 0$. $\tilde{\Delta} < 0$ is a sufficient condition for $\Delta > 0$. To prove $\tilde{\Delta} < 0$, we only need to show that

$$\frac{1}{\sigma-1} \left(\frac{\sigma}{\sigma-1}\right)^2 \left[\frac{2\xi}{\sigma} - \left(1 - \frac{\sigma-1}{\sigma}\xi\right)^2\right] < \frac{\left(\frac{\sigma}{\sigma-1} - \xi\right) \left(\frac{\sigma}{\sigma-1} + \xi\right)}{(\sigma-1)},$$

which is equivalent to

$$2 \left(\frac{\sigma}{\sigma-1}\right)^2 (1 - \xi) > 0.$$

Notice that the above inequality must hold. Thus, we must have $\Delta > 0$. In conclusion, there are two solutions to the quadratic equation (A108). However, $\Delta > 0$ does not necessarily mean that the solutions are all in the domain of μ , i.e., $(1, \frac{\sigma}{\sigma-1})$.

In the following analysis, we prove that there exists a unique solution in the domain of μ . The two potential solutions are μ_1 and μ_2 :

$$\begin{aligned}\mu_1 &= \frac{\left(\frac{\sigma}{\sigma-1} + \xi\right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e}{2\xi} - \frac{\sqrt{\Delta}}{2\xi} \geq \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e, \\ \mu_2 &= \frac{\left(\frac{\sigma}{\sigma-1} + \xi\right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e}{2\xi} + \frac{\sqrt{\Delta}}{2\xi} \leq \frac{\sigma}{\sigma-1} \frac{1}{\xi} \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e.\end{aligned}$$

In the following analysis, we prove that $\mu_2 > \frac{\sigma}{\sigma-1}$ and $\mu_1 \in (1, \frac{\sigma}{\sigma-1}]$. To prove this, we only need to show that

$$\sqrt{\Delta} > \left| 2\xi \frac{\sigma}{\sigma-1} - \left(\frac{\sigma}{\sigma-1} + \xi \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e \right|.$$

In particular, $\mu_2 > \frac{\sigma}{\sigma-1}$ is equivalent to

$$\mu_2 = \frac{\left(\frac{\sigma}{\sigma-1} + \xi\right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e}{2\xi} + \frac{\sqrt{\Delta}}{2\xi} > \frac{\sigma}{\sigma-1},$$

i.e.,

$$\sqrt{\Delta} > 2\xi \frac{\sigma}{\sigma-1} - \left(\frac{\sigma}{\sigma-1} + \xi \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e.$$

Substituting $\mu(\theta_e)$ in the above inequality by $\mu(\theta_e) = 1 + \frac{1}{\sigma-1} \mu^*(\theta_e)$, we have

$$\begin{aligned}\sqrt{\Delta} &> 2\xi \frac{\sigma}{\sigma-1} - \left(\frac{\sigma}{\sigma-1} + \xi \right) \left[1 + \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right] \\ &= \left(\frac{\sigma+1}{\sigma-1} \xi - \frac{\sigma}{\sigma-1} \right) - \left(\frac{\sigma}{\sigma-1} + \xi \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*).\end{aligned}$$

When $\xi \leq \frac{\sigma}{\sigma+1}$, the above inequality must holds. When $\xi > \frac{\sigma}{\sigma+1}$, to prove the above inequality, we only need to show

$$\Delta - \left[\left(\frac{\sigma+1}{\sigma-1} \xi - \frac{\sigma}{\sigma-1} \right) - \left(\frac{\sigma}{\sigma-1} + \xi \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right]^2 > 0.$$

To see this, notice that

$$\begin{aligned}\Delta_M &< \Delta - \left[\left(\frac{\sigma+1}{\sigma-1} \xi - \frac{\sigma}{\sigma-1} \right) - \left(\frac{\sigma}{\sigma-1} + \xi \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right]^2 \\ &= \left(\frac{\sigma}{\sigma-1} - \xi \right)^2 + \frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1} - \xi \right)^2 \mathbf{E}_{\bar{\gamma}}(\mu^*) \\ &\quad + \frac{\left(\frac{\sigma}{\sigma-1} + \xi\right)^2}{(\sigma-1)^2} [\mathbf{E}_{\bar{\gamma}}(\mu^*)]^2 - \frac{2}{\sigma-1} \left(\frac{\sigma}{\sigma-1} \right)^2 \frac{2\xi}{\sigma} \mathbf{E}_{\bar{\gamma}}(\mu^{*2}) \\ &\quad - \left[\left(\frac{\sigma+1}{\sigma-1} \xi - \frac{\sigma}{\sigma-1} \right) - \left(\frac{\sigma}{\sigma-1} + \xi \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right]^2,\end{aligned}$$

where

$$\Delta_M = \Delta_H(\mathbf{E}_{\bar{\gamma}}(\mu^*)) - \left[\left(\frac{\sigma+1}{\sigma-1} \xi - \frac{\sigma}{\sigma-1} \right) - \left(\frac{\sigma}{\sigma-1} + \xi \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right]^2.$$

Rearranging the right side the above equation, we have

$$\Delta_M = \frac{2}{\sigma-1} \frac{2\sigma}{\sigma-1} \xi (1-\xi) [1 - \mathbf{E}_{\bar{\gamma}}(\mu^*)] \geq 0.$$

Since

$$\Delta - \left[\left(\frac{\sigma+1}{\sigma-1} \zeta - \frac{\sigma}{\sigma-1} \right) - \left(\frac{\sigma}{\sigma-1} + \zeta \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right]^2 > \Delta_M > 0,$$

we have

$$\mu_2 = \frac{\sigma}{\sigma-1} \frac{1}{\zeta} \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e > \frac{\sigma}{\sigma-1}.$$

We now prove that

$$\frac{\sigma}{\sigma-1} \geq \mu_1 = \frac{\left(\frac{\sigma}{\sigma-1} + \zeta \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e}{2\zeta} - \frac{\sqrt{\Delta}}{2\zeta} > 1$$

First, $\mu_1 \geq \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e$, where $\mu(\theta_e) \in (1, \frac{\sigma}{\sigma-1}]$ and $\bar{\gamma}(\theta_e)$ is a density with support Θ_e . Thus, $\mu_1 > 1$. Second, notice that

$$\Delta - \left[\left(\frac{\sigma+1}{\sigma-1} \zeta - \frac{\sigma}{\sigma-1} \right) - \left(\frac{\sigma}{\sigma-1} + \zeta \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right]^2 \geq 0,$$

(see, e.g., (A109)) we have

$$\begin{aligned} \sqrt{\Delta} &\geq \left(\frac{\sigma}{\sigma-1} + \zeta \right) \left[1 + \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) \right] \left(\frac{\sigma}{\sigma-1} + \zeta \right) - \frac{2\zeta\sigma}{\sigma-1} \\ &= \left(\frac{\sigma}{\sigma-1} + \zeta \right) \frac{1}{\sigma-1} \mathbf{E}_{\bar{\gamma}}(\mu^*) - \left(\frac{\sigma+1}{\sigma-1} \zeta - \frac{\sigma}{\sigma-1} \right) \end{aligned}$$

The above inequality implies

$$\sqrt{\Delta} \geq \left(\frac{\sigma}{\sigma-1} + \zeta \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e - \frac{2\zeta\sigma}{\sigma-1},$$

which gives

$$\mu_1 = \frac{\left(\frac{\sigma}{\sigma-1} + \zeta \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e}{2\zeta} - \frac{\sqrt{\Delta}}{2\zeta} \leq \frac{\sigma}{\sigma-1}.$$

In conclusion, we have

$$\mu = \frac{\left(\frac{\sigma}{\sigma-1} + \zeta \right) \int_{\theta_e} \bar{\gamma}(\theta_e) \mu(\theta_e) d\theta_e}{2\zeta} - \frac{\sqrt{\Delta}}{2\zeta}.$$

■

Corollary A1 provides a special case where we can provide explicit optimal tax formulas with given social welfare weights. Note that $\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \zeta \right) = 2$ is consistent with condition (22). Corollary A1 thus suggests that under some reasonable parameter values, there is a unique and well-defined solution to the equation system (32) and (35).

3.2. Statistic-based Optimal Tax Formulas

In the following Theorem A1, we provide optimal tax rules in terms of social welfare weights and a small number of empirical observable statistics (see Supplemental Appendix I.1.2 for the definitions of elasticities). Recall that under Assumption 1, we can define $F_{y_o}(y_o(\theta_o)) = F_o(\theta_o)$ and $f_{y_o}(y_o(\theta_o)) = F'_{y_o}(y_o(\theta_o))$ as the CDF and PDF of entrepreneurial and labor income.

THEOREM A1: Suppose the cross wage elasticities $\varepsilon_{L_w}^\varphi(\theta'_e, \theta_e)$ and $\varepsilon_{L_e}^\omega(\theta'_e, \theta_e)$ are independent θ'_e for any $(\theta'_e, \theta_e) \in \Theta_e^2$; the markup $\mu(\theta_e)$ is exogenous in the equilibrium; and the firm-level production technology is in the form of (1). The optimal tax wedges satisfy the following statistic-based tax formula for any $\theta_w \in \Theta_w$ and $\theta_e \in \Theta_e$:

$$(A110) \quad \frac{1}{1 - \tau_w(\theta_w)} = \frac{1 + [1 - \bar{g}_w(\theta_w)] \frac{1 - F_{y_w}(y_w(\theta_w))}{y_w(\theta_w) f_{y_w}(y_w(\theta_w))} \frac{1}{\varepsilon_{1-\tau_w}^{y_w}(\theta_e)}}{\mu}.$$

and

$$(A111) \quad \frac{1}{1 - \tau_e(\theta_e)} = \left[\frac{1}{\mu(\theta_e)} \left[1 + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_{y_e}(y_e(\theta_e))}{y_e(\theta_e) f_{y_e}(y_e(\theta_e))} \frac{1}{\bar{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)} \frac{\frac{1+\varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \zeta] - 1}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1} \right] + \right. \\ \left. \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e) \left[1 - \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)} \right] \left[1 - g_e(\theta_e) - \frac{[1 - \bar{g}_e(\theta_e)] \frac{1 - F_{y_e}(y_e(\theta_e))}{y_e(\theta_e) f_{y_e}(y_e(\theta_e))} \frac{1}{\bar{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1} \right] \right] \\ \times \frac{1}{1 + \left[\frac{\mu}{\mu(\theta_e)} - 1 \right] \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)}}.$$

PROOF:

Notice that (A50) and $\frac{d \ln[1 - \tau_e(\theta_e)]}{d \theta_e} \frac{d \ln y_e(\theta_e)}{d \theta_e} = -\frac{T_e''(y_e(\theta_e)) y_e(\theta_e)}{1 - T_e'(y_e(\theta_e))}$, we have:

$$H(\theta_e) = \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{d \ln y_e(\theta_e)}{d \theta_e} - \frac{\varepsilon_{Q_{ij}}^S(\theta_e) \frac{d \ln[1 - \tau_e(\theta_e)]}{d \theta_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[1 - \zeta \varepsilon_{Q_{ij}}^S(\theta_e) \right] - \varepsilon_{Q_{ij}}^S(\theta_e)} \right] \\ = \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln y_e(\theta_e)}{d \theta_e} \frac{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - \left[1 - \frac{T_e''(y_e(\theta_e)) y_e(\theta_e)}{1 - T_e'(y_e(\theta_e))} \right]}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1} \\ = \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln y_e(\theta_e)}{d \theta_e} \frac{1}{\bar{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)} \frac{1}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1}.$$

Therefore, (32) is equivalent to:

$$\frac{1}{1 - \tau_e(\theta_e)} = \frac{\frac{1}{\mu(\theta_e)} \left[1 + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_{y_e}(y_e(\theta_e))}{y_e(\theta_e) f_{y_e}(y_e(\theta_e))} \frac{1}{\bar{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)} \frac{\frac{1+\varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \zeta] - 1}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1} \right]}{1 + RE(\theta_e) \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)}} \\ + \frac{\left[1 - \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)} \right] \varepsilon_{Q_{-ij}}^{P,cross}(\theta_e) \left[\frac{[1 - g_e(\theta_e)] - \frac{1 - F_{y_e}(y_e(\theta_e))}{y_e(\theta_e) f_{y_e}(y_e(\theta_e))} \frac{1}{\bar{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)}}{[1 - \bar{g}_e(\theta_e)] \frac{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1}} \right]}{1 + RE(\theta_e) \zeta \frac{\varepsilon_{l_e}^\omega(\theta_e)}{\varepsilon_{Lw}^\omega(\theta_e)}},$$

where we use $\varepsilon_{l_e}^{Q_{ij}}(\theta_e) = 1$ and $\varepsilon_{Lw}^{Q_{ij}}(\theta_e) = \zeta$.

Notice that $\frac{\frac{1+\varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \zeta] - 1}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1} = 1$ when $I = 1$, and $\frac{\frac{1+\varepsilon_e}{\varepsilon_e} [\mu(\theta_e) - \zeta] - 1}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1} < 1$ when $I > 1$. When tax is linear $\frac{1}{\bar{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)} = \frac{1}{\frac{1+\varepsilon_e}{\varepsilon_e} \left[\frac{1}{\varepsilon_{Q_{ij}}^S(\theta_e)} - \zeta \right] - 1}$.

Last, (A110) is equivalent to (31) because by the definition of $\varepsilon_{1-\tau_w}^{y_w}(\theta_w)$, one has $\frac{1 - F_{y_w}(y_w(\theta_w))}{y_w(\theta_w) f_{y_w}(y_w(\theta_w))} \frac{1}{\bar{\varepsilon}_{1-\tau_w}^{y_w}(\theta_w)} = \frac{1 - F_w(\theta_w)}{f_w(\theta_w)} \frac{x'_w(\theta_w)}{x_w(\theta_w)} \frac{1 + \varepsilon_w}{\varepsilon_w}$. ■

Equations (A110) and (A111) are statistic-based formulas in the the sense that the statistics on the right side can be observed at each income or profit level. Theorem A1 shows in addition to the social welfare weights, income distributions, and elasticities of income, the statistic-based optimal tax formulas are related

to the markups. When the market is competitive, such that $\mu(\theta_e) = \mu = 1$ and $\varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) = 0$, (A111) is reduced to the traditional optimal statistic-based tax formula:

$$\frac{1}{1 - \tau_e(\theta_e)} = 1 + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_{y_e}(y_e(\theta_e))}{y_e(\theta_e) f_{y_e}(y_e(\theta_e))} \frac{1}{\tilde{\varepsilon}_{1-\tau_e}^{y_e}(\theta_e)}$$

However, in general cases, the statistic-based optimal profit tax in an economy with market power is different from the traditional one (see, e.g., [Piketty \(1997\)](#); [Diamond \(1998\)](#); [Saez \(2001\)](#)). Moreover, the statistic-based optimal profit tax formula is quite different from the statistic-based optimal income tax formula, unless the economy is competitive or monopoly-competitive.

3.3. Proof of Theorem 2

(I) NON-LINEAR SALES TAXES.

Before we prove Theorem 2, we present the second-best problem formally. In our benchmark model, we consider an environment with uniform linear sales tax, which restricts $\tau_s(\theta_e)$ to be constant. In this section, we remove this policy constraint and allow for non-linear sales tax as considered by [Ales, Bellofatto and Wang \(2017\)](#). To do this, we allow the planner to contract with entrepreneurs on sales income $S(\theta_e) \equiv P(\theta_e) Q_{ij}(\theta_e)$ in addition to $y_e(\theta_e)$. An entrepreneur reporting θ'_e should obtain $y_e(\theta'_e)$ in profit, $S(\theta'_e)$ in sales income, and receive $c_e(\theta'_e)$ in after-tax profit. Thus, a θ_e -type entrepreneur reporting θ'_e should choose L_w and l_e to satisfy the following two promise-keeping constraints:

$$P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) Q_{ij} = S(\theta'_e) \quad \text{and} \quad P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) Q_{ij} - WL_w = y_e(\theta'_e).$$

The two promise-keeping constraints determine the combination of L_w and l_e that are needed to complete the tasks of sales income and profit. Denote $L_w(\theta'_e|\theta_e)$ and $l_e(\theta'_e|\theta_e)$ as the labor input and effort needed to complete the tasks, respectively. Denote $Q_{ij}(x_e(\theta_e) l_e, L_w)$ as the firm-level production function. Combining these two constraints we immediately have

$$L_w(\theta'_e|\theta_e) = \frac{S(\theta'_e) - y_e(\theta'_e)}{W},$$

and $l_e(\theta'_e|\theta_e)$ satisfies

$$P\left(Q_{ij}\left(x_e(\theta_e) l_e(\theta'_e|\theta_e), \frac{S(\theta'_e) - y_e(\theta'_e)}{W}\right), Q_{-ij}(\theta_e), \theta_e\right) Q_{ij}\left(x_e(\theta_e) l_e(\theta'_e|\theta_e), \frac{S(\theta'_e) - y_e(\theta'_e)}{W}\right) = S(\theta'_e).$$

Two things worth noting here. First, by enforcing tasks of sales income and profit, the planner can directly control the amount of labor inputs. That is to say, $L_w(\theta'_e|\theta_e)$ is independent of θ_e . Second, notice that $\frac{\partial [P(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) Q_{ij}]}{\partial Q_{ij}} = P(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) [1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e)]$ and $\varepsilon_{Q_{ij}}^{P, own}(\theta_e) > -1$. As long as Q_{ij} strictly increases in $l_e(\theta'_e|\theta_e)$ with $P(Q_{ij}(0, \cdot), Q_{-ij}(\theta_e), \theta_e) Q_{ij}(0, \cdot) = 0$, there exists a unique solution $l_e(\theta'_e|\theta_e)$ for any $S(\theta'_e) \geq 0$.

Therefore, under our setup, we can reformulate the entrepreneur's problem as

$$(A112) \quad V_e(\theta_e) \equiv \max_{\theta'_e} c_e(\theta'_e) - \phi_e(l_e(\theta'_e|\theta_e)).$$

Solving the above problem, as in the benchmark model, we have

$$(A113) \quad \frac{\partial l_e(\theta'_e|\theta_e)}{\partial \theta_e} = - \frac{\frac{\partial P(Q_{ij}, \theta_e)}{\partial \theta_e} Q_{ij} + P(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) [1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e)] \frac{\partial Q_{ij}}{\partial x_e(\theta_e)} x'_e(\theta_e)}{P(Q_{ij}, Q_{-ij}(\theta_e), \theta_e) [1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e)] \frac{\partial Q_{ij}}{\partial l_e(\theta'_e|\theta_e)}}$$

and

$$(A114) \quad V'_e(\theta_e) = \frac{\phi'_e(l_e(\theta_e)) Q_{ij}(\theta_e)}{\frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}} \left[\frac{\frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d \theta_e} \big|_{Q_{ij}=Q_{ij}(\theta_e)}}{1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e)} + \frac{\frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln x_e(\theta_e)} \frac{x'_e(\theta_e)}{x_e(\theta_e)}}{\frac{\partial \ln x_e(\theta_e)}{x_e(\theta_e)}} \right]$$

which is equivalent to the benchmark incentive-compatible constraint (27), because in the benchmark model $\mu(\theta) = \frac{1}{1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e)} \cdot \frac{Q_{ij}(\theta_e)}{\frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}} = l_e(\theta_e)$ and $\frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln x_e(\theta_e)} x'_e(\theta_e) = \frac{x'_e(\theta_e)}{x_e(\theta_e)}$.

The worker's problem remains the same as before. Therefore, all incentive-compatible allocations satisfying (24), (27) and the resource constraints are feasible. The planner's problem is similar to the one in the benchmark model, except that the policy constraint $\frac{d \ln \omega(\theta_e)}{d \theta_e} = 0$ is now relaxed.

(II) LAGRANGIAN PROBLEM.

We prove Theorem 2 by following the Lagrangian problem presented in Supplemental Appendix II.2.3.

Note that the expression of markup (19) is now generalized to be $\mu(\theta_e) = \frac{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)} [1 - \tau_s(\theta_e)]}{W}$. By the definition of $\omega(\theta_e)$ (see, e.g., (28)), we have $\omega(\theta_e) = P(\theta_e) \left[1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e) \right] \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)}$. Notice that $1 - \tau_s(\theta_e) = \frac{W}{\omega(\theta_e)}$, as in the benchmark model, we have $\mu(\theta_e) = \frac{1}{1 + \varepsilon_{Q_{ij}}^{P, own}(\theta_e)}$.

(i) When the uniform restriction on $\omega(\theta_e, \theta_e l_e(\theta_e), L_w(\theta_e), Q)$ is loosened and $\varphi(\theta_e) = 0$. In this case, according to the expression of $\frac{\partial \mathcal{L}}{\partial L_w(\theta_e)}$ (e.g., (A76)), we have

$$(A115) \quad \begin{aligned} P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)} &= \frac{\lambda'}{\lambda} - \frac{\kappa'(\theta_e)}{\lambda L_w(\theta_e) N_e f_e(\theta_e)} \frac{\partial \ln Q_{ij}(\theta_e)}{\partial \ln L_w(\theta_e)} \\ &= \frac{\lambda'}{\lambda} - \frac{\kappa'(\theta_e) \xi}{\lambda L_w(\theta_e) N_e f_e(\theta_e)}. \end{aligned}$$

Dividing both sides by $\frac{\varepsilon_{L_w}^{\omega}(\theta_e)}{L_w(\theta_e) N_e f_e(\theta_e)}$ and integrating across θ_e gives

$$\int_{\theta_e} P(\theta_e) Q_{ij}(\theta_e) \xi \frac{N_e f_e(\theta_e)}{\varepsilon_{L_w}^{\omega}(\theta_e)} d\theta_e = \frac{\lambda'}{\lambda} \int_{\theta_e} \frac{L_w(\theta_e) N_e f_e(\theta_e)}{\varepsilon_{L_w}^{\omega}(\theta_e)} d\theta_e - \int_{\theta_e} \frac{\kappa'(\theta_e)}{\lambda \varepsilon_{L_w}^{\omega}(\theta_e)} \xi d\theta_e,$$

Using $\kappa(\underline{\theta_e}) = \kappa(\bar{\theta_e}) = 0$ and integration by parts, we have

$$(A116) \quad \begin{aligned} \frac{\lambda'}{\lambda W} &= \frac{\xi \int_{\theta_e} P(\theta_e) Q_{ij}(\theta_e) N_e f_e(\theta_e) d\theta_e}{W \int_{\theta_e} L_w(\theta_e) N_e f_e(\theta_e) d\theta_e} \\ &= \frac{\xi \int_{\theta_e} \frac{P(\theta_e) Q_{ij}(\theta_e)}{W L_w(\theta_e)} W L_w(\theta_e) N_e f_e(\theta_e) d\theta_e}{\int_{\theta_e} W L_w(\theta_e) N_e f_e(\theta_e) d\theta_e} \\ &= \frac{\xi \int_{\theta_e} \frac{\mu(\theta_e)}{\xi [1 - \tau_s^E(\theta_e)]} W L_w(\theta_e) N_e f_e(\theta_e) d\theta_e}{\int_{\theta_e} W L_w(\theta_e) N_e f_e(\theta_e) d\theta_e} \\ &= \mu^*. \end{aligned}$$

where the third equation is derived from the generalized definition of markup (i.e., $\mu(\theta_e) = \frac{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)} [1 - \tau_s(\theta_e)]}{W}$).

By the definition of tax wedges, we can substitute $P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)}$ in (A115) with $\frac{W\mu(\theta_e)}{1-\tau_s^E(\theta_e)}$:

$$\begin{aligned}
\frac{1}{1-\tau_s^E(\theta_e)} &= \frac{\lambda'}{\lambda W\mu(\theta_e)} - \frac{\kappa'(\theta_e) \zeta}{\lambda L_w(\theta_e) N_e f_e(\theta_e) W\mu(\theta_e)} \\
&= \frac{\mu^*}{\mu(\theta_e)} - \frac{\kappa'(\theta_e) \zeta}{\lambda L_w(\theta_e) N_e f_e(\theta_e) W\mu(\theta_e)} \\
&= \frac{\mu^*}{\mu(\theta_e)} + \frac{P(\theta_e) Q_{ij}(\theta_e) [1-\tau_s^E(\theta_e)] \zeta [1-\tau_e^E(\theta_e)] \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left[\begin{aligned} &\psi_e(\theta_e) \frac{1+\varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \psi'_e(\theta_e) \\ &+ \psi_e(\theta_e) \frac{d \ln [\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e} \end{aligned} \right]}{L_w(\theta_e) W\mu(\theta_e) \lambda N_e f_e(\theta_e)} \\
&= \frac{\mu^*}{\mu(\theta_e)} + [1-\tau_e^E(\theta_e)] \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left[\frac{[1-\bar{g}_e(\theta_e)][1-F_e(\theta_e)]}{f_e(\theta_e)} \left[\frac{1+\varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{d \ln [\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e} \right] - \right],
\end{aligned}$$

The third equation is derived by the modified (A85):

$$\begin{aligned}
\kappa'(\theta_e) &= - \frac{d \left[\psi_e(\theta_e) \phi'_e(l_e(\theta_e)) l_e(\theta_e) \hat{\mu}(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \right]}{d\theta_e} \\
&= - \left[\begin{aligned} &\psi'_e(\theta_e) \phi'_e(l_e(\theta_e)) l_e(\theta_e) \hat{\mu}(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) + \\ &\psi_e(\theta_e) \phi'_e(l_e(\theta_e)) \frac{1+\varepsilon_e}{\varepsilon_e} l'_e(\theta_e) \hat{\mu}(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) + \\ &\psi_e(\theta_e) \phi'_e(l_e(\theta_e)) l_e(\theta_e) \frac{d \ln [\hat{\mu}(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e} \end{aligned} \right] \\
&= - \phi'_e(l_e(\theta_e)) l_e(\theta_e) \hat{\mu}(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left[\begin{aligned} &\psi_e(\theta_e) \frac{1+\varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \\ &\psi'_e(\theta_e) + \psi_e(\theta_e) \frac{d \ln [\hat{\mu}(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e} \end{aligned} \right]
\end{aligned}$$

and the tax wedge, $1-\tau_e^E(\theta_e) = \frac{\phi'_e(l_e(\theta_e))}{\frac{P(\theta_e)}{\mu(\theta_e)} \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)} [1-\tau_s^E(\theta_e)]}$. These two equations implies

$$\kappa'(\theta_e) = -P(\theta_e) Q_{ij}(\theta_e) [1-\tau_e^E(\theta_e)] [1-\tau_s^E(\theta_e)] \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left[\begin{aligned} &\psi_e(\theta_e) \frac{1+\varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \\ &\psi'_e(\theta_e) + \psi_e(\theta_e) \frac{d \ln [\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e} \end{aligned} \right].$$

The last equation, i.e.,

$$\begin{aligned}
\text{(A117)} \quad \frac{\tau_s^E(\theta_e)}{1-\tau_s^E(\theta_e)} &= \overbrace{\left[\frac{\mu^*}{\mu(\theta_e)} - 1 \right]}^{\widetilde{RE}(\theta_e)} + [1-\tau_e^E(\theta_e)] \varepsilon_{Q-ij}^{P,cross}(\theta_e) \overbrace{\left[\begin{aligned} &[1-\bar{g}_e(\theta_e)] - \frac{[1-\bar{g}_e(\theta_e)][1-F_e(\theta_e)]}{f_e(\theta_e)} \\ &\times \left[\frac{1+\varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{d \ln [\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)]}{d\theta_e} \right] \end{aligned} \right]}^{\widehat{IRE}(\theta_e)}
\end{aligned}$$

is derived by (A81), (A82) and $1-\tau_s^E(\theta_e) = \frac{W\mu(\theta_e)}{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)}} = \frac{WL_w(\theta_e)\mu(\theta_e)}{\bar{\zeta}P(\theta_e)Q_{ij}(\theta_e)}$.

(ii) According to the expression of $\frac{\partial \mathcal{L}}{\partial l_w(\theta_w)}$ (e.g., (A75)), we have

$$\frac{1}{\frac{\phi'_w(l_w(\theta_w))}{x_w(\theta_w)}} = \frac{\lambda}{\lambda'} \left[1 - \frac{x'_w(\theta_w)}{x_w(\theta_w)} \frac{\psi_w(\theta_w)}{\lambda N_w f_w(\theta_w)} \frac{1+\varepsilon_w}{\varepsilon_w} \right].$$

Substitute $\frac{\phi'_w(l_w(\theta_w))}{x_w(\theta_w)}$ by $[1 - \tau_w^E(\theta_w)] W$:

$$(A118) \quad \frac{1}{1 - \tau_w^E(\theta_w)} = \frac{1}{\mu^*} \left[1 + [1 - \bar{g}_w(\theta_w)] \frac{1 - F_w(\theta_w)}{f_w(\theta_w)} \frac{x'_w(\theta_w)}{x_w(\theta_w)} \frac{1 + \varepsilon_w}{\varepsilon_w} \right].$$

(iii) According to the expression of $\frac{\partial \mathcal{L}}{\partial l_e(\theta_e)}$ (e.g., (A77)), we have:

$$\begin{aligned} \frac{\tilde{\tau}_e(\theta_e)}{1 - \tilde{\tau}_e(\theta_e)} &= [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\partial \ln P(Q(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\ &\quad - \frac{1}{1 - \tilde{\tau}_e(\theta_e)} \frac{WL_w(\theta_e)}{P(\theta_e) Q_{ij}(\theta_e)} \frac{1}{\xi} \frac{\kappa'(\theta_e) \xi}{\lambda N_e f_e(\theta_e) WL_w(\theta_e)}. \end{aligned}$$

Remember that $1 - \tilde{\tau}_e(\theta_e) = \frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}} = \frac{[1 - \tau_e^E(\theta_e)][1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)}$. Using $\frac{\mu(\theta_e)}{1 - \tau_s^E(\theta_e)} = \mu^* - \frac{\kappa'(\theta_e) \xi}{\lambda WL_w(\theta_e) N_e f_e(\theta_e)}$ to substitute $\frac{\kappa'(\theta_e) \xi}{\lambda WL_w(\theta_e) N_e f_e(\theta_e)}$ and $\frac{\xi [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)}$ to substitute $\frac{WL_w(\theta_e)}{P(\theta_e) Q_{ij}(\theta_e)}$, we have:

$$\begin{aligned} \frac{1}{1 - \tilde{\tau}_e(\theta_e)} &= 1 + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\partial \ln P(Q(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\ &\quad + \frac{1}{1 - \tilde{\tau}_e(\theta_e)} \left[1 - \mu^* \frac{1 - \tau_s^E(\theta_e)}{\mu(\theta_e)} \right]. \end{aligned}$$

Lastly, substitute $1 - \tilde{\tau}_e(\theta_e)$ by $\frac{[1 - \tau_e^E(\theta_e)][1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)}$ and rearrange:

$$(A119) \quad \frac{\tau_e^E(\theta_e)}{1 - \tau_e^E(\theta_e)} = \frac{1 - \mu^* + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\partial \ln P(Q(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]}{\mu^*}.$$

(III) PART ONE: AN EXPLICIT EXPRESSION.

According to the definitions of tax wedges, we have

$$(A120) \quad \frac{\xi [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)} P(\theta_e) Q(\theta_e) = WL_w(\theta_e)$$

and

$$(A121) \quad \frac{P(\theta_e) Q_{ij}(\theta_e) [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)} [1 - \tau_e^E(\theta_e)] = l_e(\theta_e)^{1 + \frac{1}{\varepsilon_e}}.$$

Substitute $P(\theta_e)$ and $Q(\theta_e)$ in (A120) by (A1) and $Q_{ij}(\theta_e) = x_e(\theta_e) l_e(\theta_e) L_w(\theta_e)^\xi$:

$$\left([1 - \tau_s^E(\theta_e)] \frac{\chi(\theta_e) x_e(\theta_e)^{1 - \frac{1}{\sigma}}}{\mu(\theta_e)} \right) \frac{\xi A^{\frac{\sigma-1}{\sigma}} Q^{\frac{1}{\sigma}}}{N_e^{\frac{1}{\sigma}} W} l_e(\theta_e)^{1 - \frac{1}{\sigma}} L_w(\theta_e)^{\xi \frac{\sigma-1}{\sigma}} = L_w(\theta_e),$$

Rearrange the above equation to derive an expression for $L_w(\theta_e)$:

$$(A122) \quad L_w(\theta_e) = \left(\frac{X_e(\theta_e) [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)} \right)^{\frac{1}{1 - \xi \frac{\sigma-1}{\sigma}}} l_e(\theta_e)^{\frac{1 - \frac{1}{\sigma}}{1 - \xi \frac{\sigma-1}{\sigma}}} \left(\frac{\xi A^{\frac{\sigma-1}{\sigma}} Q^{\frac{1}{\sigma}}}{W N_e^{\frac{1}{\sigma}}} \right)^{\frac{1}{1 - \xi \frac{\sigma-1}{\sigma}}}.$$

On the other hand, a combination of the two first-order conditions delivers:

$$\frac{WL_w(\theta_e)}{\xi} \left[1 - \tau_e^E(\theta_e) \right] = l_e(\theta_e)^{1+\frac{1}{\epsilon_e}}.$$

The above two equations imply:

$$L_w(\theta_e) = \left(\frac{X_e(\theta_e) [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)} \right)^{\frac{1}{1-\xi\frac{\sigma-1}{\sigma}}} \left[\frac{WL_w(\theta_e)}{\xi} [1 - \tau_e^E(\theta_e)] \right]^{\frac{1-\frac{1}{\sigma}}{1-\xi\frac{\sigma-1}{\sigma}} \frac{\epsilon_e}{1+\epsilon_e}} \left(\frac{\xi A^{\frac{\sigma-1}{\sigma}} Q^{\frac{1}{\sigma}}}{WN_e^{\frac{1}{\sigma}}} \right)^{\frac{1}{1-\xi\frac{\sigma-1}{\sigma}}}$$

i.e.,

$$L_w(\theta_e) = \left[1 - \tau_s^E(\theta_e) \right]^{\frac{1+\frac{\sigma-1}{\sigma} \frac{\epsilon_e}{1+\epsilon_e}}{1-\frac{\sigma-1}{\sigma} \left(\frac{\epsilon_e}{1+\epsilon_e} + \xi \right)}} \left(\frac{X_e(\theta_e)}{\mu(\theta_e)} \right)^{\frac{1}{1-\frac{\sigma-1}{\sigma} \left(\frac{\epsilon_e}{1+\epsilon_e} + \xi \right)}} \left[\frac{W}{\xi} \right]^{\frac{\frac{\sigma-1}{\sigma} \frac{\epsilon_e}{1+\epsilon_e}}{1-\frac{\sigma-1}{\sigma} \left(\frac{\epsilon_e}{1+\epsilon_e} + \xi \right)}} \left(\frac{\xi A^{\frac{\sigma-1}{\sigma}} Q^{\frac{1}{\sigma}}}{WN_e^{\frac{1}{\sigma}}} \right)^{\frac{\frac{1-\frac{1}{\sigma}}{1-\xi\frac{\sigma-1}{\sigma}} \frac{\epsilon_e}{1+\epsilon_e}}{1-\frac{\sigma-1}{\sigma} \left(\frac{\epsilon_e}{1+\epsilon_e} + \xi \right)}}.$$

Combining the above two equations, we can solve for $L_w(\theta_e)$ and $l_e(\theta_e)$ and derive:

$$(A123) \quad \frac{d \ln L_w(\theta_e)}{d\theta_e} = \frac{\frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e) [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)} \right]}{1 - \xi \frac{\sigma-1}{\sigma}} + \frac{\frac{\sigma-1}{\sigma} \frac{l'_e(\theta_e)}{l_e(\theta_e)}}{1 - \xi \frac{\sigma-1}{\sigma}}$$

and

$$(A124) \quad \frac{d \ln l_e(\theta_e)}{d\theta_e} = \frac{\frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e) [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)} \right] + [1 - \xi \frac{\sigma-1}{\sigma}] \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e}}{\frac{1+\epsilon_e}{\epsilon_e} - \frac{\sigma-1}{\sigma} \left(1 + \frac{1+\epsilon_e}{\epsilon_e} \xi \right)}.$$

Comparing the above two equations to (A43) and (A45), one can see that they are the same except that now $X_e(\theta_e)$ is replaced by $X_e(\theta_e) [1 - \tau_s^E(\theta_e)]$. In addition, we have:

$$\begin{aligned} & \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\ &= \mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \frac{d \ln Q_{ij}(\theta_e)}{d\theta_e} \\ &= \mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \left[\frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \xi \frac{L'_w(\theta_e)}{L_w(\theta_e)} \right] \\ &= \mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[1 + \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e) \right] \frac{x'_e(\theta_e)}{x_e(\theta_e)} + \frac{\mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)}{1 - \xi \frac{\sigma-1}{\sigma}} \frac{d \ln l_e(\theta_e)}{d\theta_e} \\ & \quad + \frac{\xi \mu(\theta_e) \varepsilon_{Q_{-ij}}^{P, cross}(\theta_e)}{1 - \xi \frac{\sigma-1}{\sigma}} \frac{d \left[\ln \frac{X_e(\theta_e)}{\mu(\theta_e)} + \ln [1 - \tau_s^E(\theta_e)] \right]}{d\theta_e}, \end{aligned}$$

where the third equation is derived by (A123) and

$$\begin{aligned}
& \mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[1 + \mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e)\right] \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\
&= \mu(\theta_e) \left[\frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{\sigma-1}{\sigma} \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\
&= \mu(\theta_e) \frac{d \ln \frac{x_e(\theta_e)}{\mu(\theta_e)}}{d\theta_e} + \frac{d\mu(\theta_e)}{d\theta_e}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& \mu(\theta_e) \frac{d \ln P_{ij}(Q_{ij}, Q_{-ij}(\theta_e), \theta_e)}{d\theta_e} \Big|_{Q_{ij}=Q_{ij}(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \\
&= \frac{\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) d \ln l_e(\theta_e)}{1 - \zeta \frac{\sigma-1}{\sigma}} + \frac{\mu(\theta_e) - \zeta d \ln \left[\frac{x_e(\theta_e)}{\mu(\theta_e)} \right]}{1 - \zeta \frac{\sigma-1}{\sigma}} + \frac{\zeta \mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) d \ln [1 - \tau_s^E(\theta_e)]}{1 - \zeta \frac{\sigma-1}{\sigma}} + \frac{d\mu(\theta_e)}{d\theta_e} \\
&= \left[(\mu(\theta_e) - \zeta) \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right] \frac{\frac{d \ln [X_e(\theta_e)/\mu(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \frac{\sigma-1}{\sigma} \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)} \\
&\quad + \mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \frac{\left(1 + \zeta \frac{1 + \varepsilon_e}{\varepsilon_e} \right) \frac{d \ln [1 - \tau_s^E(\theta_e)]}{d\theta_e} + \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} - \frac{\sigma-1}{\sigma} \left(1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta \right)} + \frac{d\mu(\theta_e)}{d\theta_e},
\end{aligned}$$

where the second equation is derived by (A124). In addition, we have

$$\begin{aligned}
& \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\partial \ln P(Q(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] \\
&= \left[(\mu(\theta_e) - \zeta) \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right] H(\theta_e) + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \\
&\quad + \mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \frac{\left(1 + \zeta \frac{1 + \varepsilon_e}{\varepsilon_e} \right) \frac{d \ln [1 - \tau_s^E(\theta_e)]}{d\theta_e} + \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} \left(1 - \frac{\sigma-1}{\sigma} \zeta \right) - \frac{\sigma-1}{\sigma}}.
\end{aligned}$$

Substituting $\frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\partial \ln P(Q(\theta_e), \theta_e)}{\partial \theta_e} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right]$ and $\frac{d \ln l_e(\theta_e)}{d\theta_e}$ in (A119) and (A117) by the above equation and (A124), respectively, and rearrange the formulas, we have

$$\frac{1}{1 - \tau_e^E(\theta_e)} = \frac{1 + [1 - \bar{g}_e(\theta_e)] \left\{ \begin{aligned} & H(\theta_e) \left[(\mu(\theta_e) - \zeta) \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right] + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} + \\ & \mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{\left(1 + \zeta \frac{1 + \varepsilon_e}{\varepsilon_e} \right) \frac{d \ln [1 - \tau_s^E(\theta_e)]}{d\theta_e} + \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e}}{1 - \frac{\sigma-1}{\sigma} \left(\frac{\varepsilon_e}{1 + \varepsilon_e} + \zeta \right)} \end{aligned} \right\}}{\mu^*}$$

and

$$\frac{1}{1 - \tau_s^E(\theta_e)} = \frac{\mu^*}{\mu(\theta_e)} + [1 - \tau_e^E(\theta_e)] \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left[\begin{aligned} & \frac{[1 - g_e(\theta_e)] - \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \times}{\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{\frac{d \ln \left[\frac{x_e(\theta_e)}{\mu(\theta_e)} \right]}{d\theta_e} + \left[\frac{d \ln [1 - \tau_s^E(\theta_e)]}{d\theta_e} + \left(1 - \zeta \frac{\sigma-1}{\sigma} \right) \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e} \right]}{\frac{1 + \varepsilon_e}{\varepsilon_e} \left(1 - \zeta \frac{\sigma-1}{\sigma} \right) - \frac{\sigma-1}{\sigma}} + \frac{d \ln \left[\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \right]}{d\theta_e} \end{aligned} \right]$$

i.e.,

$$\frac{1}{1 - \tau_s^E(\theta_e)} = \frac{\mu^*}{\mu(\theta_e)} + \left[1 - \tau_e^E(\theta_e)\right] \varepsilon_{Q-ij}^{P,cross}(\theta_e) \left[\frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \left[\frac{d \ln \left[\mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \right]}{d\theta_e} - \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \right] + \frac{d \ln [1 - \tau_s^E(\theta_e)]}{d\theta_e} + (1 - \zeta) \frac{\sigma - 1}{\sigma} \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e} \right] \right],$$

$$\text{where } H(\theta_e) = \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{d \ln [X_e(\theta_e) / \mu(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} (1 - \zeta) \frac{\sigma - 1}{\sigma} - \frac{\sigma - 1}{\sigma}} + \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \right].$$

Set

$$H_0(\theta_e) = \left\{ [1 - \bar{g}_e(\theta_e)] \left\{ \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{\sigma - 1}{\sigma} \zeta \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} - \left[\frac{\sigma - 1}{\sigma} - \left(1 - \frac{\zeta}{\mu(\theta_e)}\right) \frac{1 + \varepsilon_e}{\varepsilon_e} \right] H(\theta_e) \right\} + \varepsilon_{Q-ij}^{P,cross}(\theta_e) [1 - g_e(\theta_e)] - \frac{1}{\mu(\theta_e)} \right\},$$

$$H_1(\theta_e) = \frac{1 + \left\{ [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left([\mu(\theta_e) - \zeta] \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right) + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} - [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \mu(\theta_e) \varepsilon_{Q-ij}^{P,cross}(\theta_e) \frac{1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta}{1 - \frac{\sigma - 1}{\sigma} (1 - \frac{\varepsilon_e}{1 + \varepsilon_e} + \zeta)} \frac{d \ln H_0(\theta_e)}{d\theta_e} \right\}}{\mu^*}$$

$$= \frac{1 + \left\{ [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[[\mu(\theta_e) - \zeta] \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right] + [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} - H_2(\theta_e) \frac{1 + \frac{1 + \varepsilon_e}{\varepsilon_e} \zeta}{\frac{1 + \varepsilon_e}{\varepsilon_e} \zeta} \mu^* \frac{d \ln H_0(\theta_e)}{d\theta_e} \right\}}{\mu^*},$$

$$H_2(\theta_e) = [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{\mu(\theta_e)}{\mu^*} \varepsilon_{Q-ij}^{P,cross}(\theta_e) \frac{\frac{1 + \varepsilon_e}{\varepsilon_e} \zeta}{1 - \frac{\sigma - 1}{\sigma} \left(\frac{\varepsilon_e}{1 + \varepsilon_e} + \zeta \right)},$$

$$\text{where } H(\theta_e) = \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{d \ln [X_e(\theta_e) / \mu(\theta_e)]}{d\theta_e}}{\frac{1 + \varepsilon_e}{\varepsilon_e} (1 - \zeta) \frac{\sigma - 1}{\sigma} - \frac{\sigma - 1}{\sigma}} + \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \right]. \text{ We have}$$

$$\frac{1}{1 - \tau_s^E(\theta_e)} = H_0(\theta_e) [1 - \tau_e^E(\theta_e)]$$

and

$$\frac{1}{1 - \tau_e^E(\theta_e)} = H_1(\theta_e) - H_2(\theta_e) \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e}.$$

Solving the above differential equation, we have

$$1 - \tau_e^E(\theta_e) = \left[1 - \tau_e^E(\bar{\theta}_e)\right] e^{-\int_{\bar{\theta}_e}^{\theta_e} \frac{H_1(s)}{H_2(s)} ds} + \int_{\theta_e}^{\bar{\theta}_e} e^{-\int_{\bar{\theta}_e}^s \frac{H_1(u)}{H_2(u)} du} \frac{1}{H_2(s)} ds.$$

Last, according to (A119), when $\bar{\theta}_e$ is finite, such that $1 - F_e(\bar{\theta}_e) = 0$, $\frac{\tau_e^E(\bar{\theta}_e)}{1 - \tau_e^E(\bar{\theta}_e)} = \frac{1 - \mu^*}{\mu^*}$.

(IV) PART TWO: A RESULT FOR COMPARISON.

Notice that

$$\begin{aligned} \frac{\tau_s^E(\theta_e)}{1 - \tau_s^E(\theta_e)} &= \overbrace{\left[\frac{\mu^*}{\mu(\theta_e)} - 1 \right]}^{\widehat{RE}(\theta_e)} + [1 - \tau_e^E(\theta_e)] \varepsilon_{Q-ij}^{P, cross}(\theta_e) \overbrace{\left[\begin{aligned} &\left[1 - g_e(\theta_e) \right] - \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \\ &\times \left[\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{l'_e(\theta_e)}{l_e(\theta_e)} + \frac{d \ln [\mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e)]}{d\theta_e} \right] \end{aligned} \right]}^{\widehat{IRE}(\theta_e)} \\ &= \overbrace{\left[\frac{\mu^*}{\mu(\theta_e)} - 1 \right]}^{\widehat{RE}(\theta_e)} + [1 - \tau_e^E(\theta_e)] \varepsilon_{Q-ij}^{P, cross}(\theta_e) \left[\begin{aligned} &\frac{[1 - g_e(\theta_e)] - \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \times}{\frac{1 + \varepsilon_e}{\varepsilon_e} \frac{d}{d\theta_e} \left[\ln \frac{X_e(\theta_e) [1 - \tau_s^E(\theta_e)]}{\mu(\theta_e)} \right] + [1 - \zeta] \frac{\sigma - 1}{\sigma} \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e}} \\ &+ \frac{d \ln [\mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e)]}{d\theta_e} \end{aligned} \right]. \end{aligned}$$

When $\tau_s^{E'}(\theta_e) = \tau_e^{E'}(\theta_e) = 0$,

$$\frac{\tau_s^E(\theta_e)}{1 - \tau_s^E(\theta_e)} = \left[\frac{\mu^*}{\mu(\theta_e)} - 1 \right] + [1 - \tau_e^E(\theta_e)] \varepsilon_{Q-ij}^{P, cross}(\theta_e) [[1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] H(\theta_e)].$$

Notice that

$$\frac{1}{1 - \tau_e^E(\theta_e)} = \frac{1}{\mu^*} \left[1 + [1 - \bar{g}_e(\theta_e)] \left\{ \begin{aligned} &\left[(\mu(\theta_e) - \zeta) \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right] H(\theta_e) + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} + \\ &\mu(\theta_e) \varepsilon_{Q-ij}^{P, cross}(\theta_e) \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{(1 + \zeta) \frac{1 + \varepsilon_e}{\varepsilon_e} \frac{d \ln [1 - \tau_s^E(\theta_e)]}{d\theta_e} + \frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e}}{1 - \frac{\sigma - 1}{\sigma} \left(\frac{\varepsilon_e}{1 + \varepsilon_e} + \zeta \right)} \end{aligned} \right\} \right].$$

We have the following results when $\frac{d \ln [1 - \tau_e^E(\theta_e)]}{d\theta_e} = \frac{d \ln [1 - \tau_s^E(\theta_e)]}{d\theta_e} = 0$:

$$\frac{1}{1 - \tau_e^E(\theta_e)} = \frac{1 + [1 - \bar{g}_e(\theta_e)] \left\{ H(\theta_e) \left[(\mu(\theta_e) - \zeta) \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right] + \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \frac{d \ln [\mu(\theta_e) - \zeta]}{d\theta_e} \right\}}{\mu^*}.$$

If in addition $\mu'(\theta_e) = 0$:

$$\frac{1}{1 - \tau_e^E(\theta_e)} = \frac{1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[(\mu(\theta_e) - \zeta) \frac{1 + \varepsilon_e}{\varepsilon_e} - 1 \right]}{\mu^*}.$$

3.4. Proof of Proposition 3

(i) Rewriting the general tax formula for the entrepreneurs in equation (32) for the case with uniform markups, we have part one of Proposition 3. By the definitions of $\varepsilon_{1 - \tau_e}^{y_e}(\theta_e)$ and $\varepsilon_{Q-ij}^{P, cross}(\theta_e)$ (see Supplemental Appendix I.1.2 for details), we have:

$$\begin{aligned} \frac{1}{1 - \tau_e(\theta_e)} &= \frac{1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[\frac{1 + \varepsilon_e}{\varepsilon_e} [\mu - \zeta] - 1 \right]}{\mu} + \\ &\quad \frac{\frac{\sigma}{\sigma - 1}}{\frac{\sigma}{\sigma - 1} - \zeta} \left[\frac{\sigma - 1}{\sigma} - \frac{1}{\mu} \right] \{ [1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] H(\theta_e) \}. \end{aligned}$$

(ii) Supposing $g_o(\cdot)$ (such that $\bar{g}_o(\cdot)$) is exogenous, we have:

$$\begin{aligned} \frac{d \left[\frac{1}{1-\tau_e(\theta_e)} \right]}{d\mu} &= \frac{1}{\mu^2} \left[\mu [1 - \bar{g}_e(\theta_e)] H(\theta_e) \frac{1+\varepsilon_e}{\varepsilon_e} - [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[(\mu - \zeta) \frac{1+\varepsilon_e}{\varepsilon_e} - 1 \right] \right. \\ &\quad \left. + \frac{\frac{\sigma}{\sigma-1}}{\zeta} \{ [1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] H(\theta_e) \} - 1 \right] \\ &= \frac{1}{\mu^2} \left[[1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[\zeta \frac{1+\varepsilon_e}{\varepsilon_e} + 1 \right] \right. \\ &\quad \left. + \frac{\frac{\sigma}{\sigma-1}}{\zeta} \{ [1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] H(\theta_e) \} - 1 \right] \\ &= \frac{1}{\mu^2} \left[[1 - \bar{g}_e(\theta_e)] H(\theta_e) \zeta \left[\frac{1+\varepsilon_e}{\varepsilon_e} - \frac{1}{\frac{\sigma}{\sigma-1} - \zeta} \right] \right. \\ &\quad \left. + [1 - g_e(\theta_e)] \frac{\frac{\sigma}{\sigma-1}}{\zeta} - 1 \right] \end{aligned}$$

and

$$\begin{aligned} \frac{d \left[\frac{1-\tau_w(\theta_w)}{1-\tau_e(\theta_e)} \right]}{d\mu} &= \frac{1 - \tau_w(\theta_w)}{\mu} \frac{d \left[\frac{\mu}{1-\tau_e(\theta_e)} \right]}{d\mu} \\ &= \frac{1 - \tau_w(\theta_w)}{\mu} \left[\frac{[1 - \bar{g}_e(\theta_e)] H(\theta_e) \frac{1+\varepsilon_e}{\varepsilon_e} + \frac{1}{\frac{\sigma}{\sigma-1} - \zeta} \{ [1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] H(\theta_e) \}}{\zeta} \right]. \end{aligned}$$

According to the above equations, $\tau_e(\theta_e)$ increases in μ if and only if:

$$[1 - \bar{g}_e(\theta_e)] H(\theta_e) \zeta \left[\frac{1+\varepsilon_e}{\varepsilon_e} - \frac{1}{\frac{\sigma}{\sigma-1} - \zeta} \right] + [1 - g_e(\theta_e)] \frac{\frac{\sigma}{\sigma-1}}{\zeta} - 1 > 0,$$

i.e.,

$$\begin{aligned} (A125) \quad g_e(\theta_e) &< 1 - \frac{[1 - \bar{g}_e(\theta_e)] H(\theta_e) \zeta \left[\frac{1+\varepsilon_e}{\varepsilon_e} - \frac{1}{\frac{\sigma}{\sigma-1} - \zeta} \right]}{\frac{\frac{\sigma}{\sigma-1}}{\zeta}} \\ &= \frac{\zeta(\sigma-1)}{\sigma} \left\{ 1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \zeta \right) - 1 \right] \right\}. \end{aligned}$$

$\frac{1-\tau_w(\theta_w)}{1-\tau_e(\theta_e)}$ increases in μ iff

$$[1 - \bar{g}_e(\theta_e)] H(\theta_e) \frac{1+\varepsilon_e}{\varepsilon_e} + \frac{1}{\frac{\sigma}{\sigma-1} - \zeta} \{ [1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] H(\theta_e) \} > 0,$$

i.e.,

$$(A126) \quad g_e(\theta_e) < 1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \zeta \right) - 1 \right],$$

or equivalently

$$g_e(\theta_e) - 1 + [1 - \bar{g}_e(\theta_e)] H(\theta_e) < [1 - \bar{g}_e(\theta_e)] H(\theta_e) \frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \zeta \right),$$

which must be true when $[1 - g_e(\theta_e)] - [1 - \bar{g}_e(\theta_e)] H(\theta_e) > 0$ and $H(\theta_e) > 0$.

(iii) When $g_e(\theta_e) = \bar{g}_e(\theta_e)$ and $H(\theta_e) > 0$, inequality (A125) is equivalent to

$$g_e(\theta_e) < \frac{\frac{\xi(\sigma-1)}{\sigma} \left\{ 1 + H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1 \right] \right\}}{1 + \frac{\xi(\sigma-1)}{\sigma} H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1 \right]},$$

where $\frac{1+H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1 \right]}{\frac{\xi(\sigma-1)}{\sigma} + H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1 \right]}$ increases in $H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1 \right]$, because $\frac{\sigma}{\xi(\sigma-1)} > 1$. Besides, under condition (22), we have $\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1 > 0$ and that $H(\theta_e) \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1 \right]$ increases in $H(\theta_e)$. Thus, when $H(\theta_e) > 0$, the above inequality holds if

$$g_e(\theta_e) \leq \frac{\xi(\sigma-1)}{\sigma}.$$

Analogously, when $g_e(\theta_e) = \bar{g}_e(\theta_e) < 1$ and $H(\theta_e) > 0$, inequality (A126) must hold. ■

IV. Discussion and Robustness

Below, we explore the relationship between our results and several distinct issues that have been extensively studied in the literature. We frame these discussions through the lens of our model, leveraging its unique features and notation where necessary to derive new insights.

4.1. Top Profit Tax

(I) MARKET POWER AND TOP PROFIT TAX.

The taxes paid by top earners constitute the majority of income tax revenue, making the design of optimal tax rates for high-income individuals critical. Proposition A1 offers an analytic formula for determining the optimal top profit tax rate, with specific application to high-profit entrepreneurs (denoted as those with $\theta_e \geq \hat{\theta}_e$). We assume that the profit and sales tax rates in the real economy for the top entrepreneurs are linear and lower than one. Then, we have the following result:

PROPOSITION A1: Suppose that there exist $\hat{\theta}_e \in \Theta_e$ such that for $\theta_e \geq \hat{\theta}_e$, $H(\theta_e) = \hat{H}$, $\mu(\theta_e) = \hat{\mu}$ and $g_e(\theta_e) = \hat{g}_e$ are constants. Then we have the following results in the nested-CES economy:

(i) The optimal profit tax rate for $\theta_e \geq \hat{\theta}_e$ is constant and satisfy:

$$(A127) \quad \frac{1}{1 - \hat{\tau}_e} = \frac{\frac{1+(1-\hat{g}_e)\hat{H}_e \left[\frac{1+\varepsilon_e}{\varepsilon_e} (\hat{\mu} - \xi) - 1 \right]}{\hat{\mu}} + \frac{1 - \frac{\sigma}{\sigma-1} \frac{1}{\hat{\mu}}}{\frac{\sigma}{\sigma-1} - \xi} (1 - \hat{g}_e) (1 - \hat{H})}{1 - \frac{\xi}{\sigma-1} \frac{\mu - \hat{\mu}}{\hat{\mu}}},$$

(ii) Given σ , $\hat{\tau}_e$ decreases in I (increases in $\hat{\mu}$) if:

$$(A128) \quad \hat{g}_e < 1 - \frac{1}{2 \cdot \hat{H}}.$$

PROOF:

(i) Notice that (A50). Denote by y_e^i , $F_{y_e}^i(y_e^i)$, and $f_{y_e}^i(y_e^i)$ the initial profit, CDF, and PDF of initial profit, respectively. $\frac{1-F_{y_e}^i(y_e^i(\hat{\theta}_e))}{y_e^i(\hat{\theta}_e)f_{y_e}^i(y_e^i(\hat{\theta}_e))}$ is the initial hazard ratio of profit. When the profit and sales taxes are linear in the initial, the initial hazard ratio of profit $\frac{1-F_{y_e}^i(y_e^i(\hat{\theta}_e))}{y_e^i(\hat{\theta}_e)f_{y_e}^i(y_e^i(\hat{\theta}_e))} = H(\theta_e)$. Thus $H(\theta_e)$ is constant for $\theta_e \geq \hat{\theta}_e$, if the initial hazard ratio of profit is constant for $\theta_e > \hat{\theta}_e$.

(ii) Notice that if for any $\theta_e > \hat{\theta}_e$, $g_e(\theta_e) = \hat{g}_e$ is constant, $\bar{g}_e(\theta_e)$ is also constant and equal to $\hat{g}_e$ for any $\theta_e \geq \hat{\theta}_e$. With the above two findings, part two of Proposition A1 follow directly from Theorem 1.

(iii) To prove part three of Proposition A1, first note that $\frac{1}{1-\hat{\tau}_e}$ and thus $\hat{\tau}_e$ increases with μ . Notice that for any $\theta_e \in \Theta_e$, $\mu(\theta_e)$ non-decreases with the decrease of I , μ as a weighted average of $\mu(\theta_e)$ must non-decreases with the decrease of I . Thus a condition guarantees that, given μ , $\frac{1}{1-\hat{\tau}_e}$ increases with $\hat{\mu}$ is a sufficient condition for $\hat{\tau}_e$ to be increasing with the decrease of I . We now show that (A129) is such a condition.

To see this, we first treat $\hat{g}_e$ and μ as given and take the derivative of $\frac{1}{1-\hat{\tau}_e}$ with respect to $\hat{\mu}$:

$$\frac{d\left(\frac{1}{1-\hat{\tau}_e}\right)}{d\hat{\mu}} = \frac{-\frac{1+(1-\hat{g}_e)\hat{H}\left[\frac{1+\varepsilon_e}{\varepsilon_e}(\hat{\mu}-\zeta)-1\right]}{\hat{\mu}^2} + \frac{(1-\hat{g}_e)\hat{H}\frac{1+\varepsilon_e}{\varepsilon_e}}{\hat{\mu}} + \frac{\frac{\sigma}{\sigma-1}\frac{1}{(\hat{\mu})^2}}{\frac{\sigma}{\sigma-1}-\zeta}(1-\hat{g}_e)(1-\hat{H})}{1 - \frac{\zeta}{\frac{\sigma}{\sigma-1}-\zeta}\frac{\mu-\hat{\mu}}{\hat{\mu}}}$$

$$- \frac{\frac{\zeta}{\frac{\sigma}{\sigma-1}-\zeta}\frac{\mu}{\hat{\mu}^2} \left[\frac{1}{\hat{\mu}} \left[1 + (1-\hat{g}_e)\hat{H}\left[\frac{1+\varepsilon_e}{\varepsilon_e}(\hat{\mu}-\zeta)-1\right] \right] + \frac{1-\frac{\sigma}{\sigma-1}\frac{1}{\hat{\mu}}}{\frac{\sigma}{\sigma-1}-\zeta}(1-\hat{g}_e)(1-\hat{H}) \right]}{\left[1 - \frac{\zeta}{\frac{\sigma}{\sigma-1}-\zeta}\frac{\mu-\hat{\mu}}{\hat{\mu}} \right]^2}$$

Rearranging the right side of the above equation, we find that $\frac{d}{d\hat{\mu}}\left(\frac{1}{1-\hat{\tau}_e}\right) > 0$ if and only if

$$(1-\hat{g}_e)\hat{H} \left[\left(\frac{\sigma}{\sigma-1} - \zeta \right) \left[\left(\hat{\mu} - \zeta \mu^{\frac{\sigma-1}{\sigma}} \right) \frac{1+\varepsilon_e}{\varepsilon_e} - \frac{1}{\varepsilon_{1-\hat{\tau}_e}^{ye,top}} \right] + \left(\frac{\sigma}{\sigma-1} - \zeta \mu^{\frac{\sigma-1}{\sigma}} \right) \right] > \frac{\sigma}{\sigma-1} - \zeta.$$

Note that the term in the bracket is positive, because for any $\theta_e \in \Theta_e$, $\mu(\theta_e) \leq \frac{\sigma}{\sigma-1}$. Thus, we have

$$1 - \hat{g}_e > \frac{\frac{\sigma}{\sigma-1} - \zeta}{\hat{H} \left[\left(\frac{\sigma}{\sigma-1} - \zeta \right) \left[\left(\hat{\mu} - \zeta \mu^{\frac{\sigma-1}{\sigma}} \right) \frac{1+\varepsilon_e}{\varepsilon_e} - \frac{1}{\varepsilon_{1-\hat{\tau}_e}^{ye,top}} \right] + \left(\frac{\sigma}{\sigma-1} - \zeta \mu^{\frac{\sigma-1}{\sigma}} \right) \right]}.$$

Dividing both the numerator and denominator of the right side of the above inequality by $\frac{\sigma}{\sigma-1} - \zeta$ and substituting $\frac{1}{\varepsilon_{1-\hat{\tau}_e}^{ye,top}}$ by $\frac{1+\varepsilon_e}{\varepsilon_e}(\hat{\mu}-\zeta)-1$, we have

$$(A129) \quad \hat{g}_e < 1 - \frac{1}{\left[\zeta \left(1 - \mu^{\frac{\sigma-1}{\sigma}} \right) \frac{1+\varepsilon_e}{\varepsilon_e} + 1 + \frac{\frac{\sigma}{\sigma-1}-\zeta\mu^{\frac{\sigma-1}{\sigma}}}{\frac{\sigma}{\sigma-1}-\zeta} \right] \hat{H}}.$$

(A129). Notice that $\mu \leq \frac{\sigma}{\sigma-1}$, the term in the bracket of the right side of (A127) is not less than 2. Thus, condition (A127) satisfies if $\hat{g}_e < 1 - \frac{1}{2\hat{H}}$.

Last, since for any $I > 1$, (A128) is a sufficient condition for $\hat{\tau}_e$ to be increasing with the marginal decrease of I (just suppose that I is continuous), and $\hat{\tau}_e$ is twice continuously differentiable with respect to I , according to the mean value theorem, (A128) is also a sufficient condition for $\hat{\tau}_e$ to be increasing with the decrease of I from $n+1$ to n , where $n \in \mathbb{N}_+$. ■

Formula (A127) extends the traditional top income tax formula, generalizing the top tax rate result of Saez (2001) (where $\zeta = 0$, $\mu = 1$, $I = 1$, $\sigma \rightarrow \infty$) by incorporating more complex technology and market structures. Compared to Corollary 5 of Sachs, Tsyvinski and Werquin (2020) (where $\zeta = 0$, $\mu = 1$, and $I = 1$), it highlights the influence of market structure and superstar effects on the optimal top tax rate. Similarly, relative to Scheuer and Werning (2017) (where $\mu = 1$, $I = 1$, and $\sigma \rightarrow \infty$), it emphasizes the interplay between superstar effects and market structure.

Proposition A1 is particularly valuable as it suggests that, under reasonable assumptions, observable statistics like $\hat{H}_e$ can be used to derive the optimal top profit tax rate and assess whether it should be adjusted in response to changes in technology and market structure. A crucial question here is: should the government increase the top profit tax rate as market power rises? Condition (A127) indicates that whether the optimal top profit tax rate increases depends heavily on the value of $\hat{H}_e$. Specifically, if $\hat{g}_e \rightarrow 0$, condition (A128) becomes equivalent to $\hat{H}_e > \frac{1}{2}$. This condition is sufficient, though not necessary, for $\hat{\tau}_e$ to increase with $\hat{\mu}$.⁵⁷ We provide a more relaxed sufficient condition: $\hat{g}_e < 1 - \frac{1}{\left[\xi \left(1 - \mu^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1+\varepsilon_e}{\varepsilon_e}} + 1 + \frac{\frac{\sigma-1}{\sigma} - \xi \mu^{\frac{\sigma-1}{\sigma}}}{\frac{\sigma-1}{\sigma} - \xi} \right] \hat{H}}$, where the

bracketed term is at least 2, as $\mu \leq \frac{\sigma}{\sigma-1}$.

Empirical evidence supports this conclusion. The hazard ratio of top income in the United States was approximately 0.5 in 1992 and 1993 (Saez (2001)) and about $\frac{2}{3}$ in 2005 (Diamond and Saez (2011)). By 2007, the hazard ratios for top labor, capital, and total incomes were approximately 0.62, 0.76, and 0.71, respectively (Saez and Stantcheva (2018)). These findings suggest that condition (A128) is met in the U.S., indicating that the top profit tax rate should rise as market power increases.

(II) SPAN OF CONTROL AND TOP PROFIT TAX.

Similar to our model, Scheuer and Werning (2017) also explore optimal income taxation within a Lucas (1978) span of control framework. They examine the superstar effect, which arises from positive assortative matching between the abilities of entrepreneurs and the scale of their firms. In their model, the span of control reflects the magnitude of the superstar effect, representing the extent to which entrepreneurs can amplify the productivity of their workers.

In our framework, Proposition A1 provides valuable insights into how the span of control influences optimal taxation. Drawing from Scheuer and Werning (2017), an increase in ξ impacts both the elasticity of profits with respect to the tax rate and the hazard ratio of profit. Interestingly, when $I = 1$, these two effects offset each other, rendering the top profit tax rate independent of ξ . To understand this, note that under $I = 1$, the markup is uniform ($\mu = \hat{\mu} = \frac{\sigma}{\sigma-1}$), simplifying formula (A127) to:

$$(A130) \quad \frac{1}{1 - \hat{\tau}_e} = \frac{1 + (1 - \hat{g}_e) \hat{H} \left[\frac{1+\varepsilon_e}{\varepsilon_e} (\hat{\mu} - \xi) - 1 \right]}{\hat{\mu}} = \frac{1}{\hat{\mu}} \left[1 + (1 - \hat{g}_e) \frac{1 - F_e(\hat{\theta}_e)}{f_e(\hat{\theta}_e)} \frac{d \ln X_e(\hat{\theta}_e)}{d \hat{\theta}_e} \frac{\sigma - 1}{\sigma} \frac{1 + \varepsilon_e}{\varepsilon_e} \right],$$

where the second equality follows from the fact that $\hat{H} = \frac{1 - F_e(\hat{\theta}_e)}{f_e(\hat{\theta}_e)} \frac{\frac{\sigma-1}{\sigma} \frac{1+\varepsilon_e}{\varepsilon_e} \frac{d \ln X_e(\hat{\theta}_e)}{d \hat{\theta}_e}}{\frac{1+\varepsilon_e}{\varepsilon_e} \left(\frac{\sigma}{\sigma-1} - \xi \right) - 1}$. The equation above establishes that $\hat{\tau}_e$ is independent of ξ in a monopoly competitive economy. In other words, the “neutrality” of the span of control found by Scheuer and Werning (2017) still holds.

The “neutrality” of the span of control does not necessarily hold in a more general setting. As an illustration, we consider a special case where the top markup is $\hat{\mu} = \frac{\sigma}{\sigma-1}$. We consider this case because $\frac{\sigma}{\sigma-1}$ is the theoretical maximum value of markup in our model. Substitute $\hat{H}_e$ in formula (A127) by $\hat{H} = \frac{1}{\hat{\mu} - \xi - \frac{\varepsilon_e}{1+\varepsilon_e}} \frac{\gamma'_e(\hat{\theta}_e)}{\gamma_e(\hat{\theta}_e)} \frac{1 - F_e(\hat{\theta}_e)}{f_e(\hat{\theta}_e)}$. As the markup for $\theta_e > \hat{\theta}_e$ is assumed to be constant and $\hat{\mu} = \frac{\sigma}{\sigma-1}$, one has:

$$\frac{1}{1 - \hat{\tau}_e} = \frac{\frac{1}{\hat{\mu}} \left[1 + (1 - \hat{g}_e) \frac{1 - F_e(\hat{\theta}_e)}{f_e(\hat{\theta}_e)} \frac{d \ln X_e(\hat{\theta}_e)}{d \hat{\theta}_e} \frac{\sigma - 1}{\sigma} \frac{1 + \varepsilon_e}{\varepsilon_e} \right]}{1 + \left[\frac{\frac{\sigma}{\sigma-1}}{\frac{\sigma}{\sigma-1} - \xi} - 1 \right] \frac{\hat{\mu} - \mu}{\hat{\mu}}}.$$

It can be seen that in this case $\hat{\tau}_e$ decreases in ξ , because the rising span of control enlarges the influence of reallocation effect, which can be seen from the multiplier on $RE(\theta_e)$ in formula (31), i.e., $\frac{\xi}{\frac{\sigma}{\sigma-1} - \xi}$.

⁵⁷In Supplemental Appendix III.3.1

4.2. Alternative Market Structures

(I) MARKET STRUCTURE, INDIRECT REDISTRIBUTION, AND OPTIMAL TAXATION.

An important finding of this paper is the pivotal role of market structure in determining the optimal tax, particularly in shaping the indirect redistribution effect of taxation. In many prior studies on endogenous prices and optimal taxation, taxes have a first-order effect on relative prices, which can help ease incentive constraints and improve income distribution (see, e.g., [Naito \(1999\)](#); [Stiglitz \(2018\)](#); [Sachs, Tsyvinski and Werquin \(2020\)](#); [Cui, Gong and Li \(2021\)](#)). Specifically, when the marginal productivity of labor (wage) decreases with labor inputs, the planner can compress the wage distribution by reducing the marginal tax rate on high-skilled agents, thereby encouraging their labor supply. [Saez \(2004\)](#) contends that the indirect redistribution effect of taxes diminishes when agents make endogenous human capital investments, as they effectively set their own wages, reducing the tax's influence on prices to a second-order effect. In response, [Naito \(2004\)](#) shows that when human capital is imperfectly substitutable, the indirect redistribution effect remains relevant.

Our findings advance this discussion by showing that the extent of indirect redistribution depends critically on market structure. Even in the presence of market power, indirect redistribution persists, but its magnitude decreases as market power intensifies. In fact, the IRE vanishes entirely under monopoly, where firms independently set their prices. Conversely, under competition or oligopoly, where prices are determined fully or partially outside the firm, IRE still plays a role. This insight reconciles the seemingly conflicting conclusions of [Saez \(2004\)](#) and [Naito \(2004\)](#): [Saez \(2004\)](#) examines a scenario where agents act as monopolistic suppliers of their own factors, while [Naito \(2004\)](#) assumes a competitive labor market.

(II) MONOPOLISTIC COMPETITION WITH KIMBALL AGGREGATION AND ENDOGENOUS MARKUPS.

In some of our results, we employ a constant elasticity of substitution (CES) technology. Although markups are endogenous, tax policies do not influence the equilibrium markup under CES. Here, we consider an alternative setup using Kimball aggregation, which introduces non-constant elasticity of substitution. This modification allows taxes to affect markups. For tractability, we focus on monopolistic competition and the second-best allocation.

The technology is described below:

$$(A131) \quad 1 = \int_{\theta_e} \chi(\theta_e) Y\left(\frac{Q(\theta_e)}{Q/N_e}\right) dF_e(\theta_e),$$

where $Q(\theta_e) = x_e(\theta_e) l_e(\theta_e) L_w(\theta_e)^\xi$, $Y(\cdot)$ is a twice differentiable function, and Q is the quantity of final goods. Under the above technology

$$(A132) \quad P(\theta_e) = \frac{\chi(\theta_e) Y'\left(\frac{Q(\theta_e)}{Q/N_e}\right)}{\int_{\theta_e} \chi(\theta_e) Y'\left(\frac{Q(\theta_e)}{Q/N_e}\right) \frac{Q(\theta_e)}{Q/N_e} dF_{\theta_e}(\theta_e)}$$

$$(A133) \quad \text{and } \mu(\theta_e) = \frac{\varepsilon(\theta_e)}{\varepsilon(\theta_e) - 1} \text{ with } \varepsilon(\theta_e) = -\frac{Y''\left(\frac{Q(\theta_e)}{Q}\right) \frac{Q(\theta_e)}{Q}}{Y'\left(\frac{Q(\theta_e)}{Q}\right)}.$$

The markup $\mu(\theta_e)$ is a function of $\frac{Q(\theta_e)}{Q}$. According to Lemma 1, the incentive compatible condition of the entrepreneur is:

$$(A134) \quad V'_e(\theta_e) = \phi'_e(l_e(\theta_e)) l_e(\theta_e) \left[\mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right], \forall \theta_e \in \Theta_e.$$

The planner chooses $\{l_e(\theta_e), L_w(\theta_e), V_e(\theta_e), l_w(\theta_w), V_w(\theta_w), Q\}_{\theta_e \in \Theta_e, \theta_w \in \Theta_w}$ to maximize (5) subject to the resource constraints (13), the labor market clear condition (14), and the incentive conditions (A134) and (24).

As a comparison to the optimal profit tax under monopolistic competition in the benchmark model, we now have the following proposition:

PROPOSITION A2: *Under monopolistic competition with Kimball aggregation, the effective tax rate on entrepreneurial effort satisfies:*

$$(A135) \quad \frac{1 - \frac{1 - \tau_e(\theta_e)}{\mu(\theta_e)}}{\frac{1 - \tau_e(\theta_e)}{\mu(\theta_e)}} = [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] + \mu(\theta_e) \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)} \right],$$

for any $\theta_e \in \Theta_e$.

PROOF:

Set up the Lagrangian function as

$$\begin{aligned} & \mathcal{L}(L_w, l_w, l_e, V_w, V_e, Q; \lambda, \lambda', \psi_w, \psi_e, \kappa, \varphi) \\ = & \sum_{o \in \{w, e\}} N_o \int_{\theta_o} G(V_o(\theta_o)) \tilde{f}_o(\theta_o) d\theta_o + \lambda \left[Q - \sum_{o \in \{w, e\}} N_o \int_{\theta_o} [V_o(\theta_o) + \phi_o(l_o(\theta_o))] f_o(\theta_o) d\theta_o - R \right] \\ & + \lambda' \left[N_w \int_{\theta_w} x_w(\theta_w) l_w(\theta_w) f_w(\theta_w) d\theta_w - N_e \int_{\theta_e} L_w(\theta_e) f_e(\theta_e) d\theta_e \right] \\ & + \lambda'' \left[1 - \int_{\theta_e} \chi(\theta_e) Y \left(\frac{Q(\theta_e)}{Q/N_e} \right) dF_{\theta_e}(\theta_e) \right] \\ & + \int_{\theta_w} \psi_w(\theta_w) \left[l_w(\theta_w) \phi'_w(l_w(\theta_w)) \frac{x'_w(\theta_e)}{x_w(\theta_e)} - V'_w(\theta_w) \right] d\theta_w \\ & + \int_{\theta_e} \psi_e(\theta_e) \left[\phi'_e(l_e(\theta_e)) l_e(\theta_e) \left[\mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] - V'_e(\theta_e) \right] d\theta_e, \end{aligned}$$

where $\mu(\theta_e)$ is a function of $\frac{Q(\theta_e)}{Q}$.

Taking partial integrals yields the following

$$- \int_{\theta_e} \psi_o(\theta_e) V'_o(\theta_e) d\theta_e = V_o(\underline{\theta}_o) \psi_o(\underline{\theta}_o) - V_o(\bar{\theta}_o) \psi_o(\bar{\theta}_o) + \int_{\theta_o} \psi'_o(\theta_o) V_o(\theta_o) d\theta_o.$$

The derivatives with respect to the endpoint conditions yield boundary conditions:

$$\psi_o(\underline{\theta}_o) = \psi_o(\bar{\theta}_o) = 0, o \in \{w, e\}.$$

Thus,

$$\int_{\theta_o} \psi'_o(\theta_o) d\theta_o = 0,$$

Substituting the above conditions into the Lagrangian function yields the following first-order conditions:

$$(A136) \quad \frac{\partial \mathcal{L}}{\partial Q} = \lambda + N_e \lambda'' \int_{\theta_e} \chi(\theta_e) Y \left(\frac{Q(\theta_e)}{Q/N_e} \right) \frac{Q(\theta_e)}{Q^2} dF_e(\theta_e) = 0, o \in \{w, e\},$$

$$(A137) \quad \frac{\partial \mathcal{L}}{\partial V_o(\theta_o)} = G'(V_o(\theta_o)) N_o \tilde{f}_o(\theta_o) + \psi'_o(\theta_o) - \lambda N_o f_o(\theta_o) = 0,$$

$$(A138) \quad \frac{\partial \mathcal{L}}{\partial l_w(\theta_w)} = -\lambda N_w \phi'_w(l_w(\theta_w)) f_w(\theta_w) + \lambda' N_w x_w(\theta_w) f_w(\theta_w) + \psi_w(\theta_w) \frac{\phi'_w(l_w(\theta_w))}{x_w(\theta_w)} \frac{1 + \varepsilon_w}{\varepsilon_w} = 0,$$

$$(A139) \quad \frac{\partial \mathcal{L}}{\partial L_w(\theta_e)} = \left[-\lambda'' N_e \chi(\theta_e) Y' \left(\frac{Q(\theta_e)}{Q/N_e} \right) \frac{1}{Q} \frac{\partial Q(\theta_e)}{\partial L_w(\theta_e)} f_e(\theta_e) - \lambda' N_e f_e(\theta_e) + \psi_e(\theta_e) \phi'_e(l_e(\theta_e)) l_e(\theta_e) \frac{\partial \mu(\theta_e)}{\partial Q(\theta_e)} \frac{\partial Q(\theta_e)}{\partial L_w(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)} \right] = 0,$$

and

$$(A140) \quad \frac{\partial \mathcal{L}}{\partial l_e(\theta_e)} = \psi_e(\theta_e) \phi'_e(l_e(\theta_e)) \frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] + \psi_e(\theta_e) \phi'_e(l_e(\theta_e)) l_e(\theta_e) \frac{\partial \mu(\theta_e)}{\partial Q(\theta_e)} \frac{\partial Q(\theta_e)}{\partial l_e(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \left[-\lambda'' N_e \chi(\theta_e) Y' \left(\frac{Q(\theta_e)}{Q} \right) \frac{1}{Q} \frac{\partial Q(\theta_e)}{\partial l_e(\theta_e)} f_e(\theta_e) - \lambda \phi'_e(l_e(\theta_e)) N_e \right] f_e(\theta_e) = 0.$$

Substitute λ'' in (A139) by (A136) and use (A132):

$$P(\theta_e) \frac{\partial Q(\theta_e)}{\partial L_w(\theta_e)} = \frac{\lambda'}{\lambda} + \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \phi'_e(l_e(\theta_e)) l_e(\theta_e) \frac{\partial \mu(\theta_e)}{\partial Q(\theta_e)} \frac{\partial Q(\theta_e)}{\partial L_w(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)},$$

where we substitute $\frac{\psi_e(\theta_e)}{\lambda N_e f_e(\theta_e)}$ by (A82). Substituting $P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial L_w(\theta_e)}$ by $\frac{\mu(\theta_e)W}{1 - \tau_s(\theta_e)}$ gives:

$$(A141) \quad \frac{1}{1 - \tau_s(\theta_e)} = \frac{\frac{\lambda'}{\lambda}}{\mu(\theta_e)W} + \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \frac{\phi'_e(l_e(\theta_e)) l_e(\theta_e)}{WL_w(\theta_e)} \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \zeta \frac{\chi'(\theta_e)}{\chi(\theta_e)},$$

Equivalently,

$$\frac{1}{1 - \tau_s(\theta_e)} = \frac{\frac{\lambda'}{\lambda}}{\mu(\theta_e)W} + \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q(\theta_e)}{\partial l_e(\theta_e)}} \frac{Q(\theta_e) P(\theta_e)}{WL_w(\theta_e)} \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \zeta \frac{\chi'(\theta_e)}{\chi(\theta_e)}.$$

Substitute $\frac{Q(\theta_e)P(\theta_e)}{WL_w(\theta_e)}$ by $\frac{P(\theta_e)Q(\theta_e)}{WL_w(\theta_e)} \zeta = \frac{\mu(\theta_e)}{1 - \tau_s(\theta_e)}$ and substitute $\frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q(\theta_e)}{\partial l_e(\theta_e)}}$ by $\frac{1 - \tau_e(\theta_e)}{\mu(\theta_e)}$:

$$\frac{1}{1 - \tau_s(\theta_e)} = \frac{\frac{\lambda'}{\lambda}}{\mu(\theta_e)W} + \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} \frac{1 - \tau_e(\theta_e)}{1 - \tau_s(\theta_e)} \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)}.$$

Rearrange the above equation to derive

$$(A142) \quad 1 - \tau_s(\theta_e) = \frac{1 - \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} [1 - \tau_e(\theta_e)] \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)}}{\frac{\lambda'}{\lambda \mu(\theta_e)W}}.$$

Therefore,

$$\frac{1 - \tau_s(\theta'_e)}{1 - \tau_s(\theta_e)} = \frac{1 - \frac{[1 - \bar{g}_e(\theta'_e)][1 - F_e(\theta'_e)]}{f_e(\theta'_e)} [1 - \tau_e(\theta'_e)] \frac{\partial \ln \mu(\theta'_e)}{\partial \ln Q(\theta'_e)} \frac{\chi'(\theta'_e)}{\chi(\theta'_e)}}{1 - \frac{[1 - \bar{g}_e(\theta_e)][1 - F_e(\theta_e)]}{f_e(\theta_e)} [1 - \tau_e(\theta_e)] \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)}}.$$

Substitute $\lambda'' N_e$ in (A140) by (A136) and use (A132):

$$(A143) \quad \frac{1 - \frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q(\theta_e)}{\partial l_e(\theta_e)}}}{\frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q(\theta_e)}{\partial l_e(\theta_e)}}} = [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{1 + \varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] + \mu(\theta_e) \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)} \right]$$

where we substitute $\frac{\psi_e(\theta_e)}{\lambda N_e f_e(\theta_e)}$ by (A82). Using $\frac{1-\tau_e(\theta_e)}{\mu(\theta_e)} = \frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}}$ to substitute $\frac{\phi'_e(l_e(\theta_e))}{P(\theta_e) \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}}$ in (A143), delivers

$$(A144) \quad \frac{1 - \frac{1-\tau_e(\theta_e)}{\mu(\theta_e)}}{\frac{1-\tau_e(\theta_e)}{\mu(\theta_e)}} = [1 - \bar{g}_e(\theta_e)] \frac{1 - F_e(\theta_e)}{f_e(\theta_e)} \left[\frac{1+\varepsilon_e}{\varepsilon_e} \left[\mu(\theta_e) \frac{\chi'(\theta_e)}{\chi(\theta_e)} + \frac{x'_e(\theta_e)}{x_e(\theta_e)} \right] + \mu(\theta_e) \frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} \frac{\chi'(\theta_e)}{\chi(\theta_e)} \right].$$

■

Compared to the monopoly competitive case in the benchmark model (see, e.g., equation (38)), the optimal tax here accounts for the endogeneity of markups. When $\frac{\partial \ln \mu(\theta_e)}{\partial \ln Q(\theta_e)} > 0$, the additional term in the tax formula is positive, suggesting that higher markups generally necessitate higher tax rates. Therefore, our main finding – that the top profit tax rate should increase with rising markups – remains robust under varying elasticities of substitution.

(III) FREE ENTRY.

In this paper, we assume that the number of incumbents in a market is exogenous, meaning our analysis does not capture the impact of taxes on markups and efficiency through the extensive margin of firm entry. This omission is important because entry plays a significant role in shaping real-world policy outcomes. However, fixing the number of incumbents provides a reasonable starting point, as introducing an extensive margin complicates the analysis with an entry game that is challenging both analytically and computationally.

For example, [De Loecker, Eeckhout and Mongey \(2019\)](#) analyze a model with endogenous markups ‘a la [Atkeson and Burstein \(2008\)](#) and firm entry. Their entry game builds on [Berry \(1992\)](#) and involves complex iterations to compare all possible entry configurations. Combining such a framework with incomplete information and optimal taxation would likely render the solution intractable. Future research could extend our model to include firm entry, thereby analyzing the effects of taxation on both the intensive and extensive margins for more nuanced policy insights. There are several ways to introduce the extensive margin. Various approaches could be adopted to incorporate the extensive margin. For instance, [Scheuer \(2014\)](#) and [Rothschild and Scheuer \(2013\)](#) introduce occupational choice, while [Edmond, Midrigan and Xu \(2023\)](#) consider an economy with free entry under Kimball demand. Another approach might involve adding a fringe of small businesses to each market.

(IV) PERFORMANCE PAY AND OPTIMAL PROFIT TAX.

In the real economy, entrepreneurs may only receive a portion of the profits through performance pay. We now examine how profit sharing affects optimal taxation. Suppose that a fraction s of the company’s profits is paid to entrepreneurs as performance pay. The entrepreneur’s problem becomes:

$$\begin{aligned} V_{e,ij}(\theta_e) &\equiv \max_{l_{e,ij}, L_{w,ij}} c_e - \phi_e(l_e) \\ \text{s.t. } c_{e,ij} &= [y_{e,ij} - T_e(y_{e,ij})] \cdot s y_{e,ij} \\ &= (1 - t_s) P_{ij} \left(Q_{ij}, \{Q_{-ij}(\theta_e)\}_{-i \neq i}, \theta_e \right) Q_{ij} - W L_{w,ij}, \end{aligned}$$

where s is the entrepreneur’s share of the profits, and the remaining profits are distributed evenly among taxpayers (or households).

In this setup, the planner’s problem and the constrained optimal allocation remain unchanged. However, the optimal taxation formula is adjusted to incorporate performance pay. Specifically, the tax wedges satisfy: $\tau_s(\cdot) = t_s$, $\tau_w(\theta_w) = T'_w(y_w(\theta_w))$ and $\tau_e(\theta_e) = 1 - (1 - t_s) [1 - T'_e(y_e(\theta_e))] \cdot s$. Therefore, introducing s won’t change the effective tax rate on the effort of the entrepreneur but proportionally increase $1 - T'_e(y_e(\theta_e))$. Consequently, our main results remain valid.

4.3. Alternative Technology, Preference, and Policy Specifications

(I) CAPITAL INVESTMENT.

Our benchmark model does not explicitly incorporate capital, but it can be adapted to do so. The framework remains valid if entrepreneurial effort is replaced with capital investment, particularly when some factor costs are non-deductible before taxation. Consider an economy where entrepreneurs choose labor inputs L_w and capital K , instead of effort:

$$\max_{K, L_w} P_{ij} (Q_{ij} (K, L_w), Q_{-ij} (\theta_e), \theta_e) Q_{ij} (K, L_w) - WL_w - rK - \phi_K (K, \theta_e) - T_e (y_e)$$

where $Q_{ij} (K, L_w)$ is the firm-level production function of capital and labor inputs, r is the market price of capital,⁵⁸ and $\phi_K (K, \theta_e)$ is the unobservable cost of investment.

In practice, while the market price of capital (r) is observable, the total opportunity costs – such as fundraising or management – are often unobservable, captured by $\phi_K (K, \theta_e)$.⁵⁹ An alternative explanation for $\phi_K (K, \theta_e)$ is the preference for asset (wealth). In that case, $\phi_K (K, \theta_e)$ can be negative, which means investment directly generates positive utility. The common ground in these situations is that the elasticity of investment may be finite, which is the key point of [Saez and Stantcheva \(2018\)](#), in which case, $y_e = P_{ij} (Q_{ij} (K, L_w), Q_{-ij} (\theta_e), \theta_e) Q_{ij} (K, L_w) - WL_w - rK$.

In reality, incomplete deductibility of investment costs is common. For instance, while debt interest is deductible, equity investments often are not. Equity investments affect the cash flow of shareholders and generate costs, in which case, $y_e = P_{ij} (Q_{ij} (K, L_w), Q_{-ij} (\theta_e), \theta_e) Q_{ij} (K, L_w) - WL_w$. Then, even if $\phi_K = 0$ there are non-deductible capital costs before tax. This distinction suggests that profits tax can effectively function as a tax on capital (see [Myles \(2008\)](#), Chapter 8).

Our core results remain robust under these scenarios. The fundamental issue in the incentive problem is the unobservability of certain inputs, and our model can accommodate both deductible and non-deductible factors. Consequently, the optimal profit tax formula derived here applies broadly, regardless of specific input configurations.

(II) UNIFORM INCOME TAX.

Thus far, we have considered differentiated tax policies for labor income and profits. However, in practice, governments may not perfectly distinguish between these two income sources. This raises the question of optimal policy under a uniform tax on both profit and labor income.

Introducing this policy constraint complicates the problem significantly. As an illustration, let $\theta_w (\theta_e)$ denote the ability of a worker whose income matches that of an entrepreneur with ability $y_e (\theta_e)$ (i.e., $y_e (\theta_e) = y_w (\theta_w (\theta_e))$). The first-order conditions for both types imply:

$$(A145) \quad 1 - T'_w (y_e (\theta_e)) = \frac{\phi'_w \left(\frac{y_e (\theta_e)}{W \chi_w (\theta_w (\theta_e))} \right)}{W \chi_w (\theta_w (\theta_e))}, \quad \text{and} \quad 1 - T'_e (y_e (\theta_e)) = \frac{\phi'_e (l_e (\theta_e))}{\frac{P(\theta_e)}{\mu(\theta_e)} \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}}.$$

Thus, under a uniform tax, the following policy constraint applies:

$$\frac{\phi'_w \left(\frac{y_e (\theta_e)}{W \chi_w (\theta_w (\theta_e))} \right)}{W \chi_w (\theta_w (\theta_e))} = \frac{\phi'_e (l_e (\theta_e))}{\frac{P(\theta_e)}{\mu(\theta_e)} \frac{\partial Q_{ij}(\theta_e)}{\partial l_e(\theta_e)}},$$

where $y_e (\theta_e) = P (Q_{ij} (\theta_e), \theta_e) Q_{ij} (\theta_e) - WL_w (\theta_e)$. The planner must treat $\theta_w (\theta_e)$ as an additional decision variable (see [Fu, Gong and Li \(2021\)](#) for solutions to this type of problem).

⁵⁸In a dynamic or open economy model, r could be exogenous or determined by capital production technology, making the inclusion of K more intuitive (e.g., see [Cui, Gong and Li \(2021\)](#)).

⁵⁹Under this illustration, $\phi_K (K, \theta_e)$ can still be treated as the utility cost of entrepreneurial effort, where the entrepreneurs use their knowledge to manage the factor inputs (more generally, one can take $\phi_K (K, L_w, \theta_e)$).

Beyond the complexity, there is also a conceptual reason for differentiating labor and profit taxes in our benchmark model. Entrepreneurial income often includes corporate profits, which are distinct from labor income. Even after profits are declared as personal income, they remain conceptually different from wages. Therefore, differentiated taxation is a realistic and practical starting point, especially when analyzing the impact of rising market power.

Both differentiated and uniform taxation approaches are well-explored in the literature (see [Scheuer \(2014\)](#) and [Rothschild and Scheuer \(2013\)](#)).⁶⁰ Given the real-world tax distinctions, we prioritize differentiated taxation in our analysis to better understand its role in addressing rising market power.

(III) ENDOGENOUS SOCIAL WELFARE WEIGHTS.

In much of the analysis above, we assume exogenous social welfare weights, thereby isolating the influence of market power on optimal taxation through the four key components. This abstraction allows us to focus on the direct effects of market power.

Our analysis of the laissez-faire economy suggests that entrepreneurs' gross utility generally increases with rising market power. An increase in market power, accompanied by higher markups, redistributes income from workers to firms (primarily through lower wages) and reduces overall welfare. If similar outcomes hold under optimal taxation, the marginal social welfare weights for entrepreneurs will decline, leading to higher optimal profit tax rates. This result aligns with findings in prior studies (e.g., [Kushnir and Zubrickas \(2019\)](#)). Additionally, a non-linear tax system can facilitate transfers between entrepreneurs and workers, contingent on the social welfare function. Under a utilitarian framework, the burden distribution is indeterminate, but it becomes determinate under a concave welfare function, where market power plays a critical role.

Segregating the impact of endogenous social welfare weights is essential because generalized welfare weights may depend on factors beyond gross utility, such as market structure and revenue contributions (see [Saez and Stantcheva \(2016\)](#); [Scheuer \(2014\)](#)). For example, welfare weights might reflect the social versus private value of entrepreneurship. By isolating this endogeneity, Proposition 3 remains applicable. Empirically, we show that the optimal top profit tax rate increases with rising markups, supporting prior studies' conclusions.

V. Quantitative Analysis

Figure A1 shows profit tax rates in terms of θ :

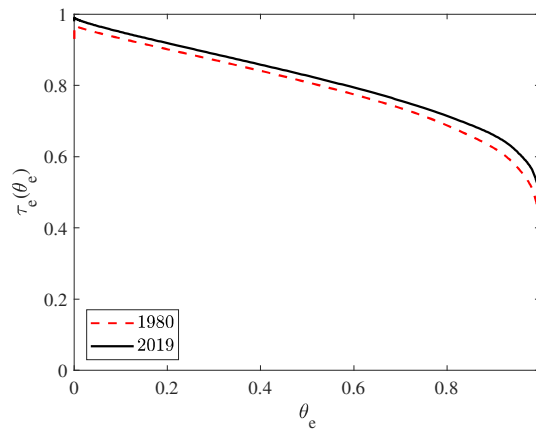


Figure A1. : Profit Tax Wedges

Figure A2 gathers results regarding the changes with k .

Figure A3 gathers changes in profit tax rates with the changes in ζ and σ .

Figure A4 gathers changes in labor income tax rates with the changes in ζ and σ .

⁶⁰[Scheuer \(2014\)](#) consider both cases, referring to the uniform policy constraint as a no-discrimination constraint.

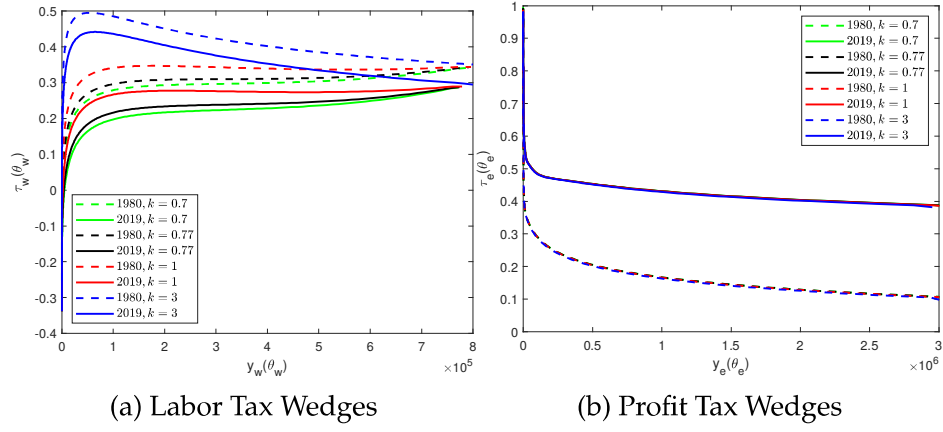


Figure A2. : Tax Wedges Change with k

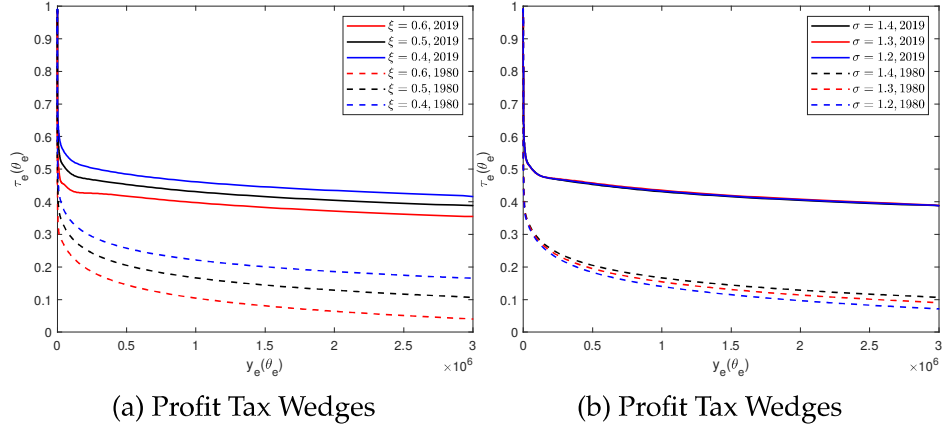


Figure A3. : Profit Tax Wedges Change with ζ and σ

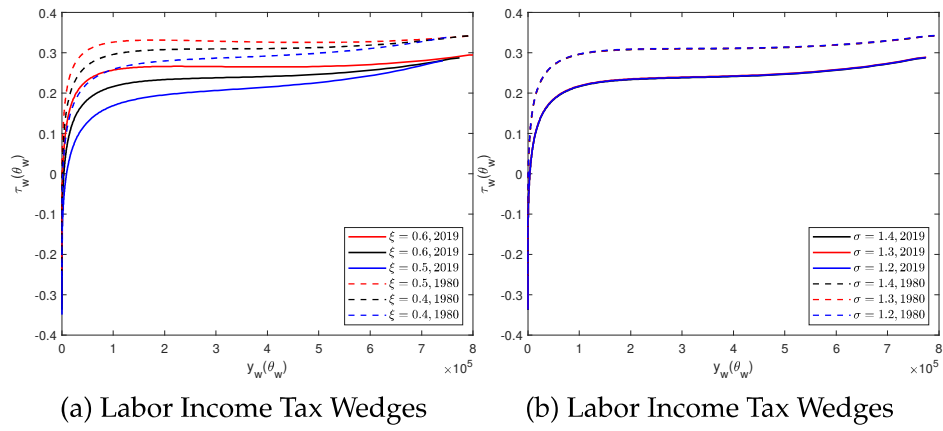


Figure A4. : Labor Income Tax Wedges Change with ζ and σ