ONLINE APPENDIX

Profits, Scale Economies, and the Gains from Trade and Industrial Policy

Ahmad Lashkaripour
Indiana University

Volodymyr Lugovskyy
Indiana University

August 2023

Contents

A The Redundancy of Consumption Taxes

Without loss of generality suppose country $i \in \mathbb{C}$ imposes a full set of tax instruments, while the rest of the world is passive. Now, consider any arbitrary combination of taxes (indexed by A) that includes (i) industrial (or domestic production) subsidies, $s^A_{i,k}$, (ii) domestic consumption taxes, $\tau^A_{i,k}$, (iii) import taxes, $t^A_{ji,k}$, and (iv) export subsidies, $x^A_{ij,k}$. This set of tax instruments –which includes consumption taxes– produces the following wedges between producer and consumer price indexes for various product varieties:

$$\tilde{P}^A_{ii,k} = \frac{1+\tau^A_{i,k}}{1+s^A_{i,k}} P_{ii,k}; \qquad \tilde{P}^A_{ji,k} = (1+t^A_{ji,k})(1+\tau^A_{i,k})P_{ji,k}; \qquad \tilde{P}^A_{ij,k} = \frac{1}{(1+x^A_{ij,k})(1+s^A_{i,k})} P_{ij,k}; \quad (j \neq i)$$

Our claim here is that the same wedges can be replicated without appealing to consumption taxes, $\tau_{i,k}$. This claim can be established by considering an alternative tax schedule, B, which excludes consumption taxes (i.e., $1+\tau^B_{i,k} = 0$), but includes the following set of production subsidies, export subsidies, and import taxes:

$$1+s^B_{i,k} = \frac{1+s^A_{i,k}}{1+\tau^A_{i,k}}; \qquad 1+t^B_{ji,k} = \left(1+t^A_{ji,k}\right)\left(1+\tau^A_{i,k}\right); \qquad 1+x^B_{ij,k} = (1+x^A_{ij,k})(1+\tau^A_{i,k})$$

It is straightforward to see that schedule B can reproduce the same wedge between producer and consumer prices as the original schedule A (i.e., $\tilde{\mathbf{P}}^A = \tilde{\mathbf{P}}^B$). In particular,

$$\begin{aligned}
\tilde{P}^B_{ii,k} &= \frac{1}{1+s^B_{i,k}} P_{ii,k} = \frac{1+\tau^A_{i,k}}{1+s^A_{i,k}} P_{ii,k} = \tilde{P}^A_{ii} \\
\tilde{P}^B_{ji,k} &= (1+t^B_{ji,k})P_{ji,k} = (1+t^A_{ji,k})(1+\tau^A_{i,k})P_{ji,k} = \tilde{P}^A_{ji,k} \\
\tilde{P}^B_{ij,k} &= \frac{1}{(1+x^B_{ij,k})(1+s^B_{i,k})} P_{ij,k} = \frac{1}{(1+x^A_{ij,k})(1+s^A_{i,k})} P_{ij,k} = \tilde{P}^A_{ij,k}.
\end{aligned}$$

It also follows trivially that $\tilde{P}^B_{nj,k} = P_{nj,k} = \tilde{P}^A_{nj,k}$ if $n, j \neq i$, because the rest of the world does not impose taxes.[1] This equivalence indicates that consumption taxes are redundant if the government has access to the other three sets of instruments. Note that the same can be said about production subsidies. More specifically, the effect of industry-level production subsidies can be perfectly replicated with a combination of consumption taxes, import taxes, and export subsidies. However, due to *product differentiation*, if two (of the $2(N-1)+2$) tax instruments are restricted, the replication argument fails. That is, if both production subsidies and consumption taxes are restricted, export subsidies and import taxes cannot fully replicate their effect.

[1] Note that the rest of the world imposing or not imposing taxes, does not matter for the redundancy of consumption taxes. The above argument can be easily extrapolated to the case where all countries impose arbitrary taxes.

B Proof of Lemma 1

Consider two policy-wage combinations, $\mathbf{T} = (\mathbf{s}, \mathbf{t}, \mathbf{x}; \mathbf{w})$, and $\mathbf{T}' = (\mathbf{s}', \mathbf{t}', \mathbf{x}'; \mathbf{w}')$, that differ in uniform shifters a and $\tilde{a} \in \mathbb{R}_+$:

$$\begin{cases} \mathbf{1} + \mathbf{x}'_i = a\left(\mathbf{1} + \mathbf{x}_i\right) & \mathbf{1} + \mathbf{x}'_{-i} = \mathbf{1} + \mathbf{x}_{-i} \\ \mathbf{1} + \mathbf{t}'_i = a\left(\mathbf{1} + \mathbf{t}_i\right) & \mathbf{1} + \mathbf{t}'_{-i} = \mathbf{1} + \mathbf{t}_{-i} \\ \mathbf{1} + \mathbf{s}'_i = \left(\mathbf{1} + \mathbf{s}_i\right)/\tilde{a} & \mathbf{1} + \mathbf{s}'_{-i} = \mathbf{1} + \mathbf{s}_{-i} \\ w'_i = \left(a/\tilde{a}\right) w_i & \mathbf{w}'_{-i} = \mathbf{w}_{-i} \end{cases}.$$

Our goal is to prove that (i) if $\mathbf{T} \in \mathbb{F}$ then $\mathbf{T}' \in \mathbb{F}$, and (ii) $W_n(\mathbf{T}) = W_n(\mathbf{T}')$ for all $n \in \mathbb{C}$. To prove these claims, we appeal to two intermediate lemmas. The first lemma establishes the following: Suppose equilibrium quantities are identical under policy-wage vectors $\mathbf{T}$ and $\mathbf{T}'$ (*i.e.*, $Q_{jn,k}(\mathbf{T}') = Q_{jn,k}(\mathbf{T})$ for all jn, k). Then, the implied nominal income and price levels under $\mathbf{T}$ and $\mathbf{T}'$ are the same up to a scale. The second lemma is a standard result from consumer theory: It indicates the nominal income and price levels implied by the first lemma confirm the original assumption that $Q_{jn,k}(\mathbf{T}') = Q_{jn,k}(\mathbf{T})$ for all jn, k. Below, we state and prove the first of these lemmas for any $a \in \mathbb{R}_+$.

Lemma 1. *$Q_{jn,k}(\mathbf{T}') = Q_{jn,k}(\mathbf{T})$ for all jn, k* $\Longrightarrow \begin{cases} \tilde{\mathbf{P}}_i\left(\mathbf{T}'\right) = a\tilde{\mathbf{P}}_i\left(\mathbf{T}\right); & \tilde{\mathbf{P}}_{-i}\left(\mathbf{T}'\right) = \tilde{\mathbf{P}}_{-i}\left(\mathbf{T}\right) \\ Y_i\left(\mathbf{T}'\right) = aY_i\left(\mathbf{T}\right); & \mathbf{Y}_{-i}\left(\mathbf{T}'\right) = \mathbf{Y}_{-i}\left(\mathbf{T}\right) \end{cases}$

Proof. Our goal is to compute nominal income and consumer prices under $\mathbf{T}$ and $\mathbf{T}'$ starting from the assumption that $Q_{jn,k}(\mathbf{T}') = Q_{jn,k}(\mathbf{T})$ for all jn, k. We start our proof with nominal prices: To simplify the notation, define $\delta_{jn,k}(\mathbf{T}) \equiv \overline{\rho}_{jn,k}\mathcal{Q}_{j,k}(\mathbf{T})^{-\frac{\mu_k}{1+\mu_k}}$. Note that –by assumption– $\delta_{jn,k}(\mathbf{T}) = \delta_{jn,k}(\mathbf{T}') = \overline{\delta}_{jn,k}$. First, consider the price of a typical good ji, k imported by i from origin $j \neq i$. Using Equations 6 and 7, the consumer price of ji, k under combination $\mathbf{T}'$ can be related to its price under $\mathbf{T}$ as follows:

$$\tilde{P}_{ji,k}(\mathbf{T}') = \overline{\delta}_{ji,k}\frac{1 + t'_{ji,k}}{(1 + x'_{ji,k})(1 + s'_{j,k})}w'_j = \overline{\delta}_{ji,k}\frac{a(1 + t_{ji,k})}{(1 + x_{ji,k})(1 + s_{j,k})}\, w_j = a\tilde{P}_{ji,k}(\mathbf{T}),$$

where the third equality follows from the fact that $1 + t'_{ji,k} = a(1 + t_{ji,k})$, while $w'_j = w_j$, $x'_{ji,k} = x_{ji,k}$, and $s'_{j,k} = s_{j,k}$ (since $w_j \in \mathbf{w}_{-i}$, $x_{ji,k} \in \mathbf{x}_{-i}$, and $s_{j,k} \in \mathbf{s}_{-i}$). Second, consider a typical good ii, k that is produced and consumed locally in country i. The consumer price of ii, k under combination $\mathbf{T}'$ can be related to its price under $\mathbf{T}$ as follows

$$\tilde{P}_{ii,k}(\mathbf{T}') = \overline{\delta}_{ii,k}\frac{1}{1 + s'_{i,k}}w'_i = \overline{\delta}_{ii,k}\frac{1}{\frac{1}{\tilde{a}}(1 + s_{i,k})} \times \frac{a}{\tilde{a}}w_i = a\tilde{P}_{ii,k}(\mathbf{T}),$$

where the third equality follows from the fact that $1 + s'_{i,k} = (1 + s_{i,k})/\tilde{a}$ and $w'_i = aw_i/\tilde{a}$. Third, consider the price of a typical good ij, k export by i to destination market $j \neq i$. The consumer price of ij, k under combination $\mathbf{T}'$ can be related to its price under $\mathbf{T}$ as follows:

$$\tilde{P}_{ij,k}(\mathbf{T}') = \overline{\delta}_{ij,k}\frac{1 + t'_{ij,k}}{(1 + x'_{ij,k})(1 + s'_{i,k})}w'_i = \overline{\delta}_{ij,k}\frac{1 + t_{ij,k}}{a(1 + x'_{ij,k}) \times \frac{1}{\tilde{a}}(1 + s'_{i,k})} \times \frac{a}{\tilde{a}}w_i = \tilde{P}_{ij,k}(\mathbf{T}),$$

where the third equality follows from the fact that $1 + x'_{ij,k} = a(1 + x_{ij,k})$, $1 + s_{i,k} = (1 + s'_{i,k})/\tilde{a}$, and $w'_i = aw_i/\tilde{a}$; while $t'_{ij,k} = t_{ji,k}$ since $t_{ji,k} \in \mathbf{t}_{-i}$. Lastly, is follows trivially that $\tilde{P}_{jn,k}(\mathbf{T}') = \tilde{P}_{jn,k}(\mathbf{T})$ if j and $n \neq i$. Considering that $\tilde{\mathbf{P}}_i = \left\{\tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ii}\right\}$, the above equations establish that

$$\tilde{\mathbf{P}}_i\left(\mathbf{T}'\right) = a\tilde{\mathbf{P}}_i\left(\mathbf{T}\right), \qquad \tilde{\mathbf{P}}_{-i}\left(\mathbf{T}'\right) = \tilde{\mathbf{P}}_{-i}\left(\mathbf{T}\right).$$

Next, we turn to our claim about nominal income levels. To simplify the presentation, we hereafter use $X \equiv X(\mathbf{T})$ and $X' \equiv X(\mathbf{T}')$ to denote the value of a generic variable X under policy-wage combinations $\mathbf{T}$ and $\mathbf{T}'$. Keeping in mind this choice of notation, country i's nominal income under $\mathbf{T}'$, i.e.,

$Y_i' \equiv Y_i(\mathbf{T}')$ is given by:

$$Y_i' = w_i' L_i + \sum_k \left[\left(\frac{1}{1+s_{i,k}'} - 1 \right) P_{ii,k}' Q_{ii,k}' \right] + \sum_k \sum_{j \neq i} \left(\frac{t_{ji,k}'}{(1+x_{ji,k}')(1+s_{j,k}')} P_{ji,k}' Q_{ji,k}' + \left[\frac{1}{(1+x_{ij,k}')(1+s_{i,k}')} - 1 \right] P_{ij,k}' Q_{ij,k}' \right)$$

$$= w_i' L_i + \sum_k \left[\left(1 - [1+s_{i,k}'] \right) \tilde{P}_{ii,k}' Q_{ii,k}' \right] + \sum_k \sum_{j \neq i} \left(\left(1 - \frac{1}{1+t_{ji,k}'} \right) \tilde{P}_{ji,k}' Q_{ji,k}' + \left[\frac{1}{1+t_{ij,k}'} - \frac{(1+x_{ij,k}')(1+s_{i,k}')}{1+t_{ij,k}'} \right] \tilde{P}_{ij,k}' Q_{ij,k}' \right).$$

Note that, by assumption, policy-wage combinations $\mathbf{T}$ and $\mathbf{T}'$ exhibit the same output schedule, i.e., $Q_{ii,k}' = Q_{ii,k}$, $Q_{ji,k}' = Q_{ji,k}$, and $Q_{ij,k}' = Q_{ij,k}$. Also, recall that ($\mathbf{T}$ and $\mathbf{T}'$ are constructed such that) $1 + t_{ji,k}' = a(1+t_{ji,k})$, $1 + x_{ij,k}' = a(1+x_{ij,k})$, $1 + s_{i,k} = (1+s_{i,k}')/\tilde{a}$, and $w_i' = a w_i/\tilde{a}$, $t_{ij,k}' = t_{ji,k}$. Considering these relationships and plugging our earlier result that (i) $\tilde{P}_{ii,k} = aP_{ii,k}$, (ii) $P_{ji,k}' = a\tilde{P}_{ji,k}$, and (iii) $\tilde{P}_{ij,k}' = \tilde{P}_{ij,k}$ into the above equation, yields the following expression for Y_i' :

$$Y_i' = \frac{a}{\tilde{a}} w_i L_i + \sum_k \left[\left(1 - \frac{1}{\tilde{a}}(1+s_{i,k}) \right) a\tilde{P}_{ii,k} Q_{ii,k} \right]$$
$$+ \sum_{j,k} \left[\left(1 - \frac{1}{a(1+t_{ji,k})} \right) a\tilde{P}_{ji,k} Q_{ji,k} + \left[\frac{1}{1+t_{ij,k}} - \frac{a(1+x_{ij,k}) \times \frac{1}{\tilde{a}}(1+s_{i,k})}{1+t_{ij,k}} \right] \tilde{P}_{ij,k} Q_{ij,k} \right].$$

Appealing to the balanced trade condition, $\sum_k \sum_{j \neq i} \left(\frac{1}{1+t_{ji,k}} \tilde{P}_{ji,k} Q_{ji,k} - \frac{1}{1+t_{ij,k}} \tilde{P}_{ij,k} Q_{ij,k} \right) = 0$, and observing that $(1+s_{i,k})\tilde{P}_{ii,k} = P_{ii,k}$ and $\frac{(1+x_{ij,k})(1+s_{i,k})}{1+t_{ij,k}} \tilde{P}_{ij,k} = P_{ij,k}$, the above equation reduces to

$$Y_i' = \frac{a}{\tilde{a}} w_i L_i + a \sum_k \left[\tilde{P}_{ii,k} Q_{ii,k} + \sum_{j \neq i} \tilde{P}_{ji,k} Q_{ji,k} \right] - \frac{a}{\tilde{a}} \sum_k \left[P_{ii,k} Q_{ii,k} + \sum_{j \neq i} P_{ij,k} Q_{ij,k} \right].$$

Invoking the labor market clearing condition, $w_i L_i - \sum_k \sum_n P_{ijn,k} Q_{in,k} = 0$, the above equation further simplifies as follows

$$Y_i' = a \sum_k \left[\tilde{P}_{ii,k} Q_{ii,k} + \sum_{j \neq i} \tilde{P}_{ji,k} Q_{ji,k} \right] = a \left[w_i L_i + \mathcal{R}_i \right] = a Y_i,$$

where $\mathcal{R}_i \equiv \mathcal{R}_i(\mathbf{T})$ denotes country i's tax revenues under $\mathbf{T}$. To bel clear, the third line, in the above equation, follows from country i's balanced budget condition (i.e., total expenditure = total income). Turning to the rest of the world: The fact that $Y_n(\mathbf{T}') = Y_n(\mathbf{T})$ for all $n \neq i$ follows trivially from a similar line of arguments–hence, establishing our claim about nominal income levels:

$$Y_i\left(\mathbf{T}'\right) = aY_i\left(\mathbf{T}\right); \qquad \mathbf{Y}_{-i}\left(\mathbf{T}'\right) = \mathbf{Y}_{-i}\left(\mathbf{T}\right)$$

□

Lemma 1 (proved above) starts from the assumption that $Q_{jn,k}(\mathbf{T}') = Q_{jn,k}(\mathbf{T})$ for all jn, k. Our next lemma indicates that this assumption is validated by the nominal income and price levels implied by $\mathbf{T}$ and $\mathbf{T}'$. Below, we state this lemma noting that it follows trivially from the Marshallian demand function, $Q_{ji,k} = \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$, being homogeneous of degree zero.

Lemma 2. *$\forall a \in \mathbb{R}_+$: $\begin{cases} \tilde{\mathbf{P}}_i\left(\mathbf{T}'\right) = a\tilde{\mathbf{P}}_i\left(\mathbf{T}\right) \\ Y_i(\mathbf{T}') = aY_i(\mathbf{T}) \end{cases} \implies Q_{ji,k}(\mathbf{T}') = Q_{ji,k}(\mathbf{T})$ for all ji, k*

Together, Lemmas 2 and 1 establish that equilibrium quantities should be indeed identical under policy-wage combinations $\mathbf{T}$ and $\mathbf{T}'$—i.e., $Q_{jn,k}(\mathbf{T}') = Q_{jn,k}(\mathbf{T})$ for all jn, k. Hence, if $\mathbf{T} \in \mathbb{F}$ it follows immediately that (i) $\mathbf{T}' \in \mathbb{F}$, and (ii) $W_n(\mathbf{T}) = W_n(\mathbf{T}')$ for all $n \in \mathbb{C}$, which is the claim of Lemma 1.

C Nested-Eaton and Kortum (2002) Framework

Here we show that the nested CES import demand function specified by Assumption (A1), can also arise from within-product specialization à la Eaton and Kortum (2002). To this end, suppose that each industry k is comprised of a continuum of homogenous goods indexed by ν. The sub-utility of the representative consumer in country i with respect to industry k is a log-linear aggregator across the continuum of goods in that industry:

$$Q_{i,k} = \int_0^1 \ln \tilde{q}_{i,k}(\nu) d\nu$$

As in our main model, country j hosts $\overline{M}_{j,k}$ firms indexed by ω, with $\Omega_{j,k}$ denotes the set of all firms serving industry k from country j.[2] Each firm ω supplies good ν to market i at the following *quality-adjusted* price:

$$\tilde{p}_{ji,k}(\nu;\omega) = \tilde{p}_{ji,k}(\omega) / \varphi(\nu;\omega),$$

where $\tilde{p}_{ji,k}(\omega)$ is a nominal price (driven by production costs) that applies to all goods supplied by firm ω in industry k, while the quality component, $\varphi(\nu;\omega)$, is good×firm-specific. Suppose for any given good ν, firm-specific qualities are drawn independently from the following nested Fréchet joint distribution:

$$F_k(\boldsymbol{\varphi}(\boldsymbol{\nu})) = \exp\left[-\sum_{i=1}^{N}\left(\sum_{\omega\in\Omega_{i,k}} \varphi(\nu;\omega)^{-\vartheta_k}\right)^{\frac{\theta_k}{\vartheta_k}}\right],$$

The above distribution generalizes the basic Fréchet distribution in Eaton and Kortum (2002). In particular, it relaxes the restriction that productivities are perfectly correlated across firms within the same country. Instead, it allows for sub-national productivity differentiation and also for the degrees of cross- and sub-national productivity differentiations (ϑ_k and θ_k, respectively) to diverge. A special case of the distribution where $\vartheta_k \longrightarrow \infty$ corresponds to the standard Eaton and Kortum (2002) specification.

The above distribution also has deep theoretical roots. The Fisher–Tippett–Gnedenko theorem states that if ideas are drawn from a (normalized) distribution, in the limit the distribution of the best draw takes the form of a general extreme value (GEV) distribution, which includes the above Fréchet distribution as a special case. A special application of this result can be found in Kortum (1997) who develops an idea-based growth model where the limit distribution of productivities is Fréchet, with $\varphi_{\omega,k}$ reflecting the stock of technological knowledge accumulated by firms ω in category k.

Given the vector of effective prices, the representative consumer in county i (who is endowed with income Y_i) maximizes their real consumption of each good, $\tilde{q}_{i,k}(\nu) = e_{i,k}Y_i/\tilde{p}_{i,k}(\nu)$, by choosing $\tilde{p}_{i,k}(\nu) = \min_\omega \left\{\tilde{p}_{ji,k}(\omega)\right\}$. That being the case, the consumer's discrete choice problem for each good ν can be expressed as:

$$\min_{\omega} \; \tilde{p}_{ji,k}(\omega) / z(\nu;\omega) \sim \max_{\omega} \; \ln z(\nu;\omega) - \ln \tilde{p}_{ji,k}(\omega).$$

To determine the share of goods for which firm ω is the most competitive supplier, we can invoke the theorem of "General Extreme Value." Specifically, define $G(\tilde{\boldsymbol{p}}_i)$ as follows

$$G_k(\tilde{\boldsymbol{p}}_i) = \sum_{j=1}^{N}\left(\sum_{\omega\in\Omega_{j,k}} \exp(-\vartheta_k \ln \tilde{p}_{ji,k}(\omega))\right)^{\frac{\theta_k}{\vartheta_k}} = \sum_{j=1}^{N}\left(\sum_{\omega\in\Omega_{j,k}} \tilde{p}_{ji,k}(\omega)^{-\vartheta_k}\right)^{\frac{\theta_k}{\vartheta_k}}.$$

Note that $G_k(.)$ is a continuous and differentiable function of vector $\tilde{\boldsymbol{p}}_i \equiv \left\{\tilde{p}_{ji,k}(\omega)\right\}$ and has the following properties:

i. $G_k(.) \geq 0$;

ii. $G_k(.)$ is a homogeneous function of rank θ_k: $G_k(\rho\tilde{\mathbf{p}}_i) = \rho^{\theta_k} G_k(\tilde{\mathbf{p}}_i)$ for any $\rho \geq 0$;

iii. $\lim_{\tilde{p}_{ji,k}(\omega)\to\infty} G_k(\tilde{\mathbf{p}}_i) = \infty$, $\forall\omega$;

iv. the m'th partial derivative of $G_k(.)$ with respect to a generic combination of m variables $\tilde{p}_{ji,k}(\omega)$, is non-negative if m is odd and non-positive if m is even.

[2] The implicit assumption here is that entry is restricted, so that $\overline{M}_{j,k}$ is exogenous.

Manski and McFadden (1981) prove that if $G_k(.)$ satisfies the above conditions, and $\varphi(\nu;\omega)$'s are drawn from distribution,

$$F_k(\boldsymbol{\varphi}(\nu)) = \exp\left(-G_k(e^{-\ln \boldsymbol{\varphi}})\right) = \exp\left(-\sum_{j=1}^{N}\left(\sum_{\omega\in\Omega_{j,k}} \varphi(\nu;\omega)^{-\vartheta_k}\right)^{\frac{\theta_k}{\vartheta_k}}\right),$$

then the probability of choosing variety ω (from origin j in industry k) is given by

$$\pi_{ji,k}(\omega) = \frac{\left(\frac{\tilde{p}_{ji,k}(\omega)}{\theta_k}\right)\frac{\partial G_k(\tilde{\mathbf{p}}_i)}{\partial p_{ji,k}(\omega)}}{G_k(\tilde{\mathbf{p}}_i)} = \frac{\tilde{p}_{ji,k}\left(\omega\right)\tilde{p}_{ji,k}\left(\omega\right)^{\vartheta_k-1}\left(\sum_{\omega'\in\Omega_{j,k}}\tilde{p}_{ji,k}\left(\omega'\right)^{-\vartheta_k}\right)^{\frac{\theta_k}{\vartheta_k}-1}}{\sum_{n=1}^{N}\left(\sum_{\omega'\in\Omega_{j,k}}\tilde{p}_{ji,k}\left(\omega'\right)^{-\vartheta_k}\right)^{\frac{\theta_k}{\vartheta_k}}}.$$

Rearranging the above equation yields the following expression for probability shares,

$$\pi_{ji,k}(\omega) = \left(\frac{\tilde{p}_{ji,k}\left(\omega\right)}{\tilde{P}_{ji,k}}\right)^{-\vartheta_k}\left(\frac{\tilde{P}_{ji,k}}{\tilde{P}_{i,k}}\right)^{-\theta_k},$$

where $\tilde{P}_{ji,k} \equiv \left[\sum_{\omega'\in\Omega_{j,k}}\tilde{p}_{ji,k}\left(\omega'\right)^{-\vartheta_k}\right]^{-1/\vartheta_k}$ and $\tilde{P}_{i,k} \equiv \left[\sum \tilde{P}_{ji,k}^{-\theta_k}\right]^{-\frac{1}{\theta_k}}$. Given that the probability shares coincide with the share of goods sourced from firm ω, total sales of firm ω to market i, in industry k can be calculated as:

$$\tilde{p}_{ji,k}(\omega)q_{ji,k}(\omega) = \tilde{p}_{ji,k}(\omega)\,\frac{\pi_{ji,k}(\omega)e_{i,k}Y_i}{\tilde{p}_{ji,k}(\omega)} = \left(\frac{\tilde{p}_{ji,k}\left(\omega\right)}{\tilde{P}_{ji,k}}\right)^{-\vartheta_k}\left(\frac{\tilde{P}_{ji,k}}{\tilde{P}_{i,k}}\right)^{-\theta_k} e_{i,k}Y_i$$

which is identical to the nested-CES function specified by Assumption (A1), with corresponding substitution parameters $\gamma_k - 1 = \vartheta_k$ and $\sigma_k - 1 = \theta_k$.

D Firm-Selection under Melitz-Pareto

In this appendix, we outline the isomorphism between our baseline model and one that admits selection effects. In doing so, we borrow heavily from Kucheryavyy, Lyn, and Rodríguez-Clare (2023a) (KLR, hereafter). We rely on three key assumptions, hereafter:

i. Within-industry demand is governed by the same nested-CES utility function presented under Assumption (A1). As in the baseline mode, σ_k and γ_k respectively denote the upper- and lower-tier elasticities of substitution.

ii. The firm-level productivity distribution, $G_{i,k}(z)$, is Pareto with shape parameter, ϑ_k.

iii. The fixed "marketing" cost is paid in terms of labor in the destination market.

iv. Taxes are applied before the markup, and operate as a cost-shifter.

Following KLR, we also assume that cross-industry utility aggregator is Cobb-Douglas, with $e_{i,k}$ denoting the constant share of country i's expenditure on industry k. Following the derivation in KLR, we can define the effective supply of production labor in country i as

$$\widetilde{L}_i = \left[1 - \sum_k e_{i,k}\left(\frac{\vartheta_k - \gamma_k + 1}{\vartheta_k\gamma_k}\right)\right] L_i.$$

The labor market clearing condition is, accordingly, given by $\sum w_i L_{i,k} = w_i \widetilde{L}_i$. With regards to aggregate markup levels, we can appeal to the well-known result that the profit margin in each industry is constant and given by the following expression:

$$\text{mark-up} \sim \frac{\sum_n P_{in,k}Q_{in,k}}{w_i L_{i,k}} = \frac{\gamma_k\vartheta_k}{\left(\gamma_k - 1\right)\left(\vartheta_k + 1\right) - \vartheta_k}.$$

With regards to aggregate demand functions, we can follow the derivation in Appendix B.2 of KLR to express demand for national-level variety ji,k as

$$Q_{ji,k}=\left(\frac{\tilde{P}_{ji,k}}{\tilde{P}_{i,k}}\right)^{-\sigma_k^{\text{Melitz}}}Q_{i,k},$$

where $\sigma_k^{\text{Melitz}}\equiv 1+\vartheta_k\left[1+\vartheta_k\left(\frac{1}{\sigma_k-1}-\frac{1}{\gamma_k-1}\right)\right]^{-1}$ denotes the trade elasticity under firm-selection. Moreover, we can show that national-level producer price indexes are given by the following formulation:

$$P_{ij,k}^{\text{Melitz}}=\begin{cases}\bar{\varrho}_{ij,k}w_i & \text{if entry is restricted}\\ \bar{\varrho}'_{ij,k}w_iQ_{i,k}^{-\frac{\vartheta_k}{1+\vartheta_k}} & \text{if entry is free}\end{cases},$$

where $\bar{\varrho}_{ij,k}$ and $\bar{\varrho}'_{ij,k}$ are composed of structural parameters that are invariant to policy–this includes ϑ_k that regulates firm selection.[3] Abstracting from taxes, $\tilde{P}_{i,k}=\left(\sum P_{ji,k}^{1-\sigma_k}\right)^{\frac{1}{1-\sigma_k}}$ is the CES industry-level consumer price index that shows up in indirect utility $V_i(.)$. Referring to our earlier result about constant markup margins, aggregate profits in country i given by

$$\Pi_i^{\text{Melitz}}=\begin{cases}\sum_k\sum_j\left(\frac{\frac{\gamma_k\vartheta_k}{(\gamma_k-1)(\vartheta_k+1)-\vartheta_k}}{1+\frac{\gamma_k\vartheta_k}{(\gamma_k-1)(\vartheta_k+1)-\vartheta_k}}P_{ij,k}Q_{ij,k}\right) & \text{if entry is restricted}\\ 0 & \text{if entry is free}\end{cases}.$$

To fixe ideas, recall that we used μ_k to denote both (1) the scale elasticity under free entry, and (2) the profit margin under restricted entry in the baseline model. This overlapping choice of notation was motivated by the observation that in the generalized Krugman model, the scale elasticity (under free entry) and the profit margin (under restricted entry) are identical and equal to $\mu_k=\frac{1}{\gamma_k-1}$. This equivalence, though, was not used to derive any of our theorems. Instead, it was only invoked to simplify the presentation of our theorems. Evidently, under the Melitz-Pareto model the equivalence between the scale elasticity and the profit margin crumbles. Taking note of this nuance, the Melitz-Pareto model is isomorphic to our baseline model with the following reinterpretation of parameters:

$$1+\mu_k^{\text{Melitz}}=\begin{cases}1+\frac{1}{\vartheta_k} & \text{if entry is free}\\ \frac{\gamma_k\vartheta_k}{(\gamma_k-1)(\vartheta_k+1)-\vartheta_k} & \text{if entry is restricted}\end{cases};\qquad \sigma_k^{\text{Melitz}}=1+\frac{\vartheta_k}{1+\vartheta_k\left(\frac{1}{\sigma_k-1}-\frac{1}{\gamma_k-1}\right)}.$$

The Marshallian demand elasticities in the Melitz-Pareto model are accordingly given by the following equations as a function σ_k^{Melitz} and expenditure shares:

$$\varepsilon_{ji,k}^{(ji,k)}=-1-(\sigma_k^{\text{Melitz}}-1)\left(1-\lambda_{ji,k}\right);\qquad \varepsilon_{ji,k}^{(ji,k)}=\sigma_k^{\text{Melitz}}\lambda_{ji,g}$$

In the above expressions, γ_k and σ_k can be taken directly from our firm-level demand estimation. Doing so, identifies the Melitz-Pareto model's key parameters up to a Pareto shape parameter, ϑ_k. To obtain an estimate for ϑ_k, we can estimate the trade elasticity, $\sigma_k^{\text{Melitz}}-1$, using macro-level trade data and standard techniques from the literature. Given the estimated trade elasticities, we can simply recover ϑ_k by plugging our micro-level estimates for γ_k and σ_k into the expression for σ_k^{Melitz}.

D.1 The Case where Taxes are Applied After Markups

Our derivation, above, assumed that taxes are applied before the markup, and act as a cost shifter. Below, we discuss how relaxing this assumption may affect the arguments listed above. To this end, we focus on the spacial case where preferences are non-nested. Namely,

$$\text{non-nested preferences}\sim\sigma_k=\gamma_k,\qquad\forall k\in\mathbb{K}.$$

Following the Online Appendix 5 in Costinot and Rodríguez-Clare (2014), the trade elasticity in the Melitz-Pareto model with non-nested preferences is described by the following formulation:

[3] Unlike $\tilde{P}_{i,k}$, the national-level indexes, $\tilde{P}_{ji,k}$, are not the same as the CES price indexes defined in the main text, but this is not problematic from the point of the isomorphism result we are seeking.

$$\sigma_k^{\text{Melitz}} = \begin{cases} 1+\vartheta_k & \text{tax applied before markup} \\ \frac{\sigma}{\sigma-1}\vartheta & \text{tax applied after markup} \end{cases}.$$

Appealing to the above formulation, we can show that *Theorem 1* nests, as a special case, the optimal tariff formula derived by Demidova and Rodriguez-Clare (2009) for a small open economy in a *single-industry×two-country* Melitz-Pareto model. To demonstrate this, drop the industry subscript k and reduce the global economy into two countries, i.e., $\mathbb{C} = \{i, j\}$. Noting that $1 - \lambda_{ij} = \lambda_{jj}$ in the two-country case, we can deduce from the above formulation and Theorem 1 that

$$\frac{1+t_{ji}^*}{1+x_{ij}^*} = 1 + \frac{1}{(\sigma^{\text{Melitz}}-1)\lambda_{jj}} = \frac{1}{(\frac{\sigma}{\sigma-1}\vartheta - 1)\lambda_{jj}}.$$

By the Lerner symmetry, export and import taxes are equivalent in the single-industry model.[4] Hence, without loss of generality, we can set $x_{ij}^* = 0$. Moreover, if country i is a small open economy, then $\lambda_{jj} \approx 1$. Combining these two observations, we can arrive at the familiar-looking optimal tariff formula in Demidova and Rodriguez-Clare (2009):

$$t_{ji}^* = \frac{\frac{\sigma-1}{\sigma}}{\vartheta - \frac{\sigma-1}{\sigma}} \sim \text{small open economy w/ one traded sector.}$$

E Proof of Theorem 1

Our proof proceeds in five steps. The first four steps characterize the optimal tax/subsidy schedule for country $i \in \mathbb{C}$ under *free entry*. The last step demonstrates that this characterization can be extrapolated to the case with *restricted entry*.

Step #1: Express Equilibrium Variables as function of $\tilde{\mathbb{P}}_i$ and w

Our goal is to characterize optimal policy for country $i \in \mathbb{C}$ assuming the rest of the world is passive in their use of taxes: $\mathbf{t}_{-i} = \mathbf{x}_{-i} = \mathbf{s}_{-i} = \mathbf{0}$. To simplify the proof, we reformulate country i's optimal policy problem as one where the government chooses the optimal consumer prices (rather than the actual taxes) associated with its economy. By construction, country i's optimal tax schedule can be recovered from its optimal consumer-to-producer price ratios. The first step in reformulating the optimal policy problem is to express equilibrium variables (e.g., $Q_{ji,k}$, Y_i, etc.) as a function of (1) the vector of consumer prices associated with economy i, $\tilde{\mathbb{P}}_i \equiv \left\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij}\right\}$, where

$$\tilde{\mathbf{P}}_{ii} \equiv \left\{P_{ii,k}\right\}_k; \qquad \tilde{\mathbf{P}}_{ji} \equiv \left\{P_{ji,k}\right\}_{j\neq i,k}; \qquad \tilde{\mathbf{P}}_{ij} \equiv \left\{P_{ij,k}\right\}_{j\neq i,k} \tag{E.1}$$

and (2) the vector of national-level wage rates across the world,

$$\mathbf{w} = \{w_1, ..., w_N\}.$$

The following lemma shows that our desired formulation of equilibrium variables follows from (a) treating $\tilde{\mathbb{P}}_i$ and $\mathbf{w}$ as given, and (b) solving a system that satisfies all equilibrium conditions excluding the labor market clearing condition.

Lemma 3. *All equilibrium outcomes (excluding $\tilde{\mathbb{P}}_i$ and $\mathbf{w}$) can be uniquely determined as a function of $\tilde{\mathbb{P}}_i \equiv \left\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij}\right\}$, and $\mathbf{w}$.*

Proof. As noted above, the proof follows from solving all equilibrium conditions excluding the equilibrium expression for consumer prices, $\tilde{P}_{ji,k}$ (which are encompassed by $\tilde{\mathbb{P}}_i$), and the country-specific balanced trade conditions (which pin down $\mathbf{w}$).[5] Stated formally, we need to solve the following system treating $\tilde{\mathbb{P}}_i$, and $\mathbf{w}$ as given:

[4] The Lerner symmetry is a special case of the equivalence result presented under Lemma 1. Also, note that the market equilibrium is efficient in the single industry Krugman model studied by Gros (1987). As such, the optimal industrial subsidy can be normalized to zero, i.e., $s_i^* = 0$.

[5] Note that by Walras' law, the balanced trade condition is equivalent to the labor market clearing condition in each country.

$$
\begin{aligned}
&\text{[optimal pricing]} && P_{jn,k} = \bar{\rho}_{ji,k} w_j \left[\sum_i \bar{a}_{ji,k} Q_{ji,k}\right]^{-\frac{\mu_k}{1+\mu_k}} \\
&\text{[optimal consumption]} && Q_{jn,k} = \mathcal{D}_{jn,k}(Y_n, \tilde{\mathbf{P}}_{1n}, ...\tilde{\mathbf{P}}_{Nn}) \\
&\text{[RoW imposes zero taxes]} && \tilde{P}_{jn,k} = P_{jn,k} \quad (\tilde{P}_{jn,k} \notin \tilde{\mathbb{P}}_i); \qquad Y_n = w_n L_n \quad (n \neq i) \\
&\text{[Balanced Budget in } i\text{]} && Y_i = w_i L_i + \left(\tilde{\mathbf{P}}_{ii} - \mathbf{P}_{ii}\right) \cdot \mathbf{Q}_{ii} + \left(\tilde{\mathbf{P}}_{ij} - \mathbf{P}_{ij}\right) \cdot \mathbf{Q}_{ij} + \left(\tilde{\mathbf{P}}_{ji} - \mathbf{P}_{ji}\right) \cdot \mathbf{Q}_{ji},
\end{aligned}
$$

where "$\cdot$" denotes the inner product operator for equal-sized vectors (i.e., $\mathbf{a} \cdot \mathbf{b} = \sum_n a_n b_n$). Since there is a unique equilibrium, the above system is exactly identified in that it uniquely determines $P_{jn,k}$, $Q_{jn,k}$, and Y_n as a function of $\tilde{\mathbb{P}}_i$ and $\mathbf{w}$. □

Following Lemma 3, we can express total income in country i, Y_i, as well as the entire demand schedule in that country as follows:

$$
Y_i \equiv Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}); \qquad Q_{ji,k} \equiv Q_{ji,k}(\tilde{\mathbb{P}}_i; \mathbf{w}) = \mathcal{D}_{ji,k}\left(Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}), \tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}\right).
$$

Recall that $\mathcal{D}_{ji,k}(.)$ denotes the Marshallian demand function facing variety ji, k. Taking note of the above representation, our main objective is to reformulate country i's policy problem as one where the government chooses $\tilde{\mathbb{P}}_i$ (as opposed to directly choosing tax rates). This reformulation, though, needs to take into account that $\mathbf{w}$ is an equilibrium outcome that implicitly depends on the choice of $\tilde{\mathbb{P}}_i$. To track this constraint, we define the $(\tilde{\mathbb{P}}_i; \mathbf{w})$ combinations that are feasible as follows.

Definition 1. *A policy-wage combination $(\tilde{\mathbb{P}}_i; \boldsymbol{w})$ is **feasible** iff given $\tilde{\mathbb{P}}_i$, the vector of wages, $\boldsymbol{w}$, satisfies the balanced trade condition in every country $n \in \mathbb{C}$. In particular,*

$$
(\tilde{\mathbb{P}}_i; \mathbf{w}) \in \mathbb{F}_P \iff \begin{cases} \sum_{j \neq n} \sum_{k \in \mathbb{K}} \left[P_{jn,k}(\tilde{\mathbb{P}}_i; \mathbf{w}) Q_{jn,k}(\tilde{\mathbb{P}}_i; \mathbf{w}) - P_{nj,k}(\tilde{\mathbb{P}}_i; \mathbf{w}) Q_{nj,k}(\tilde{\mathbb{P}}_i; \mathbf{w})\right] = 0 & \text{if } n \neq i \\ \sum_{j \neq n} \sum_{k=1}^{K} \left[P_{ji,k}(\tilde{\mathbb{P}}_i; \mathbf{w}) Q_{jn,k}(\tilde{\mathbb{P}}_i; \mathbf{w}) - \tilde{P}_{ij,k} Q_{nj,k}(\tilde{\mathbb{P}}_i; \mathbf{w})\right] = 0 & \text{if } n = i \end{cases}.
$$

To elaborate on the above definition: The balanced trade condition for a generic country $n \in \mathbb{C}$ can be expresses as $\sum_{j \neq n,k} \left[\frac{1}{1+t_{jn,k}} \tilde{P}_{nj,k} Q_{jn,k} - \frac{1}{1+t_{nj,k}} \tilde{P}_{nj,k} Q_{nj,k}\right]$. The expression for the balanced trade condition, above, follows from the assumption that only country i imposes taxes and the rest of the world is passive. We should emphasize one more time that by Walras' law the satisfaction of the balanced trade condition is analogous to the satisfaction of the labor market clearing condition in each country. Relatedly, take note of the equivalence between $\mathbb{F}_P$ and $\mathbb{F}$–with the latter being defined in the main text under Definition (D2). Taking note of these implicit details, we now proceed to reformulate the optimal policy problem (P1).

Step #2: Reformulate the Optimal Tariff Problem

Before proceeding with the second step of the proof, we formally present our notation for partial derivatives. We will rely heavily on this choice of notation, especially in the subsequent steps of the proof where we derive the first-order conditions.

Notation [*Partial Derivative*] *Let $f(x_1, x_2)$ be a function of two variables, where $x_2 = g(x_1)$ is possibly an implicit function of x_1. We henceforth use*

$$
\left(\frac{\partial f(x_1, x_2)}{\partial x_1}\right)_{x_2} = \frac{\partial f(x_1, \bar{x}_2)}{\partial x_1}
$$

to denote the derivative of $f(.)$ w.r.t. x_1, treating $x_2 = \bar{x}_2$ as a constant.[6]

Moving on with Step 2, recall the original formulation of the optimal policy problem (P1) from Section I:

$$
\max_{\mathbb{T}_i} \; W_i(\mathbb{T}_i; \mathrm{w}) \qquad s.t. \;\; (\mathbb{T}_i; \mathrm{w}) \in \mathbb{F} \qquad \text{(P1)}
$$

[6]Based on the above notation and the chain rule, the full derivative of $f(.)$ w.r.t. x_1 is given by

$$
\frac{\mathrm{d}f(x_1, x_2)}{\mathrm{d}x_1} = \left(\frac{\partial f(x_1, x_2)}{\partial x_1}\right)_{x_2} + \left(\frac{\partial f(x_1, x_2)}{\partial x_2}\right)_{x_2} \frac{\mathrm{d}g(x_1)}{\mathrm{d}x_1}
$$

In the above formulation, $\mathbb{T}_i \equiv (\mathbf{t}_i, \mathbf{x}_i, \mathbf{s}_i)$ denotes country i's vector of taxes and $\mathbb{F}$ is defined according to Definition (D2, Section I) and analogously to $\mathbb{F}_P$. Our next intermediate result shows that Problem (P1) can be alternatively cast as one where the government chooses the optimal vector of consumer prices $\tilde{\mathbb{P}}_i$ associated with its economy. After determining $\tilde{\mathbb{P}}_i$, the optimal tax vectors, $\mathbf{t}_i^*$, $\mathbf{x}_i^*$, and $\mathbf{s}_i^*$ can be automatically recovered from the optimal consumer-to-producer price ratios.

Lemma 4. *Country i's vector of optimal taxes, $\{\mathbf{t}_i^*, \mathbf{x}_i^*, \mathbf{s}_i^*\}$, can be determined by solving the following problem instead of (P1):*

$$\max_{\tilde{\mathbb{P}}_i} \; W_i(\tilde{\mathbb{P}}_i; \mathrm{w}) \equiv V_i(Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}), \tilde{\mathbf{P}}_i) \quad s.t. \begin{cases} (\tilde{\mathbb{P}}_i; \mathbf{w}) \in \mathbb{F}_P \\ \mathbf{w}_{-i} = \overline{\mathbf{w}}_{-i}, \end{cases} \qquad (\widetilde{\text{P1}}),$$

where $\overline{\mathbf{w}}_{-i}$ denotes the vector wages in the rest of the world under the status quo.

Proof. The proof consists of two parts. First, we can verify that there is a one-to-one correspondence between the optimal choice w.r.t. $\tilde{\mathbb{P}}_i \equiv \left\{\tilde{\mathbf{P}}_{ii}^*, \tilde{\mathbf{P}}_{ji}^*, \tilde{\mathbf{P}}_{ij}^*\right\}$ and $\mathbb{T}_i^* \equiv \{\mathbf{t}_i^*, \mathbf{x}_i^*, \mathbf{s}_i^*\}$. More specifically, given information on $\tilde{\mathbb{P}}_i$ (and the accompanying wage vector $\mathbf{w}^*$), we can uniquely recover the optimal tax/subsidy rates using the following set of equations:

$$1 + t_{ji,k}^* = \frac{\tilde{P}_{ji,k}^*}{P_{ji,k}(\tilde{\mathbb{P}}_i, \mathbf{w}^*)}; \qquad 1 + x_{ij,k}^* = \frac{P_{ji,k}(\tilde{\mathbb{P}}_i, \mathbf{w}^*)/\tilde{P}_{ji,k}^*}{P_{ii,k}(\tilde{\mathbb{P}}_i, \mathbf{w}^*)/\tilde{P}_{ii,k}^*}; \qquad 1 + s_{i,k}^* = \frac{P_{ii,k}(\tilde{\mathbb{P}}_i, \mathbf{w}^*)}{\tilde{P}_{ii,k}^*}.$$

The correspondence presented above, indicates an equivalence between choosing $\tilde{\mathbb{P}}_i$ versus choosing $\mathbb{T}_i$ directly. That is,

$$\max_{\tilde{\mathbb{P}}_i} \; W_i(\tilde{\mathbb{P}}_i; \mathrm{w}) \;\; s.t. \; (\tilde{\mathbb{P}}_i; \mathbf{w}) \in \mathbb{F}_P \sim \max_{\mathbb{T}_i} \; W_i(\mathbb{T}_i; \mathrm{w}) \;\; s.t. \; (\mathbb{T}_i; \mathbf{w}) \in \mathbb{F}.$$

Second, we must rationalize the constraint on foreign wages, $\mathbf{w}_{-i} = \overline{\mathbf{w}}_{-i}$. This constraint in the two-country case ($N = 2$) follows directly from Walras' law. Beyond that, it follows from Walras' law *and* the assumption that cooperative buffers in the rest of the world preserve relative wages between other countries, i.e., $\mathrm{d}\ln\left(w_n/w_j\right) = 0$ for all $n, j \neq i$ (see Appendix G for details).[7] □

Step #3. Deriving and Simplifying the System of First-Order Conditions

This step derives and solves the system of first-order necessary conditions (F.O.C.s) associated with Problem $\widetilde{\text{P1}}$. This system of F.O.C.s can be formally expressed as follows:

$$\nabla_{\tilde{P}} W_i(\tilde{\mathbb{P}}_i; \mathbf{w}) + \nabla_{\mathbf{w}} W_i \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\tilde{P}}\right)_{(\tilde{\mathbb{P}}_i; \mathbf{w}) \in \mathbb{F}_P} = 0, \qquad \forall \tilde{P} \in \tilde{\mathbb{P}}_i.$$

where recall that $\tilde{\mathbb{P}}_i = \left\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ij}, \tilde{\mathbf{P}}_{ji}\right\}$ includes all consumer price variables associated with economy i. To elaborate the right-hand side of the above equation consists of two terms, as implied by the chain rule: The first term accounts for the change in welfare holding $\mathbf{w}$ fixed. The second term account for the change in $\mathbf{w}$ w.r.t. $\tilde{P} \in \tilde{\mathbb{P}}_i$ in order to satisfy feasibility.

Our characterization of optimal policy employs the dual approach, the presentation of which relies heavily on Marshallian demand elasticities. So, for future reference, we formally define these elasticities below.

[7] As noted in Appendix G, the constraint $\mathbf{w}_{-i} = \overline{\mathbf{w}}_{-i}$, holds in some canonical special cases of our framework irrespective of cooperative wage buffers in the RoW. One can also show that the constraint $\mathbf{w}_{-i} = \overline{\mathbf{w}}_{-i}$ is non-binding *at the optimum* if trade is bilaterally balanced. In particular, specify country i's welfare as $W_i\left(\tilde{\mathbb{P}}_i, w_i, \mathbf{w}_{-i}\right) = V_i\left(w_i L_i + \mathcal{R}_i\left(\tilde{\mathbb{P}}_i, w_i, \mathbf{w}_{-i}\right), \tilde{\mathbf{P}}_i\right)$, where recall that $\tilde{\mathbf{P}}_i \subset \tilde{\mathbb{P}}_i$. Taking partial derivatives w.r.t. $w_n \in \mathbf{w}_{-i}$ and noting that $Y_n = w_n \bar{L}_n$, yields $\frac{\partial W_i(\tilde{\mathbb{P}}_i, w_i, \mathbf{w}_{-i})}{\partial w_n} = \frac{\partial \mathcal{R}_i(\tilde{\mathbb{P}}_i, w_i, \mathbf{w}_{-i})}{\partial w_n} - \sum_k \left[P_{ji,k} Q_{ji,k}\right] + \sum_k \left[\left(\tilde{P}_{ij,k} - P_{ij,k}\right) Q_{ij,k} \eta_{ij,k}\right] + \sum_{n \neq i} \sum_{k,g} \left[\left(\tilde{P}_{in,k} - P_{in,k}\right) Q_{in,k} \varepsilon_{in,k}^{(jn,g)}\right]$. Now, suppose the gross trade matrix is bilaterally balanced, $\sum_k \left[P_{ji,k} Q_{ji,k} - \tilde{P}_{ij,k} Q_{ij,k}\right]$. Then, one can invoke the optimality condition w.r.t. $\tilde{\mathbf{P}}_{in}$ (Equation E.21) and appeal to the Slutsky symmetry, $\varepsilon_{in,k}^{(jn,g)} = e_{jn,g} \varepsilon_{jn,g}^{(in,k)} / e_{in,k}$ and the demand function's homogeneity of degree zero property $\varepsilon_{ij,k}^{(ij,k)} + = \eta_{ij,k} + \sum_{n,g \neq i,k} e_{ij,k}^{(nj,g)}$ to get $\frac{\partial W_i(.)}{\partial w_n} |_{\tilde{\mathbb{P}}_i = \tilde{\mathbb{P}}_i^*} = 0$.

Notation [*Marshallian Demand Elasticities*] *Let $Q_{ji,k} \equiv \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$ denote the Marshallian demand function facing variety ji, k. This demand function is characterized by the following set of demand elasticities:*

$$\varepsilon_{ji,k}^{(ni,g)} \equiv \frac{\partial \ln \mathcal{D}_{ji,k}(Y_i, \tilde{\boldsymbol{P}}_i)}{\partial \ln \tilde{P}_{ni,g}} \sim \textit{price elasticity}$$

$$\eta_{ji,k} \equiv \frac{\partial \ln \mathcal{D}_{ji,k}(Y_i, \tilde{\boldsymbol{P}}_i)}{\partial \ln Y_i} \sim \textit{income elasticity,}$$

where $\tilde{\mathbf{P}}_i = \left\{\tilde{\mathbf{P}}_{1i}, \tilde{\mathbf{P}}_{2i}, ..., \tilde{\mathbf{P}}_{Ni}\right\}$ corresponds to the entire of vector of consumer prices in market i. Also, recall from the main text that $V(Y_i, \tilde{\mathbf{P}}_i)$ denotes the indirect utility associated with the Marshallian demand function, $\mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$.

In what follows, we appeal the above definition to characterize the first-order condition w.r.t. each element of $\tilde{\mathbb{P}}_i$. We start with country i's import prices, $\tilde{\mathbf{P}}_{ji}$, and then proceed to domestic and export price instruments, $\tilde{\mathbf{P}}_{ii}$, and $\tilde{\mathbf{P}}_{ij}$.

***Step 3.A*: Deriving the F.O.C. w.r.t. $P_{ji,k} \in \tilde{\mathbb{P}}_i$.**

Consider the price of import variety ji, k, supplied by *origin j–industry k* (where $j \neq i$). To present the first-order necessary condition (F.O.C.) w.r.t. the price of ji, k, we use $\mathbb{P}_{-ji,k}$ to denote all elements of $\tilde{\mathbb{P}}_i$ excluding $\tilde{P}_{ji,k}$:

$$\mathbb{P}_{-ji,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ji,k}\} \sim \text{entire policy vector excluding } \tilde{P}_{ji,k}$$

Next, recall that $W_i(\tilde{\mathbb{P}}_i; \mathbf{w}) \equiv V_i(Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}), \tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji})$ where income, $Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}) = \bar{w}_i L_i + \mathcal{R}_i(\tilde{\mathbb{P}}_i; \mathbf{w})$, is dictated by the balanced budget condition. Applying the chain rule to $W_i(\tilde{\mathbb{P}}_i; \mathbf{w})$, the F.O.C. w.r.t. $\tilde{P}_{ji,k}$ (holding the remaining elements of $\tilde{\mathbb{P}}_i$ constant) can be stated as follows: [8]

$$\left(\frac{\mathrm{d}W_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\mathrm{d}\ln \tilde{P}_{ji,k}}\right)_{\mathbb{P}_{-ji,k}} = \overbrace{\frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ji,k}} + \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial Y_i}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}, \mathbb{P}_{-ji,k}}}^{\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \tilde{P}_{ji,k}}\right)_{\mathbf{w}, \mathbb{P}_{-ji,k}}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ji,k}}\right)_{\mathbb{P}_{-ji,k}} = 0 \tag{E.2}$$

The first term on the right-hand side of the above equation accounts for the direct welfare effects of a change in the price of good ji, k (holding Y_i and $\tilde{\mathbb{P}}_{-ji,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ji,k}\}$ constant). The second term accounts for welfare effects that channel through revenue-generation (holding $\mathbf{w}$ and $\tilde{\mathbb{P}}_{-ji,k}$ constant). The last term accounts for general equilibrium wage effects. Below, we characterize each of these elements one-by-one.

The term accounting for direct price effects can be simplified by appealing to Roy's identity, $\frac{\partial V_i / \tilde{P}_{ji,k}}{\partial V_i / \partial Y_i} = -Q_{ji,k}$, which indicates that

$$\text{[Roy's identity]} \qquad \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ji,k}} = -\tilde{P}_{ji,k} Q_{ji,k} \left(\frac{\partial V_i}{\partial Y_i}\right). \tag{E.3}$$

---

[8] We can alternatively formulate the above optimization problem using the method of Lagrange multipliers, and by appealing to Lagrange sufficiency theorem. In that case the objective function can be formulated as follows:

$$\max_{\tilde{\mathbb{P}}_i, Y_i} \mathcal{L}_i(\tilde{\mathbb{P}}_i; \mathbf{w}) = V_i(Y_i, \tilde{\mathbf{P}}_i) + \lambda_y \left(Y_i - \bar{w}_i L_i - \mathcal{R}_i(\tilde{\mathbb{P}}_i; \mathbf{w})\right).$$

The F.O.C.with respect to Y_i entails that $\lambda_Y = \frac{\partial V_i(.)}{\partial Y_i}$. Hence, the F.O.C. with respect to $\tilde{P}_{ji,k} \in \tilde{\mathbb{P}}_i$ can be expressed as

$$\frac{\mathrm{d}\mathcal{L}_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\mathrm{d}\ln \tilde{P}_{ji,k}} = \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ji,k}} + \lambda_y \left(\frac{\partial(\bar{w}_i L_i + \mathcal{R}_i(\tilde{\mathbb{P}}_i; \mathbf{w}))}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}, \mathbb{P}_{-ji,k}} + \left(\frac{\partial \mathcal{L}_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ji,k}}\right)_{\mathbb{P}_{-ji,k}} = 0,$$

which is equivalent to the F.O.C. expressed above.

To characterize $\left(\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\mathbb{P}_{-ji,k}}$, note that total income in country i (which dictates total expenditure) is the sum of wage payments plus tax revenues:[9]

$$Y_i(\tilde{\mathbb{P}}_i;\mathbf{w}) = w_i L_i + \underbrace{\sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni}\right)\cdot\mathbf{Q}_{ni}\right]}_{\text{import tax revenues}} + \underbrace{\left(\tilde{\mathbf{P}}_{ii}-\mathbf{P}_{ii}\right)\cdot\mathbf{Q}_{ii}}_{\text{production tax revenues}} + \underbrace{\sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{in}-\mathbf{P}_{in}\right)\cdot\mathbf{Q}_{in}\right]}_{\text{export tax revenues}},$$

Holding w and $\tilde{\mathbb{P}}_{-ji,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ji,k}\}$ fixed, $\tilde{P}_{ji,k}$ has no effect on wage payments: $\left(\partial\left(w_i L_i\right)/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = 0$. The effect of $\tilde{P}_{ji,k}$ on import tax revenues can be unpacked as follows:

$$\left(\frac{\partial \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni}\right)\cdot\mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ji,k}}\right) = \tilde{P}_{ji,k}Q_{ji,k} + \sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-P_{ni,g}\right)Q_{ni,g}\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]$$
$$-\sum_g\sum_{n\neq i}\left[P_{ni,g}Q_{ni,g}\sum_{j\neq i}\left[\frac{P_{ji,g}Q_{ji,g}}{P_{ni,g}Q_{ni,g}}\left(\frac{\partial \ln P_{ji,g}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]. \tag{E.4}$$

The first term in the above expression accounts for the direct, arithmetic effect of $\tilde{P}_{ji,k}$ on import tax revenues. The second term accounts for the change in revenue due to the change in country i's import demand schedule as a result of changing $\tilde{P}_{ji,k} \in \tilde{\mathbf{P}}_i$. The change in demand can itself be decomposed into two components:

$$\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \underbrace{\frac{\partial \ln \mathcal{D}_{ni,g}(\tilde{\mathbf{P}}_i,Y_i)}{\partial \ln \tilde{P}_{ji,k}}}_{\text{price effect}} + \underbrace{\frac{\partial \ln \mathcal{D}_{ni,g}(\tilde{\mathbf{P}}_i,Y_i)}{\partial \ln Y_i}\left(\frac{\partial \ln Y_i}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}}_{\text{income effect}} = \varepsilon^{(ji,k)}_{ni,g} + \eta_{ni,g}\left(\frac{\partial \ln Y_i}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}, \tag{E.5}$$

where $\varepsilon^{(ji,k)}_{ni,g}$ and $\eta_{ni,g}$ denote the Marshallian price and income elasticities of demand. The fact that adjustments to demand quantities depend on $\left(\partial \ln Y_i/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$ reflects the circular nature of our general equilibrium setup. We will not unpack this term for now, as we demonstrate later that income effects are welfare-neutral at the optimum an need not be explicitly specified.

The last term in Equation E.4, accounts for scale effects: Noting that $P_{ni,g} = \bar{\varrho}_{ni,g} w_n \left[\sum_\iota \tau_{n\iota,g} Q_{n\iota,g}\right]^{-\frac{\mu_g}{1+\mu_g}}$, a change in the export supply of good ni,g (due to a change in $\tilde{P}_{ji,k}$) alters the scale of production in *origin n–industry g* and the producer prices associated with that location. Due to cross demand effects, this change also impacts the producer price of domestic suppliers as well as foreign suppliers in location $\jmath \neq i$ outside of n.[10] Using the above definition for $\omega_{ni,g}$, we can simplify Equation E.4 as

[9]The operator "$\cdot$" denotes the inner product of two equal-sized vectors. Also, since we are focused on the free entry case, for now, the profit-adjusted wage rate is equal to the actual (unadjusted) wage rate, i.e., $\bar{w}_i = w_i$.

[10]To give an example, the producer price of goods supplied by country i in industry g ($P_{ij,g}$) respond to a reduction in $Q_{ni,g}$ through the following chain of effects:

$$Q_{ni,g}\downarrow \xrightarrow{\text{scale effects }(n,g)} P_{n\ell,g}\uparrow \xrightarrow{\text{cross-demand effects }(\ell\neq i)} Q_{i\ell,g}\uparrow \xrightarrow{\text{scale effects }(i,g)} P_{ij,g}\downarrow$$

Stated verbally, a reduction in $Q_{ni,g}$ lowers the producer price of *origin n–industry g* goods in all markets including $\ell \neq i$. Since consumer prices in location $\ell \neq i$ are not regulated by policy, an increase in $P_{n\ell,g}$ is fully passed onto consumer prices (provided that $n \neq i$), leading to a increase in $Q_{i\ell,g}$ through cross-substitution (or cross-demand) effects. The increase in $Q_{i\ell,g}$, in turn, lowers the producer price of goods supplied by o*rigin i–industry g* to all markets.

follows:

$$\left(\frac{\partial \sum_{n\neq i}\left[(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni})\cdot\mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \tilde{P}_{ji,k}Q_{ji,k}+\sum_{g}\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-\left[1+\sum_{j\neq i}\frac{P_{ji,g}Q_{ji,g}}{P_{ni,g}Q_{ni,g}}\left(\frac{\partial \ln P_{ji,g}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]P_{ni,g}\right)Q_{ni,g}\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right] \tag{E.6}$$

Moving on, the effect of a change in $\tilde{P}_{ji,k}$ on country i's production and export tax revenues can be unpacked as

$$\left(\frac{\partial \sum_{n}\left[(\tilde{\mathbf{P}}_{in}-\mathbf{P}_{in})\cdot\mathbf{Q}_{in}\right]}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \sum_{g}\left[\left(\tilde{P}_{ii,g}-P_{ii,g}\right)Q_{ii,g}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]$$
$$-\sum_{g}\sum_{n}\left[P_{in,g}Q_{in,g}\left(\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}+\sum_{\ell\neq i}\left(\frac{\partial \ln P_{in,g}}{\partial \ln Q_{\ell i,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\left(\frac{\partial \ln Q_{\ell i,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right)\right]. \tag{E.7}$$

The first term in the above equation accounts for revenue effects that channel through a change in the demand for domestic varieties (i.e., ii,g). The second term accounts for scale effects—i.e., a change in $Q_{ii,g}$ alters the scale of production in origin i–industry k, and the producer prices associated with country i in all export markets. Likewise, the change in the price of competing variaties (e.g., $Q_{\ell i,g}$, where $\ell \neq i$) affects the scale of production in country i through cross-scale effects—as explained in Footnote 10. We can simplify the first sum in the second line of the above equation by invoking the *Free Entry* condition. In particular,[11]

$$\sum_{n}\left[P_{in,g}Q_{in,g}\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right] = \sum_{n}\left[\frac{P_{in,g}Q_{in,g}}{P_{ii,g}Q_{ii,g}}\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]P_{ii,g}Q_{ii,g} = \tag{E.9}$$
$$= \sum_{n}\left[\frac{r_{in,g}}{r_{ii,g}}\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]P_{ii,g}Q_{ii,g} = -\frac{\mu_g}{1+\mu_g}P_{ii,g}Q_{ii,g},$$

where the last line follows from the fact that *(a)* $\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} = -\frac{\mu_g}{1+\mu_g}r_{ii,g}$, and *(b)* $\sum_n r_{in,g} = 1$. We can plug the above equation back into Equation E.7 to simplify it as follows:

$$\left(\frac{\partial \sum_{n}\left[(\tilde{\mathbf{P}}_{in}-\mathbf{P}_{in})\cdot\mathbf{Q}_{in}\right]}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \sum_{g}\left[\left(\tilde{P}_{ii,g}-\left[1-\frac{\mu_g}{1+\mu_g}\right]P_{ii,g}\right)Q_{ii,g}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]+\sum_{\ell\neq i}\sum_{n,g}\left[P_{in,g}Q_{in,g}\left(\frac{\partial \ln P_{in,g}}{\partial \ln Q_{\ell i,g}}\right)\left(\frac{\partial \ln Q_{\ell i,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]. \tag{E.10}$$

Note that $\left(\partial \ln Q_{ii,g}/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$ encompasses price and income effects as indicated by Equation E.5. To keep track of the general equilibrium scale effects in Equations E.6 and E.10, we use $\omega_{ni,g}$ to

---

[11] In particular, note that $P_{in,g} = \tau_{in,g}P_{ii,g}$, where by Free Entry, $P_{ii,g} = \bar{\rho}_{ii,g}w_i\mathcal{Q}_{i,g}^{-\frac{\mu_g}{1+\mu_g}}$, with $\mathcal{Q}_{i,g} = \sum_n \bar{a}_{in,g}Q_{in,g}$ denoting country i's effective output in industry g. Hence, holding $\mathbf{w}$ and $\tilde{\mathbb{P}}_{-ji,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ji,k}\}$ constant, we can show that

$$\sum_{n}\left[\left(\frac{\partial \ln P_{in,g}}{\partial \ln Q_{ij,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\frac{r_{in,g}}{r_{ij,g}}\right] = \sum_{n}\left[\left(\frac{\partial \ln P_{ii,g}}{\partial \ln \mathcal{Q}_{i,g}}\frac{\partial \ln \mathcal{Q}_{i,g}}{\partial \ln Q_{ij,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\frac{r_{in,g}}{r_{ij,g}}\right] = \frac{\partial \ln P_{ii,g}}{\partial \ln \mathcal{Q}_{i,g}} = -\frac{\mu_g}{1+\mu_g}, \tag{E.8}$$

where the second line follows from the fact that $\partial \ln \mathcal{Q}_{i,g}/\partial \ln Q_{ij,g} = r_{ij,g}$, by definition.

denote the general equilibrium inverse export supply elasticity associated with ni,g's . Namely,

$$\begin{aligned}\omega_{ni,g} &\equiv \sum_{\ell\in\mathbb{C}}\left[\frac{P_{i\ell,g}Q_{i\ell,g}}{P_{ni,g}Q_{ni,g}}\left(\frac{\partial\ln P_{i\ell,g}}{\partial\ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]+\sum_{j\neq i}\left[\frac{P_{ji,g}Q_{ji,g}}{P_{ni,g}Q_{ni,g}}\left(\frac{\partial\ln P_{ji,g}}{\partial\ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]\\ &= \frac{1}{r_{ni,g}\rho_{n,g}}\sum_g\left[\frac{\bar{w}_iL_i}{\bar{w}_nL_n}\rho_{i,g}\left(\frac{\partial\ln P_{ii,g}}{\partial\ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}+\sum_{j\neq i}\frac{\bar{w}_jL_j}{\bar{w}_nL_n}r_{ji,g}\rho_{j,g}\left(\frac{\partial\ln P_{ji,g}}{\partial\ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]\sim\text{export supply elasticity}\end{aligned} \tag{E.11}$$

The second line in the above definition derives from the fact that $\left(\frac{\partial\ln P_{i\ell,g}}{\partial\ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}=\left(\frac{\partial\ln P_{ii,g}}{\partial\ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}$ for all $\ell\in\mathbb{C}$ (as the price of origin i's good sold to different locations differ in only a constant iceberg cost shifter) and that sales shares for each origin $n\in\mathbb{C}$ are defined as follows:

$$r_{ni,g}\equiv\frac{P_{ni,g}Q_{ni,g}}{\sum_{\iota\in\mathbb{C}}\left(P_{n\iota,g}Q_{n\iota,g}\right)}\sim\text{good-specific sales share;}\qquad \rho_{n,g}=\frac{\sum_{\iota\in\mathbb{C}}\left(P_{n\iota,g}Q_{n\iota,g}\right)}{\bar{w}_nL_n}\sim\text{industry-wide sales share.}$$

For now, we do not unpack the supply elasticity, $\omega_{ni,g}$. We relegate this task instead to Step #4 of the proof, where we solve our full system of F.O.C.s. Next, we simplify Equations E.6 and E.10 by invoking our definition for $\omega_{ni,g}$, and combine the resulting expressions to produce the following equation describing the effect of policy on total tax revenues:

$$\begin{aligned}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} &= \tilde{P}_{ji,k}Q_{ji,k}+\sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-[1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\varepsilon_{ni,g}^{(ji,k)}\right]\\ &+\sum_g\left[\left(\tilde{P}_{ii,g}-\frac{1}{1+\mu_g}P_{ii,g}\right)Q_{ii,g}\varepsilon_{ii,g}^{(ji,k)}\right]+\Delta_i(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}.\end{aligned} \tag{E.12}$$

The uniform term $\Delta_i(\tilde{\mathbb{P}}_i)$ regulates the net force of (circular) general equilibrium income effects. It correspondingly depends on the Marshallian income elasticities of demand:

$$\Delta_i(\tilde{\mathbb{P}}_i;\mathbf{w})\equiv\sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-[1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\eta_{ni,g}\right]+\sum_g\left[\left(\tilde{P}_{ii,g}-\frac{1}{1+\mu_g}P_{ii,g}\right)Q_{ii,g}\eta_{ii,g}\right]. \tag{E.13}$$

To characterize the general equilibrium wage effects in the F.O.C. (i.e., the last term on the right-hand side of Equation E.2), we invoke our earlier result under Lemma 4: By the targeting principle $\mathbf{w}_{-i}$ is welfare neutral at the optimum (i.e., $\tilde{\mathbb{P}}_i=\tilde{\mathbb{P}}_i$), which implies that

$$\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{d\mathbf{w}}{d\ln\tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}}=\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial w_i}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_i}\left(\frac{dw_i}{d\ln\tilde{P}_{ji,k}}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_{-ji,k}}.$$

That is, we can characterize the term that encompasses wage effects, treating $\mathbf{w}_{-i}$ as given. Accordingly, the term $\left(dw_i/d\ln\tilde{P}_{ji,k}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_{-ji,k}}$ can be calculated by applying the Implicit Function Theorem to country i's balanced trade condition,[12]

$$[\text{Balanced Trade}]\qquad \mathrm{T}_i\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)\equiv\sum_{n\neq i}\left[\mathbf{P}_{ni}(\tilde{\mathbb{P}}_i;\mathbf{w})\cdot\mathbf{Q}_{ni,g}(\tilde{\mathbb{P}}_i;\mathbf{w})-\tilde{\mathbf{P}}_{ni}\cdot\mathbf{Q}_{ni,g}(\tilde{\mathbb{P}}_i;\mathbf{w})\right],$$

while treating $\mathbf{w}_{-i}=\bar{\mathbf{w}}_{-i}$ as if it were given. This step yields the following equation

[12]To be clear about the notation, we can write country i's balanced trade condition without appealing to the inner product operator as follows: $\mathrm{T}_i\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)\equiv\sum_g\sum_{n\neq i}\left(P_{ni,g}(\tilde{\mathbb{P}}_i,\mathbf{w})Q_{ni,g}(\tilde{\mathbb{P}}_i,\mathbf{w})-\tilde{P}_{in,g}Q_{in,g}(\tilde{\mathbb{P}}_i,\mathbf{w})\right)=0.$

$$\left(\frac{d\ln w_i}{d\ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_{-ji,k}} = -\left(\frac{\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i,\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} / \left(\frac{\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i,\mathbf{w})}{\partial \ln w_i}\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}$$

$$= \frac{-\sum_{n\neq i}\left[(\mathbf{P}_{ni}\odot\mathbf{Q}_{ni})\cdot\left(\frac{\partial \ln \mathbf{Q}_{ni}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} + (\mathbf{P}_{ni}\odot\mathbf{Q}_{ni})\cdot\left(\mathbf{\Omega}_{ni}\odot\frac{\partial \ln \mathbf{Q}_{ni}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]}{\left(\frac{\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i,\mathbf{w})}{\partial \ln w_i}\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}}. \quad \text{(E.14)}$$

where $\mathbf{\Omega}_{ni} \equiv \{\omega_{ni,k}\}_k$ is a vector composed of export supply elasticities (as defined under Equation E.11) and $\odot$ denotes the element-wise product of two equal-sized vectors (i.e., $\mathbf{a}\odot\mathbf{b} = [a_n b_n]_n$). The second line in the above equation follows from the fact that $\left(\partial \ln Q_{in,g}(\tilde{\mathbb{P}}_i,\mathbf{w})/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = 0$ if $n \neq i$. That is, if we fix the vector of wages, $\mathbf{w}$, the choice of $\tilde{P}_{ji,k}$ has no effect on the demand schedule in the rest of the world. In other words, the only way the effect of $\tilde{P}_{ji,k}$ transmits to foreign markets is through its effect on $\mathbf{w}$.[13] Now, define the importer-wide term, $\bar{\tau}_i$, as follows:

$$\bar{\tau}_i \equiv \frac{\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial w_i}\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}\left(\frac{\partial V_i(.)}{\partial Y_i}\right)^{-1}}{\left(\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i,\mathbf{w})/\partial \ln w_i\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}}. \quad \text{(E.15)}$$

Importantly, note that $\bar{\tau}_i$ does not feature an industry-specific subscript. Combining Equation E.14 with the expression for $\bar{\tau}_i$, we can summarize the wage effects in the F.O.C. (Equation E.2) as follows

$$\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{d\mathbf{w}}{d\ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = -\sum_g\sum_{n\neq i}\left[[1+\omega_{ni,g}]\bar{\tau}_i P_{ni,g}Q_{ni,g}\varepsilon_{ni,g}^{(ji,k)}\right]$$
$$-\sum_g\sum_{n\neq i}\left[[1+\omega_{ni,g}]\bar{\tau}_i P_{ni,g}Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}. \quad \text{(E.16)}$$

Finally, plugging Equations E.3, E.12, and E.16 back into the F.O.C. (Equation E.2); yields the following optimality condition w.r.t. to price instrument $\tilde{P}_{ji,k}\in\tilde{\mathbb{P}}_i$:

$$[\text{FOC w.r.t. } \tilde{P}_{ji,k}] \qquad \sum_{n\neq i}\sum_g\left[\left(\frac{\tilde{P}_{ni,g}}{P_{ni,g}}-(1+\bar{\tau}_i)(1+\omega_{ni,g})\right)P_{ni,g}Q_{ni,g}\varepsilon_{ni,g}^{(ji,k)}\right]$$
$$+\sum_g\left[\left(\frac{\tilde{P}_{ii,g}}{P_{ii,g}}-\frac{1}{1+\mu_g}\right)P_{ii,g}Q_{ii,g}\varepsilon_{ii,g}^{(ji,k)}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i;\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = 0. \quad \text{(E.17)}$$

The uniform term $\widetilde{\Delta}_i(.)$ is defined analogously to $\Delta_i(.)$, but adjusts for the interaction of general equilibrium wage and income effects:

$$\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i;\mathbf{w}) \equiv \sum_g\sum_{n\neq i}\left[\left(\frac{\tilde{P}_{ni,g}}{P_{ni,g}}-(1+\omega_{ni,g})(1+\bar{\tau}_i)\right)P_{ni,g}Q_{ni,g}\eta_{ni,g}\right]+\sum_g\left[\left(\frac{\tilde{P}_{ii,g}}{P_{ii,g}}-\frac{1}{1+\mu_g}\right)P_{ii,g}Q_{ii,g}\eta_{ii,g}\right]. \quad \text{(E.18)}$$

Before moving forward, a remark on the uniform term $\bar{\tau}_i$ is in order. We do not unpack this term because the multiplicity of country i's optimal tax schedule (per Lemma 1) will render the exact value assigned to $\bar{\tau}_i$ as redundant. We will elaborate more on this point when we combine the F.O.C.s *w.r.t.* all tax instruments in step #4 of the proof.

[13] The partial derivative in the numerator of Equation E.14 is subject to the implicit restriction that $\sum_n \mathbf{P}_{in,k}\cdot\mathbf{Q}_{in,k} = w_i L_i$ is held constant. We mechanically ensure this constraints is satisfied by differentiating a reformulated balanced trade condition, $T_i = \sum_{n\neq i}\left[\mathbf{P}_{ni}(\tilde{\mathbb{P}}_i;\mathbf{w})\cdot\mathbf{Q}_{ni,g}(\tilde{\mathbb{P}}_i;\mathbf{w}) - \tilde{\mathbf{P}}_{ni}\cdot\mathbf{Q}_{ni,g}(\tilde{\mathbb{P}}_i;\mathbf{w})\right] + \left(\sum_n \mathbf{P}_{in,k}\cdot\mathbf{Q}_{in,k} - w_i L_i\right) = 0$, where the last term in parenthesis is zero.

***Step 3.B*: Deriving the F.O.C. w.r.t. $P_{ii,k} \in \tilde{\mathbb{P}}_i$.**

Next, we derive the F.O.C. w.r.t. to a locally produced and locally consumer variety ii,k. Recall that the objective function can is given by $W_i = V_i(Y_i(\tilde{\mathbb{P}}_i;\mathbf{w}), \tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji})$. The F.O.C. w.r.t. $\tilde{P}_{ii,k}$, holding the remaining elements of $\tilde{\mathbb{P}}_i$ (namely, $\mathbb{P}_{-ii,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ii,k}\}$) constant, can be stated as

$$\left(\frac{\mathrm{d}W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\mathrm{d}\ln \tilde{P}_{ii,k}}\right)_{\mathbb{P}_{-ii,k}} = \frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ii,k}} + \frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial Y_i}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ii,k}}\right)_{\mathbf{w},\mathbb{P}_{-ii,k}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ii,k}}\right)_{\mathbb{P}_{-ii,k}} = 0. \tag{E.19}$$

Each element of the right-hand side can be characterized in a manner identical to *Step 3.A*. Specifically, the first term can be simplified using Roy's identity. The second term, which accounts for revenue-raising effects can be characterized using cross-demand elasticities w.r.t. $\tilde{P}_{ii,k}$ instead of $\tilde{P}_{ji,k}$. The same goes for the last term accounting for general equilibrium wage effects. Repeating the derivations in *Step 3.A*, the F.O.C. characterized by Equation E.19 can be unpacked as follows:

$$[\text{FOC w.r.t. } \tilde{P}_{ii,k}] \qquad \sum_{n\neq i}\sum_g\left[\left(\frac{\tilde{P}_{ni,g}}{P_{ni,g}} - (1+\bar{\tau}_i)(1+\omega_{ni,g})\right)P_{ni,g}Q_{ni,g}\varepsilon_{ni,g}^{(ii,k)}\right]$$
$$+\sum_g\left[\left(\frac{\tilde{P}_{ii,g}}{P_{ii,g}} - \frac{1}{1+\mu_g}\right)P_{ii,g}Q_{ii,g}\varepsilon_{ii,g}^{(ii,k)}\right] + \widetilde{\Delta}_i(\tilde{\mathbb{P}}_i;\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ii,k}}\right)_{\mathbf{w},\mathbb{P}_{-ii,k}} = 0, \tag{E.20}$$

where the uniform terms, $\widetilde{\Delta}_i(.)$, and $\bar{\tau}_i$, have the same definition as that introduced under Equations E.18 and E.15.

***Step 3.C*: Deriving the F.O.C. w.r.t. $P_{ij,k} \in \tilde{\mathbb{P}}_i$.**

Finally, we derive the F.O.C. w.r.t. to export variety ij,k, which is sold to destination $j \neq i$ in industry k. Note again that the objective function is given by $W_i = V_i(Y_i(\tilde{\mathbb{P}}_i;\mathbf{w}), \tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji})$. The F.O.C. w.r.t. $\tilde{P}_{ij,k}$, holding the remaining elements of $\tilde{\mathbb{P}}_i$ (namely, $\tilde{\mathbb{P}}_{-ij,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ij,k}\}$) constant, can be stated as

$$\left(\frac{\mathrm{d}W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\mathrm{d}\ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = \frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ij,k}} + \frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial Y_i}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = 0. \tag{E.21}$$

The first term as before accounts for the direct effect of a price change on consumer surplus. This term is trivially equal to zero in this case, since $\tilde{P}_{ij,k} \notin \tilde{\mathbf{P}}_i$. That is, since ij,k is not part of the domestic consumption bundle, raising its price has no direct effect on consumer surplus in country i:

$$\tilde{P}_{ij,k} \notin \tilde{\mathbf{P}}_i \implies \frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ij,k}} = 0. \tag{E.22}$$

The second term in Equation E.21 accounts for the revenue-raising effects of a change in $\tilde{P}_{ij,k} \in \tilde{\mathbb{P}}_i$. To unpack this term note that total income (or expenditure) in country i is dictated by the sum of wage payments and tax revenues:

$$Y_i(\tilde{\mathbb{P}}_i;\mathbf{w}) = w_iL_i + \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni}\right)\cdot\mathbf{Q}_{ni}\right] + \left(\tilde{\mathbf{P}}_{ii} - \mathbf{P}_{ii}\right)\cdot\mathbf{Q}_{ii} + \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{in} - \mathbf{P}_{in}\right)\cdot\mathbf{Q}_{in}\right],$$

Hence, holding wages $\mathbf{w}$ constant, the change in country i's income amounts to the change in import, domestic, and export tax revenues. The effect on import tax revenues can be unpacked as follows:

$$\left(\frac{\partial \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni}\right)\cdot\mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = \sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g} - P_{ni,g}\right)Q_{ni,g}\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right]$$
$$-\sum_g\sum_{n\neq i}\left[P_{ni,g}Q_{ni,g}\omega_{nj,g}\left(\frac{\partial \ln Q_{nj,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right]. \tag{E.23}$$

where $\omega_{nj,g}$ is the export supply elasticity as defined by E.11. The first term on the right-hand side accounts for general equilibrium income effects: Specifically, a change in $\tilde{P}_{ij,k}$ can raise country i's income Y_i through higher tax revenues, and alter the entire demand schedule, $Q_{ni,g} = \mathcal{D}_{ni,g}(\tilde{\mathbf{P}}_i, Y_i)$, in the local market. The second term accounts for scale effects: To elaborate, a change in $\tilde{P}_{ij,k}$ distorts origin i's export supply schedule in market $j \in \mathbb{C}$. This change alters the scale of production and the producer prices associated with *origin n–industry g* that serves market j (this includes $P_{ni,g}$ which is associated with economy i). It also changes the scale of production and producer prices from foreign suppliers through cross-demand effects. These changes in international producer prices, impacts country i's terms-of-trade by changing its import tax revenues. Also, note that since the rest of the world (including country j) is passive in terms of taxation, their income is pinned to their wage rate and vector $\mathbf{w}$. Hence, $\left(\partial Y_j / \partial \ln \tilde{P}_{ij,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = 0$, which implies that $\left(\partial \ln Q_{nj,g} / \partial \ln \tilde{P}_{ij,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = \partial \ln \mathcal{D}_{nj,g}(\bar{Y}_j, \tilde{\mathbf{P}}_j) / \partial \ln \tilde{P}_{ij,k} = \varepsilon^{(ij,k)}_{nj,g}$. Likewise, since $\tilde{P}_{ij,k} \notin \tilde{\mathbf{P}}_i$, its only effect on the demand schedule in the local market i is through general equilibrium income effects. Putting these results together, we can posit that

$$\left(\frac{\partial \ln Q_{n\iota,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = \begin{cases} \varepsilon^{(ij,k)}_{nj,g} & \text{if } \iota = j \\ \eta_{ni,g}\left(\frac{\partial \ln Y_i}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} & \text{if } \iota = i \end{cases}.$$

Considering the above expressions and noting our earlier definition for $\omega_{ni,g}$ under Equation E.11, Equation E.23 can be simplified as

$$\begin{aligned}\left(\frac{\partial \sum_{n\neq i}\left[(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni})\cdot \mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} &= -\sum_g\sum_{n\neq i}\left[\omega_{nj,g}P_{ni,g}Q_{ni,g}\varepsilon^{(ij,k)}_{nj,g}\right] + \sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g} - [1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial \ln Y_i}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} \\ &= -\sum_g\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon^{(ij,k)}_{nj,g}\right] + \sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g} - [1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial \ln Y_i}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}.\end{aligned} \tag{E.24}$$

The last line in the above equation follows from (1) the definition of ω, which entails that $\omega_{nj,g}r_{ni,g} = \omega_{ni,g}r_{nj,g}$, and (2) the fact that $r_{ni,g}/r_{nj,g} = P_{ni,g}Q_{ni,g}/P_{nj,g}Q_{nj,g}$, since the markup is uniform across output sold to different destinations in the same industry.

The effect of $\tilde{P}_{ij,k}$ on country i's production and export tax revenues can be unpacked as follows:[14]

$$\begin{aligned}\left(\frac{\partial}{\partial \ln \tilde{P}_{ij,k}}\sum_n\left[(\tilde{\mathbf{P}}_{in} - \mathbf{P}_{in})\cdot \mathbf{Q}_{in}\right]\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} &= \sum_g\left[\left(\tilde{P}_{ii,g} - P_{ii,g}\right)Q_{ii,g}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right] \\ + \tilde{P}_{ij,k}Q_{ij,k} + \sum_g\left[\left(\tilde{P}_{ij,g} - P_{ij,g}\right)Q_{ij,g}\left(\frac{\partial \ln Q_{ij,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right] &- \sum_g\sum_n\left[P_{in,g}Q_{in,g}\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ij,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\left(\frac{\partial \ln Q_{ij,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right].\end{aligned} \tag{E.25}$$

The first term (in the 1st line) represents the effect on domestic tax revenues that channel through general equilibrium income effects. The second term on the right-hand side ($\tilde{P}_{ij,k}Q_{ij,k}$) represents the direct, arithmetic effect of $\tilde{P}_{ij,k}$ on export tax revenues. The third term represents revenue effects that channel through a change in the demand for all varieties sold to destination j (i.e., ij, g). The last term accounts for scale effects—i.e., a change in $Q_{ij,g}$ alters the scale of production in *origin i–industry g*, and modifies all the producer prices associated with that industry. As noted in *Step 3.A*, the last term in Equation E.25 can be simplified using the *free-entry* condition, which entails that (See Equation E.9):

$$\sum_{n\in\mathbb{C}}\left[P_{in,g}Q_{in,g}\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ij,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right] = -\frac{\mu_g}{1+\mu_g}P_{ij,g}Q_{ij,g},$$

[14] $\sum_n\left[(\tilde{\mathbf{P}}_{in} - \mathbf{P}_{in})\cdot \mathbf{Q}_{in}\right] = (\tilde{\mathbf{P}}_{ii} - \mathbf{P}_{ii})\cdot \mathbf{Q}_{ii,g} + \sum_{n\neq i}\left[(\tilde{\mathbf{P}}_{in} - \mathbf{P}_{in})\cdot \mathbf{Q}_{in}\right]$ is the sum of domestic and export tax revenues.

Also, recall from our earlier discussion that since country $j \neq i$ collects no tax revenues by assumption, $\left(\partial Y_j / \partial \ln \tilde{P}_{ij,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = 0$, which implies that $\left(\partial \ln Q_{nj,g} / \partial \ln \tilde{P}_{ij,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = \partial \ln \mathcal{D}_{nj,g}(\bar{Y}_j, \tilde{\mathbf{P}}_j) / \partial \ln \tilde{P}_{ij,k} = \varepsilon^{(ij,k)}_{nj,g}$. Plugging these expressions back into Equation E.25 simplifies it as

$$\left(\frac{\partial}{\partial \ln \tilde{P}_{ij,k}} \sum_n [(\tilde{\mathbf{P}}_{in} - \mathbf{P}_{in}) \cdot \mathbf{Q}_{in}]\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = \tilde{P}_{ij,k} Q_{ij,k} + \sum_g \left[\left(\tilde{P}_{ij,g} - [1 - \frac{\mu_g}{1+\mu_g}] P_{ij,g}\right) Q_{ij,g} \varepsilon^{(ij,k)}_{ij,g}\right]$$
$$+ \sum_g \left[\left(\tilde{P}_{ii,g} - \frac{1}{1+\mu_g} P_{ii,g}\right) Q_{ii,g} \eta_{ii,g}\right] \left(\frac{\partial \ln Y_i}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}. \tag{E.26}$$

Combining Equations E.24 and E.26, we can express the sum of tax revenue-related effects as

$$\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = \tilde{P}_{ij,k} Q_{ij,k} + \sum_g \left[\left(\tilde{P}_{ij,g} - [1 - \frac{\mu_g}{1+\mu_g}] P_{ij,g}\right) Q_{ij,g} \varepsilon^{(ij,k)}_{ij,g}\right]$$
$$- \sum_g \sum_{n \neq i} \left[\omega_{ni,g} P_{nj,g} Q_{nj,g} \varepsilon^{(ij,k)}_{nj,g}\right] + \Delta_i(\tilde{\mathbb{P}}_i; \mathbf{w}) \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}, \tag{E.27}$$

where $\Delta_i(.)$ encompasses the terms accounting for circular income effects and is given by Equation E.13. No we turn to characterizing the general equilibrium wage effects in the F.O.C.—namely, the last term on the right-hand side of Equation E.2. To this end, we invoke our observation based on the *targeting principle* (as stated under Lemma 4) that $\left(\frac{\partial W_i(.)}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d} \ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = \left(\frac{\partial W_i(.)}{\partial w_i}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_i} \left(\frac{\mathrm{d} w_i}{\mathrm{d} \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_{-ij,k}}$.
The term $\left(\frac{\mathrm{d} w_i}{\mathrm{d} \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_{-ji,k}}$ can be calculated by applying the Implicit Function Theorem to country i's balanced trade condition,

$$[\text{Balanced Trade}] \qquad \mathrm{T}_i\left(\tilde{\mathbb{P}}_i, \mathbf{w}\right) \equiv \sum_{n \neq i} \left[\mathbf{P}_{ni}(\tilde{\mathbb{P}}_i; \mathbf{w}) \cdot \mathbf{Q}_{ni,g}(\tilde{\mathbb{P}}_i; \mathbf{w}) - \tilde{\mathbf{P}}_{ni} \cdot \mathbf{Q}_{ni,g}(\tilde{\mathbb{P}}_i; \mathbf{w})\right],$$

while treating $\mathbf{w}_{-i} = \bar{\mathbf{w}}_{-i}$ as given. This application yields the following equation (*Notation*: $\boldsymbol{\Omega}_{nj} \equiv \left\{\omega_{nj,k}\right\}_k$ is a vector composed of export supply elasticities, while $\odot$ and $\cdot$ denotes the element-wise and inner products of two equal-sized vectors):

$$\left(\frac{\mathrm{d} \ln w_i}{\mathrm{d} \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w}_{-i},\tilde{\mathbb{P}}_{-ij,k}} = -\left(\frac{\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i, \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} / \left(\frac{\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i, \mathbf{w})}{\partial \ln w_i}\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}$$
$$= \frac{-\tilde{P}_{ij,k} Q_{ij,k} - \left(\tilde{\mathbf{P}}_{ij} \odot \mathbf{Q}_{ij}\right) \cdot \left(\frac{\partial \ln \mathbf{Q}_{ij}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} + \sum_{n \neq i} \left[\left(\mathbf{P}_{ni} \odot \mathbf{Q}_{ni}\right) \cdot \left(\frac{\partial \ln \mathbf{Q}_{ni}}{\partial \ln \tilde{P}_{ij,k}} + \boldsymbol{\Omega}_{nj} \odot \frac{\partial \ln \mathbf{Q}_{nj}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right]}{\left(\frac{\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i, \mathbf{w})}{\partial \ln w_i}\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}} \tag{E.28}$$

The numerator in the second line of the above equation is composed of three terms: The first term accounts for the arithmetic effect of $\tilde{P}_{ji,k}$ on country i's trade balance. The second term account for own- and cross-price effects that are specific to market j—the market to which good ij,k is being exported. The last term accounts for scale effects: Specifically, a change in $\tilde{P}_{ij,k}$ interacts with the *balanced trade condition* by modifying the producer of a generic good ni,g imported from *origin* i–

industry g. As before, define the uniform importer-wide term, $\bar{\tau}_i$, as follows

$$\bar{\tau}_i \equiv \frac{\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial w_i}\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}\left(\frac{\partial V_i(.)}{\partial Y_i}\right)^{-1}}{\left(\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i,\mathbf{w})/\partial \ln w_i\right)_{\bar{\mathbf{w}}_{-i},\tilde{\mathbb{P}}_i}}. \tag{E.29}$$

Combining Equation E.28 with the expression for $\bar{\tau}_i$, we can summarize the wage effects in the F.O.C. (Equation E.2) as follows:

$$\left(\frac{\partial W_i(.)}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{d\mathbf{w}}{d\ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = \bar{\tau}_i\tilde{P}_{ij,k}Q_{ij,k}+\sum_g\left[\bar{\tau}_i\tilde{P}_{ij,g}Q_{ij,g}\varepsilon^{(ij,k)}_{ij,g}\right]$$
$$-\sum_g\sum_{n\neq i}\left[\bar{\tau}_i\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon^{(ij,k)}_{nj,g}\right]-\sum_g\sum_{n\neq i}\left[[1+\omega_{ni,g}]\bar{\tau}_iP_{ni,g}Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}. \tag{E.30}$$

Finally, plugging Equations E.22, E.27, and E.30 back into the F.O.C. (Equation E.21); and dividing all the expressions by $(1+\bar{\tau}_i)$ yields the following optimality condition w.r.t. to price instrument $\tilde{P}_{ij,k}\in\tilde{\mathbb{P}}_i$:

$$[\text{FOC w.r.t. } \tilde{P}_{ij,k}]\qquad \tilde{P}_{ij,k}Q_{ij,k}+\sum_g\left[\left(1-\frac{1}{(1+\bar{\tau}_i)(1+\mu_g)}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right)\tilde{P}_{ij,g}Q_{ij,g}\varepsilon^{(ij,k)}_{ij,g}\right]$$
$$-\sum_g\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon^{(ij,k)}_{nj,g}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=0, \tag{E.31}$$

where $\widetilde{\Delta}_i(.)$ is defined as in Equation E.18. Also, we are not unpacking the term $\bar{\tau}_i$, for the same reasons discussed under *Step 3.A*.

Step #4: Solving the System of F.O.C.s and Establishing Uniqueness

To determine the optimal tax schedule we need to collect the each of first order conditions and simultaneously solve them under one system. For the ease of reference, we will restate the F.O.C. w.r.t. to each element of $\tilde{\mathbb{P}}_i$ below. Following Equations E.17 and E.20, the F.O.C. w.r.t. $\tilde{P}_{\ell i,k}\in\tilde{\mathbb{P}}_i$ (where $\ell=i$ or $\ell=j\neq i$) is given by the following equation:

$$(1)\quad \sum_{n\neq i}\sum_g\left[\left(1-(1+\bar{\tau}_i)(1+\omega_{ni,g})\frac{P_{ni,g}}{\tilde{P}_{ni,g}}\right)e_{ni,g}\varepsilon^{(\ell i,k)}_{ni,g}\right]+$$
$$\sum_g\left[\left(1-\frac{1}{1+\mu_g}\frac{P_{ii,g}}{\tilde{P}_{ii,g}}\right)e_{ii,g}\varepsilon^{(\ell i,k)}_{ii,g}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i;\mathbf{w})\left(\frac{\partial \ln Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{\ell i,k}}\right)_{\mathbf{w},\mathbb{P}_{-\ell i,k}}=0.$$

where $e_{ni,g}=\tilde{P}_{ni,g}Q_{ni,g}/Y_i$ denotes the (unconditional) expenditure share on good ni,g. Likewise, dividing Equation E.31 by $\tilde{P}_{ij,k}Q_{ij,k}$, the F.O.C. w.r.t. export price $\tilde{P}_{ij,k}\in\tilde{\mathbb{P}}_i$ is given by the following equation:

$$(2)\quad 1+\sum_g\left[\left(1-\frac{1}{(1+\mu_g)(1+\bar{\tau}_i)}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right)\frac{e_{ij,g}}{e_{ij,k}}\varepsilon^{(ij,k)}_{ij,g}\right]$$
$$-\sum_{n\neq i}\sum_g\left[\omega_{ni,g}\frac{e_{nj,g}}{e_{ij,k}}\varepsilon^{(ij,k)}_{nj,g}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w})\frac{Y_i}{Y_j}\left(\frac{\partial \ln Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=0.$$

To set the stage for what follows, let us emphasize four points:

(1) In accordance with the tax-neutrality result presented under Lemma 1, the optimal policy schedule is unique only up-to two arbitrary tax shifters.[15] That is, there are multiple optimal policy

[15] To be clear, the pseudo-uniqueness of the *optimal policy formula* is different from the uniqueness of the *optimal policy*

schedules that are welfare-equivalent but differ in the average level assigned to domestic and trade taxes—we come back to this point when finalizing our optimal policy formulas.

(2) The system of F.O.C.s labeled (1) can be solved independently of (2) to recover the optimal export- and import-side price wedges.

(3) The trivial solution to system (1) satisfies $\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i; \mathbf{w}) = 0$. Moreover, we can invoke Lemma 1 to show that if there exists an optimal policy schedule for which $\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i; \mathbf{w}) \neq 0$, that policy choice is welfare-equivalent to another that satisfies $\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i; \mathbf{w}) = 0$—see Appendix E.1 for an formal proof. These two observations together affirm that we can identify the full set of (welfare-equivalent) optimal policy schedules by setting $\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i; \mathbf{w}) = 0$ in the F.O.C.s. This particular proposition can be alternatively stated as an envelope-type result: If the government is afforded sufficient policy instruments, the system of F.O.C.s can be derived and solved *as if* the Marshallian demand functions were income inelastic.

(4) We focus on interior solutions that do not assign a prohibitive price to any good (i.e., $e_{ni,g} > 0$, $\forall\, ni, g$). Since prohibitive prices exclude goods from the system of F.O.C.s., one may worry that a non-interior solution that prohibits some goods but satisfies the necessary first-order conditions w.r.t. the other goods is optimal. Appendix E.2 rules out the optimality of prohibitive taxes/prices by appealing to the Inada conditions, which is standard in the literature.

All in all, these points indicate that System (1) can be solved independent of (2) and by restricting attention to interior solutions that satisfy $\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i; \mathbf{w}) = 0$. Doing so leads us to a unique trivial solution from which we can infer the remaining optimal tax schedules—all of which deliver the same welfare outcome. To establish this claim, set $\widetilde{\Delta}_i(.) = 0$ and express System (1) in matrix notation as follows:

$$\underbrace{\begin{bmatrix} e_{1i,1}\varepsilon_{1i,1}^{(1i,1)} & \cdots & e_{Ni,1}\varepsilon_{Ni,1}^{(1i,1)} & \cdots & e_{1i,K}\varepsilon_{1i,K}^{(1i,1)} & \cdots & e_{Ni,K}\varepsilon_{Ni,K}^{(1i,1)} \\ \vdots & & \ddots & & \ddots & & \vdots \\ e_{1i,1}\varepsilon_{1i,1}^{(Ni,K)} & \cdots & e_{Ni,1}\varepsilon_{Ni,1}^{(Ni,K)} & \cdots & e_{1i,K}\varepsilon_{1i,K}^{(Ni,K)} & \cdots & e_{Ni,K}\varepsilon_{Ni,K}^{(Ni,K)} \end{bmatrix}}_{\widetilde{\mathbf{E}}_i} \begin{bmatrix} 1-(1+\omega_{1i,k})(1+\bar{\tau}_i)\frac{P_{1i,1}}{\tilde{P}^{\star}_{1i,k}} \\ \vdots \\ 1-\frac{1}{1+\mu_k}\frac{P_{ii,k}}{\tilde{P}^{\star}_{ii,k}} \\ \vdots \\ 1-(1+\omega_{Ni,k})(1+\bar{\tau}_i)\frac{P_{Ni,k}}{\tilde{P}^{\star}_{Ni,k}} \end{bmatrix}_k = \mathbf{0}.$$

To prove that the above equation exhibits a unique, trivial solution it suffices to show that the expenditure-adjusted elasticity matrix, $\mathbf{E}_i = \left[e_{ji,k}\varepsilon_{ji,k}^{(ni,g)}\right]_{jk,ng}$ is non-singular. The following intermediate lemma establishes this result using the primitive properties of Marshallian demand functions.

Lemma 5. *The $NK \times NK$ matrix $\widetilde{\mathbf{E}}_i \equiv \left[e_{ji,k}\varepsilon_{ji,k}^{(ni,g)}\right]_{jk,ng}$ is non-singular.*

Proof. We can appeal to Proposition 2.E.2 in Mas-Colell, Whinston, Green, et al. (1995), which indicates that the Marshallian demand function satisfies $e_{ji,k} =| e_{ji,k}\varepsilon_{ji,k}^{(ji,k)} | - \sum_{n,g\neq j,k} | e_{ni,g}\varepsilon_{ni,g}^{(ji,k)} |$—a property often referred to as Cournot aggregation. Since $e_{ji,k} > 0$ (as we have ruled out prohibitive prices), Cournot aggregation ensures the matrix $\widetilde{\mathbf{E}}_i$ is strictly diagonally dominant. The Lèvy-Desplanques Theorem (Horn and Johnson (2012)), accordingly, ensures that $\widetilde{\mathbf{E}}_i$ is non-singular. The lower bound on $\det(\widetilde{\mathbf{E}}_i)$ follows trivially from Gerschgorin's circle theorem. Specifically, following Ostrowski (1952),

$$|\det\left(\widetilde{\mathbf{E}}_i\right)| \geq \prod_{j\in\mathbb{C}}\prod_{k\in\mathbb{K}}\left(\left|e_{ji,k}\varepsilon_{ji,k}^{(ji,k)}\right| - \sum_{(n,g)\neq(j,k)}\left|e_{ni,g}\varepsilon_{ni,g}^{(ji,k)}\right|\right) = \prod_{j\in\mathbb{C}}\prod_{k\in\mathbb{K}} e_{ji,k} > 0.$$

□

equilibrium. Establishing the latter is a daunting task well beyond the scope of this paper (see Kucheryavyy et al. (2023a)).

Appealing to above lemma, it is immediate that the unique solution to the above matrix equation is indeed the trivial solution, given by:

$$\frac{\tilde{P}^*_{ji,k}}{P_{ji,1}} = (1+\omega_{ji,k})(1+\bar{\tau}_i); \qquad \frac{\tilde{P}^*_{ii,k}}{P_{ii,k}} = \frac{1}{1+\mu_k}. \tag{E.32}$$

It is straightforward to check that the above solution constitutes a global maximum by contradiction. To present the logic: Since $\lim_{\tilde{\mathbf{P}}_i\to\infty} W_i(\tilde{\mathbb{P}}_i,\mathbf{w}) \to 0$, the above solution identifies a vector of consumer prices at home, $\tilde{\mathbf{P}}^*_i \in \tilde{\mathbb{P}}_i$, that yields a strictly higher welfare level than prohibitive prices. As such, $\tilde{\mathbb{P}}_i$ cannot constitute a global minimum. Lastly, it is straightforward to see that if the domestic price elements in $\tilde{\mathbb{P}}_i$ satisfy E.32, then

$$\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i;\mathbf{w}) \equiv \sum_g \sum_{n\neq i}\left[\left(\frac{\tilde{P}^*_{ni,g}}{P_{ni,g}} - (1+\omega_{ni,g})(1+\bar{\tau}_i)\right) P_{ni,g}Q_{ni,g}\eta_{ni,g}\right] + \sum_g \left[\left(\frac{\tilde{P}^*_{ii,g}}{P_{ii,g}} - \frac{1}{1+\mu_g}\right) P_{ii,g}Q_{ii,g}\eta_{ii,g}\right] = 0.$$

That is, the term accounting for general equilibrium income effects amounts to zero in the neighborhood of the optimum, *as if* demand functions were income inelastic (i.e., $\eta_{ni,g} = \eta_{ii,g} = 0$) Capitalizing on this result, we can proceed to solving System (2), knowing that $\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w}) = 0$. To this end, let us economize on the notation by defining $\mathscr{X}$ as follows:

$$\mathscr{X}_{ij,k} \equiv \frac{1}{(1+\mu_g)(1+\bar{\tau}_i)}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}.$$

Invoking this minor switch of notation, the F.O.C. specified by System (2) implies the following optimality condition:

$$1+\sum_g\left[(1-\mathscr{X}_{ij,g})\,\frac{e_{ij,g}\varepsilon^{(ij,k)}_{ij,g}}{e_{ij,k}}\right] - \sum_{n\neq i}\sum_g\left[\omega_{ni,g}\frac{e_{nj,g}\varepsilon^{(ij,k)}_{nj,g}}{e_{ij,k}}\right] = 0. \tag{E.33}$$

To simplify the above expression we will appeal to the Cournot aggregation property–a well-known primitive property of Marshallian demand as discussed earlier (see Mas-Colell et al. (1995)):

$$[\text{Cournot aggregation}] \qquad 1+\sum_g\left[\frac{e_{ij,g}}{e_{ij,k}}\varepsilon^{(ij,k)}_{ij,g}\right] = -\sum_{n\neq i}\sum_g\left[\frac{e_{nj,g}}{e_{ij,k}}\varepsilon^{(ij,k)}_{nj,g}\right].$$

Next, combine the above expression with Equation E.33, while noting that by Slutsky's equation $\frac{e_{nj,g}}{e_{ij,k}}\varepsilon^{(ij,k)}_{nj,g} = \varepsilon^{(nj,g)}_{ij,k}$ if $\eta_{ni,g} = 1$ for all ni,g. Performing these steps yields the following:

$$-\sum_g\left[\mathscr{X}_{ij,g}\varepsilon^{(ij,g)}_{ij,k}\right] - \sum_{n\neq i}\sum_g\left[(1+\omega_{ni,g})\varepsilon^{(nj,g)}_{ij,k}\right] = 0 \qquad \forall(ij,k).$$

We can rewrite the above equation in matrix algebra as follows:

$$-\mathbf{E}_{ij}\mathbf{X}_{ij} - \mathbf{E}^{(-ij)}_{ij}\left(\mathbf{1}_{(N-1)K} + \mathbf{\Omega}_i\right) = 0, \tag{E.34}$$

where $\mathbf{X}_{ij} \equiv \left[\mathscr{X}_{ij,k}\right]_k$ is a $K\times 1$ vector. The $K\times K$ matrix $\mathbf{E}_{ij} \sim \mathbf{E}^{(ij)}_{ij} \equiv \left[\varepsilon^{(ij,g)}_{ij,k}\right]$ encompasses the own- and cross-price elasticities between the different varieties sold by origin i to market j—see Definition (D1). Analogously, $\mathbf{E}^{(-ij)}_{ij} \equiv \left[\varepsilon^{(nj,g)}_{ij,k}\right]_{k,n\neq i,g}$ is a $K\times(N-1)K$ matrix summarizing the cross-price elasticity of market j's demand between varieties sold by origin i and all other (non-i) origin countries. $\mathbf{\Omega}_i \equiv \left[\omega_{ni,g}\right]_{n,g}$ is a $(N-1)K\times 1$ vector of all *import good-specific* inverse supply elasticities. To invert the above system we need to establish that $\mathbf{E}_{ij}$ is non-singular, which is done under the following lemma.

Lemma 6. *The $K\times K$ matrix $\mathbf{E}_{ij} \equiv \left[\varepsilon^{(ij,g)}_{ij,k}\right]_{k,g}$ is non-singular.*

Proof. The proof proceeds similar to Lemma 6: The Marshallian demand function's homogeneity of degree zero implies that $\mid \varepsilon^{(ij,k)}_{ij,k} \mid = \eta_{ij,k} + \sum_{n,g\neq i,k} \mid \varepsilon^{(nj,g)}_{ij,k} \mid$. Based on this property, since $\eta_{ij,k} > 0$, the matrix $\mathbf{E}_{ij}$ is strictly diagonally dominant. The Lèvy-Desplanques Theorem (Horn and Johnson (2012)), therefore, ensures that $\mathbf{E}_{ij}$ is non-singular. □

Following the above lemma we can invert the system specified by Equation E.34 to obtain the optimal level of $\mathbf{X}_{ij} = \left[\mathscr{X}_{ij,k}\right]_k$:

$$\mathbf{X}^*_{ij} = -\mathbf{E}_{ij}^{-1}\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1}_{(N-1)K} + \mathbf{\Omega}_i\right). \tag{E.35}$$

Next, there remains the task of recovering the optimal tax/subsidy rates from the optimal price wedges implies by Equations E.32 and E.35. Noting the following relationship between taxes/subsidies and price wedges,

$$1 + t^*_{ji,k} = \frac{\tilde{P}_{ji,k}}{P_{ji,k}}; \qquad 1 + s^*_{i,k} = \frac{P_{ii,k}}{\tilde{P}^*_{ii,k}}; \qquad 1 + x_{ij,k} = \frac{P_{ij,k}/\tilde{P}^*_{ij,k}}{P_{ii,k}/\tilde{P}^*_{ii,k}};$$

country i's unilaterally optimal tax schedule can be expressed as follows:

$$\begin{aligned}
&[\text{domestic subsidy}] \quad 1 + s^*_{i,k} = 1 + \mu_k \\
&[\text{import tariff}] \quad 1 + t^*_{ji,k} = (1 + \omega_{ji,k})(1 + \bar{\tau}_i) \\
&[\text{export subsidy}] \quad \mathbf{1} + \mathbf{x}^*_{ij} = -\mathbf{E}_{ij}^{-1}\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1} + \mathbf{t}^*_i\right).
\end{aligned} \tag{E.36}$$

The last step is to invoke the multiplicity of optimal tax schedules as indicated by Lemma 1. Doing so indicates that the uniform term $\bar{\tau}_i$ is redundant and need not be unpacked. To elaborate, Lemma 1 indicates that any policy schedule that includes an import tax equal to $(1 + \bar{t}_i \in \mathbb{R}_+)$

$$1 + t_{ji,k} = (1 + \omega_{ji,k})(1 + \bar{\tau}_i)(1 + \bar{t}_i)$$

is also optimal, since it delivers an identical level of welfare to the original optimal policy schedule specified by E.36. As such, the exact value assigned to $\bar{\tau}_i$ is redundant for a welfare standpoint. This is why we did not unpack the term $\bar{\tau}_i$ earlier in Step #3. Lemma 1 indicates that there is another dimension of multiplicity, whereby any uniform shift in domestic production subsidies (paired with a proportional adjustment to w_i) preserves the equilibrium. Considering these points, the optimal policy schedule (after accounting for all dimensions of multiplicity) is given by:

$$\begin{aligned}
&[\text{domestic subsidy}] \quad 1 + s^*_{i,k} = (1 + \mu_k)(1 + \bar{s}_i) \\
&[\text{import tariff}] \quad 1 + t^*_{ji,k} = (1 + \omega_{ji,k})(1 + \bar{t}_i) \\
&[\text{export subsidy}] \quad \mathbf{1} + \mathbf{x}^*_{ij} = -\mathbf{E}_{ij}^{-1}\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1} + \mathbf{t}^*_i\right),
\end{aligned}$$

where $1 + \bar{s}_i = 1 + \bar{t}_i \in \mathbb{R}_+$ are arbitrary tax shifters. What remains is a formal characterization of the good-specific supply elasticity, $\omega_{ji,k}$, which is presented below.

Characterizing the (Inverse) Export Supply Elasticity, $\omega_{ji,k}$. To fix ideas, it is helpful to repeat the definition of the export supply elasticity presented earlier:

$$\omega_{ji,k} \equiv \frac{1}{r_{ji,k}\rho_{j,k}} \sum_g \left[\frac{\bar{w}_i L_i}{\bar{w}_j L_j}\rho_{i,g}\left(\frac{\partial \ln P_{ii,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} + \sum_{n \neq i} \frac{\bar{w}_n L_n}{\bar{w}_j L_j} r_{ni,g}\rho_{n,g}\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right], \tag{E.37}$$

where $r_{ni,g} = P_{ni,g}Q_{ni,g}/\sum_{\iota \in \mathbb{C}}\left(P_{n\iota,g}Q_{n\iota,g}\right)$ and $\rho_{n,g} = \sum_{\iota \in \mathbb{C}}\left(P_{n\iota,g}Q_{n\iota,g}\right)/\bar{w}_n L_n$ respectively denote the good ni, g-specific and industry-wide sales shares associated with origin $n \in \mathbb{C}$. Also, note that the producer price of good ni, g under free entry is given by $P_{ni,g} = \tau_{ni,g}P_{nn,g}$, where

$$P_{nn,g} = \bar{\varrho}_{nn,g} w_n \sum_{\iota \in \mathbb{C}}\left[\tau_{n\iota,g}Q_{n\iota,g}\right]^{-\frac{\mu_g}{1+\mu_g}} \qquad \forall (n, g)$$

To characterize $\omega_{ji,k}$, we need to characterize $\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\mathbb{T}_i} = \left(\frac{\partial \ln P_{ii,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\mathbb{T}_i}$ for each origin n–industry g. To this end we can apply the Implicit Function Theorem to the following function:

$$F_{ni,g}(Q_{1i,g},...,Q_{Ni,g},P_{11,g},...,P_{NN,g}) = P_{nn,g} - \bar{\varrho}_{nn,g} w_n \left[\tau_{ni,g} Q_{ni,g} + \sum_{\ell \neq i} \tau_{n\ell,g} Q_{n\ell,g} (\underbrace{\boldsymbol{\tau}_{-i\ell} \odot \mathbf{P}_{-i}}_{\tilde{\mathbf{P}}_{-i\ell}}) \right]^{-\frac{\mu_g}{1+\mu_g}} = 0.$$

where $\boldsymbol{\tau}_{-in} \odot \mathbf{P}_{-i} \sim \{\tau_{jn,g} P_{jj,g}\}_{j \neq i,g}$ denotes the vector of consumer prices in market $n \neq i$ from all origins aside from i. The above function implicitly characterizes the producer prices in each origin j–industry g as a function of export supply levels to market i (i.e., $Q_{1i,g},...,Q_{Ni,g}$). Importantly, the above function treats both $\bar{\mathbb{P}}_i$ and $\mathbf{w}$ as given, as all elements of $\bar{\mathbb{P}}_i$ are chosen directly the by the government in i. Accordingly, the function $Q_{n\iota,g}(.)$ on the right-hand side derives from the Marshallian demand function,

$$Q_{jn,g}(\underbrace{\boldsymbol{\tau}_{-in} \odot \mathbf{P}_{-i}}_{\tilde{\mathbf{P}}_{-in}}) = \mathcal{D}_{n\iota,g}(\tilde{\mathbf{P}}_{-in}, \overline{\tilde{\mathbf{P}}_{in}}, \underbrace{\overline{w_n L_n}}_{Y_n}),$$

treating $\overline{\tilde{\mathbf{P}}_{in}} \in \tilde{\mathbb{P}}_i$ and $w_n \in \mathbf{w}$ as given. This function accounts for the fact that any change in the producer price of varieties associated with *origin n–industry g* will affect the consumer prices and the demand schedule in all market excluding i. The reason is that prices in international markets (excluding i) are not directly pinned down by the choice, $\tilde{\mathbb{P}}_i$. For the sake of presentation, abstract from cross-industry demand effects. Applying the Implicit Function Theorem to the system of equations specified by $F_{ni,k}(.)$, yields the following:

$$\begin{bmatrix} \left(\frac{\partial \ln P_{11,k}}{\partial \ln Q_{1i,k}}\right)_{\mathbf{w},\bar{\mathbb{P}}_i} & \cdots & \left(\frac{\partial P_{11,k}}{\partial \ln Q_{Ni,k}}\right)_{\mathbf{w},\bar{\mathbb{P}}_i} \\ \vdots & \ddots & \vdots \\ \left(\frac{\partial \ln P_{NN,k}}{\partial \ln Q_{1i,k}}\right)_{\mathbf{w},\bar{\mathbb{P}}_i} & \cdots & \left(\frac{\partial P_{NN,k}}{\partial \ln Q_{Ni,k}}\right)_{\mathbf{w},\bar{\mathbb{P}}_i} \end{bmatrix} = - \underbrace{\begin{bmatrix} \frac{\partial F_{1i,k}(.)}{\partial \ln P_{11,k}} & \cdots & \frac{\partial F_{1i,k}(.)}{\partial \ln P_{NN,k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{Ni,k}(.)}{\partial \ln P_{11,k}} & \cdots & \frac{\partial F_{Ni,k}(.)}{\partial \ln P_{NN,k}} \end{bmatrix}^{-1}}_{\mathbf{A}_i} \begin{bmatrix} \frac{\partial F_{1i,k}(.)}{\partial \ln Q_{1i,k}} & \cdots & \frac{\partial F_{1i,k}(.)}{\partial \ln Q_{Ni,k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{Ni,k}(.)}{\partial \ln Q_{1i,k}} & \cdots & \frac{\partial F_{Ni,k}(.)}{\partial \ln Q_{Ni,k}} \end{bmatrix}. \tag{E.38}$$

The elements of the matrixes on the right-hand side of the above equation are given by

$$\frac{\partial F_{ni,k}(.)}{\partial \ln P_{jj,k}} = \mathbb{1}_{j=n} + \mathbb{1}_{j \neq i} \times \frac{\mu_k}{1+\mu_k} \sum_{\iota \neq i} r_{n\iota,k} \varepsilon_{n\iota,k}^{(j\iota,k)}; \qquad \frac{\partial F_{ni,k}(.)}{\partial \ln Q_{ji,k}} = \mathbb{1}_{j=n} \frac{\mu_k}{1+\mu_k} r_{ji,k}.$$

Notice that the off-diagonal elements of $\mathbf{A}_i$ are near-zero (i.e., $r_{n\iota,k}\varepsilon_{n\iota,k}^{(j\iota,k)} \propto r_{n\iota,k}\lambda_{j\iota,k} \approx 0$ if $n \neq j \neq \iota$). So, we can apply the method proposed by Wu, Yin, Vosoughi, Studer, Cavallaro, and Dick (2013) to characterize $\mathbf{A}_i^{-1}$ to a first-order approximation around $r_{j\iota,k} \approx \lambda_{j\iota,k} \approx 0$ (for $j \neq \iota$). This procedure is detailed in Appendix E.3 and yields the following expression based on the matrix Equation E.38:

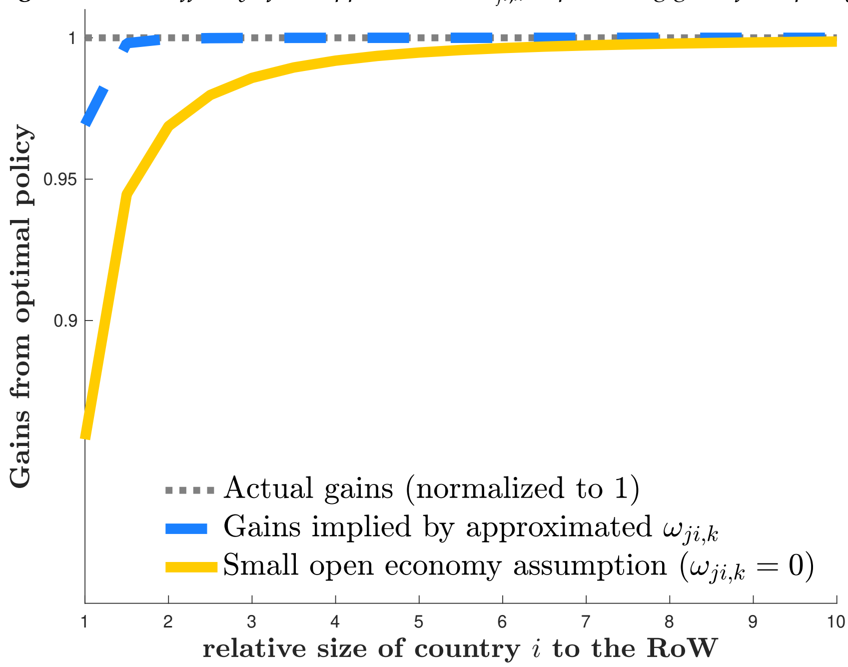
$$\left(\frac{\partial \ln P_{nn,k}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\bar{\mathbb{P}}_i} \approx \begin{cases} \frac{-\frac{\mu_k}{1+\mu_k} r_{ni,k}}{1+\frac{\mu_k}{1+\mu_k}\sum_{\iota \neq i} r_{n\iota,k}\varepsilon_{n\iota,k}^{(n\iota,k)}} & n = j \\ \frac{\frac{\mu_k}{1+\mu_k} r_{ji,k}}{1+\frac{\mu_k}{1+\mu_k}\sum_{\iota \neq i} r_{n\iota,k}\varepsilon_{n\iota,k}^{(n\iota,k)}} \left(\frac{\mu_k}{1+\mu_k}\sum_{\iota \neq i} r_{n\iota,k}\varepsilon_{n\iota,k}^{(j\iota,k)}\right) & n \neq j \end{cases}$$

Plugging the above expression back into the definition specified by Equation E.37, while noting that $r_{ni,k} \times r_{ji,k} \approx 0$ if $j \neq i$ and $n \neq i$, yields the following approximation for the export supply elasticity:

$$\omega_{ji,k} \approx \frac{-\frac{\mu_k}{1+\mu_k} r_{ji,k}}{1+\frac{\mu_k}{1+\mu_k}\sum_{\iota \neq i} r_{j\iota,k}\varepsilon_{j\iota,k}} \left[1 - \frac{\mu_k}{1+\mu_k}\frac{w_i L_i}{w_j L_j}\sum_{n \neq i}\frac{\rho_{i,k} r_{in,k}}{\rho_{j,k} r_{ji,k}}\varepsilon_{in,k}^{(jn,k)}\right].$$

For the sake of clarity, note that $w_i = \bar{w}_i$ under free entry—so, we can replace w_i with $\bar{w}_i$ everywhere in the above approximation. Figure E.1 illustrates the goodness of our approximated $\omega_{ji,k}$ using a rather conservative numerical example. We simulate a two-country×two-industry economy in which trade is relatively open and the tax-imposing country is relatively large compared to the rest of the world.

Figure E.1: *The efficacy of the approximated $\omega_{ji,k}$ at predicting gains from policy*



Note: the above simulation is based on a two country–two industry model with the following specifications: (2) $\sigma_1 = \sigma_2 = 5$, (2) $\mu_1 = 0.25$ and $\mu_2 = 0.5\mu_1$; (3) expenditure shares are assigned the following values $\lambda_{21,1} = 0.6$, $\lambda_{12,1} = 0.25/\delta$, $\lambda_{21,2} = 0.25$; $\lambda_{12,2} = 0.4/\rho$ where ρ is relative size.

We compute the actual gains from optimal policy for the tax-imposing country i, and compare them to gains implied by (1) our approximated $\omega_{ji,k}$ as well (2) the small open economy approximation, $\omega_{ji,k} \approx 0$. Evidently, our approximated value for $\omega_{ji,k}$ yields indistinguishable results relative to approximation-free benchmark.[16]

Step #5. Extending the Derivation to the Restrict Entry Case

Equipped with a full characterization of optimal policy under free entry, we now switch attention to the case of restricted entry. The main difference between the two cases is in how producer prices vary with export supply: Under restricted entry, holding $\mathbf{w} = \{\bar{w}_n\}$ fixed, contacting the export supply of good ni, g affects the producer prices associated with origin n through a uniform reduction in the average markup $\bar{\mu}_n$. Namely,

$$P_{ni,g} = \bar{\varrho}'_{ni,g}\frac{1+\mu_k}{1+\bar{\mu}_n}\bar{w}_n \implies \left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} = -\left(\frac{\partial \ln(1+\bar{\mu}_n)}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i},$$

where economy n's (endogenously-determined) average profit margin is given by

$$1+\bar{\mu}_n = \frac{\sum_{\iota\in\mathbb{C}}\sum_{k\in\mathbb{K}}\left[P_{n\iota,k}Q_{n\iota,k}\right]}{\sum_{\iota\in\mathbb{C}}\sum_{k\in\mathbb{K}}\left[\frac{1}{1+\mu_k}P_{n\iota,k}Q_{n\iota,k}\right]}.$$

Another difference is that non-tax-revenue income in country i is the sum of wage payments plus profits. Stated formally, total income in country i can be specified as follows (*notation*: the operator "$\cdot$" denotes the inner product of two equal-sized vectors):

$$Y_i(\tilde{\mathbb{P}}_i;\mathbf{w}) = \underbrace{(1+\bar{\mu}_i)w_iL_i}_{\bar{w}_iL_i} + \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni}\right)\cdot\mathbf{Q}_{ni}\right] + \left(\tilde{\mathbf{P}}_{ii} - \mathbf{P}_{ii}\right)\cdot\mathbf{Q}_{ii} + \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{in} - \mathbf{P}_{in}\right)\cdot\mathbf{Q}_{in}\right], \quad \text{(E.39)}$$

In the above formulation, $\bar{w}_iL_i = (1+\bar{\mu}_i)w_iL_i$, stands for the sum of wage payments plus profits.

[16] To be clear, the above approximation is only intended for the quantitative applications. It should not be viewed as a limitation of our theory. The optimal tax formula derived earlier in combination with Equation E.38 deliver an exact theoretical specification for the first-best optimal policy schedule.

With the background information provided above, we can recycle our earlier derivations from the free entry case to characterize the F.O.C. w.r.t. each price instrument in $\tilde{\mathbb{P}}_i$.

First-Order Condition w.r.t. $\tilde{P}_{ji,k}$ and $\tilde{P}_{ii,k} \in \tilde{\mathbb{P}}_i$. To fix ideas, recall from Step #3 of the proof that the F.O.C. w.r.t. $\tilde{P}_{ji,k} \in \tilde{\mathbb{P}}_i$ (where possibly $j = i$) is given by

$$\left(\frac{\mathrm{d}W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\mathrm{d}\ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = \frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ji,k}} + \frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial Y_i}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = 0. \tag{E.40}$$

As before, $\tilde{\mathbb{P}}_{-ji,k} \equiv \tilde{\mathbb{P}}_i - \left\{\tilde{P}_{ji,k}\right\}$ denotes the vector of country i's price instruments excluding $\tilde{P}_{ji,k}$. Each term on the right-hand can be unpacked as in the free entry case, with one difference: holding $\mathbf{w}$ constant, a change in good ji,k's export supply affects the entire vector of prices from origin j. Specifically, noting that $P_{ji,g} = \bar{\varrho}'_{ji,g}\frac{1+\mu_g}{1+\bar{\mu}_j}w_j$, indicates that

$$\left(\frac{\partial \ln P_{ji,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} = -\left(\frac{\partial \ln(1+\bar{\mu}_j)}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} \qquad \forall g \in \mathbb{K}.$$

Noting this distinction, we now repeat the steps present earlier to unpack each term on the right-hand side of Equation E.40. By Roy's identity, the first term on the right-hand side can be unpacked as follows:

$$\frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ji,k}} = -\tilde{P}_{ji,k}Q_{ji,k}\left(\frac{\partial V_i}{\partial Y_i}\right).$$

Recall that the second term on the right-hand side of Equation E.40 accounts for the revenue-raising effects of policy. Specifically, taking note of Equation E.39, the effect on import tax revenues can be unpacked as follows:

$$\left(\frac{\partial \sum_{n\neq i}\left(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni}\right)\cdot \mathbf{Q}_{ni}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \tilde{P}_{ji,k}Q_{ji,k} + \sum_{g}\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-P_{ni,g}\right)Q_{ni,g}\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]$$
$$-\sum_{g\in\mathbb{K}}\sum_{n\neq i}\left[P_{ni,g}Q_{ni,g}\sum_{s\in\mathbb{K}}\left[\sum_{j\neq i}\frac{P_{ji,s}Q_{ji,s}}{P_{ni,g}Q_{ni,g}}\left(\frac{\partial P_{ji,s}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} + \sum_{\ell\in\mathbb{C}}\frac{P_{i\ell,s}Q_{i\ell,s}}{P_{i\ell,g}Q_{i\ell,g}}\left(\frac{\partial P_{i\ell,s}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]. \tag{E.41}$$

As in the free entry case, $\left(\partial \ln Q_{ni,g}/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$ encompasses demand adjustments that channel through both price and income effects—see Equation E.5. We can simplify the last term on the right-hand side of above equation, by appealing to our definition of the export supply elasticity:

$$\omega_{ni,g} \equiv \sum_{\ell\in\mathbb{C}}\left[\frac{P_{i\ell,g}Q_{i\ell,g}}{P_{ni,g}Q_{ni,g}}\left(\frac{\partial \ln P_{i\ell,g}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right] + \sum_{j\neq i}\left[\frac{P_{ji,g}Q_{ji,g}}{P_{ni,g}Q_{ni,g}}\left(\frac{\partial \ln P_{ji,g}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right]$$
$$= \frac{1}{r_{ni,g}\rho_{n,g}}\sum_{g}\left[\frac{\bar{w}_iL_i}{\bar{w}_nL_n}\rho_{i,g}\left(\frac{\partial \ln(1+\bar{\mu}_i)}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} + \sum_{j\neq i}\frac{\bar{w}_jL_j}{\bar{w}_nL_n}r_{ji,g}\rho_{j,g}\left(\frac{\partial \ln(1+\bar{\mu}_n)}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right] \tag{E.42}$$

The second line indicates our focus on the restricted entry case, wherein $\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} = \left(\frac{\partial \ln(1+\bar{\mu}_n)}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}$ for all g. That is, holding $\mathbf{w}$ constant, producer prices from each origin change equal-proportionally across all industries with the aggregate profit margin, $1+\bar{\mu}_i$. Plugging the above expression back into Equation E.41 yields the following expression that summarizes the (conditional) effect of policy on import tax revenues:

$$\left(\frac{\partial \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni}\right)\cdot\mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \tilde{P}_{ji,k}Q_{ji,k}+\sum_{g}\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-[1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]. \tag{E.43}$$

The effect of policy on export and domestic tax revenues can be unpacked as in Equation E.7, which was derived earlier for the free entry case. To simplify this equation under restricted entry, we can use the following observation:

$$\sum_{n}\left[P_{in,g}Q_{in,g}\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{i}}\right] = -\sum_{s\in\mathbb{K}}\sum_{n\in\mathbb{C}}\left[\frac{P_{in,s}Q_{in,s}}{P_{ii,g}Q_{ii,g}}\left(\frac{\partial \ln(1+\overline{\mu}_i)}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{i}}\right]P_{ii,g}Q_{ii,g} =$$

$$= \sum_{s\in\mathbb{K}}\sum_{n\in\mathbb{C}}\left[\frac{r_{in,s}\rho_{i,s}}{r_{ii,g}\rho_{i,g}}\left(\frac{\overline{\mu}_i-\mu_g}{1+\mu_g}r_{ii,g}\rho_{i,g}\right)\right]P_{ii,g}Q_{ii,g} = -\left(1-\frac{1+\overline{\mu}_i}{1+\mu_g}\right)P_{ii,g}Q_{ii,g},$$

To explain, the second line on the above equation follows from that fact that all prices associated with economy i are included in the set $\tilde{\mathbb{P}}_i$. So, holding $\tilde{\mathbb{P}}_i$ and wages $\mathbf{w}$ constant, the policy-induced change in $Q_{ii,g}$ has only a direct arithmetic effect on country i's aggregate profit margin, i.e., $\left(\frac{\partial \ln(1+\overline{\mu}_i)}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} = \frac{\overline{\mu}_i-\mu_g}{1+\mu_g}r_{ii,g}\rho_{i,g}$.[17] Plugging the above equation back into Equation E.7 yields the following equation describing the (conditional) effects of policy on export and domestic tax revenues:

$$\left(\frac{\partial}{\partial \ln \tilde{P}_{ji,k}}\left\{\sum_{n}\left[\left(\tilde{\mathbf{P}}_{in}-\mathbf{P}_{in}\right)\cdot\mathbf{Q}_{in}\right]\right\}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \sum_{g}\left[\left(\tilde{P}_{ii,g}-\frac{1+\overline{\mu}_i}{1+\mu_g}P_{ii,g}\right)Q_{ii,g}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]. \tag{E.44}$$

Recall that $\left(\partial \ln Q_{ni,g}/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$, in the above equations, encompasses price- and income-related demand adjustments—see Equation E.5. Taking note of this detail, we can combine Equations E.43 and E.44 to arrive at the following expression that summarizes the (conditional) effect of raising $\tilde{P}_{ji,k}$ on country i's tax revenues:

$$\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \tilde{P}_{ji,k}Q_{ji,k}+\sum_{g}\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-[1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\varepsilon^{(ji,k)}_{ni,g}\right]$$

$$+\sum_{g}\left[\left(\tilde{P}_{ii,g}-\frac{1+\overline{\mu}_i}{1+\mu_g}P_{ii,g}\right)Q_{ii,g}\varepsilon^{(ji,k)}_{ii,g}\right]+\Delta_i(\tilde{\mathbb{P}}_i;\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}},$$

where $\Delta_i(.)$, as before, encapsulated the circular income effects. The expression for $\Delta_i(.)$ is specified analogously to Equation E.13 with two amendments: (1) $\omega_{ni,g}$ is redefined according to E.42; and (2) $1/(1+\mu_g)$ replaced with $(1+\overline{\mu}_i)/(1+\mu_g)$.[18] Next, we unpack the last term on the right-hand side of Equation E.40, which accounts for general equilibrium wage effects. Repeating the steps presented

[17] Note that this argument does not extend to the aggregate profit margin in other countries. Changing the export supply of say good ji,k with policy has a circular effect on origin j's profit margin, $\overline{\mu}_j$, which occurs because the prices associated with economy $j\neq i$ are not pegged to $\tilde{\mathbb{P}}_i$. Specifically, a change in $Q_{ji,k}$ affects the entire vector of origin j's prices outside of market i. This change in prices affects the industrial composition of origin j's output and $\overline{\mu}_j$ in a circular fashion.

[18] To be more specific, $\Delta_i(.)$ is described by the following equation:

$$\Delta_i(\tilde{\mathbb{P}}_i;\mathbf{w}) \equiv \sum_{g}\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-[1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\eta_{ni,g}\right]+\sum_{g}\left[\left(\tilde{P}_{ii,g}-\frac{1+\overline{\mu}_i}{1+\mu_g}P_{ii,g}\right)Q_{ii,g}\eta_{ii,g}\right],$$

where $\overline{\mu}_i>0$ and $\omega_{ni,g}$ is given by Equation E.42 for the case of restricted entry.

for the free entry case, while noting the differences discussed above, yields the following:

$$\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln\tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = -\sum_g\sum_{n\neq i}\left[\bar{\tau}_i\left(1+\omega_{ni,g}\right)P_{ni,g}Q_{ni,g}\varepsilon_{ni,g}^{(ji,k)}\right],$$
$$-\sum_g\sum_{n\neq i}\left[\bar{\tau}_i\left(1+\omega_{ni,g}\right)P_{ni,g}Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$$

where $\bar{\tau}_i$ is given by E.15. Note that the above expression differs from the analogous expression derived under free entry in the economic forces that regulate export supply elasticity, $\omega_{ni,g}$. Under restricted entry, the export supply elasticity governs the change in aggregate profit margins in response to distortions to export supply. Combining the various terms on the right-hand side of Equation E.40, yields the following simplified representation of the F.O.C. w.r.t. $\tilde{P}_{ji,k}\in\tilde{\mathbb{P}}_i$ under restricted entry:

$$\sum_{n\neq i}\sum_g\left[\left(\frac{\tilde{P}_{ni,g}}{P_{ni,g}}-(1+\bar{\tau}_i)\right)P_{ni,g}Q_{ni,g}\varepsilon_{ni,g}^{(ji,k)}\right]$$
$$+\sum_g\left[\left(\frac{\tilde{P}_{ii,g}}{P_{ii,g}}-\frac{1+\overline{\mu}_i}{1+\mu_g}\right)P_{ii,g}Q_{ii,g}\varepsilon_{ii,g}^{(ji,k)}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i;\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}=0.$$

The uniform term $\widetilde{\Delta}_i(.)$ is described by Equation E.18, but with $\omega_{ni,g}$ redefined according to E.42 and $1/(1+\mu_g)$ replaced with $(1+\overline{\mu}_i)/(1+\mu_g)$.

First-Order Condition w.r.t. $\tilde{P}_{ij,k}$ ($j\neq i$). Now consider the F.O.C. w.r.t. the price of a generic export good ij,k (where $j\neq i$). Recall from Step #3 that the F.O.C. w.r.t. $\tilde{P}_{ij,k}\in\tilde{\mathbb{P}}_i$ is given by

$$\left(\frac{\mathrm{d}W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\mathrm{d}\ln\tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}}=\frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial\ln\tilde{P}_{ij,k}}+\frac{\partial V_i(Y_i,\tilde{\mathbf{P}}_i)}{\partial Y_i}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}+\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln\tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}}=0. \tag{E.45}$$

where $\tilde{\mathbb{P}}_{-ij,k}\equiv\tilde{\mathbb{P}}_i-\left\{\tilde{P}_{ij,k}\right\}$ denotes the vector of country i's price instruments excluding $\tilde{P}_{ij,k}$. Building on our previous discussion, each term on the right-hand side is characterized by the same formulas derived in Step #3, with two qualification: (1) The formulation assigned to $\omega_{ni,g}$ should be revised to account for restricted entry (see Equation 10), (2) all equations should be adjusted to admit a non-zero $\overline{\mu}_j$, as is required by restricted entry (see Equation 5).

Without repeating all the details from Step 3, we can unpack the terms on the right-hand side of Equation E.45 as follows: Since $\tilde{P}_{ij,k}\notin\tilde{\mathbf{P}}_i$ is not part of the domestic consumer price index, $\partial V_i(Y_i,\tilde{\mathbf{P}}_i)/\partial\ln\tilde{P}_{ji,k}=0$. The second-term on the right-hand side of Equation E.45 is given by:

$$\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=\tilde{P}_{ij,k}Q_{ij,k}+\sum_g\left[\left(\tilde{P}_{ij,g}-[1-\frac{\mu_g}{1+\mu_g}]P_{ij,g}\right)Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$-\sum_g\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right]+\Delta_i(\tilde{\mathbb{P}}_i;\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}},$$

where $\omega_{ni,g}$ is defined as in Equation E.42, while $\Delta_i(.)$ is given by Equation E.13, with the necessary adjustments described earlier.The last term on the right-hand side of Equation E.45 , which accounts for general equilibrium wage effects, can be unpacked as

$$\left(\frac{\partial W_i(.)}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{d\mathbf{w}}{d\ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = \bar{\tau}_i \tilde{P}_{ij,k} Q_{ij,k} + \sum_g \left[\bar{\tau}_i \tilde{P}_{ij,g} Q_{ij,g} \varepsilon_{ij,g}^{(ij,k)}\right]$$
$$- \sum_g \sum_{n\neq i} \left[\bar{\tau}_i \omega_{ni,g} P_{nj,g} Q_{nj,g} \varepsilon_{nj,g}^{(ij,k)}\right] - \sum_g \sum_{n\neq i} \left[[1+\omega_{ni,g}]\bar{\tau}_i P_{ni,g} Q_{ni,g} \eta_{ni,g}\right] \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}.$$

where $\bar{\tau}_i$ is given by E.15. To be clear, the above formula differs from the one derived under free entry in only how $\omega_{ni,g}$ is defined—see Equation E.42. Combining the various terms on the right-hand side of Equation E.45, yields the following simplified representation of the F.O.C. w.r.t. $\tilde{P}_{ij,k} \in \tilde{\mathbb{P}}_i$:

$$\tilde{P}_{ij,k} Q_{ij,k} + \sum_{g\in\mathbb{K}} \left[\left(1 - \frac{1+\overline{\mu}_i}{(1+\bar{\tau}_i)(1+\mu_g)} \frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right) \tilde{P}_{ij,g} Q_{ij,g} \varepsilon_{ij,g}^{(ij,k)}\right]$$
$$\sum_{g\in\mathbb{K}} \sum_{n\neq i} \left[\omega_{ni,g} P_{nj,g} Q_{nj,g} \varepsilon_{nj,g}^{(ij,k)}\right] + \tilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w}) \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = 0,$$

The uniform term $\widetilde{\Delta}_i(.)$ is described by Equation E.18, but with $\omega_{ni,g}$ redefined according to Equation E.42 and $1/(1+\mu_g)$ replaced with $(1+\overline{\mu}_i)/(1+\mu_g)$.

Solving the system of F.O.C. Given the tight correspondence between the F.O.C.s derived under the restricted and free entry cases, we can repeat the arguments as in step #4 to solve the system of F.O.C.s and establish the uniqueness of the resulting solution. Doing so yield the following formula for optimal taxes/subsidies under restricted entry:

$$\begin{aligned} &[\text{domestic subsidy}] \quad 1+s_{i,k}^* = (1+\mu_k)/(1+\overline{\mu}_i) \\ &[\text{import tariff}] \quad 1+t_{ji,k}^* = (1+\omega_{ji,k})(1+\bar{\tau}_i) \\ &[\text{export subsidy}] \quad \mathbf{1}+\mathbf{x}_{ij}^* = -\mathbf{E}_{ij}^{-1}\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1}+\mathbf{t}_i^*\right). \end{aligned}$$

Recall from Lemma 1 that there are two degrees of multiplicity associated with optimal policy schedule. As a result, we need not to unpack the uniform terms $\bar{\tau}_i$ and $\overline{\mu}_i$. Instead, for any arbitrary choice of tax shifters $1+\bar{s}_i$ and $1+\bar{t}_i \in \mathbb{R}_+$, the following tax/subsidy schedule represents an optimal solution:

$$\begin{aligned} &[\text{domestic subsidy}] \quad 1+s_{i,k}^* = (1+\mu_k)(1+\bar{s}_i) \\ &[\text{import tariff}] \quad 1+t_{ji,k}^* = (1+\omega_{ji,k})(1+\bar{t}_i) \\ &[\text{export subsidy}] \quad \mathbf{1}+\mathbf{x}_{ij}^* = -\mathbf{E}_{ij}^{-1}\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1}+\mathbf{t}_i^*\right). \end{aligned}$$

The above formula is identical to that derived under free entry, with one qualification. The (inverse) export supply elasticity $\omega_{ji,k}$ has a different interpretation under restricted entry, and is given by E.42. So, to conclude the proof, we characterize $\omega_{ji,k}$ under restricted entry next.

Characterizing the (Inverse) Export Supply Elasticity. Following Equation E.42, the inverse of the export supply elasticity under restricted entry is defined as

$$\omega_{ji,k} = \frac{-1}{r_{ji,k}\rho_{j,k}} \left[\frac{\varpi_i L_i}{\varpi_j L_j} \left(\frac{\partial \ln(1+\overline{\mu}_i)}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} + \sum_{n\neq i} \sum_g \left(\frac{\varpi_n L_n}{\varpi_j L_j} r_{ni,g}\rho_{n,g}\right) \left(\frac{\partial \ln(1+\overline{\mu}_n)}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\right], \quad \text{(E.46)}$$

where the second line follows from the fact that $P_{ni,s} = \bar{\varrho}'_{ni,s}\frac{1+\mu_s}{1+\overline{\mu}_n}\varpi_n$, which implies that $\left(\frac{\partial \ln P_{ni,s}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} = -\left(\frac{\partial \ln(1+\overline{\mu}_n)}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}$. To unpack the above equation, note that (for a given $\tilde{\mathbb{P}}_i$ and $\mathbf{w}$) the aggregate profit margin implicitly solves the following equation:

$$F_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni}) = (1+\overline{\mu}_n) - \underbrace{\frac{\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} + \sum_{\iota\neq i}\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\boldsymbol{\mu}}_{-i})}{\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} + \sum_{\iota\neq i}\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\boldsymbol{\mu}}_{-i})}}_{g_{ni}(\overline{\mu}_n,\mathbf{Q}_{ni})} = 0.$$

As before, $\odot$ and $\cdot$ respectively denote the *inner* and *element-wise* products of equal-sized vectors (i.e., $\mathbf{a}\cdot\mathbf{b} = \sum_n a_n b_n$ and $\mathbf{a}\odot\mathbf{b} = [a_n b_n]_n$), while with a slight abuse of notation, $\frac{1}{1+\boldsymbol{\mu}} \equiv \left[\frac{1}{1+\mu_k}\right]_k$. The vector $\mathbf{Q}_{ni}$ represents the export supply of goods from origin $n\neq i$ to market i (which is fully determined by $\tilde{\mathbb{P}}_i$ and $\mathbf{w}$). Outside of market i, consumer prices are not directly pegged to $\tilde{\mathbb{P}}_i$. So, holding $\bar{w}_n \in \mathbf{w}$ and $\tilde{\mathbf{P}}_{i\iota} \in \tilde{\mathbb{P}}_i$ constant, a change in $\overline{\mu}_i$ affects the producer and consumer price of goods supplied by origin n to any market $\iota\neq i$. Accordingly, $\mathbf{Q}_{n\iota}(\overline{\boldsymbol{\mu}}_{-i}) \equiv \left\{Q_{n\iota,k}(\overline{\boldsymbol{\mu}}_{-i})\right\}_k$ in the above equation is determined by the Marshallian demand function (treating $\bar{w}_n \in \mathbf{w}$ and $\tilde{\mathbf{P}}_{i\iota} \in \tilde{\mathbb{P}}_i$ as given):

$$Q_{n\iota,k}(\overline{\boldsymbol{\mu}}_{-i}) = \mathcal{D}_{n\iota,k}(\overline{\bar{w}_\iota L_\iota}, \overline{\tilde{\mathbf{P}}_{i\iota}}, \tilde{\mathbf{P}}_{-i\iota}(\overline{\boldsymbol{\mu}}_{-i}))$$

Taking note of this detail, we can compute $\left(\partial\ln(1+\overline{\mu}_n)/\partial\ln Q_{ni,g}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}$ by applying the Implicit Function Theorem to the system of equations specified by $F_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni})$. Namely,

$$\begin{bmatrix} \frac{\partial\ln(1+\overline{\mu}_1)}{\partial\ln\mathbf{Q}_{1i}} & \cdots & \frac{\partial(1+\overline{\mu}_1)}{\partial\ln\mathbf{Q}_{Ni}} \\ \vdots & \ddots & \vdots \\ \frac{\partial\ln(1+\overline{\mu}_N)}{\partial\ln\mathbf{Q}_{1i}} & \cdots & \frac{\partial(1+\overline{\mu}_N)}{\partial\ln\mathbf{Q}_{Ni}} \end{bmatrix} = -\begin{bmatrix} \frac{\partial F_{1i}(.)}{\partial\ln(1+\overline{\mu}_1)} & \cdots & \frac{\partial F_{1i}(.)}{\partial\ln(1+\overline{\mu}_N)} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{Ni}(.)}{\partial\ln(1+\overline{\mu}_1)} & \cdots & \frac{\partial F_{Ni}(.)}{\partial\ln(1+\overline{\mu}_N)} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_{1i}}{\partial\ln\mathbf{Q}_{1i}} & \cdots & \frac{\partial F_{1i}}{\partial\ln\mathbf{Q}_{Ni}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{Ni}}{\partial\ln\mathbf{Q}_{1i}} & \cdots & \frac{\partial F_{Ni}}{\partial\ln\mathbf{Q}_{Ni}} \end{bmatrix}. \tag{E.47}$$

Next, we characterize the elements of the matrixes on the right-hand side of the above equation. Considering that $F_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni}) = (1+\overline{\mu}_n) - g_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni})$, we can unpack the elements of $\left[\frac{\partial F_{ni}(.)}{\partial\ln(1+\overline{\mu}_j)}\right]_{n,j}$ as follows. Using vector algebra we can show that if $n\neq i$, then

$$\frac{\partial g_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni})}{\partial\ln(1+\overline{\mu}_n)} = \underbrace{\frac{-\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} - \sum_{\iota\neq i}\left[\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}{\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} + \sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}}_{-(1+\overline{\mu}_n)} + \underbrace{\frac{-\sum_{\iota\neq i}\left[\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\odot\varepsilon_{n\iota}\right]}{\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} + \sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}}_{-(1+\overline{\mu}_n)\sum_\iota[\mathbf{r}_{n\iota}\cdot\varepsilon_{n\iota}]}$$

$$-(1+\overline{\mu}_n)\left[\underbrace{\frac{-\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} - \sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}{\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} + \sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}}_{-1} - \underbrace{\frac{\sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\odot\varepsilon_{n\iota}\right]}{\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} + \sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}}_{(1+\overline{\mu}_n)\sum_\iota\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{r}_{n\iota}\cdot\varepsilon_{n\iota}\right]}\right].$$

where $\boldsymbol{\varepsilon}_{ni} \equiv \left[\varepsilon_{ni,g}^{(ni,g)}\right]_g$ is a $K\times 1$ vector of own-price elasticities of demand. $\mathbf{r}_{ni} \equiv \left[r_{ni,g}\rho_{n,g}\right]_g$ is a $K\times 1$ vector of sales shares. The above derivation appeals to the definition of sales shares, whereby $r_{ni,k}\rho_{n,k} = \frac{P_{ni,k}Q_{ni,k}}{\sum_j\sum_g P_{nj,g}Q_{nj,g}}$. Likewise, for any n and $\ell\neq i$, we can

$$\frac{\partial g_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni})}{\partial\ln(1+\overline{\mu}_\ell)} = \frac{-\sum_{\iota\neq i}\left[\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\odot\varepsilon_{n\iota}^{(\ell\iota)}\right] + (1+\overline{\mu}_n)\sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\odot\varepsilon_{n\iota}^{(\ell\iota)}\right]}{\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{ni}(\overline{\mu}_n)\cdot\mathbf{Q}_{ni} + \sum_{\iota\neq i}\left[\frac{1}{1+\boldsymbol{\mu}}\odot\mathbf{P}_{n\iota}(\overline{\mu}_n)\cdot\mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}.$$

Combining the above two equations we can characterizes each element of the matrix $\left[\frac{\partial F_{ni}(.)}{\partial\ln(1+\overline{\mu}_\ell)}\right]_{n,\ell}$ as follows:

$$\frac{\partial F_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni})}{\partial\ln(1+\overline{\mu}_\ell)} = (1+\overline{\mu}_n)\left[\mathbb{1}_{\ell=n} + \mathbb{1}_{\ell\neq i}\sum_k\sum_{\iota\neq i}\left[\left(1-\frac{1+\overline{\mu}_n}{1+\mu_k}\right)r_{n\iota,k}\rho_{n,k}\varepsilon_{n\iota,k}^{(\ell\iota,k)}\right]\right]$$

The elements of the matrix $\left[\frac{\partial F_{ni}}{\partial \ln \mathbf{Q}_{\ell i}}\right]_{n,\ell}$ can be unpacked with a similar logic. Specifically, if $n \neq \ell$ then $\frac{\partial F_{ni}}{\partial \ln \mathbf{Q}_{\ell i}} = 0$. Otherwise, for any $n \in \mathbb{C}$ we can derive the following expression:

$$\frac{\partial g_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni})}{\partial \ln Q_{ni,k}} = \underbrace{\frac{P_{ni,k}Q_{ni,k}}{\frac{1}{1+\boldsymbol{\mu}} \odot \mathbf{P}_{ni}(\overline{\mu}_n) \cdot \mathbf{Q}_{ni} + \sum_{\iota \neq i}\left[\frac{1}{1+\boldsymbol{\mu}} \odot \mathbf{P}_{n\iota}(\overline{\mu}_n) \cdot \mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}}_{(1+\overline{\mu}_n)r_{ni,k}\rho_{n,k}} - \underbrace{\frac{(1+\overline{\mu}_n)\frac{1}{1+\mu_k}P_{ni,k}Q_{ni,k}}{\frac{1}{1+\boldsymbol{\mu}} \odot \mathbf{P}_{ni}(\overline{\mu}_n) \cdot \mathbf{Q}_{ni} + \sum_{\iota \neq i}\left[\frac{1}{1+\boldsymbol{\mu}} \odot \mathbf{P}_{n\iota}(\overline{\mu}_n) \cdot \mathbf{Q}_{n\iota}(\overline{\mu}_n)\right]}}_{(1+\overline{\mu}_n)\frac{1+\overline{\mu}_n}{1+\mu_k}r_{ni,k}\rho_{n,k}},$$

which, in turn, characterizes every element of matrix $\left[\frac{\partial F_{ni}}{\partial \ln \mathbf{Q}_{\ell i}}\right]_{n,\ell}$ as follows:

$$\frac{\partial F_{ni}(\overline{\boldsymbol{\mu}}, \mathbf{Q}_{ni})}{\partial \ln Q_{\ell i,k}} = \mathbb{1}_{\ell=n}(1+\overline{\mu}_n)\left[\left(1 - \frac{1+\overline{\mu}_n}{1+\mu_k}\right) r_{ni,k}\rho_{n,k}\right].$$

As in the free entry case, the off-diagonal elements of $\tilde{\mathbf{A}}_i \equiv \left[\frac{\partial F_{ni}(.)}{\partial \ln(1+\overline{\mu}_j)}\right]_{n,j}$ are near zero. So, we can once again invoke the first-order approximation proposed by Wu et al. (2013) to characterize $\tilde{\mathbf{A}}_i^{-1}$. Doing so and plugging the implied values of $\frac{\partial \ln(1+\overline{\mu}_n)}{\partial \ln \mathbf{Q}_{ji}}$ back into Equation E.46, implies the following approximation for the export supply elasticity under restricted entry:

$$\omega_{ni,g} \approx \frac{-\left(1 - \frac{1+\overline{\mu}_n}{1+\mu_g}\right)\sum_k r_{ni,k}\rho_{n,k}}{1 + \sum_k \sum_{\iota \neq i}\left[1 + \left(1 - \frac{1+\overline{\mu}_n}{1+\mu_k}\right) r_{n\iota,k}\rho_{n,k}\varepsilon_{n\iota,k}\right]}.$$

E.1 Redundancy of Solutions for which $\Delta_i \neq 0$

To finalize the proof, we appeal to the multiplicity of optimal taxes to show the following: If there exists an optimal tax vector for which $\widetilde{\Delta}_i \neq 0$, that tax vector can be recovered from an equivalent optimal tax vector that satisfies $\widetilde{\Delta}_i = 0$. This can be shown building on two intermediate points: *First*, following Lemma 1, if $\mathbb{T}_i^* = \left\{t_{ji,k}^*, x_{ij,k}^*, s_{i,k}^*\right\}$ is an optimal policy choice, then policy $\mathbb{T}_i^{\star}(a) = \left\{t_{ji,k}^{\star}(a), x_{ij,k}^{\star}(a), s_{ii,k}^{\star}(a)\right\}$ is also optimal for any $a \in \mathbb{R}$, where

$$1 + t_{ji,k}^{\star}(a) = \left(1 + t_{ji,k}^*\right)(1+a)^{-1}; \qquad 1 + x_{ij,k}^{\star}(a) = \left(1 + x_{ij,k}^*\right)(1+a); \qquad 1 + s_{i,k}^{\star}(a) = \left(1 + s_{i,k}^*\right)(1+a)^{-1}.$$

Second, the aggregate term $\widetilde{\Delta}_i$, specified below in terms of taxes, appears identically in the F.O.C.s associated with every policy instrument:

$$\widetilde{\Delta}_i = \sum_g \sum_{n \neq i}\left[\left(1 - \frac{(1+\omega_{ni,g})(1+\bar{\tau}_i)}{1+t_{ni,g}}\right) e_{ni,g}\eta_{ni,g}\right] + \sum_g \left[\left(1 - \frac{1+s_{i,g}}{1+\mu_g}\right) e_{ii,g}\eta_{ii,g}\right].$$

Suppose there exists an optimal policy choice, $\mathbb{T}_i^*$ for which $\widetilde{\Delta}_i^* \neq 0$. Analogously, let $\widetilde{\Delta}_i^{\star}(a)$ denote the aggregate term collecting income effects under the equivalent policy choice, $\mathbb{T}_i^{\star}(a)$—with $\widetilde{\Delta}_i^* = \widetilde{\Delta}_i^{\star}(1)$, by construction. In particular,

$$\widetilde{\Delta}_i^{\star}(a) = \sum_g \sum_{n \neq i}\left[\left(1 - \frac{(1+\omega_{ni,g})(1+\bar{\tau}_i)}{(1+a)\left(1+t_{ni,g}^*\right)}\right) e_{ni,g}\eta_{ni,g}\right] + \sum_g \left[\left(1 - \frac{1+s_{i,g}}{(1+\mu_g)(1+a)}\right) e_{ii,g}\eta_{ii,g}\right],$$

where all equilibrium variables are evaluated at $\mathbb{T}_i^{\star}(a)$. Based on Lemma 1, all variables in the above equation (e.g., $e_{ni,g}$, $\omega_{ni,g}$, $\eta_{ni,g}$, etc.) are independent of a—since varying a preserves real equilibrium outcomes based on Lemma 1. Considering this, it should be the case that

$$\lim_{a \to -1} \widetilde{\Delta}_i^{\star}(a) < 0; \qquad \lim_{a \to \infty} \widetilde{\Delta}_i^{\star}(a) > 0.$$

So, following the Intermediate Value Theorem, there exists an $a \in (-1, \infty)$ such that $\widetilde{\Delta}_i^{\star}(a) = 0$. That is, if we suspect there to be an optimal policy $\mathbb{T}_i^*$ satisfying $\widetilde{\Delta}_i^* \neq 0$, that policy can be recovered by re-scaling an optimal policy, $\mathbb{T}_i^{\star}(a)$, that is welfare-equivalent to $\mathbb{T}_i^*$but satisfies $\widetilde{\Delta}_i^{\star}(a) = 0$.

E.2 Non-Optimality of Prohibitive Taxes

Since prohibitive taxes exclude goods from the system of F.O.C.s, we must prove that a tax schedule that prohibits say good ji,k but satisfies the F.O.C.s w.r.t. all other goods is not optimal. We prove this point separately for taxes applied to domestically-consumed goods and taxes applied to export goods.

Prohibitive tax on domestically-consumed good ji,k—We first provide a generic proof starting from the first principles to communicate the logic behind the non-optimality of prohibitive taxes. Then, we provide an alternative proof invoking the system of F.O.C.s derived earlier. To articulate our generic proof, suppose without loss of generality that good ii,k is the good not subjected to a prohibitive tax. Utility maximization entails that $\partial U_i(\mathbf{Q}_i)/\partial Q_{ji,k} = \lambda_i^{\mathcal{L}} \tilde{P}_{ji,k}$ for all ji,k, where $\lambda_i^{\mathcal{L}}$ is the Lagrange multiplier associated with the representative consumer's budget constraint ($\tilde{\mathbf{P}}_i \cdot \mathbf{Q}_i \leq Y_i$). Assuming that $U_i(.)$ satisfies the Inada conditions, utility maximization implies that the marginal utility associated with good ji,k at the prohibitive price is infinitely large; and so is the marginal rate of substitution between goods ji,k and ii,k:

$$\lim_{\frac{\tilde{P}_{ji,k}}{\tilde{P}_{ii,k}} \to \infty} \frac{\partial U_i/\partial Q_{ji,k}}{\partial U_i/\partial Q_{ii,k}} = \infty.$$

Let $F_k\left(\mathbf{Q}_i; \tilde{\mathbf{P}}_i, \tilde{\mathbf{P}}_{-i}\right) = 0$ denote the transformation frontier for country i, which reflects country i's production possibility frontier (PPF) augmented for its ability to transform exports to imports subject to balanced trade. Following Dixit and Norman (1980), the relative marginal rate of transformation, $\frac{\partial F_i/\partial Q_{ii,k}}{\partial F_i/Q_{ji,k}}$, is finite if the utility and production functions satisfy the Inada conditions and $Q_{ii,k}$ is strictly positive—which is the case since good ii,k is not subjected to a prohibitive price. As a result,

$$\frac{\tilde{P}_{ji,k}}{\tilde{P}_{ii,k}} \to \infty \quad \Longrightarrow \quad \frac{\partial U_i/\partial Q_{ji,k}}{\partial U_i/\partial Q_{ii,k}} > \frac{\partial F_i/\partial Q_{ii,k}}{\partial F_i/Q_{ji,k}},$$

indicating that (when $\tilde{P}_{ji,k} = \infty$, $Q_{ji,k} = 0$) the marginal rate of substitution between $Q_{ji,k}$ and $Q_{ii,k}$ exceeds the marginal rate of transformation. Hence, welfare ($U_i \sim W_i$) can be improved by increasing the consumption of good ji,k relative to ii,k, which entails lowering the price of ji,k from its prohibitive level. Prohibitive taxes, as such, cannot be optimal unless the scale or substitution elasticities are unbounded. This result echoes the *Grinols-Wong* theorem that a piecemeal reduction in prohibitive tariffs is welfare improving (see Feenstra (2015, P. 198)). The logic is that a prohibitively-taxed good exhibits an infinitely large marginal utility. So, the gains from restoring its consumption dominate the possible efficiency loss from cross-substitution and a lower scale-of-production on other goods.

The above point can be alternatively proven by appealing to the F.O.C.s specified by Equation E.17. Suppose all prices other than $\tilde{P}_{ji,k}$ are set to their non-prohibitive optimal level. let $\underline{\tilde{\mathbb{P}}}_i^* = \left\{P_{ji,k}, \tilde{\mathbb{P}}^*_{-ji,k}\right\}$ denote the policy vector representing this choice of prices. Following Equation E.17, the marginal welfare effects of adjusting $\tilde{P}_{ji,k}$ in the neighborhood of $\underline{\tilde{\mathbb{P}}}_i^*$ is

$$\frac{\partial W_i}{\partial \ln \tilde{P}_{ji,k}} |_{\underline{\tilde{\mathbb{P}}}_i^*} = \left(\frac{\tilde{P}_{ji,k}}{P_{ji,k}} - (1+\bar{\tau}_i)\left(1+\omega_{ji,k}\right)\right) P_{ni,g} Q_{ni,g} \left[\varepsilon_{ji,k}^{(ji,k)} + \eta_{ji,k}\left(\frac{\partial Y_i}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right].$$

Notice that $\varepsilon_{ji,k}^{(ji,k)} < 0$ and $\eta_{ji,k} > 0$, since the demand function is assumed to be well-behaved. Also, the tax revenues from good ji,k (namely,$T_{ji,k}$) approach zero *from above* as $\tilde{P}_{ji,k} \to \infty$. Hence, for sufficiently large values of $\tilde{P}_{ji,k}$, it must be the case that

$$\left(\frac{\partial \ln Y_i}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \left(\frac{\partial \ln T_{ji,k}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} < 0.$$

The above equation, correspondingly, implies that $\frac{\partial W_i}{\partial \ln \tilde{P}_{ji,k}} \mid_{\tilde{\mathbb{P}}_i^*}$ is negative when the tax rate on good ji,k is sufficiently large (or nearly-prohibitive):

$$\frac{\tilde{P}_{ji,k}}{P_{ji,k}} \gg (1+\bar{\tau}_i)\left(1+\omega_{ji,k}\right) \implies \frac{\partial W_i}{\partial \ln \tilde{P}_{ji,k}} \mid_{\tilde{\mathbb{P}}_i^*} < 0.$$

The above result reveals that lowering $\tilde{P}_{ji,k}/P_{ji,k}$ starting from a prohibitive price/tax rate will improve welfare (W_i)—asserting that a prohibitive tax, which excludes good ji,k from the system of F.O.C.s, cannot be optimal. The same logic can be applied to show that a prohibitive tax on two-or-more goods is not optimal either.

Prohibitive tax on export good ij,k—The price of export good, ij,k, does not explicitly enter the representative consumer's indirect utility function, $V_i\left(Y_i, \tilde{\mathbf{P}}_i\right)$. So, given the government's choice w.r.t. $\tilde{\mathbf{P}}_i \subset \tilde{\mathbb{P}}_i$, the choice of $\tilde{P}_{ij,k} \in \tilde{\mathbb{P}}_i$influences welfare solely through its effect on tax revenues, which contribute to income, Y_i. Under a prohibitive export tax rate, i.e., $\tilde{P}_{ij,k} = \infty$, the export tax revenues associated with good ij,k are trivially zero, i.e., $T_{ij,k} = 0$. Lowering $\tilde{P}_{ij,k}$ from its prohibitive level elevates $T_{ij,k}$ to a positive value and, thus, raises total tax revenues, T_i.[19] Lowering $\tilde{P}_{ij,k}$ from its prohibitive level, thus, raises income and thereby welfare given that $\partial V_i(.)/\partial Y_i > 0$. This chain of arguments asserts that prohibitive export taxes cannot be optimal since they yield the lowest possible tax revenue—resonating with the conventional Laffer curve argument. The same point can be demonstrated using the F.O.C.s specified by Equation E.31. In particular, suppose all prices other than $\tilde{P}_{ij,k}$ are set to their non-prohibitive optimal level. Equation E.31 indicates that the marginal welfare effect of lowering $\tilde{P}_{ij,k}$ is strictly positive *if* the initial value assigned to $\tilde{P}_{ij,k}$ is arbitrarily large.

E.3 Approximate Export Supply Elasticity: *Derivation Details*

This appendix provides a detailed derivation of our approximate export supply elasticity formula. Our derivation, recall, relies on the approximate matrix inversion technique developed by Wu et al. (2013). Note that the inverse export supply elasticity (when country i is the tax-imposing authority) is

$$\omega_{ji,k} \equiv \frac{1}{r_{ji,k}\rho_{j,k}} \sum_{g\in\mathbb{K}} \left[\frac{w_i L_i}{w_j L_j} \rho_{i,g} \left(\frac{\partial \ln P_{ii,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} + \sum_{n\neq i} \frac{w_n L_n}{w_j L_j} r_{ni,g}\rho_{n,g} \left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} \right], \tag{E.48}$$

where each of the price derivative $\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}$ can be characterized using the following system:

$$\begin{bmatrix} \left(\frac{\partial \ln P_{11,k}}{\partial \ln Q_{1i,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} & \cdots & \left(\frac{\partial P_{11,k}}{\partial \ln Q_{Ni,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} \\ \vdots & \ddots & \vdots \\ \left(\frac{\partial \ln P_{NN,k}}{\partial \ln Q_{1i,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} & \cdots & \left(\frac{\partial P_{NN,k}}{\partial \ln Q_{Ni,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} \end{bmatrix} = -\mathbf{A}_i^{-1}\mathbf{C}_i, \tag{E.49}$$

where, from Appendix X, $\mathbf{A}_i$ and $\mathbf{C}_i$ are square matrixes whose elements are

$$\left(\mathbf{A}_i\right)_{\iota,\jmath} = \mathbb{1}_{\jmath=\iota} + \mathbb{1}_{\jmath\neq i}\frac{\mu_k}{1+\mu_k}\sum_{n\neq i}\left[r_{\iota n,k}\varepsilon_{\iota n,k}^{(\jmath n,k)}\right]; \qquad \left(\mathbf{C}_i\right)_{\iota,\jmath} = \mathbb{1}_{\jmath=\iota}\frac{\mu_k}{1+\mu_k} r_{\iota i,k}.$$

Our goal is to apply Wu et al.'s (2013) approach to derive a first-order approximation for $\mathbf{A}_i$, which is then used to compute $\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}$ and $\omega_{ji,k}$. For this, decompose $\mathbf{A}_i$ into its diagonal, $\mathbf{D}_i$, and off-diagonal, $\mathbf{E}_i$, such that

$$\mathbf{A}_i = \mathbf{D}_i + \mathbf{E}_i.$$

The elements of the diagonal matrix associated with $\mathbf{A}_i$ are given by

$$\left(\mathbf{D}_i\right)_{\iota,\iota} = 1 + \mathbb{1}_{\iota\neq i}\frac{\mu_k}{1+\mu_k}\sum_{n\neq i}\left[r_{\iota n,k}\varepsilon_{\iota n,k}\right].$$

[19] Beyond the Cobb-Douglas case, lowering the price of good ij,k can alter the revenue raised from other goods through cross-demand effects. But total revenue always increase, in response, given Cournot aggregation.

The elements of the off-diagonal matrix associated with $\mathbf{A}_i$ are, correspondingly,

$$(\mathbf{E}_i)_{\iota,j} = \begin{cases} 0 & \text{if } (\iota = i) \vee (\iota = j) \\ \frac{\mu_k}{1+\mu_k} \sum_{n \neq i} \left(r_{\iota n,k} \varepsilon_{\iota n,k}^{(jn,k)} \right) & \text{if } (\iota \neq i) \wedge (\iota \neq j) \end{cases}.$$

Following Wu et al. (2013), if the off-diagonal elements of $\mathbf{A}_i$ are small, we can appeal to the Neumann Series to approximate the inverse of $\mathbf{A}_1$as

$$\mathbf{A}_i^{-1} \approx \mathbf{D}_i^{-1} - \mathbf{D}_i^{-1}\mathbf{E}_i\mathbf{D}_i^{-1}.$$

Based on the above equation, each element of the inverse of $\mathbf{A}_i$ can be written (approximately) in closed-form in terms of the diagonal and off-diagonal elements of $\mathbf{A}_i$:

$$\mathbf{A}_i^{-1} \approx \begin{bmatrix} \frac{1}{(\mathbf{D}_i)_{11}}\left(1 - \frac{(\mathbf{E}_i)_{11}}{(\mathbf{D}_i)_{11}}\right) & \frac{-(\mathbf{E}_i)_{12}}{(\mathbf{D}_i)_{11}(\mathbf{D}_i)_{22}} & \cdots & \frac{-(\mathbf{E}_i)_{1N}}{(\mathbf{D}_i)_{11}(\mathbf{D}_i)_{NN}} \\ \frac{-(\mathbf{E}_i)_{21}}{(\mathbf{D}_i)_{22}(\mathbf{D}_i)_{11}} & \frac{1}{(\mathbf{D}_i)_{22}}\left(1 - \frac{(\mathbf{E}_i)_{22}}{(\mathbf{D}_i)_{22}}\right) & \cdots & \frac{-(\mathbf{E}_i)_{2N}}{(\mathbf{D}_i)_{22}(\mathbf{D}_i)_{NN}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{-(\mathbf{E}_i)_{N1}}{(\mathbf{D}_i)_{NN}(\mathbf{D}_i)_{11}} & \frac{-(\mathbf{E}_i)_{N2}}{(\mathbf{D}_i)_{NN}(\mathbf{D}_i)_{22}} & \cdots & \frac{1}{(\mathbf{D}_i)_{NN}}\left(1 - \frac{(\mathbf{E}_i)_{NN}}{(\mathbf{D}_i)_{NN}}\right) \end{bmatrix}.$$

Invoking the above approximation, Equation E.49 yields the following approximation for $\left(\frac{\partial \ln P_{\iota\iota,k}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}$:

$$\left(\frac{\partial \ln P_{\iota\iota,k}}{\partial \ln Q_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i} \approx \begin{cases} \frac{\frac{\mu_k}{1+\mu_k} r_{ji,k}}{1+\mathbb{1}_{\iota \neq i} \times \frac{\mu_k}{1+\mu_k} \sum_{n \neq i}\left(r_{\iota n,k}\varepsilon_{\iota n,k}\right)} & \text{if } \iota = j \\ \frac{\frac{\mu_k}{1+\mu_k} \sum_{n \neq i}\left[r_{\iota n,k}\varepsilon_{\iota n,k}^{(jn,k)}\right]\left(\frac{\mu_k}{1+\mu_k} r_{ji,k}\right)}{\left(1+\frac{\mu_k}{1+\mu_k} \sum_{n \neq i} r_{jn,k}\varepsilon_{jn,k}\right)\left(1+\mathbb{1}_{\iota \neq i} \times \frac{\mu_k}{1+\mu_k} \sum_{n \neq i} r_{\iota n,k}\varepsilon_{\iota n,k}\right)} & \text{if } \iota \neq j \end{cases}$$

Plugging the price derivatives specified above into Equation E.48, while noting that $r_{ni,k} \times r_{ji,k} \approx 0$ if $j \neq i$ and $n \neq i$, yields our approximation for the export supply elasticity:

$$\omega_{ji,k} \approx \frac{-\frac{\mu_k}{1+\mu_k} r_{ji,k}}{1 + \frac{\mu_k}{1+\mu_k} \sum_{\iota \neq i} r_{j\iota,k}\varepsilon_{j\iota,k}} \left[1 - \frac{\mu_k}{1+\mu_k} \frac{w_i L_i}{w_j L_j} \sum_{n \neq i} \frac{\rho_{i,k} r_{in,k}}{\rho_{j,k} r_{ji,k}} \varepsilon_{in,k}^{(jn,k)}\right].$$

E.4 Unilaterally Optimal Policy Net of ToT Considerations

Suppose we redo the entire proof under one restriction: Country i treats the entire vector of international prices as given. This includes (a) all consumer prices, $\tilde{\mathbf{P}}_{-i} = \left\{\tilde{P}_{nj,k}\right\}_{j \neq i}$, unassociated with domestic consumers and (b) all producer prices, $\mathbf{P}_{-i} \equiv \left\{P_{nj,k}\right\}_{n \neq i}$, unassociated with domestic firms. The idea here is that government presumes that cannot manipulate the consumer, $\tilde{P}_{ij,k}$, of export goods. Nor can it influence the producer price, $P_{ji,k}$, of import (or non-imported foreign) goods. Going back to the derivations presented above, we can solve this new terms-of-trade-blind problem by discarding the F.O.C.s relating to export prices (which amounts to setting $\tilde{P}_{ij,k} = P_{ij,k}$), and setting the inverse export supply elasticity to zero everywhere. Performing these alterations yields the following formula for optimal taxes:

$$\begin{aligned} &[\text{domestic subsidy}] && 1 + s_{i,k}^* = (1+\mu_k)/(1+\overline{\mu}_i) \\ &[\text{import tariff}] && 1 + t_{ji,k}^* = 1 + \bar{\tau}_i \\ &[\text{export subsidy}] && \mathbf{1} + \mathbf{x}_{ij}^* = 1 + \bar{\tau}_i. \end{aligned}$$

Normalizing the tax-shifters to zero (i.e., $\bar{\tau}_i = \overline{\mu}_i = 0$) yields the cooperative optimal tax structures consisting of zero trade taxes and Pigouvian industrial subsidies. So, if welfare-maximizing governments were (i) blind to terms-of-trade gains from policy and (ii) granted a complete vector of domestic policy instruments, they would adopt the cooperative policy choice. This point can be restated as follows: When welfare-maximizing governments have sufficient policy instruments at their disposal, their non-cooperative choice only inflicts a terms-of-trade externality on partners. So, the sole purpose of *shallow trade* agreements is to remedy the terms-of-trade externality. This result is a strict generalization of Bagwell and Staiger (2001, 2004). Notice a *shallow agreement* cannot resolve the problem of policy implementation, which is highlighted in Section III and quantified in Section VI.

F Efficient Policy from a Global Standpoint

The central planner seeks to maximize a weighted sum of national-level welfare using two sets of policy instruments. (1) A vector of good-specific taxes which grants them the ability to set consumers prices, $\tilde{\mathbf{P}}$, in every location. (2) A vector of inter-country lump-sum transfers, which enables them to control the share of each country's income from global income. Accordingly, the planner's choice of inter-country transfers is summarized by $\boldsymbol{\alpha} = \{\alpha_i\}_i$, where α_i denotes country i's share from global income ($Y = \sum_{i,n,k} \tilde{P}_{ni,k} Q_{ni,k}$) after transfers.

Recall that country i's welfare is summarized by the indirect utility function, $W_i \sim V_i\left(\tilde{Y}_i, \tilde{\mathbf{P}}_i\right)$, where country i's income under the planner's choice of policy is $\tilde{Y}_i = \alpha_i Y$. Considering this choice of notation, the global planner's problem can be formulated as

$$\max_{\tilde{\mathbf{P}},\boldsymbol{\alpha}} \sum_i \delta_i \ln \underbrace{V_i\left(\alpha_i Y, \tilde{\mathbf{P}}_i\right)}_{W_i},$$

where δ_i denotes the Pareto weight assigned to country i. Notice that Y_i (the equilibrium income raised by country i) does not explicitly appear in the objective function since the planner can obtain any desired vector of national-level incomes, $\{\tilde{Y}_i\}_i \sim \{\alpha_i Y\}_i$, with an appropriate choice of transfers subject to $\sum_i \tilde{Y}_i = Y$, where Y is the sum of equilibrium wage payments and tax revenues across all countries.[20] Namely,

$$Y(\mathbb{P}; \mathbf{w}, Y) = \sum_n w_n L_n + \sum_{i,n} \sum_k \left(\tilde{P}_{ni,k} - P_{ni,k}\right) Q_{ni,k}(\mathbb{P}; Y),$$

where $\mathbb{P} \equiv \left\{\tilde{\mathbf{P}}, \boldsymbol{\alpha}\right\}$ denotes the complete set of policy instruments available to the central planner and $Q_{ni,k}(\mathbb{P}; Y) = \mathcal{D}_{ni,k}\left(\tilde{\mathbf{P}}_i, \alpha_i Y\right)$ with $\tilde{\mathbf{P}}_i \subset \tilde{\mathbf{P}}$.[21] Following the logic presented in Section I, we can specify the planner's objective function, $W(\mathbb{P}; \mathbf{w}, Y) = \sum_i \delta_i \ln W_i(\mathbb{P}; \mathbf{w}, Y)$, as an explicit function of policy, $\mathbb{P}$, wages, $\mathbf{w}$, and global income, Y—noting that $\mathbf{w}$ and Y are feasible if they satisfy equilibrium conditions given $\mathbb{P}$. The first-order condition w.r.t. price instrument, $\tilde{P}_{ji,g} \in \tilde{\mathbf{P}} \subset \mathbb{P}$, can be written as

$$\frac{\partial W(\mathbb{P}; \mathbf{w}, Y)}{\partial \ln \tilde{P}_{ji,g}} = \sum_n \left[\delta_n \left(\frac{\partial \ln V_n(.)}{\partial \ln \tilde{Y}_n}\right) \frac{\partial \ln Y}{\partial \ln \tilde{P}_{ji,g}}\right] + \delta_i \frac{\partial \ln V_i(.)}{\partial \ln \tilde{P}_{ji,g}} = 0. \tag{F.1}$$

where the right-hand side uses the fact that $\tilde{Y}_n = \alpha_n Y$, implying that $\frac{\partial \ln \tilde{Y}_n}{\partial \ln \tilde{P}_{ji,g}} = \frac{\partial \ln Y}{\partial \ln \tilde{P}_{ji,g}}$ since α_n is a policy choice. Borrowing from our earlier derivation leading to the proof of Theorem 1, we can specify the change in global income in response to $\tilde{P}_{ji,k}$ as

$$\begin{aligned}\frac{\partial \ln Y(\mathbb{P}; \mathbf{w}, Y)}{\partial \ln \tilde{P}_{ji,g}} &= \frac{1}{Y}\left\{\frac{\partial Y(.)}{\partial \ln \tilde{P}_{ji,g}} + \sum_n \left[\frac{\partial Y(.)}{\partial \ln w_n} \frac{\mathrm{d} \ln w_n}{\mathrm{d}\ln \tilde{P}_{ji,g}}\right] + \frac{\partial Y(.)}{\partial \ln Y} \frac{\mathrm{d} \ln Y}{\mathrm{d}\ln \tilde{P}_{ji,g}}\right\} \\ &= \frac{1}{Y}\left\{\tilde{P}_{ji,g} Q_{ji,g} + \sum_n \sum_k \left(\left[\tilde{P}_{ni,k} - \frac{1}{1+\mu_k} P_{ni,k}\right] Q_{ni,k} \varepsilon^{ji,k}_{ni,k}\right)\right. \\ &\left. + \sum_n \left(\left[w_n L_n - \sum_{\iota,k} P_{n\iota,k} Q_{n\iota,k}\right] \frac{\mathrm{d}\ln w_n}{\mathrm{d}\ln \tilde{P}_{ji,g}}\right) + \sum_{n,\iota} \sum_k \left(\left[\tilde{P}_{ni,k} - \frac{1}{1+\mu_k} P_{ni,k}\right] Q_{ni,k} \eta_{ni,k}\right) \frac{\mathrm{d} \ln Y}{\mathrm{d}\ln \tilde{P}_{ji,g}}\right\},\end{aligned} \tag{F.2}$$

where notice that the first term in the second line is zero given the labor market clearing condition, $w_n L_n = \sum_{\iota,k} P_{n\iota,k} Q_{n\iota,k}$. Appealing to Roy's identity we can formulate the mechanical consumption

[20] To put it differently, the planner's problem (as we specify it) separates the issue of restoring production efficiency from inter-national redistribution. The former objective is attained with the proper choice of $\tilde{\mathbf{P}}$; the latter is attained with the proper choice of $\boldsymbol{\alpha}$—a point we discuss more later.

[21] The notation $Y = Y(\mathbb{P}; \mathbf{w}, Y)$ makes explicit the circular nature of income effects in general equilibrium.

loss from raising $\tilde{P}_{ji,k}$ as

$$\delta_i \frac{\partial \ln V_i(.)}{\partial \ln \tilde{P}_{ji,g}} = -\frac{\delta_i}{\tilde{Y}_i} \tilde{P}_{ji,g} Q_{ji,g} \frac{\partial \ln V_i(.)}{\partial \ln Y_i} |_{Y_i=\tilde{Y}_i}, \tag{F.3}$$

where note that $\tilde{Y}_i = \alpha_i Y$ by the definition of α_i, which is the country i's optimal share of global income given lump-sum transfers. Combining Equations F.1-F.3, and dividing by the final expression by $\sum_n \delta_n \frac{\partial \ln V_n}{\partial \ln Y_n} > 0$ yields the following F.O.C. with respect to price instrument $\tilde{P}_{ji,k}$:

$$\left(1 - \frac{1}{\alpha_i} \frac{\delta_i \frac{\partial \ln V_i}{\partial \ln Y_i}}{\sum_n \delta_n \frac{\partial \ln V_n}{\partial \ln Y_n}}\right) \tilde{P}_{ji,k} Q_{ji,k} + \sum_n \sum_k \left(\left[1 - \frac{1}{1+\mu_k} \frac{P_{ni,k}}{\tilde{P}_{ni,k}}\right] \tilde{P}_{ni,k} Q_{ni,k} \varepsilon_{ni,k}^{(ji,k)} \right)$$
$$+ \sum_{n,\iota} \sum_k \left(\left[1 - \frac{1}{1+\mu_k} \frac{P_{n\iota,k}}{\tilde{P}_{n\iota,k}}\right] \tilde{P}_{n\iota,k} Q_{n\iota,k} \eta_{n\iota,k} \right) \frac{\mathrm{d} \ln Y}{\mathrm{d} \ln \tilde{P}_{ji,g}} = 0. \tag{F.4}$$

Optimal Policy Implementation—Next, we specify the taxes that deliver the optimal consumer-to-producer price wedges. We also specify the optimal income shares, which implicitly determine the optimal international lump-sum transfers. The trivial solution to Equation F.4 involves price wedges equal to $\tilde{P}_{ni,k}/P_{ni,k} = 1/(1+\mu_k)$ and income shares that ensure the term in the first parenthesis to zero. The optimal price wedges, notice, are trade blind, indicating the optimal price wedges can be implemented with production subsidies alone. Appealing to this basic point, the production tax and transfers that satisfy the system of F.O.C.s consist of zero trade taxes, production subsidies that are proportional to the industry-level scale elasticities, and lump-sum transfers such that each country's share of global income reflects its Pareto weight and marginal utility from income. Stated formally,

$$1 + s_{i,k}^\star = \frac{P_{ni,k}^\star}{\tilde{P}_{ni,k}^\star} = 1 + \mu_k; \qquad x_{ij,k}^\star = t_{ji,k}^\star = 0; \qquad \alpha_i^\star = \frac{\delta_i \frac{\partial \ln V_i}{\partial \ln Y_i}}{\sum_n \delta_n \frac{\partial \ln V_n}{\partial \ln Y_n}}.$$

The Logic Behind Efficient Policy Formulas—The notion of optimal policy in our framework (as in much of the trade policy literature) is formulated to deliver the first-best outcome from the planner's standpoint. The planner, in particular, is afforded sufficient policy instruments to achieve both production efficiency and their desired level of redistribution. Production subsidies that restore marginal cost pricing are used to achieve production efficiency, while efficient lump-sum transfers are used to attain redistributive objectives based on Pareto weights. To elucidate this point, suppose lump-sum transfers where unavailable. Then, implementing the efficient tax schedule $\mathbb{T}^\star = (\mathbf{t}^\star, \mathbf{x}^\star, \mathbf{s}^\star)$ without transfers would deliver a Kaldor-Hicks (Kaldor (1939); Hicks (1939)) improvement but *not* necessarily a Pareto improvement. Still, the resulting equilibrium would be Hicks-optimal and, therefore, Pareto efficient. To ensure Pareto improvements (relative to Laissez-Faire) without lump-sum transfers, the optimal policy must also include non-zero trade taxes that redistribute the welfare gains from restoring marginal cost pricing across countries. But when efficient lump-sum transfers are available, the planner avoids redistribution via trade taxes as they undermine production efficiency.

Efficient Policy vs. Cooperative Tariffs—It is important to distinguish between efficient policies and cooperative tariffs of the sort examined by Ossa (2014) and Lashkaripour (2021). Efficient policies deliver the global planner's first-best outcome. Cooperative tariffs, on the other hand, maximize global welfare in second-best scenarios, where efficient production subsidies and transfers are unavailable. More formally, cooperative tariffs are given by

$$\mathbf{t}^{\star\star} = \arg\max \sum_i \delta_i \log W_i(\mathbf{t}) \qquad s.t. \begin{cases} \mathbf{s} = 0 \\ \alpha_{i,k} = Y_i(\mathbf{t}) / Y(\mathbf{t}) \end{cases}$$

So, $\mathbf{t}^{\star\star}$, by design, mimics the first-best (or efficient) subsidies and transfers to deliver the second-best.[22] In other words, cooperative tariffs seek to improve allocative efficiency by restricting trade (and thus global output) in low-μ industries. They also seek to redistribute inter-nationally, taking

[22] As we argue shortly, cooperative tariffs also internalize political economy pressures if any. Put together, these points

into account the Pareto weights in the planner's objective function or the bargaining weights in the Nash bargaining formulation of the same problem.[23]

Efficient Policy Under Political Economy Considerations—Our baseline model and the efficient policies implied by it abstract from political economy considerations. What if governments assign different political weights to say profits collected from different industries? In that case, efficient production subsidies should take into account not only the industry-level scale elasticity (or markup) but also its political weight. We can, moreover, refer to the discussion in Section IV to specify the politically-adjusted efficient policy. Recall, in particular, that the political economy model is isomorphic to an augmented version of our baseline model wherein markups are politically-adjusted and given by $\mu^{\mathscr{P}}_{i,k} = \frac{\mu_k}{\pi_{i,k}-(1-\pi_{i,k})\mu_k}$, where $\pi_{i,k}$ is the political economy weight assigned by the planner to profits collected from industry k in origin i. The efficient production subsidy, accordingly, becomes $s^{\star}_{i,k} = \mu^{\mathscr{P}}_{i,k}$.

G Internal Cooperation in the Rest of the World

When characterizing country i's unilaterally optimal policy, we treat the rest of the world as internally cooperative. Our notion of cooperation is based on the WTO's core principles: *reciprocity* and *non-discrimination* (see Bagwell and Staiger (2004)). The principle of reciprocity entails that cooperative countries maintain the balance of market access concessions internally. In our model, where labor is the sole factor of production, any change in relative market access is equivalent to a change in relative wages (see Footnote 25). Hence, to maintain the balance of concessions, cooperative countries must adopt policy buffers that neutralize relative wage disruptions among each other. Otherwise, the subset of countries whose relative wage improves in response to country i's policy reap terms-of-trade (or market access) gains at the expense of others whose relative wage deteriorates.

To formalize these arguments, we first specify the change in country n's welfare in response to country i's policy, $\{\mathrm{d}\ln(\mathbf{1}+\mathbf{x}_i), \mathrm{d}\ln(\mathbf{1}+\mathbf{t}_i), \mathrm{d}\ln(\mathbf{1}+\mathbf{s}_i)\}$. Suppose consumer preferences in country $n \neq i$ are homothetic. Appealing to Roy's identity, the welfare impacts of country i's policy shock on country n's welfare, can be expresses as $\mathrm{d}\ln W_n = \mathrm{d}\ln Y_n - \sum_j \lambda_{jn,k} e_{n,k} \mathrm{d}\ln \tilde{P}_{jn,k}$. Next, we characterize $\mathrm{d}\ln Y_n$ focusing on restricted entry for expositional purposes. Nominal income in country i is the sum of wage income adjusted for profit payments—namely,$Y_n = (1+\overline{\mu}_n) w_n L_n$, where $\overline{\mu}_n = \sum_k \mu_k \rho_{n,k}$ is the employment-weighted average markup in country n.[24] Taking full derivatives from the expression for Y_n, yields

$$\mathrm{d}\ln Y_n = \mathrm{d}\ln\left(\sum_k (1+\mu_k)\rho_{n,k}\right) + \mathrm{d}\ln w_n = \sum_k \left[\rho_{n,k}\cdot\left(\frac{1+\mu_k}{1+\overline{\mu}_n}\right)\mathrm{d}\ln\rho_{n,k}\right] + \mathrm{d}\mathrm{ln} w_n,$$

where $\mathrm{d}\ln\rho_{n,k}$ and $\mathrm{d}\ln w_n$ respectively denote the change in country n's employments shares and wage rate in response to country i's tax policy. To economize on the notation let $\mathbb{E}_\rho[.]$ and $Cov_\rho(.)$ denote cross-industry mean and covariance operators with weights, $\{\rho_{i,k}\}$. As a matter of accounting, $\sum_k \rho_{n,k}\mathrm{d}\ln\rho_{n,k} \sim \mathbb{E}_\rho[\mathrm{dln}\,\rho_{n,k}] = 0$, indicating that the first term in the last line of the above equation can be specified as

$$\sum_k \left[\rho_{n,k}\cdot\left(\frac{1+\mu_k}{1+\overline{\mu}_n}\right)\mathrm{d}\ln\rho_{n,k}\right] \sim Cov_\rho\left(\frac{1+\mu_k}{1+\overline{\mu}_n}, \mathrm{dln}\,\rho_{n,k}\right).$$

echo Ossa' (2016) verbal argument that second-best cooperative tariffs pursue three objectives: First, they seek to improve allocative efficiency by mimicking efficient production subsidies. Second, they seek to redistribute welfare inter-nationally based on Pareto or bargaining weights. Third, they seek to promote politically-organized industries. Though, they are not the first-best instrument for reaching either objective.

[23] Correspondingly, if the baseline economy is efficient, there exists a set of Pareto/bargaining weights (δ) for which $\mathbf{t}^{\star\star} = 0$ —see e.g., the analytic formula for $\mathbf{t}^{\star\star}$ in Lashkaripour (2021).

[24] Recall for Section I, that $\overline{\mu}_n \equiv \frac{\sum_{k,j}\frac{\mu_k}{1+\mu_k}P_{nj,k}Q_{nj,k}}{\sum_{k,j}\frac{1}{1+\mu_k}P_{nj,k}Q_{nj,k}}$. Noting that $w_n L_{n,k} = \frac{1}{1+\mu_k}P_{nj,k}Q_{nj,k}$, we can rewrite $\overline{\mu}_n$ as

$$\overline{\mu}_n = \frac{\sum_{k,j}\frac{\mu_k}{1+\mu_k}P_{nj,k}Q_{nj,k}}{\sum_{k,j}\frac{1}{1+\mu_k}P_{nj,k}Q_{nj,k}} = \sum_{k,j}\mu_k\frac{L_{n,k}}{L_n} \sim \sum_{k,j}\mu_k\rho_{n,k},$$

where $\rho_{n,k} \equiv L_{n,k}/L_n$ is the employment share assocaited with industry k in country n.

Next, we specify the welfare effects due to changes in consumer prices. The change in good-specific consumers prices for goods originating from $j \neq i$ is determined by the underlying wage change, i.e., $\mathrm{d}\ln \tilde{P}_{jn,k} = \mathrm{d}\ln w_j$ for all $j \neq i$. The change in consumers prices for goods originating from country i is the sum of the direct tax change and the indirect wage effects, i.e. $\mathrm{d}\ln \tilde{P}_{in,k} = \mathrm{d}\ln w_i + \mathrm{d}\ln(1 + x_{in,k})$. Putting the pieces together, we can write the change in country n's welfare in response to country i's non-cooperative policy as

$$\mathrm{d}\ln W_n = Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_n}, \mathrm{d}\ln\rho_{n,k}\right) + (1-\lambda_{ii})\,\mathrm{d}\ln w_n$$
$$- \sum_{j\neq i,n}\sum_k\left(\lambda_{jn,k}e_{n,k}\mathrm{d}\ln w_j\right) - \sum_k\left(\lambda_{in,k}e_{n,k}\left[\mathrm{d}\ln w_i + \mathrm{d}\ln(1+x_{in,k})\right]\right).$$

We can rearrange the above equation and decompose the various welfare terms as

$$\mathrm{d}\ln W_n = -\overbrace{\left(\lambda_{in}\mathrm{d}\ln(w_i/w_n) + \sum_k \lambda_{ni,k}e_{n,k}\mathrm{d}\ln(1+x_{in,k})\right)}^{\text{Terms-of-Trade vis-a-vis Country } i}$$
$$+ \underbrace{Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_n}, \mathrm{d}\ln\rho_{n,k}\right)}_{\text{Allocative Efficiency}} - \underbrace{\sum_{j\neq i}\left[\lambda_{jn}\mathrm{d}\ln(w_j/w_n)\right]}_{\text{Terms-of-Trade vis-a-vis RoW}}, \qquad \text{(G.1)}$$

where $\lambda_{jn} \equiv \sum_k \lambda_{ji,k}e_{n,k}$ denotes aggregate expenditure shares.[25] Following Baqaee and Farhi (2017), *Allocative Efficiency* effects are defined as the welfare change net of Hulten (1978) in response to policy shocks that do not raise revenues for closed-economy n. The remaining terms are, by design, terms of trade (ToT) effects. Theses can be divided into changes in ToT vis-à-vis country i and changes to ToT vis-à-vis the rest of the world, with which country n maintains cooperation.

Extraterritorial Terms of Trade Effects—Following Equation G.1, country i's non-cooperative policy can disrupt country n's ToT and balance of concessions vis-à-vis countries other than i. Consider country $j \neq i$ who is cooperative with country n. If $\mathrm{d}\ln(w_j/w_n) < 0$, country n's ToT improves relative to j in response to country i's policy. Or stated differently, the bilateral balance of market access concessions tilts in favor of country n, which violates *reciprocity*. We call these *"Extraterritorial Terms of Trade Effects,"* as they disrupt the ToT and balance of concessions between countries in the rest of the world. To restore *reciprocity*—one of WTO's core principles—the rest of the world must exert wage buffers that neutralize the *extraterritorial ToT effects* associated with country i's policy.

Neutralizing Extraterritorial ToT Effects with Cooperative Wage Buffers—The rest of the world can institute cooperative wage buffers to neutralize the extraterritorial ToT effects associated with country i's policy, ensuring that $\Delta\ln(w_j/w_n) = 0$ for all n, $j \neq i$. Ideally, these policies must satisfy WTO's *non-discrimination* principle and be efficient, which effectively rules out trade tax measures. A policy option that satisfies these requirements is a wage tax-cum-subsidy that is either revenue-neutral or financed via an efficient lump-sum tax on residents of all countries outside of i. To elaborate, let $\underline{w}_n^*$ denote the wage rate in country n after the implementation of country i's optimal policy, $\mathbb{P}_i$, *if*

[25] We can re-formulate the above decomposition in the spirit of Arkolakis, Costinot, and Rodriguez-Clare (2012), to clarify that market access is fully-determined by relative wages. In particular, appealing to the CES demand system whereby $\mathrm{d}\ln\left(\tilde{P}_{jn,k}/\tilde{P}_{nn,k}\right) = \frac{1}{1-\sigma_k}\mathrm{d}\ln\left(\lambda_{jn,k}/\lambda_{nn,k}\right)$, we can alternatively express the welfare effects of an external shock to economy n as

$$\mathrm{d}\ln W_n = \mathrm{d}\ln Y_n - \sum_k\left[e_{i,k}\mathrm{d}\ln\tilde{P}_{nn,k}\right] + \sum_{j,k}\left[\frac{e_{i,k}}{1-\sigma_k}\lambda_{jn,k}\mathrm{d}\ln\left(\frac{\lambda_{jn,k}}{\lambda_{nn,k}}\right)\right]$$
$$= Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_n}, \mathrm{d}\ln\rho_{n,k}\right) + \sum_k\left[\frac{e_{i,k}}{1-\sigma_k}\mathrm{d}\ln\lambda_{ii,k}\right],$$

where the last line follows from the adding up constraint, $\sum_j \lambda_{jn,k}\mathrm{d}\ln\lambda_{ji,k} = 0$. Notice that $\lambda_{ii,k}$, by definition, summarizes an open economy's market access. Comparing the above representation to Equation G.1 indicates that the change in market access can be alternatively summarized by changes to relative wages.

no policy buffers were in place. The country-specific wage subsidy, $\boldsymbol{\tau} \equiv \{\tau_n^w\}_{n \neq i}$, is allotted such that $\tau_n^w/\tau_j^w < 1$ if $\underline{w}_n^*/\underline{w}_j^* < w_n/w_j$—to the point that the post-subsidy effective relative wage rates $(w_n^*(\boldsymbol{\tau})/w_j^*(\boldsymbol{\tau}))$ remains equal to their status-quo level. Namely, $w_n^*/w_j^* = w_n/w_j$ for all n, $j \neq i$.

Treating the RoW as Internally Cooperative vs. Merely Passive—Treating the rest of the world as cooperative or passive is, after all, a theoretical formality since for all practical purposes, country i's good-specific taxes have little-to-no effect on aggregate relative wages in the rest of the world. We demonstrate this point numerically in Appendix H using multiple simulations of our model—including some where country i is large. But even in theory, one can envision many settings in which the rest of the world being cooperative or passive is immaterial. Let us provide one such example. Suppose there is a traded homogeneous sector, k_0, operating under constant-returns to scale technologies (i.e., $\sigma_{k0} \approx \gamma_{k0} \to \infty$). Moreover, assume that sector k_0 has a strictly positive employment share, i.e., $\rho_{n,k_0} > 0$, in every country n (with the possible exception of i). Assuming that $e_{n,k0}$ is sufficiently large such that $\hat{\rho}_{n,k_0} \neq 0$ in response to country i's policy, ensures that $w_n/w_{n'}$ remains constant for all n, $n' \neq i$—even if the rest of the world is passive.[26] Notice that this is a strictly weaker version of the common assumption adopted by Fajgelbaum, Grossman, and Helpman (2011) and Ossa (2011), among others. In particular, these studies assume that neither country i's nor the rest of the world's employment in the homogeneous sector reduces to zero in response to country i's policy (i.e., $\hat{\rho}_{i,k_0} \neq 0$ and $\hat{\rho}_{n,k_0} \neq 0$, $\forall n \neq i$). Our example, in contrast, only requires that the rest of the world's employment in the homogeneous sector does not collapse to zero.

H Numerical Examination of Optimal Policy Formulas

This appendix illustrates the accuracy and speed of our theoretical optimal policy formulas by benchmarking against results obtained from numerical optimization. We, more specifically, demonstrate two points. First, our formulas often outperform numerical optimization as they identify an optimal policy schedule that is strictly superior to that specified by standard numerical optimization routines. The improved accuracy is especially notable when analyzing a global economy with many countries. Second, our theoretical formulas are orders of magnitude faster than numerical optimization at detecting optimal policy.

We must underscore two points to set the stage for our numerical analysis. First, throughout this section, we use our approximate formula for the inverse export supply elasticity. Second, we treat the rest of the world as passive rather than internally cooperative. As such, our numerical analysis reveals two additional points. First, that our approximation of the export supply elasticity exhibits great numerical precision. Second, treating the rest of the world as internally cooperative vs. passive is virtually inconsequential. Since individual policy instruments have negligible impacts on relative wages in the rest of the world, our optimal policy formulas retain accuracy even if the rest of the world is passive—even if the tax-imposing country is relatively large.

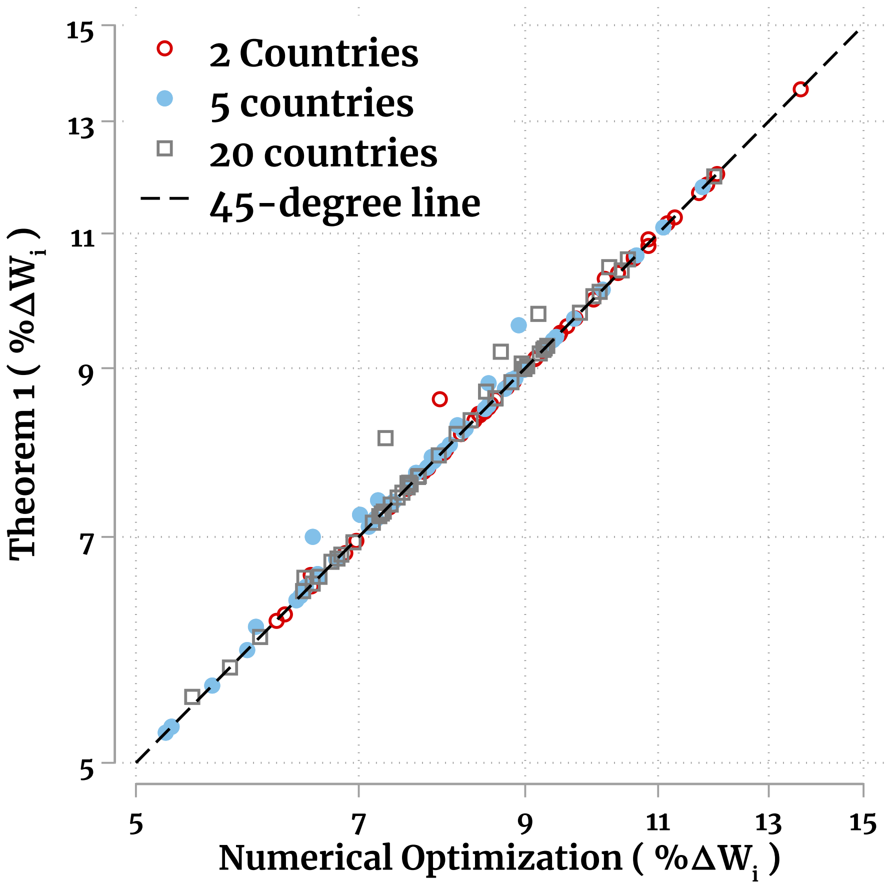
Details of Numerical Simulation—We examine three hypothetical economies with $N = 2$, 5, and 20 countries, each containing $S = 10$ industries. We assume that preferences across industries are Cobb-Douglas, with $e_{i,k}$ denoting the Cobb-Douglas weight on industry k in country i. To compute the optimal policy, we need to assign values to the following vector of parameters/endowments, $\Theta = \{\mu_k, \sigma_k, e_{i,k}, L_i\}_{i,k}$. The information relating to other parameters is implicit in the value assigned to the matrix of bilateral expenditure shares and national income levels, $\mathbf{X} = \left\{\lambda_{ij,k}, Y_i\right\}_{i,j,k}$. We normalize $Y_i = 100$ for all countries and randomly draw the remaining parameters/variables from a uniform distribution using the RAND function in MATLAB. We repeat this 50-time for each case, resulting in 150 simulations of the global economy under randomly-selected parameters. For each choice of parameters, we numerically solve for the optimal policy equilibrium using (a) our theoretical formulas relying on the optimization-free approach described under Proposition 1 in Section 7 and (b) using numerical optimization relying on the MPEC approach described in Ossa (2014). The latter is the standard approach when theoretical formulas are unavailable, so we benchmark our formulas' numerical accuracy and speed against it. Our implementation of MPEC, as in Ossa (2014), uses MATLAB's standard optimization routine, FMINCON.

[26] To elaborate, the price of the homogeneous good k_0 must be equalized across origins. Let a_{n,k_0} denote the constant unit labor requirement for producing good k_0 in origin n. Price equalization entails that $w_n/w_{n'} = a_{n',k_0}/a_{n,k_0}$, which is constant.

Accuracy of Theoretical Formulas

Figure H.1 compares the welfare gains implied by our optimal policy formulas to those obtained from numerical optimization (MPEC). Each dot corresponds to one of our 150 simulations. A dot lying on the 45-degree line in Figure H.1 indicates that our theoretical formulas identify the same optimal policy schedule numerical optimization. Dots lying above the 45-degree line correspond to simulations where our theoretical formulas (Theorem 1) outperform numerical optimization (MPEC)—that is, they identify an optimal policy schedule that strictly dominates in terms of implied welfare gains, which is the policy objective. The reverse is true for dots below the 45-degree line. Keep in mind that the simulations in Figure H.1 use our approximation of the export supply elasticity and treat the rest of the world as passive—each of which can possibly compromise the performance of our theoretical formulas relative to numerical optimization.

Figure H.1: *Gains from optimal policy: theoretical formula vs. numerical optimization*

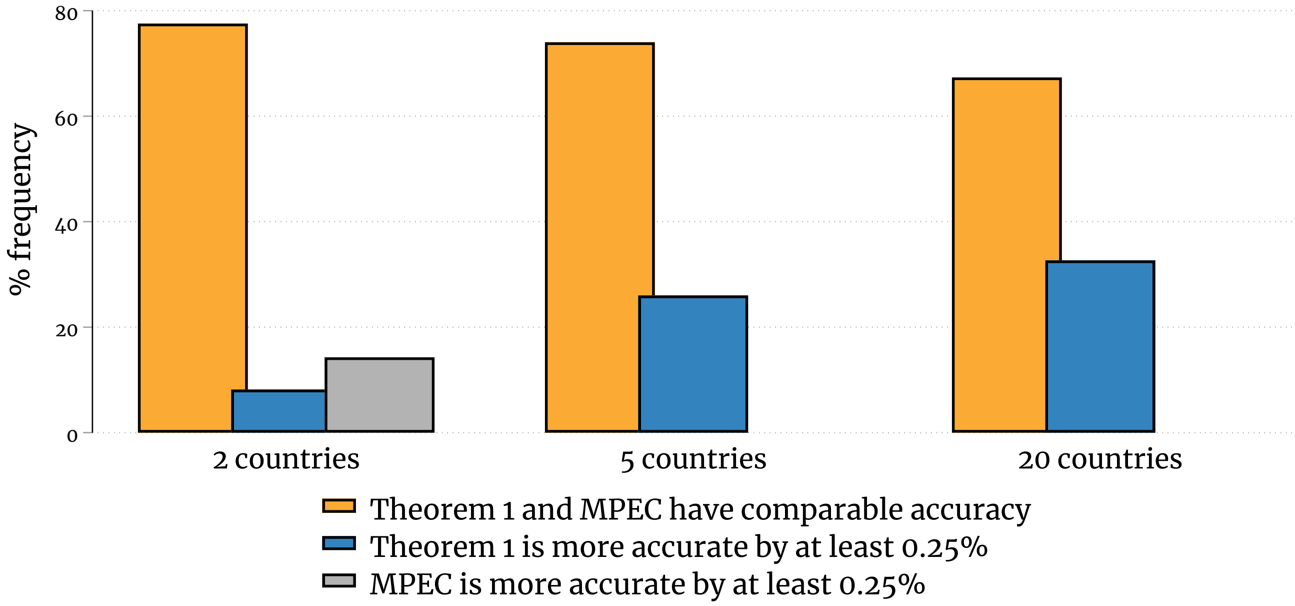


Note: This figure reports results from 150 simulations in which parameters are randomly sampled for three cases of our model—namely, $N = 2, 5, 10$, with $K = 10$. The y-axis reports the pre-cent welfare gains predicted by our optimal policy formulas (Theorem 1) in a simulated model. The x-axis reports reports the welfare gains obtained from numerical optimization conducted using MATLAB's FMINCON routine.

Figure H.1 reveals that the prediction of our theory is virtually identical to numerical optimization in most cases. On several occasions, our theoretical formulas outperform numerical optimization by a non-trivial margin. These are more frequent when we simulate a global economy consisting of more countries. We summarize this point more clearly in Figure H.2. We divide simulation outcomes into three categories:

i. [*Orange*] Simulations where our theoretical formulas and numerical optimization (MPEC) predicted comparable gains from optimal policy, i.e., $\frac{\left|\Delta W_i^{\text{theory}} - \Delta W_i^{\text{MPEC}}\right|}{\min\left\{\Delta W_i^{\text{theory}}, \Delta W_i^{\text{MPEC}}\right\}} \leq 0.0025$.

__Figure H.2:__ Performance of theoretical formulas vs. numerical optimization (MPEC)



Note: This figure summarizes results from 150 simulations with randomly-sampled parameters. The yellow bars represent the frequency of simulations where our optimal policy formulas and MPEC (numerical optimization) predict welfare gains that are within 0.25% of one another in terms of magnitude. The blue bars represent the frequency of simulations where our optimal policy formulas predict welfare gains that at least 0.25% greater than those implied by numerical optimization (MPEC). The grey bars represents the frequency of simulations where our optimal policy formulas predict welfare gains that at least 0.25% greater than those implied by numerical optimization (MPEC).

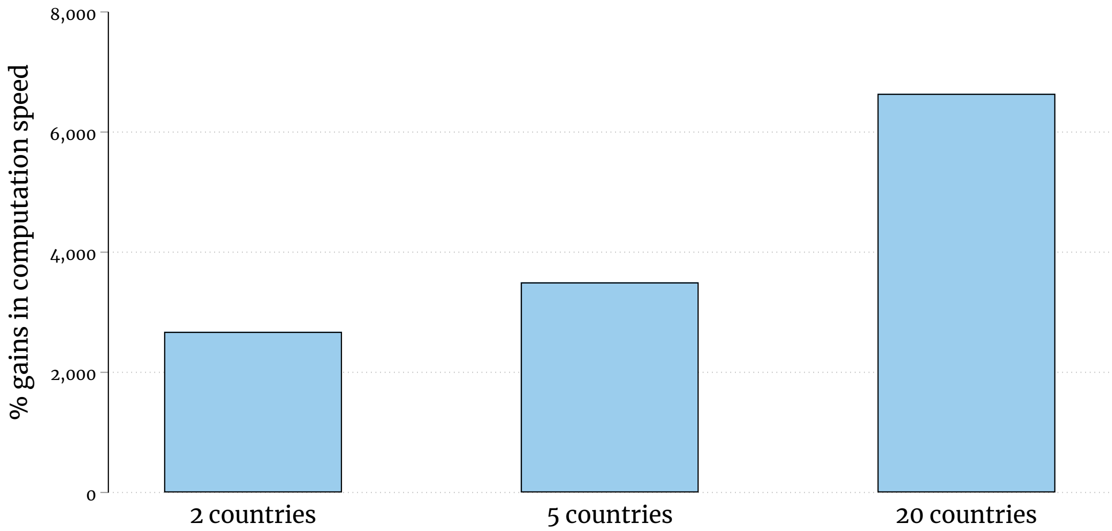
ii. [*Blue*] Simulations where our theoretical formulas outperform numerical optimization (MPEC) by at least 0.25%, i.e., $\Delta W_i^{\text{theory}} > 1.0025 \times \Delta W_i^{\text{MPEC}}$.

iii. [*Grey*] Simulations where numerical optimization (MPEC) outperforms our theoretical formulas by at least 0.25%, i.e., $\Delta W_i^{\text{MPEC}} > 1.0025 \times \Delta W_i^{\text{theory}}$.

A clear takeaway is that –for all practical purposes– our theoretical formulas either deliver comparable accuracy or outperform numerical optimizations. Numerical optimization exhibits great accuracy when dealing with only two countries. With 20 countries, however, our theoretical formulas outperform numerical optimization by at least 0.25% in more than thirty percent of the simulations. This improvement is noteworthy in practice, as we are often interested in cases where country i implements trade policy in relation to tens if not hundreds of trading partners. In these cases, numerical optimization must identify an optimal vector of policies consisting of hundreds and thousands of free-moving policy instruments—which can compromise accuracy depending on the properties of the underlying objective function.

Why does numerical optimization become less accurate with many countries?—Figure H.1 reveals that, when dealing with many countries, our theoretical formulas occasionally outperform numerical optimization by a non-trivial margin. The reason is that with many countries and free-moving tax instruments, numerical optimization may detect a near-prohibitive good-specific tax rate that is non-optimal but artificially satisfies the first-order conditions to a good approximation. Even though numerical optimization identifies the appropriate policy vis-à-vis most goods in these cases, it fails with respect to one or more goods for which it converges to a high and non-optimal tax rate. One can perhaps navigate this pitfall by setting bounds on feasible tax choices, but it is unclear what these bounds should be without theory. Relatedly, Figure H.2 suggests that our *approximated* optimal policy formulas occasionally underperform numerical optimization when dealing with two countries. This is a mere reflection of our export-supply-elasticity-approximation error, which can be non-trivial when country i is excessively large relative to the rest of the world. To elaborate, our simulation assigns the

Figure H.3: *Improvement in computation speed from using theoretical formulas*



Note: The figure compares the per-cent increase in computation speed when using our optimal policy formulas over numerical optimization. Each bar represents the average increase over 50 simulations with randomly-sampled parameters.

same size to all countries. Correspondingly, country i is similar in size to the entire rest of the world in the simulation with two countries—thereby, the possibly large approximation error.[27]

The RoW being Internally Cooperative or Passive is Immaterial—When performing numerical optimization, we purposely treat the rest of the world as passive—i.e., we do not restrict relative wages to remain constant in the rest of the world. Yet our optimal policy formulas (which treat the rest of the world as internally cooperative) deliver predictions that are virtually identical to those obtained from numerical optimization (which treats the rest of the world as passive)—even though the tax-imposing country is large relative to the rest of the world in our simulations. These outcomes all but corroborate our previous assertion that the rest of the world being passive vs. cooperative is a theoretical formality and virtually immaterial from a quantitative standpoint.

Computational Speed of Theoretical Formulas

Figure H.3 reveals that our theoretical formulas deliver orders of magnitude improvements in computation speed relative to numerical optimization (MPEC). Let t^{theory} denote the time it takes to compute the optimal policy equilibrium with the aid of our theoretical formulas. Correspondingly, let t^{MPEC} denote the computational time required to run numerical optimization. The y-axis in Figure H.3 corresponds to $100 \times \left(t^{\text{MPEC}}/t^{\text{thoery}}\right)$, which is the per-cent improvement in computation speed when using our theoretical formulas over numerical optimization. Our theory delivers a more than 20-fold improvement in speed with two countries, and a more than 60-fold improvement with 20 countries. The gains will be, accordingly, greater in real-world scenarios involving many countries (like those examined in Section 7). The improvement in computation speed is especially crucial when determining the Nash equilibrium of a non-cooperative policy game, wherein each country's optimal policy must be solved iteratively as a function of others' policies. We perform such an analysis in Section 7, where it takes us a few minutes to identify the Nash equilibrium versus many hours if we had relied on numerical optimization.

I Proof of Theorem 2

The proof of Theorem 2 has the same basic foundation as Theorem 1. We reformulate the optimal policy problem, expressing equilibrium variables (e.g., $Q_{ji,k}$, Y_i, etc.) as a function of (1) the vector of

[27] To be clear, in the two-country case, the rest of the world being passive (rather than internally cooperative) is irrelevant and our approximation of the export supply elasticity is the only source of numerical error.

consumer prices associated with economy i, excluding $\tilde{\mathbf{P}}_{ii}$, i.e., $\tilde{\mathbb{P}}_i \equiv \{\tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij}\}$;[28] and (2) the vector of national-level wage rates all over the world, $\mathbf{w} = \{w_1, ..., w_N\}$. To implement this reformulation of equilibrium variables, we need to solve the following system treating $\tilde{\mathbb{P}}_i$, and $\mathbf{w}$ as given:

$$\begin{aligned}
&[\text{optimal pricing}] && P_{jn,k} = \bar{\rho}_{ji,k} w_j \\
&[\text{optimal consumption}] && Q_{jn,k} = \mathcal{D}_{jn,k}(Y_n, \tilde{\mathbf{P}}_{1n}, ...\tilde{\mathbf{P}}_{Nn}) \\
&[\text{RoW imposes zero taxes}] && \tilde{P}_{jn,k} = P_{jn,k} \quad (\tilde{P}_{jn,k} \notin \tilde{\mathbb{P}}_i); \qquad Y_n = \overbrace{(1+\overline{\mu}_n) w_n L_n}^{w_n L_n + \Pi_n} \quad (n \neq i) \\
&[\text{Balanced Budget in } i] && Y_i = (1+\overline{\mu}_i) w_i L_i + \left(\tilde{\mathbf{P}}_{ij} - \mathbf{P}_{ij}\right) \cdot \mathbf{Q}_{ij} + \left(\tilde{\mathbf{P}}_{ji} - \mathbf{P}_{ji}\right) \cdot \mathbf{Q}_{ji} \\
&[\text{avg. profit margin in } j] && 1 + \overline{\mu}_j = \frac{\sum_{n \in \mathbb{C}} \left[\mathbf{P}_{jn} \cdot \mathbf{Q}_{jn}\right]}{\sum_{n \in \mathbb{C}} \left[\mathbf{P}_{jn} \cdot \left(\mathbf{Q}_{jn} \oslash (\mathbf{1} + \boldsymbol{\mu})\right)\right]}
\end{aligned}$$

where "$\cdot$" denotes the inner product operator for vectors of equal size. "$\oslash$" denotes element-wise division of equal-sized vectors, with $\boldsymbol{\mu} \equiv \{\mu_k\}_k$. Since there is a unique equilibrium, the above system is exactly identified in that it uniquely determines $P_{jn,k}(\tilde{\mathbb{P}}_i; \mathbf{w})$, $Q_{jn,k}(\tilde{\mathbb{P}}_i; \mathbf{w})$, $Y_n(\tilde{\mathbb{P}}_i; \mathbf{w})$, and $\overline{\mu}_i(\tilde{\mathbb{P}}_i; \mathbf{w})$ as a function of $\tilde{\mathbb{P}}_i$ and $\mathbf{w}$. Appealing to the above reformulation of the equilibrium, we can reformulate the original optimal policy problem (P2) as follows.

Lemma 7. *Country i's vector of second-best trade taxes, $\{\boldsymbol{t}_i^{**}, \boldsymbol{x}_i^{**}\}$, can be determined by solving the following problem:*

$$\max_{\tilde{\mathbb{P}}_i} \; W_i(\tilde{\mathbb{P}}_i; \mathbf{w}) \equiv V_i(Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}), \tilde{\mathbf{P}}_i) \quad s.t. \; (\tilde{\mathbb{P}}_i; \mathbf{w}) \in \mathbb{F}_P \qquad (\widetilde{\text{P2}}),$$

where the feasibility constraint is satisfied if, given $\tilde{\mathbb{P}}_i$, the wage vector $\boldsymbol{w}$ satisfies balanced trade in each country:

$$(\tilde{\mathbb{P}}_i; \boldsymbol{w}) \in \mathbb{F}_P \iff \begin{cases} \sum_{j \neq n} \sum_{k \in \mathbb{K}} \left[P_{jn,k}(\tilde{\mathbb{P}}_i; \boldsymbol{w}) Q_{jn,k}(\tilde{\mathbb{P}}_i; \boldsymbol{w}) - P_{nj,k}(\tilde{\mathbb{P}}_i; \boldsymbol{w}) Q_{nj,k}(\tilde{\mathbb{P}}_i; \boldsymbol{w})\right] = 0 & \text{if } n \neq i \\ \sum_{j \neq n} \sum_{k=1}^{K} \left[P_{ji,k}(\tilde{\mathbb{P}}_i; \boldsymbol{w}) Q_{jn,k}(\tilde{\mathbb{P}}_i; \boldsymbol{w}) - \tilde{P}_{ij,k} Q_{nj,k}(\tilde{\mathbb{P}}_i; \boldsymbol{w})\right] = 0 & \text{if } n = i \end{cases}.$$

The system of F.O.C.'s underlying Problem $(\widetilde{\text{P2}})$ can be expressed as follows.

$$\nabla_{\tilde{P}} W_i(\tilde{\mathbb{P}}_i; \mathbf{w}) + \nabla_{\mathbf{w}} W_i \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\tilde{P}}\right)_{(\tilde{\mathbb{P}}_i; \mathbf{w}) \in \mathbb{F}_P} = \mathbf{0}, \qquad \forall \tilde{P} \in \tilde{\mathbb{P}}_i = \left\{\tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij}\right\}.$$

In what follows we characterize and simplify the system of F.O.C., building heavily on the results presented in Appendix E.

Deriving the First-Order Condition w.r.t. $\tilde{\mathbf{P}}_{ji}$

Consider the consumer price index $\tilde{P}_{ji,k} \in \tilde{\mathbb{P}}_i$ associated with a good imported by i from *origin j–industry k*. The F.O.C. w.r.t. this price instrument can be stated as follows:

$$\left(\frac{\mathrm{d}W_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\mathrm{d}\ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ji,k}} + \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial Y_i} \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}, \tilde{\mathbb{P}}_{-ji,k}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = 0, \tag{I.1}$$

where $\tilde{\mathbb{P}}_{-ji,k} \equiv \tilde{\mathbb{P}}_i - \left\{\tilde{P}_{ji,k}\right\}$ denotes the vector of price instruments excluding $\tilde{P}_{ji,k}$. The above equation is similar to what we characterized in Appendix E under restricted entry, with two distinctions: First, country i's government does not control the price of domestically produced and domestically consumed varieties, i.e., $\tilde{\mathbf{P}}_{ii} \notin \tilde{\mathbb{P}}_i$. Second, country i's income does not include domestic tax revenues:

$$Y_i = (1+\overline{\mu}_i) w_i L_i + \left(\tilde{\mathbf{P}}_{ij} - \mathbf{P}_{ij}\right) \cdot \mathbf{Q}_{ij} + \left(\tilde{\mathbf{P}}_{ji} - \mathbf{P}_{ji}\right) \cdot \mathbf{Q}_{ji}.$$

[28] Recall that vectors $\tilde{\mathbf{P}}_{ji} \equiv \left\{\tilde{P}_{ji,k}\right\}_{j \neq i,k}$ and $\tilde{\mathbf{P}}_{ij} \equiv \left\{\tilde{P}_{ij,k}\right\}_{j \neq i,k}$ encompass only the export/import prices linked to economy i.

Taking note of these two differences, we can build on the derivation in Appendix E to simplify Equation I.1. By Roy's identity, the first term on the right-hand side of Equation I.1 can be stated as

$$\frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ji,k}} = -\tilde{P}_{ji,k} Q_{ji,k} \left(\frac{\partial V_i}{\partial Y_i}\right).$$

Without repeating the derivations, the second term on the right-hand side of Equation I.1 reduces to

$$\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = \tilde{P}_{ji,k} Q_{ji,k} + \sum_g \sum_{n \neq i} \left[\left(\tilde{P}_{ni,g} - (1+\omega_{ni,g}) P_{ni,g}\right) Q_{ni,g} \varepsilon_{ni,g}^{(ji,k)}\right]$$
$$\sum_g \left[\left(1 - \frac{1+\bar{\mu}_i}{1+\mu_g}\right) P_{ii,g} Q_{ii,g} \varepsilon_{ii,g}^{(ji,k)}\right] + \Delta_i'(\tilde{\mathbb{P}}_i, \mathbf{w}) \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$$

where $\Delta_i'(\tilde{\mathbb{P}}_i;\mathbf{w})$ is a uniform term (without industry subscripts) and is given by

$$\Delta_i'(\tilde{\mathbb{P}}_i;\mathbf{w}) \equiv \sum_g \sum_{n \neq i} \left[\left(\tilde{P}_{ni,g} - (1+\omega_{ni,g}) P_{ni,g}\right) Q_{ni,g} \eta_{ni,g}\right] + \sum_g \left[\left(1 - \frac{1+\bar{\mu}_i}{1+\mu_g}\right) P_{ii,g} Q_{ii,g} \eta_{ii,g}\right]. \tag{I.2}$$

To be clear, the above expressions can be derived by repeating the steps in Appendix E, while dropping domestic tax revenues from the expression for income Y_i. Likewise, the third term on the right-hand side of Equation I.1 can be stated as

$$\left(\frac{\partial W_i(.)}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d} \ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = -\sum_g \sum_{n \neq i} \left[(1+\omega_{ni,g}) \bar{\tau}_i P_{ni,g} Q_{ni,g} \varepsilon_{ni,g}^{(ji,k)}\right] - \sum_g \sum_{n \neq i} \left[(1+\omega_{ni,g}) \bar{\tau}_i P_{ni,g} Q_{ni,g} \eta_{ni,g}\right] \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$$

where $\bar{\tau}_i$ is given by J.1. Combining the above equations the F.O.C. specified by Equation I.1 can be simplified as

$$\sum_{n \neq i} \sum_g \left[\left(\frac{\tilde{P}_{ni,g}}{P_{ni,g}} - (1+\bar{\tau}_i)(1+\omega_{ni,g})\right) P_{ni,g} Q_{ni,g} \varepsilon_{ni,g}^{(ji,k)}\right]$$
$$+ \sum_g \left[\left(1 - \frac{1+\bar{\mu}_i}{1+\mu_g}\right) P_{ii,g} Q_{ii,g} \varepsilon_{ii,g}^{(ji,k)}\right] + \widetilde{\Delta}_i'(\tilde{\mathbb{P}}_i, \mathbf{w}) \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}} = 0, \tag{I.3}$$

where $\widetilde{\Delta}_i'(\tilde{\mathbb{P}}_i;\mathbf{w})$ is specified analogously to $\Delta_i'(\tilde{\mathbb{P}}_i, \mathbf{w})$, but features an adjustment for general equilibrium wage effects:

$$\widetilde{\Delta}_i'(\tilde{\mathbb{P}}_i;\mathbf{w}) \equiv \sum_g \sum_{n \neq i} \left[\left(\tilde{P}_{ni,g} - (1+\bar{\tau}_i)(1+\omega_{ni,g}) P_{ni,g}\right) Q_{ni,g} \eta_{ni,g}\right] + \sum_g \left[\left(1 - \frac{1+\bar{\mu}_i}{1+\mu_g}\right) P_{ii,g} Q_{ii,g} \eta_{ii,g}\right]. \tag{I.4}$$

Deriving the First-Order Condition w.r.t. $\tilde{\mathbf{P}}_{ij}$

Now, consider the consumer price index $\tilde{P}_{ij,k} \in \tilde{\mathbb{P}}_i$ associated with a good exported by i from *destination j–industry k.* The F.O.C. w.r.t. this price instrument can be stated as follows:

$$\left(\frac{\mathrm{d}W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\mathrm{d} \ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ij,k}} + \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial Y_i} \left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d} \ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = 0. \tag{I.5}$$

where $\tilde{\mathbb{P}}_{-ij,k} \equiv \tilde{\mathbb{P}}_i - \left\{\tilde{P}_{ij,k}\right\}$ denotes the vector of price instruments excluding $\tilde{P}_{ij,k}$. As with the previous subsection, The above equation is similar to what we characterized in Appendix E, with two distinctions: First, country i's government does not control the price of domestically produced and domestically consumed varieties, i.e., $\tilde{\mathbf{P}}_{ii} \notin \tilde{\mathbb{P}}_i$. Second, country i's income does not include domestic tax revenues. Noting these two distinctions, we can borrow from the derivation in Appendix E to simplify Equation I.5.

Namely, since $\tilde{P}_{ij,k} \notin \tilde{\mathbf{P}}_i$ is not part of the domestic consumer price index in i, $\partial V_i(Y_i, \tilde{\mathbf{P}}_i)/\partial \ln \tilde{P}_{ji,k} =$

0. So, the first term on the right-hand side of Equation I.5 collapses to zero. Without repeating the derivations from Appendix E, the second term on the right-hand side of Equation I.5 reduces to

$$\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = \tilde{P}_{ij,k}Q_{ij,k} + \sum_g \left[\left(\tilde{P}_{ij,g} - \frac{1+\overline{\mu}_i}{1+\mu_g}P_{ij,g}\right)Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right] - \sum_g\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right] + \Delta_i'(\tilde{\mathbb{P}}_i;\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}$$

where $\Delta_i'(\tilde{\mathbb{P}}_i;\mathbf{w})$ is a uniform term without industry subscripts, as defined by Equation I.2. To elaborate, the above expression can be derived by repeating the steps in Appendix E, while dropping domestic tax revenues from the expression for income Y_i. Likewise, the third term on the right-hand side of Equation I.5 can be stated as

$$\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = \bar{\tau}_i\tilde{P}_{ij,k}Q_{ij,k} + \sum_g\left[\bar{\tau}_i\tilde{P}_{ij,g}Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$-\sum_g\sum_{n\neq i}\left[\bar{\tau}_i\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right] - \sum_g\sum_{n\neq i}\left[[1+\omega_{ni,g}]\bar{\tau}_iP_{ni,g}Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}},$$

where $\bar{\tau}_i$ is given by J.1. Combining the above equations the F.O.C. specified by Equation I.5 can be simplified as

$$\tilde{P}_{ij,k}Q_{ij,k} + \sum_{g\in\mathbb{K}}\left[\left(1 - \frac{1+\overline{\mu}_i}{(1+\bar{\tau}_i)(1+\mu_g)}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right)\tilde{P}_{ij,g}Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$\sum_{g\in\mathbb{K}}\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right] + \widetilde{\Delta}_i'(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} = 0, \quad \text{(I.6)}$$

where $\widetilde{\Delta}_i'(\tilde{\mathbb{P}}_i;\mathbf{w})$ is given by Equation I.4.

Solving the System of First-Order Conditions

First, note that we can solve the system specified by Equation I.3 independent of I.6. To solve the system of Equations I.3, we can rely on the intermediate observation that if

$$\left(\mathbf{1} - \frac{1+\overline{\mu}_i}{\mathbf{1}+\boldsymbol{\mu}}\right)\odot\mathbf{P}_{ii}\odot\mathbf{Q}_{ii}\cdot\boldsymbol{\varepsilon}_{ii}^{(ji,k)} + \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni} - (1+\bar{\tau}_i)(\mathbf{1}+\boldsymbol{\Omega}_{ni})\odot\mathbf{P}_{ni}\right)\odot\mathbf{Q}_{ni}\cdot\boldsymbol{\varepsilon}_{ni}^{(ji,k)}\right] = 0, \quad \text{(I.7)}$$

then, to a first-order approximation around $\mu_k \approx \overline{\mu}_i$, $\widetilde{\Delta}_i'(\boldsymbol{\mu}) \approx 0$. So, the optimal choice of $\tilde{\mathbf{P}}_{ji}^{**}$ (and the implied tariff vector) can be determined by solving Equation I.7 instead of I.3.[29] Before moving forward, though, let us clarify the vector notation used to express Equation I.7. The vector operators "$\cdot$" and "$\odot$" are respectively the inner product and element-wise product operators. The $K\times 1$ vector $\frac{1+\overline{\mu}_i}{1+\boldsymbol{\mu}} = \left[\frac{1+\overline{\mu}_i}{1+\mu_k}\right]_k$ is composed of industry-level The $K\times 1$ vectors $\tilde{\mathbf{P}}_{ni} = \left\{\tilde{P}_{ni,k}\right\}_k$ and $\mathbf{Q}_{ni} = \left\{Q_{ni,k}\right\}_k$ encompass the consumer price and quantity associated with all of country i's import goods for origin $n \neq i$. Analogously, $\boldsymbol{\varepsilon}_{ni}^{(ji,k)} = \left\{\varepsilon_{ni,g}^{(ji,k)}\right\}_g$ encompasses the elasticity of demand for each the goods imported from n w.r.t. the price of ji,k.

We simplify Equation I.7 in three steps: First, by noting that $\tilde{\mathbf{P}}_{ii} = \mathbf{P}_{ii}$ and appealing to Cournot's aggregation, $\sum_{J\in\mathbb{C}}\left[\tilde{\mathbf{P}}_{ji}\odot\mathbf{Q}_{ji}\cdot\boldsymbol{\varepsilon}_{ji}^{(ji,k)}\right] = -\tilde{P}_{ji,k}Q_{ji,k}$, we can rewrite Equation I.7 as

$$\frac{1+\overline{\mu}_i}{1+\boldsymbol{\mu}}\odot\tilde{\mathbf{P}}_{ii}\odot\mathbf{Q}_{ii}\cdot\boldsymbol{\varepsilon}_{ii}^{(ji,k)} + (1+\bar{\tau}_i)\sum_{n\neq i}\left[(\mathbf{1}+\boldsymbol{\Omega}_{ni})\odot\mathbf{P}_{ni}\odot\mathbf{Q}_{ni}\cdot\boldsymbol{\varepsilon}_{ni}^{(ji,k)}\right] + \tilde{P}_{ji,k}Q_{ji,k} = 0. \quad \text{(I.8)}$$

Second, we invoke the *Slutsky Equation*,[30] to rewrite the first two term in the above equation. Specifi-

[29] Note that Equation I.7 is essentially I.3 with $\widetilde{\Delta}_i'(.)$ set to zero.

[30] Recalling that $e_{ji,k} = \tilde{P}_{ji,k}Q_{ji,k}/Y_i$ denotes the share of expenditure on ji,k, the Slutsky equation can be formally stated

cally, taking note that

$$\eta_{ii,g} = \eta_{ji,k} = 1 \quad \xrightarrow{\text{Slutsky Equation}} \quad \tilde{P}_{ni,g} Q_{ni,g} \varepsilon_{ni,g}^{(ji,k)} = \tilde{P}_{ji,k} Q_{ji,k} \varepsilon_{ji,k}^{(ni,g)}.$$

We can reduces the F.O.C. described under Equation I.8 to

$$1 + \sum_g \left[\frac{1+\mu_g}{1+\bar{\mu}_i} \varepsilon_{ji,k}^{(ii,g)} \right] + (1+\bar{\tau}_i) \sum_g \sum_{n \neq i} \left[(1+\omega_{ni,g}) \frac{P_{ni,g}}{\tilde{P}_{ni,g}} \varepsilon_{ji,k}^{(ni,g)} \right] = 0. \tag{I.9}$$

Lastly, we use the Marshallian demand function's *homogeneity of degree zero* property, whereby $\eta_{ji,k} + \sum_{J,g} \varepsilon_{ji,k}^{(Ji,g)} = 1 + \sum_{J,g} \varepsilon_{ji,k}^{(Ji,g)} = 0$. Invoking this property we rewrite Equation I.9 as follows

$$\sum_g \left[\left(1 - \frac{1+\mu_g}{1+\bar{\mu}_i} \right) \varepsilon_{ji,k}^{(ii,g)} \right] + \sum_g \sum_{n \neq i} \left[\left(1 - (1+\bar{\tau}_i)(1+\omega_{ni,g}) \frac{P_{ni,g}}{\tilde{P}_{ni,g}} \right) \varepsilon_{ji,k}^{(ni,g)} \right] = 0.$$

The above equation, which should hold for all $ji,k \neq ii,k$ specifies a system of FOCS that can be expressed in matrix no notation as

$$\underbrace{\begin{bmatrix} \varepsilon_{1i,1}^{(ii,1)} & \cdots & \varepsilon_{Ni,K}^{(ii,1)} \\ \vdots & \ddots & \vdots \\ \varepsilon_{1i,1}^{(ii,K)} & \cdots & \varepsilon_{Ni,K}^{(ii,K)} \end{bmatrix}}_{\mathbf{E}_{-ii}^{(ii)}} \begin{bmatrix} 1 - \frac{\mu_1}{\bar{\mu}_i} \\ \vdots \\ 1 - \frac{\mu_K}{\bar{\mu}_i} \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon_{1i,1}^{(1i,1)} & \cdots & \varepsilon_{i-1i,k}^{(1i,1)} & \varepsilon_{i+1i,k}^{(1i,1)} & \cdots & \varepsilon_{Ni,K}^{(1i,1)} \\ \vdots & \ddots & & & \ddots & \vdots \\ \varepsilon_{1i,1}^{(Ni,K)} & \cdots & \varepsilon_{i-1i,k}^{(Ni,K)} & \varepsilon_{i+1i,k}^{(Ni,K)} & \cdots & \varepsilon_{Ni,K}^{(Ni,K)} \end{bmatrix}}_{\mathbf{E}_{-ii}} \begin{bmatrix} 1 - (1+\bar{\tau}_i)(1+\omega_{ni,g}) \frac{P_{1i,1}}{\tilde{P}_{1i,1}} \\ \vdots \\ 1 - (1+\bar{\tau}_i)(1+\omega_{ni,g}) \frac{P_{Ni,K}}{\tilde{P}_{Ni,K}} \end{bmatrix} = \mathbf{0}. \tag{I.10}$$

Following the proof of Lemma 6 from Appendix E, we can easily show the matrix $\mathbf{E}_{-ii}^{(ii)}$ is invertible. We can, thus, invert the system specified by Equation I.10 to produce the following formula for optimal import price wedges:

$$\left[(1+\bar{\tau}_i)(1+\omega_{ji,k}) \frac{P_{ji,k}}{\tilde{P}_{ji,k}^{**}} \right]_{j,k} = \mathbf{1} + \mathbf{E}_{-ii}^{-1} \mathbf{E}_{-ii}^{(ii)} \left[1 - \frac{1+\mu_k}{1+\bar{\mu}_i} \right]_k, \tag{I.11}$$

where, to be clear, $\mathbf{E}_{-ii} \equiv \left[\mathbf{E}_{ni}^{(ji)} \right]_{j,n \neq i}$ and $\tilde{\mathbf{E}}_{-ii}^{(ii)} \equiv \left[\mathbf{E}_{ni}^{(ii)} \right]_{n \neq i}$ are respectively $(N-1)K \times (N-1)K$ and $(N-1)K \times K$ matrixes of demand elasticities. Note that the optimal choice w.r.t. $\tilde{\mathbf{P}}_{ji}$, ensures that $\widetilde{\Delta}_i'(.) \approx 0$. Hence, the system of F.O.C. specified by Equation I.6, transforms to the exact same system we solved in Appendix E. Without repeating the details of our prior derivation, the optimal export price wedges are given by

$$\left[\frac{P_{ij,k}}{\tilde{P}_{ij,k}^{**}} (1+\bar{\tau}_i)^{-1} \right]_{j,k} = \mathbf{E}_{ij}^{-1} \mathbf{E}_{ij}^{(-ij)} \left(\mathbf{1}_{(N-1)K} + \mathbf{\Omega}_{-ii} \right), \tag{I.12}$$

where $\mathbf{1}_{(N-1)K}$ is a $N(K-1) \times 1$ column vector of ones; $\mathbf{\Omega}_{-ii} = \left[\Omega_{ni,g} \right]_{n \neq i,g}$ is a $N(K-1) \times 1$ vector of (inverse) export supply elasticities; and $\mathbf{E}_{ij}^{(-ij)}$ and $\mathbf{E}_{ij}$ have the same description as in Appendix E. The "$**$" notation is used to highlight the fact that we are solving for second-best price wedges. Next, we can recover the optimal (second-best) import tax and export subsidy rates from the optimal (second-best) price wedges implies by Equations I.11 and I.12. Specifically, noting the following relationships,

$$1 + t_{ji,k}^{**} = \frac{\tilde{P}_{ji,k}^{**}}{P_{ji,k}}; \qquad 1 + x_{ij,k}^{**} = \frac{P_{ij,k}}{\tilde{P}_{ij,k}^{**}};$$

---

as

$$\text{[Slutsky equation]} \qquad e_{ii,g} \varepsilon_{ii,g}^{(ji,k)} + e_{ji,k} e_{ii,g} \eta_{ii,g} = e_{ji,k} \varepsilon_{ji,k}^{(ii,g)} + e_{ii,g} e_{ji,k} \eta_{ji,k}.$$

country i's unilaterally second-best trade tax schedule can be expressed as follows:

$$\text{[import tariff]} \quad \mathbf{1}+\mathbf{t}_{ij}^{**} = (1+\bar{\tau}_i)\left(\mathbf{1}+\mathbf{\Omega}_{ji}\right) \oslash \left(\mathbf{1}+\mathbf{E}_{-ii}^{-1}\,\mathbf{E}_{-ii}^{(ii)}\left[1-\frac{1+\mu_k}{1+\bar{\mu}_i}\right]_k\right)$$

$$\text{[export subsidy]} \quad \mathbf{1}+\mathbf{x}_{ij}^{**} = -(1+\bar{\tau}_i)\left(\mathbf{E}_{ij}^{-1}\mathbf{E}_{ij}^{(-ij)}(\mathbf{1}+\mathbf{\Omega}_{-ii})\right) \odot \left[\frac{1+\mu_k}{1+\bar{\mu}_i}\right]_k.$$

To conclude the proof we can invoke the multiplicity of the optimal trade tax schedules (Lemma 1). As in Theorem 1, this feature indicates that the value assigned to $\bar{\tau}_i$ is redundant. In particular, following Lemma 1, we can multiply $(1+\bar{\tau}_i)$ in the above equation with any non-negative tax shifter $1+\bar{t}_i \in \mathbb{R}_+$, and maintain optimality. That being the case, the exact value assigned to $\bar{\tau}_i$ is redundant and the following describes all possible optimal tax schedules:aa

$$\text{[import tariff]} \quad \mathbf{1}+\mathbf{t}_{ij}^{**} = (1+\bar{t}_i)\left(\mathbf{1}+\mathbf{\Omega}_{ji}\right) \oslash \left(\mathbf{1}+\mathbf{E}_{-ii}^{-1}\,\mathbf{E}_{-ii}^{(ii)}\left[1-\frac{1+\mu_k}{1+\bar{\mu}_i}\right]_k\right)$$

$$\text{[export subsidy]} \quad \mathbf{1}+\mathbf{x}_{ij}^{**} = -(1+\bar{t}_i)\left(\mathbf{E}_{ij}^{-1}\mathbf{E}_{ij}^{(-ij)}(\mathbf{1}+\mathbf{\Omega}_{-ii})\right) \odot \left[\frac{1+\mu_k}{1+\bar{\mu}_i}\right]_k.$$

J Proof of Theorem 3

Theorem 3 concerns the second-best case where the government in i can choose only $\tilde{\mathbf{P}}_{ji}$, which is the vector of import prices (i.e., $\tilde{\mathbb{P}}_i = \{\tilde{\mathbf{P}}_{ji}\}$). To prove this theorem we capitalize on two results from Appendix I: First, the F.O.C. derived w.r.t. $\tilde{P}_{ji,k} \in \tilde{\mathbf{P}}_{ji}$ does not change with the unavailability of $\tilde{\mathbf{P}}_{ij}$ from the government's policy set $\tilde{\mathbb{P}}_i$. Hence, the F.O.C. w.r.t. $\tilde{P}_{ji,k}$ is described by Equation I.3 even if $\tilde{P}_{ij,k} \notin \tilde{\mathbb{P}}_i$. Second, recall from Appendix that we were able to solve the system specified by I.3 independent of the F.O.C. w.r.t. $\tilde{\mathbf{P}}_{ij}$. Invoking these two observations, the formula for optimal tariff in the case studied by Theorem 3 is given by I.11:

$$1+\mathbf{t}_{ji}^{***} = (1+\bar{\tau}_i)\left(\mathbf{1}+\mathbf{\Omega}_{ji}\right) \oslash \left(\mathbf{1}+\mathbf{E}_{-ii}^{-1}\,\mathbf{E}_{-ii}^{(ii)}\left[1-\frac{1+\mu_k}{1+\bar{\mu}_i}\right]_k\right).$$

Unlike Theorem 2, through, $\bar{\tau}_i$ is no longer redundant. Since export taxes (or equivalently $\tilde{\mathbf{P}}_{ij}$) are excluded from the government's policy set, we can no longer invoke the multiplicity implied by Lemma 1. Instead, we have to formally characterize, $\bar{\tau}_i$, starting from its definition:

$$\bar{\tau}_i \equiv \frac{\left(\frac{\partial W_i(.)}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}\left(\frac{\partial V_i(.)}{\partial Y_i}\right)^{-1}}{\left(\partial \mathrm{T}_i(\tilde{\mathbb{P}}_i,\mathbf{w})/\partial \ln w_i\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}}. \tag{J.1}$$

Also, recall that $W_i(\tilde{\mathbb{P}}_i;\mathbf{w}) = V_i(Y_i(\tilde{\mathbb{P}}_i;\mathbf{w}),\tilde{\mathbf{P}}_{ii},\tilde{\mathbf{P}}_{ji})$, where $\tilde{\mathbf{P}}_{ji} \sim \tilde{\mathbf{P}}_{-ii} \equiv \left\{\tilde{P}_{ji,k}\right\}_{j\neq i,k}$ while income equals wage payments, plus profits, plus import tax revenues: $Y_i = (1+\bar{\mu}_i)w_iL_i + (\tilde{\mathbf{P}}_{ji}-\mathbf{P}_{ji})\cdot\mathbf{Q}_{ji}$. Borrowing from the results in Appendixes E and I, the numerator in Equation J.1 can be unpacked as follows:

$$\begin{aligned}
\left(\frac{\partial W_i(.)}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}\left(\frac{\partial V_i}{\partial Y_i}\right)^{-1} &= \left(\frac{\partial Y_i}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} + \left(\frac{\partial V_i}{\partial Y_i}\right)^{-1}\frac{\partial V_i(Y_{i},\tilde{\mathbf{P}}_i)}{\partial \ln \tilde{\mathbf{P}}_{ii}}\cdot\frac{\partial \ln \tilde{\mathbf{P}}_{ii}}{\partial \ln w_i} \\
&= \bar{\mu}_i w_i L_i + \left(\frac{\partial \bar{\mu}_i}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} w_i L_i + (\tilde{\mathbf{P}}_{-ii}-\mathbf{P}_{-ii})\cdot\left(\frac{\partial \mathbf{Q}_{-ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} - \mathbf{P}_{ii}\cdot\mathbf{Q}_{ii} \\
&= \sum_{n\neq i}\left[\mathbf{P}_{in}\cdot\mathbf{Q}_{in}\right] + \left(\mathbf{1}-\frac{\bar{\mu}_i}{\boldsymbol{\mu}}\right)\odot\mathbf{P}_{ii}\cdot\left(\frac{\partial \mathbf{Q}_{ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} + (\tilde{\mathbf{P}}_{ii}-\mathbf{P}_{-ii})\cdot\left(\frac{\partial \mathbf{Q}_{-ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}.
\end{aligned} \tag{J.2}$$

To be clear about the notation, $\frac{\bar{\mu}_i}{\boldsymbol{\mu}} \equiv \left[\frac{\bar{\mu}_i}{\mu_k}\right]_k$, while $\odot$ and $\cdot$ respectively denote *inner* and *element-wise* products of equal-sized vectors, i.e., $\mathbf{a}\cdot\mathbf{b} = \sum_n a_n b_n$ and $\mathbf{a}\odot\mathbf{b} = [a_n b_n]_n$. Next, we move on to characterizing the denominator of Equation J.1. Noting that $\mathrm{T}(\tilde{\mathbb{P}}_i,\mathbf{w}) \equiv \sum_{j\neq i}\left[\mathbf{P}_{ji}\cdot\mathbf{Q}_{ji}-\mathbf{P}_{ij}\cdot\mathbf{Q}_{ij}\right]$, we can borrow from the results in Appendixes E and I to unpack the aforementioned term as follows:

$$\left(\frac{\partial \mathrm{T}_i(.)}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} = \left(\frac{\partial}{\partial \ln w_i}\sum_{j\neq i}\left[\mathbf{P}_{ji}\cdot\mathbf{Q}_{ji} - \mathbf{P}_{ij}\cdot\mathbf{Q}_{ij}\right]\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} = \mathbf{P}_{-ii}\cdot\left(\frac{\partial \mathbf{Q}_{-ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} - \sum_{j\neq i}\left[\left(\frac{\partial \mathbf{P}_{ij}\cdot\mathbf{Q}_{ij}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}\right]. \tag{J.3}$$

Plugging Equations J.2 and J.3 back into the expression for $\bar{\tau}_i$ yields the following:

$$\bar{\tau}_i = \frac{\sum_{n\neq i}\left[\mathbf{P}_{in}\cdot\mathbf{Q}_{in}\right] + \left(\mathbf{1}-\frac{\overline{\mu}_i}{\boldsymbol{\mu}}\right)\odot\mathbf{P}_{ii}\cdot\left(\frac{\partial \mathbf{Q}_{ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} + \left(\tilde{\mathbf{P}}_{ii} - \mathbf{P}_{-ii}\right)\cdot\left(\frac{\partial \mathbf{Q}_{-ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}}{\mathbf{P}_{-ii}\left(\frac{\partial \mathbf{Q}_{-ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} - \sum_{j\neq i}\left[\left(\frac{\partial \mathbf{P}_{ij}\cdot\mathbf{Q}_{ij}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}\right]}. \tag{J.4}$$

We can further simplify the above expression by invoking the F.O.C. described by Equation I.8. This equation indicates that the following relationship ought to hold at the optimum $\tilde{\mathbb{P}}_i = \tilde{\mathbb{P}}_i^{***}$:

$$\sum_{j\neq i}\sum_k\left[\left(\mathbf{1}-\frac{\overline{\mu}_i}{\boldsymbol{\mu}}\right)\odot\mathbf{P}_{ii}\cdot\left(\frac{\partial \mathbf{Q}_{ii}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_i^{***},\mathbf{w}_{-i}} + \left(\tilde{\mathbf{P}}_{-ii} - (1+\bar{\tau}_i)\mathbf{P}_{-ii}\right)\cdot\left(\frac{\partial \mathbf{Q}_{-ii}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_i^{***},\mathbf{w}_{-i}}\right] = 0.$$

Now, we will rearrange and simplify the above relationship in such a way that will help us simply Equation J.4. To this, we invoke the property that the Marshallain demand function is homogeneous of degree zero. Combining this property with the fact that $\frac{\partial \ln Y_i}{\partial \ln w_i} \approx \frac{\partial \ln \tilde{P}_{ii,k}}{\partial \ln w_i} = 1$, we can simplify the above as follows:

$$\left(\mathbf{1}-\frac{\overline{\mu}_i}{\boldsymbol{\mu}}\right)\odot\mathbf{P}_{ii}\cdot\left(\frac{\partial \mathbf{Q}_{ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} + \left(\tilde{\mathbf{P}}_{ii} - (1+\bar{\tau}_i)\mathbf{P}_{-ii}\right)\cdot\left(\frac{\partial \mathbf{Q}_{-ii}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}} = 0.$$

Using the above equation, we can cancel out the mirroring expressions in the numerator and denominator of Equation J.4. Doing so reduces and simplifies the expression for $\bar{\tau}_i$ to the following:

$$\bar{\tau}_i = \frac{-\sum_{n\neq i}\left(\mathbf{P}_{in}\cdot\mathbf{Q}_{in}\right)}{\sum_{j\neq i}\left[\left(\frac{\partial \mathbf{P}_{ij}\cdot\mathbf{Q}_{ij}}{\partial \ln w_i}\right)_{\tilde{\mathbb{P}}_i,\mathbf{w}_{-i}}\right]} = \frac{-1}{\sum_{j\neq i}\left[\mathbf{X}_{ij}\cdot\left(\mathbf{I}_K + \mathbf{E}_{ij}\right)\mathbf{1}_K\right]}. \tag{J.5}$$

The $K \times 1$ vector $\mathbf{X}_{ij} = \left[\chi_{ij,k}\right]_k$ is compose of export shares, which are defined as $\chi_{ij,k} \equiv \frac{P_{ij,k}Q_{ij,k}}{\sum_{n\neq i}\mathbf{P}_{in}\cdot\mathbf{Q}_{in}}$. To provide some intuition, the denominator of the above equation corresponds to the elasticity of international demand for origin i's labor. As such, $\bar{\tau}_i$ can be interpreted as country i's optimal markup on its wage rate in international (non-i) markets.

K Optimal Policy under IO Linkages (Theorem 4)

We first present a formal description of equilibrium under input-output (IO) linkages. We use the $\mathcal{C}$ superscript to denote final consumption goods and the $\mathcal{I}$ superscript to denote intermediate inputs. To given an example: $Q^{\mathcal{C}}_{ji,k}$ denotes the quantity of a "final" goods associated with *origin j–destination i–industry k*, while $Q^{\mathcal{I}}_{ji,k}$ denotes the quantity of an "intermediate" goods associated with origin j–destination i–industry k. Without loss of generality, we assume that good ji,k exhibits the same price irrespective of whether it is used as a final good or an intermediate input good: $\tilde{P}_{ji,k} \sim \tilde{P}^{\mathcal{C}}_{ji,k} = \tilde{P}^{\mathcal{I}}_{ji,k}$.

On the production side, we impose no restrictions on how intermediate inputs are aggregated in the production process. We, however, assume that the share of labor in production is constant and equal to $1-\bar{\alpha}_{i,k}$ for each *origin i–industry k*. To track the demand for inputs, we use $\mathcal{Y}_{i,k}$ to denote the gross revenue associated with *origin i–industry k*. Correspondingly, $\bar{\alpha}_{i,k}\mathcal{Y}_{i,k}$ denotes *origin i–industry k*'s total expenditure on intermediate inputs.

Marshallian Demand under IO Linkages

We suppose that overall demand for good ji,k, which is the sum of final good demand based on utility maximization and input demand based on cost minimization, is given by the following demand function

$$Q_{ji,k} = Q^{\mathcal{I}}_{ji,k} + Q^{\mathcal{C}}_{ji,k} = \mathcal{D}_{ji,k}(E_i, \tilde{\mathbf{P}}_i),$$

where $E_i = Y_i + \sum_g \bar{\alpha}_{i,g} \mathcal{Y}_{i,g}$ denotes market i's total expenditure on final and intermediate input goods. To make the notation consistent with our previous derivations, we use $\varepsilon^{(ni,g)}_{ji,k}$ and $\eta_{ji,k}$ to denote the price and income elasticities associated with the IO-augmented Marshallian demand function $\mathcal{D}_{ji,k}(E_i, \tilde{\mathbf{P}}_i)$.

General Equilibrium under IO Linkages

As in the baseline model, we express all equilibrium outcomes (except for wages) as a function of global taxes ($\mathbf{x}$, $\mathbf{t}$, and $\mathbf{s}$), treating wages $\mathbf{w} \equiv \{w_i\}_i$ as given. This formulation derives from solving a system that imposes all equilibrium conditions aside from the labor market clearing conditions. We formally outline this formulation below.

Notation. For a given vector of taxes and wages $\boldsymbol{T} = (\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}; \boldsymbol{w})$, equilibrium outcomes $Y_i(\boldsymbol{T})$, $\mathcal{Y}_{i,k}(\mathbf{T})$, $P_{ji,k}(\boldsymbol{T})$, $\tilde{P}_{ji,k}(\boldsymbol{T})$, $Q_{ji,k}(\boldsymbol{T})$ are determined such that (i) producer prices are characterized by 13; (ii) consumer prices are given by 7; (iii) Consumption and input demand choices are given by $\mathcal{D}_{ji,k}(E_i, \tilde{\mathbf{P}}_i)$, where $E_i = Y_i + \sum_g \bar{\alpha}_{i,g} \mathcal{Y}_{i,g}$; (iv) net income (which dictates total final good expenditure by country i) equals wage payments plus tax revenues: $Y_i = w_i L_i + \mathcal{R}_i$,[31] where $\mathcal{R}_i$ are described by 8 and (v) gross industry-level revenues are given by $\mathcal{Y}_{i,g} = \sum_n P_{in,k} Q_{in,k}$.

As in the baseline model, $\mathbf{w}$ is itself an equilibrium outcome. So, a vector $\mathbf{T} = (\mathbf{t}, \mathbf{x}, \mathbf{s}; \mathbf{w})$ is feasible insofar as $\mathbf{w}$ is the equilibrium wage, consistent with $\mathbf{t}$, $\mathbf{x}$, and $\mathbf{s}$. So, to fix ideas we define the set of feasible *policy–wage* vectors as follows.

Definition (D2-IO). The set of feasible policy–wage vectors, $\mathbb{F}$, consists of any vector $\boldsymbol{T} = (\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}; \boldsymbol{w})$ where $\boldsymbol{w}$ satisfies the labor market clearing condition in every country, given $\boldsymbol{t}$, $\boldsymbol{x}$, and $\boldsymbol{s}$:

$$\mathbb{F} = \left\{ \boldsymbol{T} = (\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}; \boldsymbol{w}) \mid w_i L_i = \sum_j \sum_k Qz_{ij,k}(\boldsymbol{T}) - \sum_j \sum_k P^{\mathcal{I}}_{ji,k}(\boldsymbol{T}) Q^{\mathcal{I}}_{ji,k}(\boldsymbol{T}); \quad \forall i \in \mathbb{C} \right\}.$$

Before moving on to the proof, two important details are in order: First, we can easily verify that the labor market clearing condition specified by Definition D2-IO is equivalent to the balanced trade condition. Second, under IO linkages, the choice w.r.t. taxes (or equivalently $\tilde{\mathbb{P}}_i \equiv \{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij}\}$) may affect the entire vector of producer prices, $\left\{ P_{nj,k} \right\}$, through its effect on input prices. To track these IO-related effects, let $\alpha^{j,g}_{i,k}$ denotes the (possibly variable) cost share of intermediate inputs from *origin* $j\times$ *industry* g used in the output of *origin* $i\times$ *industry* k. By Shepherd's Lemma, the direct effect of raising input price $\tilde{P}^{\mathcal{I}}_{ji,g}$ on the producer price $P_{ij,k}$ can be expressed as follows:

$$\text{[Shepherd's Lemma]} \qquad \left(\frac{\partial \ln P_{ij,k}(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \ln \tilde{P}^{\mathcal{I}}_{ji,g}} \right)_{\mathbb{P}_{-ji,g}, \mathbf{w}} = \alpha^{j,g}_{i,k}. \quad \forall (\jmath, j, i \in \mathbb{C}); \; \forall (g, k \in \mathbb{K}).$$

We use the Shepherd's Lemma in combination with our dual approach (from Appendix E) to characterize the optimal policy schedule for each country i. Recall that the optimal policy problem in our dual approach is reformulated as

$$\max_{\tilde{\mathbb{P}}_i} \; W_i(\tilde{\mathbb{P}}_i; \mathrm{w}) \equiv V_i(Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}), \tilde{\mathbf{P}}_i) \quad s.t. \; (\tilde{\mathbb{P}}_i; \mathbf{w}) \in \mathbb{F}_P,$$

where $\tilde{\mathbb{P}}_i \equiv \left\{ \tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij} \right\}$ denotes the vector of consumer prices directly linked to economy i. The feasible set $\mathbb{F}_P$ is defined analogously to $\mathbb{F}$. Below, we derive and solve the system of F.O.C. associated with the above problem, building on the results introduced earlier under Appendix E.

Tax Neutrality under IO Linkages

Our baseline characterization of optimal policy relied on the tax neutrality result presented under Lemma 1. An analogous (but slightly different) neutrality result holds under IO linkages. To present this result, we use operator$\mathscr{C}(.)$, which configures a uniform tax-shifter depending on whether the

[31] *Note that net profits are equal top zero (i.e., $\Pi_i = 0$) as we are focusing on the case of free entry.*

taxed item is used for final consumption or intermediate input use. In particular, for an arbitrary tax-shifter, $\tilde{a} \in \mathbb{R}_+$, $\mathscr{C}(\tilde{a}) = \tilde{a}$ if the taxed item is a final good and $\mathscr{C}(\tilde{a}) = 1$ otherwise.

Lemma 8. *[Tax Neutrality under IO Linkages] For any a and $\tilde{a} \in \mathbb{R}_+$ (i) if $\mathbf{T} = (\mathbf{1}+\mathbf{t}_i, \mathbf{t}_{-i}, \mathbf{1}+\mathbf{x}_i, \mathbf{x}_{-i}, \mathbf{1}+\mathbf{s}_i, \mathbf{s}_{-i}; \bar{w}_i, \mathbf{w}_{-i}) \in \mathbb{F}$, then $\mathbf{T}' = (a(\mathbf{1}+\mathbf{t}_i)/\mathscr{C}(\tilde{a}), \mathbf{t}_{-i}, a(\mathbf{1}+\mathbf{x}_i)/\mathscr{C}(\tilde{a}), \mathbf{x}_{-i}, (\mathbf{1}+\mathbf{s}_i)\mathscr{C}(\tilde{a}), \mathbf{s}_{-i}; a\bar{w}_i, \mathbf{w}_{-i}) \in \mathbb{F}$. Moreover, (ii) welfare is preserved under $\mathbf{T}$ and $\mathbf{T}'$: $W_n(\mathbf{T}) = W_n(\mathbf{T}')$ for all $n \in \mathbb{C}$.*

The above lemma is akin to Lemma 1, but differs in one basic detail. The neutrality of uniform trade tax adjustments (i.e., the Lerner Symmetry) holds in the IO model without qualification. The neutrality of uniform domestic tax adjustments holds the consumption side but not on the production side. More specifically, a uniform increase in consumption taxes is welfare-neutral in the IO model. Accordingly, the tax adjustments that apply via $\mathscr{C}(\tilde{a})$ are constructed to mimic a uniform consumption tax hike. With the above background, we are now ready to derive and solve the system of F.O.C.s that determine optimal policy under IO linkages.

***Step* #1: Deriving the F.O.C. *w.r.t.* $\tilde{P}_{ji,k}$ and $\tilde{P}_{ii,k} \in \tilde{\mathbb{P}}_i$**

First, we derive the F.O.C. w.r.t. to import variety ji,k, supplied by origin j–industry k. Given that $W_i = V_i(Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}), \tilde{\mathbf{P}}^{\mathcal{C}}_{ii}, \tilde{\mathbf{P}}^{\mathcal{C}}_{ji})$, the F.O.C. w.r.t. $\tilde{P}_{ji,k} \sim \tilde{P}^{\mathcal{I}}_{ji,k} \sim \tilde{P}^{\mathcal{C}}_{ji,k}$, holding $\tilde{\mathbb{P}}_{-ji,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ji,k}\}$ constant, can be stated as

$$\left(\frac{\partial W_i}{\partial \ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}^{\mathcal{C}}_i)}{\partial \ln \tilde{P}_{ji,k}} + \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial Y_i}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}, \tilde{\mathbb{P}}_{-ji,k}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i, \mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d} \ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}} = 0. \tag{K.1}$$

The right-hand side of the above equation can be characterized similar to Appendix E, with two distinctions: First, total demand for good ji,k is the sum of consumption plus input demand: $Q_{ji,k} = Q^{\mathcal{C}}_{ji,k} + Q^{\mathcal{I}}_{ji,k}$. So, we have to distinguish between welfare effects that channel through consumption and those that channel through input demand. Second, we need to account for the effect of a change in input price $\tilde{P}_{ji,k} \sim \tilde{P}^{\mathcal{I}}_{ji,k}$ on the producer prices associated with economy i. To this end, we can invoke Shepherd's Lemma, which implies that

$$\left(\frac{\partial \ln P_{ij,k}(\tilde{\mathbb{P}}_i; \mathbf{w})}{\partial \ln \tilde{P}^{\mathcal{I}}_{ji,g}}\right)_{\mathbb{P}_{-ji,g}, \mathbf{w}} = \alpha^{j,g}_{i,k}. \qquad \forall \jmath, j, i \in \mathbb{C}; \;\; g, k \in \mathbb{K}.$$

Considering the above caveats, we can proceed as in Appendix E. By Roy's identity, the first term on the right-hand side of the F.O.C. (Equation K.1) can be stated as

$$\frac{\partial V_i(Y_i, \tilde{\mathbf{P}}^{\mathcal{C}}_i)}{\partial \ln \tilde{P}_{ji,k}} = -\tilde{P}_{ji,k} Q^{\mathcal{C}}_{ji,k}\left(\frac{\partial V_i}{\partial Y_i}\right).$$

Next, consider the second term on the right-hand side of Equation K.1, which accounts for income effects. Recall that total income in country i equals the sum of wage payments plus import, production and export tax revenues:

$$Y_i(\tilde{\mathbb{P}}_i; \mathbf{w}) = w_i L_i + \sum_{n \neq i}\left[(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni}) \cdot \mathbf{Q}_{ni}\right] + (\tilde{\mathbf{P}}_{ii} - \mathbf{P}_{ii}) \cdot \mathbf{Q}_{ii} + \sum_{n \neq i}\left[(\tilde{\mathbf{P}}_{in} - \mathbf{P}_{in}) \cdot \mathbf{Q}_{in}\right].$$

The effect of $\tilde{P}_{ji,k}$ on import tax revenues can be derived and express exactly as in Appendix E:

$$\left(\frac{\partial \sum_{n \neq i}\left[(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni}) \cdot \mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}, \tilde{\mathbb{P}}_{-ji,k}} = \tilde{P}_{ji,k} Q_{ji,k} + \sum_g \sum_{n \neq i}\left[\left(\tilde{P}_{ni,g} - [1+\omega_{ni,g}] P_{ni,g}\right) Q_{ni,g}\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w}, \tilde{\mathbb{P}}_{-ji,k}}\right] \tag{K.2}$$

The logic is that holding the vector of wages $\mathbf{w}$ and country i's export prices $\tilde{\mathbf{P}}_{ij} \in \tilde{\mathbb{P}}_i$ fixed, a change in $\tilde{P}_{ji,k}$ has not effect on the producer price of imports $\mathbf{P}_{ji}$ through the input-output network.

The effect of a change in $\tilde{P}_{ji,k}$ on country i's production and export tax revenues can be formulated as

$$\left(\frac{\partial}{\partial \ln \tilde{P}_{ji,k}}\left\{(\tilde{\mathbf{P}}_{ii}-\mathbf{P}_{ii})\cdot\mathbf{Q}_{ii,g}+\sum_{n\neq i}\left[(\tilde{\mathbf{P}}_{in}-\mathbf{P}_{in})\cdot\mathbf{Q}_{in}\right]\right\}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$$

$$=\sum_g\left[\left(\tilde{P}_{ii,g}-P_{ii,g}\right)Q_{ii,g}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]+\sum_g\sum_n\left[P_{in,g}Q_{in,g}\left[\left(\frac{\partial P_{in,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}+\left(\frac{\partial \ln P_{in,g}}{\partial \ln \tilde{P}^{\mathcal{I}}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]\right], \tag{K.3}$$

The above expression differs from Equation K.4 (in Appendix E) in the last term on the second line. This term accounts for the effect of raising input price $\tilde{P}_{ji,k}\sim\tilde{P}^{\mathcal{I}}_{ji,k}$ on the producer prices associated with economy i. As explained above, we can appeal to Shephard's lemma to simplify this extra term as

$$\sum_{g\in\mathbb{K}}\sum_{n\in\mathbb{C}}\left[P_{in,g}Q_{in,g}\left(\frac{\partial \ln P_{in,g}}{\partial \ln \tilde{P}^{\mathcal{I}}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k},\mathbf{w}}\right]=-\sum_{g\in\mathbb{K}}\sum_{n\in\mathbb{C}}\left(Q_{in,g}P_{in,g}\alpha^{j,k}_{i,g}\right)=-\tilde{P}_{ji,k}Q^{\mathcal{I}}_{ji,k}.$$

Plugging the above expression back into Equation K.3 and redoing the derivations covered in Appendix E, yields the following expression for the effect of $\tilde{P}_{ji,k}$ on country i's production and export tax revenues:

$$\left(\frac{\partial}{\partial \ln \tilde{P}_{ji,k}}\left\{\sum_n\left[(\tilde{\mathbf{P}}_{in}-\mathbf{P}_{in})\cdot\mathbf{Q}_{in}\right]\right\}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}=-\tilde{P}_{ji,k}Q^{\mathcal{I}}_{ji,k}+\sum_g\left[\left(\tilde{P}_{ii,g}-\left[1-\frac{\mu_g}{1+\mu_g}\right]\right)P_{ii,g}Q_{ii,g}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right]. \tag{K.4}$$

where recall that $\left(\partial \ln Q_{ii,g}/\partial \ln \tilde{P}_{ji,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}$ encompasses price and income effects à la Equation E.5 in Appendix E. Combining Equations K.2 and K.4, and noting that $\tilde{P}_{ji,k}Q_{ji,k}-\tilde{P}_{ji,k}Q^{\mathcal{I}}_{ji,k}=\tilde{P}_{ji,k}Q^{\mathcal{C}}_{ji,k}$ yields the following expression that summarizes all the revenue-related welfare effects in the F.O.C.:

$$\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}=\tilde{P}_{ji,k}Q^{\mathcal{C}}_{ji,k}+\sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-[1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\varepsilon^{(ji,k)}_{ni,g}\right]$$
$$+\sum_g\left[\left(\tilde{P}_{ii,g}-\frac{1}{1+\mu_g}P_{ii,g}\right)Q_{ii,g}\varepsilon^{(ji,k)}_{ii,g}\right]+\Delta_i(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial E_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}. \tag{K.5}$$

The uniform term $\Delta_i(.)$ accounts for circular income effects and is given by Equation E.13 in Appendix E. Finally, the last term on the right-hand side of Equation K.1, which accounts for general equilibrium wage effects, can be specified in the same exact way as in Appendix E:

$$\left(\frac{\partial W_i(\tilde{\mathbb{P}}_i,\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{d\mathbf{w}}{d\ln \tilde{P}_{ji,k}}\right)_{\tilde{\mathbb{P}}_{-ji,k}}=-\bar{\tau}_i\sum_g\sum_{n\neq i}\left[P_{ni,g}Q_{ni,g}\left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}\right],$$

where $\bar{\tau}_i$ is given by Equation E.15 in Appendix E. Combining the above expressions, the F.O.C. specified by Equation K.1 reduced to

$$[\text{FOC w.r.t. }\tilde{P}_{ji,k}]\qquad \sum_{n\neq i}\sum_g\left[\left(\frac{\tilde{P}_{ni,g}}{P_{ni,g}}-(1+\bar{\tau}_i)(1+\omega_{ni,g})\right)P_{ni,g}Q_{ni,g}\varepsilon^{(ji,k)}_{ni,g}\right]$$
$$+\sum_g\left[\left(\frac{\tilde{P}_{ii,g}}{P_{ii,g}}-\frac{1}{1+\mu_g}\right)P_{ii,g}Q_{ii,g}\varepsilon^{(ji,k)}_{ii,g}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial E_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ji,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ji,k}}=0, \tag{K.6}$$

where $\widetilde{\Delta}_i(.)$ is given by Equation E.18 in Appendix E. Note that the above equation has an identical representation to the F.O.C. in the baseline model. The intuition is that holding country i's export prices $\tilde{\mathbf{P}}_{ij} \in \tilde{\mathbb{P}}_i$ fixed, the choice w.r.t. $\tilde{P}_{ji,k}$ has no first-order effect on country i's terms-of-trade channels through the input-output network. If good ji,k is used as an input in export good in,g, any possible terms-of-trade gains from taxing $\tilde{P}_{ji,k}$ will be internalized by the optimal choice w.r.t. $\tilde{P}_{in,g}$. Furthermore, it is easy to check that Equation K.6 characterizes the F.O.C. w.r.t. $\tilde{P}_{ii,k} \in \tilde{\mathbb{P}}_i$ as long as we replace ji,k with ii,k everywhere in that equation. Finally, as in Appendix E, we do not unpack the uniform term $\bar{\tau}_i$ because the multiplicity of country i's optimal tax schedule will render the exact value assigned to $\bar{\tau}_i$ redundant.

***Step* #2: Deriving the F.O.C. *w.r.t.* $P_{ij,k} \in \tilde{\mathbb{P}}_i$**

Consider export variety ij,k, which is sold to destination $j \neq i$ in industry k. Noting that $W_i = V_i(Y_i(\tilde{\mathbb{P}}_i;\mathbf{w}), \tilde{\mathbf{P}}_{ii}^{\mathcal{C}}, \tilde{\mathbf{P}}_{ji}^{\mathcal{C}})$, the F.O.C. w.r.t. $\tilde{P}_{ij,k}$, holding $\tilde{\mathbb{P}}_{-ij,k} \equiv \tilde{\mathbb{P}}_i - \{\tilde{P}_{ij,k}\}$ constant, can be stated as

$$\left(\frac{\partial W_i}{\partial \ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ij,k}} + \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial Y_i}\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} + \left(\frac{\partial W_i(\tilde{\mathbb{P}}_i,\mathbf{w})}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i} \cdot \left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}} = 0. \tag{K.7}$$

The first term as before accounts for direct price effects. This term is trivially equal to zero since $\tilde{P}_{ij,k} \notin \tilde{\mathbf{P}}_i$. That is, since ij,k is not part of the domestic consumption bundle, raising its price has no direct effect on consumer surplus in country i:

$$\tilde{P}_{ij,k} \notin \tilde{\mathbf{P}}_i \implies \frac{\partial V_i(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ij,k}} = 0. \tag{K.8}$$

The second term in Equation K.7 accounts for welfare effects that channel through tax revenues. Specifically, Holding wages $\mathbf{w}$ fixed, the change in country i's income amounts to the change in import, domestic, and export tax revenues. The effect on import tax revenues can be expressed as follows:

$$\begin{aligned}\left(\frac{\partial \sum_{n\neq i}\left[(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni})\cdot\mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} =& \sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-P_{ni,g}\right)Q_{ni,g}\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right] \\ &- \sum_g\sum_{n\neq i}\left[P_{ni,g}Q_{ni,g}\left[\left(\frac{\partial P_{ni,g}}{\partial \ln Q_{nj,g}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_i}\left(\frac{\partial \ln Q_{nj,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} + \left(\frac{\partial \ln P_{ni,g}}{\partial \ln \tilde{P}^{\mathcal{I}}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right]\right].\end{aligned} \tag{K.9}$$

The above equation differs from Equation E.23 in Appendix E in only the last term on the second line. This term accounts for that fact that raising the price of input good ij,k can affect the entire vector of producer prices in the rest of the world through input-output networks. Given Shephard's lemma we can simplify this term by noting that

$$\Lambda_{ij,k} \equiv \sum_{n\neq i}\sum_{g\in\mathbb{K}}\left(P_{ni,g}Q_{ni,g}\left(\frac{\partial P_{ni,g}}{\partial \tilde{P}^{\mathcal{I}}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}\right) / \tilde{P}_{ij,k}Q_{ij,k}$$

denotes the share of the export value associated with good ij,k that is reimported back into economy i. Plugging the above expression back into K.9 and repeating the derivation performed in Appendix E, yields the following:

$$\begin{aligned}\left(\frac{\partial \sum_{n\neq i}\left[(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni})\cdot\mathbf{Q}_{ni}\right]}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}} =& -\Lambda_{ij,k}\tilde{P}_{ij,k}Q_{ij,k} - \sum_g\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon^{(ij,k)}_{nj,g}\right] \\ &+ \sum_g\sum_{n\neq i}\left[\left(\tilde{P}_{ni,g}-[1+\omega_{ni,g}]P_{ni,g}\right)Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial \ln E_i}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}.\end{aligned} \tag{K.10}$$

Repeating the derivation in Appendix E, the effect of a change in $\tilde{P}_{ij,k}$ on country i's production and export tax revenues can be formulated as

$$\left(\frac{\partial}{\partial \ln \tilde{P}_{ij,k}}\sum_{n}\left[(\tilde{\mathbf{P}}_{in}-\mathbf{P}_{in})\cdot\mathbf{Q}_{in}\right]\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=\tilde{P}_{ij,k}Q_{ij,k}+\sum_{g}\left[\left(\tilde{P}_{ij,g}-[1-\frac{\mu_g}{1+\mu_g}]P_{ij,g}\right)Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$+\sum_{g}\left[\left(\tilde{P}_{ii,g}-\frac{1}{1+\mu_g}P_{ii,g}\right)Q_{ii,g}\eta_{ii,g}\right]\left(\frac{\partial \ln E_i}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}. \tag{K.11}$$

To be clear, holding $\tilde{P}_{ji,k}\in\tilde{\mathbb{P}}_i$ fixed, a change in $\tilde{P}_{ij,k}$ has no effect on the input price faced by firm located in i. That is, $\left(\partial P_{ni,g}/\partial \tilde{P}^{\mathcal{I}}_{ij,k}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=0$. This point explains why the above expression is rather identical to that derived in Appendix E. Combining Equations K.10 and K.7, we can express the sum of all tax-revenue-related terms as

$$\left(\frac{\partial Y_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=\left(1-\Lambda_{ij,k}\right)\tilde{P}_{ij,k}Q_{ij,k}+\sum_{g}\left[\left(\tilde{P}_{ij,g}-[1-\frac{\mu_g}{1+\mu_g}]P_{ij,g}\right)Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$-\sum_{g}\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right]+\Delta_i(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial E_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}, \tag{K.12}$$

where $\Delta_i(.)$ encompasses the terms accounting for circular income effects and is given by Equation E.13. Taking note of the already-discussed distinctions between the present and baseline models and repeating the derivations performed earlier in Appendix E, the last term in right-hand side of Equation K.7) can be formulated as

$$\left(\frac{\partial W_i(.)}{\partial \mathbf{w}}\right)_{\tilde{\mathbb{P}}_i}\cdot\left(\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln \tilde{P}_{ij,k}}\right)_{\tilde{\mathbb{P}}_{-ij,k}}=\bar{\tau}_i(1-\Lambda_{ij,k})\tilde{P}_{ij,k}Q_{ij,k}+\sum_{g}\left[\bar{\tau}_i\tilde{P}_{ij,g}Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$-\sum_{g}\sum_{n\neq i}\left[\bar{\tau}_i\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right]-\sum_{g}\sum_{n\neq i}\left[[1+\omega_{ni,g}]\bar{\tau}_iP_{ni,g}Q_{ni,g}\eta_{ni,g}\right]\left(\frac{\partial E_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}. \tag{K.13}$$

Finally, plugging Equations K.8, K.12, and K.13 back into the F.O.C. (Equation K.7); and dividing by $(1+\bar{\tau}_i)$ yields the following optimality condition w.r.t. to price instrument $\tilde{P}_{ij,k}$:

$$[\text{FOC w.r.t. } \tilde{P}_{ij,k}]\qquad (1-\Lambda_{ij,k})\tilde{P}_{ij,k}Q_{ij,k}+\sum_{g}\left[\left(1-\frac{1}{(1+\bar{\tau}_i)(1+\mu_g)}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right)\tilde{P}_{ij,g}Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$-\sum_{g}\sum_{n\neq i}\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial E_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=0, \tag{K.14}$$

where $\widetilde{\Delta}_i(.)$ is defined as in Equation E.18. Also, we are not unpacking the term $\bar{\tau}_i$, for the same reasons discussed earlier.

Step #3: Solving the System of F.O.C.s and Establishing Uniqueness

To determine the optimal tax schedule we need to collect the system of first order conditions and simultaneously solve them under one system. For the ease of reference, we will restate the F.O.C. w.r.t. to each element of $\tilde{\mathbb{P}}_i$ below. Following Equation K.6, the F.O.C. w.r.t. $\tilde{P}_{\ell i,k}$ (where $\ell=i$ or $\ell=j\neq i$), can be expressed as

$$(1)\quad \sum_{n\neq i}\sum_{g}\left[\left(1-(1+\bar{\tau}_i)(1+\omega_{ni,g})\frac{P_{ni,g}}{\tilde{P}_{ni,g}}\right)e_{ni,g}\varepsilon_{ni,g}^{(\ell i,k)}\right]+$$
$$\sum_{g}\left[\left(1-\frac{1}{1+\mu_g}\frac{P_{ii,g}}{\tilde{P}_{ii,g}}\right)e_{ii,g}\varepsilon_{ii,g}^{(\ell i,k)}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w})\left(\frac{\partial\ln E_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{\ell i,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-\ell i,k}}=0.$$

where $e_{ni,g}=\tilde{P}_{ni,g}Q_{ni,g}/Y_i$ denotes the expenditure share on good ni,g. Following Equation K.14, the F.O.C. w.r.t. export price $\tilde{P}_{ij,k}$ is given by

$$(2)\quad 1-\Lambda_{ij,k}+\sum_{g}\left[\left(1-\frac{1}{(1+\mu_g)(1+\bar{\tau}_i)}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right)\frac{e_{ij,g}}{e_{ij,k}}\varepsilon_{ij,g}^{(ij,k)}\right]$$
$$-\sum_{n\neq i}\sum_{g}\left[\omega_{ni,g}\frac{e_{nj,g}}{e_{ij,k}}\varepsilon_{nj,g}^{(ij,k)}\right]+\widetilde{\Delta}_i(\tilde{\mathbb{P}}_i,\mathbf{w})\frac{E_i}{E_j}\left(\frac{\partial\ln E_i(\tilde{\mathbb{P}}_i;\mathbf{w})}{\partial\ln\tilde{P}_{ij,k}}\right)_{\mathbf{w},\tilde{\mathbb{P}}_{-ij,k}}=0.$$

First, note that the system of F.O.C.s (1) Appealing to above lemma, it immediately follows that the unique solution to the above equation is the trivial solution given by:

$$\frac{\tilde{P}^*_{ji,k}}{P_{ji,1}}=(1+\omega_{ji,k})(1+\bar{\tau}_i);\qquad \frac{\tilde{P}^*_{ii,k}}{P_{ii,k}}=\frac{1}{1+\mu_g}.\tag{K.15}$$

With the aid of the above result, we can proceed to solving System (2), knowing that $\widetilde{\Delta}_i(\tilde{\mathbb{P}}^*_i,\mathbf{w})=0$. To this end, let us economize on the notation by defining

$$\chi_{ij,k}\equiv\frac{1}{(1+\mu_g)(1+\bar{\tau}_i)}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}.$$

Appealing to this choice of notation the F.O.C. specified by System (2) implies the following optimality condition:

$$1-\Lambda_{ij,k}+\sum_{g}\left[(1-\chi_{ij,g})\frac{e_{ij,g}\varepsilon_{ij,g}^{(ij,k)}}{e_{ij,k}}\right]-\sum_{n\neq i}\sum_{g}\left[\omega_{ni,g}\frac{e_{nj,g}\varepsilon_{nj,g}^{(ij,k)}}{e_{ij,k}}\right]=0.\tag{K.16}$$

To simplify the above expression we will use a well-know result from consumer theory, namely, the Cournot aggregation, which implies:

$$[\text{Cournot aggregation}]\qquad 1+\sum_{g}\left[\frac{e_{ij,g}}{e_{ij,k}}\varepsilon_{ij,g}^{(ij,k)}\right]=-\sum_{n\neq i}\sum_{g}\left[\frac{e_{nj,g}}{e_{ij,k}}\varepsilon_{nj,g}^{(ij,k)}\right].$$

Combining the above expression with Equation K.16 and noting that by Slutsky's equation $\frac{e_{nj,g}}{e_{ij,k}}\varepsilon_{nj,g}^{(ij,k)}=\varepsilon_{ij,k}^{(nj,g)}$ (if $\eta_{ni,g}=1$ for all ni,g), yields the following:

$$-\sum_{g}\left[\chi_{ij,g}\varepsilon_{ij,k}^{(ij,g)}\right]-\sum_{n\neq i}\sum_{g}\left[(1+\omega_{ni,g})\varepsilon_{ij,k}^{(nj,g)}\right]=0\qquad\forall(ij,k).$$

We can formulate the above equation in matrix algebra as

$$-\mathbf{E}_{ij}\mathbf{X}_{ij}-\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1}_{(N-1)K}+\mathbf{\Omega}_i\right)-\mathbf{\Lambda}_{ij}=0,\tag{K.17}$$

where $\mathbf{X}_{ij}\equiv\left[\chi_{ij,k}\right]_k$ and $\mathbf{\Lambda}_{ij}\equiv\left[\Lambda_{ij,k}\right]_k$ are $K\times 1$ vectors. The $K\times K$ matrix $\mathbf{E}_{ij}\sim\mathbf{E}_{ij}^{(ij)}\equiv\left[\varepsilon_{ij,k}^{(ij,g)}\right]$ encompasses the own- and cross-price elasticities between the different varieties sold by origin i to market j. Analogously, $\mathbf{E}_{ij}^{(-ij)}\equiv\left[\varepsilon_{ij,k}^{(nj,g)}\right]_{k,n\neq i,g}$ is a $K\times(N-1)K$ matrix of cross-price elasticities between varieties sold by i and by all other origin countries in market j. $\mathbf{\Omega}_i\equiv\left[\omega_{ni,g}\right]_{n,g}$ is a $(N-1)K\times 1$ vector of inverse export supply elasticities associated with domestic market i. To invert the system specified by Equation K.17 we can use our result (from Appendix E) that $\mathbf{E}_{ij}$ is non-singular, which yields the following formulation for $\mathbf{X}^*_{ij} = \left[\chi^*_{ij,k}\right]_k$:

$$\mathbf{X}^*_{ij} = -\mathbf{E}_{ij}^{-1}\left[\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1}_{(N-1)K} + \mathbf{\Omega}_i\right) + \mathbf{\Lambda}_{ij}\right]. \tag{K.18}$$

Now, we can recover the optimal tax/subsidy rates from the optimal price wedges implies by Equations K.15 and K.18. Specifically, noting that

$$1 + t^*_{ji,k} = \frac{\tilde{P}_{ji,k}}{P_{ji,k}}; \qquad 1 + s^*_{i,k} = \frac{\tilde{P}^*_{ii,k}}{P_{ii,k}}; \qquad 1 + x_{ij,k} = \frac{\tilde{P}^*_{ij,k}/P_{ij,k}}{\tilde{P}^*_{ii,k}/P_{ii,k}};$$

country i's unilaterally optimal tax schedule can be expressed as follows:

$$\begin{aligned} &[\text{domestic subsidy}] \qquad 1 + s^*_{i,k} = 1 + \mu_k \\ &[\text{import tariff}] \qquad 1 + t^*_{ji,k} = (1 + \omega_{ji,k})(1 + \bar{\tau}_i) \\ &[\text{export subsidy}] \qquad \mathbf{1} + \mathbf{x}^*_{ij} = -\mathbf{E}_{ij}^{-1}\left[\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1} + \mathbf{\Omega}_i\right) + \mathbf{\Lambda}_{ij}\right](1 + \bar{\tau}_i). \end{aligned}$$

The last step, is to invoke the multiplicity of optimal tax schedules provided by Lemma 9. Given the multiplicity of optimal import tax and export subsidies, the uniform trade tax-shifter, $\bar{\tau}_i$, is redundant. Following Lemma 1, any tax schedule that satisfies $1 + \tilde{t}^*_{ji,k} = \left(1 + t^*_{ji,k}\right)\frac{1+\bar{t}_i}{1+\bar{\tau}_i}$ and $1 + \tilde{x}^*_{ij,k} = \left(1 + x^*_{ij,k}\right)\frac{1+\bar{t}_i}{1+\bar{\tau}_i}$, and where $1 + \bar{t}_i \in \mathbb{R}_+$, is also optimal. As such the exact value assigned to $\bar{\tau}_i$ is redundant. This explains why we did not unpack the term $\bar{\tau}_i$ in Step #3. There is also another dimension of multiplicity whereby any uniform shift in final good production subsidies (paired with a proportional adjustments final good import tariffs and export subsidies) preserves the equilibrium. Accounting for both dimensions of multiplicity, the optimal policy schedule is given by:

$$\begin{aligned} &[\text{domestic subsidy}] \qquad 1 + s^*_{i,k} = (1 + \mu_k)\left(1 + \bar{s}_i^{\mathscr{C}}\right)^{-1} \\ &[\text{import tax}] \qquad 1 + t^*_{ji,k} = (1 + \omega_{ji,k})\left(1 + \bar{t}_i\right)\left(1 + \bar{s}_i^{\mathscr{C}}\right) \\ &[\text{export subsidy}] \qquad \mathbf{1} + \mathbf{x}^*_{ij} = -\mathbf{E}_{ij}^{-1}\left[\mathbf{E}_{ij}^{(-ij)}\left(\mathbf{1} + \mathbf{t}^*_i\right) + \mathbf{\Lambda}_{ij}(1 + \bar{t}_i)(1 + \bar{s}_i^{\mathscr{C}})\right], \end{aligned}$$

where $\bar{s}_i^{\mathscr{C}}$ is an arbitrary tax shifter that assumes a positive value if the taxed item is a final good and zero otherwise. Also, recall that the elements of $\mathbf{\Lambda}_{ij} \equiv \left[\Lambda_{ij,k}\right]_k$ correspond to the fraction of good ij,k that is reimported via the IO network.

L Tension Between Allocative Efficiency and ToT

This appendix provides a formal proof for two intermediate claims that motivate the conjectures in presented in Section III.

Claim #1: If $Cov\left(\sigma_k, \mu_k\right) < 0$, piecemeal trade policy interventions that seek to improve the terms-of-trade (relative to Laissez-Faire) worsen misallocation and vice versa.

Claim #2: If $Cov\left(\sigma_k, \mu_k\right) < 0$, a unilateral implementation of efficient industrial subsidies in country i (without reciprocity by partners) worsens the terms-of-trade.

Proof of Claim #1. Suppose country i is initially in an equilibrium where the government has implemented a uniform (possibly zero) import tariff or export tax. If existing taxes are zero then the economy is essentially operating under Laissez-Faire. Our goal is to prove the following: If $Cov\left(\mu_k, \sigma_k\right) < 0$, then any adjustment to trade policy that seeks to improve allocative efficiency (relative to the baseline equilibrium) worsens country i's terms of trade (ToT) and *vice versa*. Since a uniform import tariff is equivalent to a uniform export tax by the Lerner symmetry, we can without loss of generality focus on the case where the initial trade policy consists of a (possibly zero) uniform export tax. We first present our proof for the case of restricted entry, but extend it later to account for free entry. To economize on the notation, we hereafter use $1 + \tilde{x} \sim (1 + x)^{-1}$ to denote the export tax counterpart of export subsidy, x.[32]

Welfare Accounting under Piecemeal Policy Change—As an intermediate step, we first characterize the change in welfare in response to a piecemeal trade policy change, decomposing the welfare impacts into changes in allocative efficiency and terms of trade. When preferences are homothetic, the change in country i's welfare in response to an adjustment to export taxes, $\left\{\mathrm{d}\ln\left(1+\tilde{x}_{i,k}\right)\right\}_k$, is the sum of corresponding income and price effects.[33] Namely,

$$\mathrm{d}\ln W_i = \mathrm{d}\ln Y_i - \sum_k \sum_n \lambda_{ni,k} e_{i,k} \mathrm{d}\ln \tilde{P}_{ni,k}. \tag{L.1}$$

To formalize the tension between allocative efficiency and terms of trade succinctly, suppose that country i is sufficiently small so that its piecemeal trade policy reform has a negligible impact on relative wages and labor allocations in the rest of the world. In that case, $\mathrm{d}\ln\tilde{P}_{ii,k} = \mathrm{d}\ln w_i$ and $\mathrm{d}\ln\tilde{P}_{ni,k} \approx 0$ for all $n \neq i$. Under theses assumptions, the price effects in Equation L.1 reduce to $\sum_k \sum_n \lambda_{ni,k} e_{i,k} \mathrm{d}\ln\tilde{P}_{ni,k} = \lambda_{ii}\mathrm{d}\ln w_i$, where $\lambda_{ii} \equiv \sum_k \lambda_{ii,k} e_{i,k}$ denotes the aggregate domestic expenditure share in country i. Next we characterize income effects, $\mathrm{d}\ln Y_i$. For this, note that nominal income in country i is the sum of wage income, profits, and net revenues associated with export tax. In particular,

$$Y_i = \sum_k \left[\left(1+\mu_k\right)\rho_{i,k}\right] w_i L_i + \sum_{n\neq i}\sum_k \left[\tilde{x}_{i,k} P_{in,k} Q_{in,k}\right],$$

where $\tilde{x}_{i,k}$ is the export tax on industry k goods, which is uniform in the baseline equilibrium (i.e., $\tilde{x}_{i,k} = \bar{x}_i$) but is subsequently adjusted to improve allocative efficiency. Taking full derivatives of the above expression yields

$$\mathrm{d}\ln Y_i = \left(1-\pi_i^{\mathscr{X}}\right)\left[\mathrm{d}\ln\left(\sum_k \left(1+\mu_k\right)\rho_{i,k}\right) + \mathrm{d}\ln w_i\right]$$
$$+\pi_i^{\mathscr{X}} \sum_{n\neq i}\sum_k \left[\frac{\bar{x}_i P_{in,k} Q_{in,k}}{\sum_{n',k'} \bar{x}_i P_{in',k'} Q_{in',k'}}\left(\frac{\partial\ln P_{in,k}}{\partial\ln w_i}\mathrm{d}\ln w_i + \frac{\partial\ln Q_{in,k}}{\partial\ln\tilde{P}_{in,k}}\mathrm{d}\ln\left(1+\tilde{x}_{i,k}\right)\right)\right],$$

where $\pi_i^{\mathscr{X}} \equiv \sum_{n\neq i}\sum_k \left[\tilde{x}_{i,k} P_{in,k} Q_{in,k}\right]/Y_i$ denotes the share of export tax revenues in total revenues. One can immediately verify that the change in the (employment-weighted) aggregate profit margin is

$$\mathrm{d}\ln\left(\sum_k \left(1+\mu_k\right)\rho_{i,k}\right) = \sum_k \left[\rho_{i,k}\cdot\left(\frac{1+\mu_k}{1+\overline{\mu}_i}\right)\mathrm{d}\ln\rho_{i,k}\right],$$

where $\overline{\mu}_i \equiv \sum_k \left[\mu_k \rho_{i,k}\right]$, recall, is our short-hand notation for the aggregate profit margin in country i. Notice that $\partial\ln P_{in,k}/\partial\ln w_i = 1$ under restricted entry and $\partial\ln Q_{in,k}/\partial\ln\tilde{P}_{in,k} = -\sigma_k$, since country i is a small open economy. Invoking these points, we can simplify the expression for $\mathrm{d}\ln Y_i$ as

$$\mathrm{d}\ln Y_i = \mathrm{d}\ln w_i + \left(1-\pi_i^{\mathscr{X}}\right)\sum_k\left[\rho_{i,k}\cdot\left(\frac{1+\mu_k}{1+\overline{\mu}_i}\right)\mathrm{d}\ln\rho_{i,k}\right] - \pi_i^{\mathscr{X}}\sum_k\left[\chi_{i,k}\cdot\sigma_k\,\mathrm{d}\ln\left(1+\tilde{x}_{i,k}\right)\right], \tag{L.2}$$

where $\chi_{i,k} \equiv P_{in,k}Q_{in,k}/\sum_{n',k'}\left[P_{in',k'}Q_{in',k'}\right]$ denotes the share of industry k goods in country i's export revenues. To economize on the notation we henceforth use $\mathbb{E}_\omega\left[.\right]$ and $Cov_\omega\left(.\right)$ to denote the cross-industry mean and covariance with weights, $\left\{\omega_{i,k}\right\}_k$, that satisfy $\sum_k \omega_{i,k} = 1$. Considering this choice of notation, suppose the piecemeal trade policy reform is mean-preserving—i.e., $\mathbb{E}_\chi\left[\mathrm{d}\ln(1+\tilde{x}_{i,k})\right] \sim \sum_k\left[\chi_{i,k}\mathrm{d}\ln\left(1+\tilde{x}_{i,k}\right)\right] = 0$. To put it verbally, the tax reform raises export taxes on some industries and lowers it on others, while preserving the sales-weighted average export tax rate. Under this presupposition, the last term in the above equation amounts to the covariance between σ_k and $\mathrm{d}\ln\left(1+\tilde{x}_{i,k}\right)$.

[32]Note that a positive export tax is akin to an negative export subsidy and vice versa—i.e., $x < 0 \implies \tilde{x} > 0$.

[33]To be clear, we assume that the baseline export tax policy does not discriminate between destination markets, with $x_{i,k}$ denoting the export tax applied to all export good in industry k.

In particular,

$$\sum_k \left[\chi_{i,k} \cdot \sigma_k \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right] \sim \mathbb{E}_\chi\left[\sigma_k \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right]$$
$$= \mathbb{E}_\chi\left[\sigma_k \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right] - \mathbb{E}_\chi\left[\sigma_k\right] \cdot \underbrace{\mathbb{E}_\chi\left[\text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right]}_{=0} \sim Cov_\chi\left(\sigma_k\, , \, \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right).$$

As a matter of accounting, $\sum_k \rho_{i,k} \text{d}\ln \rho_{i,k} = 0$, so the same logic implies that the second term on the right-hand side of Equation L.2 is also the covariance between the change in industry-level employment share, $\text{d}\ln\rho_{i,k}$, and the industry-level markup (relative to the mean):

$$\sum_k \left[\rho_{i,k}\left(\frac{1+\mu_k}{1+\bar{\mu}_i}\right)\text{d}\ln\rho_{i,k}\right] \sim Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i}\, , \, \text{d}\ln\rho_{i,k}\right).$$

Substituting for the above expressions in Equation L.2 and plugging the simplified expression for $\text{d}\ln Y_i$ back into our original welfare formula (Equation L.1) delivers,

$$\text{d}\ln W_i = \underbrace{\left(1-\pi_i^{\mathscr{X}}\right) Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i}, \text{d}\ln\rho_{i,k}\right)}_{\text{Allocative Efficiency}} + \overbrace{\left[-\pi_i^{\mathscr{X}} Cov_\chi\left(\sigma_k, \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right) + \left(1-\lambda_{ii}\right)\text{d}\ln w_i\right]}^{\text{Terms of Trade}}, \tag{L.3}$$

echoing the welfare decomposition provided by Baqaee and Farhi (2019) and Atkin and Donaldson (2021). To offer intuition for this choice of decomposition, the term labeled *Allocative Efficiency* is analogous to the deviation form Hulten's (1978) formula in an inefficient closed economy. More specifically suppose country i was a closed economy hit with a labor productivity shock, $\text{d}\ln A_{i,k}$. Following the same steps as above, the welfare impact of this shock can be decomposed as

$$\text{d}\ln W_i^{\text{closed}} = \overbrace{\sum\left[\rho_{i,k}\text{d}\ln A_{i,k}\right]}^{\text{Hulten}} + Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i}, d\ln\rho_{i,k}\right),$$

where the deviation from Hulten (1978) reflects changes to allocative efficiency. Likewise, *Terms of Trade* effects in Equation L.3 are analogous to deviations from Hulten (1978) if country i were open to trade but efficient. In accordance with this logic, the *Terms of Trade* effects in Equation L.3 disappear if Country i is closed, in which case $\pi_i^{\mathscr{X}} = 1 - \lambda_{ii} = 0$. Relatedly, *Allocative Efficiency* effects disappear if the economy is efficient, in which case $\left(1+\mu_k\right) / \left(1+\bar{\mu}_i\right) = 1$.

Tension between Allocative Efficiency & Terms of Trade—An export policy shock, $\left\{\text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right\}_k$, that seeks to improve allocative efficiency must reallocate workers from low- to high-μ industries so that $Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i}, \text{d}\ln\rho_{i,k}\right) > 0$. If demand is elastic and well-behaved, this type of reallocation requires that industry-level export tax reductions to be positively correlated with markups, i.e., $Cov_\chi\left(\frac{1+\mu_k}{1+\bar{\mu}_i}, \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right)$ 0. Accordingly, if $Cov\left(\sigma_k, \mu_k\right) < 0$, the export tax changes will be positively correlated with the trade elasticity, $Cov_\chi\left(\sigma_k, \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right) > 0$. As such, an export tax reform that improves *Allocative Efficiency* (relative to the status quo) worsens the *Terms of Trade* through the term $Cov_\chi\left(\sigma_k, \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right)$. Now, consider the remaining *Terms of Trade* term that accounts for general equilibrium wage effects. Considering that $Cov_\chi\left(\sigma_k, \text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right) > 0$, the desired export tax alteration consists of raising taxes in high-trade elasticity industries (where export sales are more-sensitive to tax hikes) paired with an proportional tax reduction in low-trade elasticity industries (where export sales are less-responsive to tax cuts). This alternation will, by design, lower overall export sales by country i—resonating with the conventional Ramsey rule. The reduction in export sales will in turn deflate demand for country i's labor and its wage rate relative to rest of the world (i.e., $\text{d}\ln w_i < 0$). That is, the second *Terms of Trade* term is also negative when the export tax reform attempts to improve *Allocative Efficiency* relative to the status quo. To take stock: Suppose $Cov\left(\sigma_k, \mu_k\right) < 0$ and country i is initially in a equilibrium involving uniform (or zero) export/import taxes. In that case, improving *Allocative Efficiency* via piecemeal trade policy adjustments, $\left\{\text{d}\ln\left(1+\tilde{x}_{i,k}\right)\right\}$, coincides with a worsening of the terms of trade.

Second-Best Trade Policies are Industry-Blind in Krugman (1980)—When preferences are Cobb-Douglas across industries, country i's second-best trade policy is given by (see Section II):

$$\text{[2nd-best import tariff]} \qquad 1+t_{ji,k}^{**}=\frac{1+\left(\sigma_k-1\right)\lambda_{ii,k}}{1+\frac{1+\bar{\mu}_i}{1+\mu_k}\left(\sigma_k-1\right)\lambda_{ii,k}}\left(1+\omega_{ji,k}\right)\left(1+\bar{t}_i\right)$$

$$\text{[2nd-best export subsidy]} \qquad 1+x_{ij,k}^{**}=\frac{1+\mu_k}{1+\bar{\mu}_i}\left(\frac{\left(\sigma_k-1\right)\left(1-\lambda_{ij,k}\right)}{1+\left(\sigma_k-1\right)\left(1-\lambda_{ij,k}\right)}\right)\left(1+\bar{t}_i\right),$$

where $\bar{t}_i$ us a uniform trade tax shifter that accounts for the multiplicity of optimal policy schedules (see Lemma 1). Following Alvarez and Lucas (2007), if country i is a small open economy, then $\lambda_{ij,k}=\lambda_{ii,k}=\omega_{ji,k}\to 0$. Moreover, if we assume that the firm- and country-level degrees of market power are identical à la Krugman (1980), we have $1+\mu_k=\frac{\sigma_k}{\sigma_k-1}$. Consolidating these two points, we get the following formula for the 2nd-best trade policy of a small open economy in the multi-industry Krugman (1980) model:

$$1+t_{ji,k}^{**}=1+\bar{t}_i; \qquad\qquad 1+x_{ij,k}^{**}=\frac{\sigma_k}{\sigma_k-1}\left(\frac{\sigma_k-1}{\sigma_k}\right)\frac{1+\bar{t}_i}{1+\bar{\mu}_i}=\left(1+\bar{t}_i\right)\left(1-\frac{1}{\bar{\sigma}_i}\right),$$

where $\frac{1}{\bar{\sigma}_i}=\sum_k \rho_{i,k}\frac{1}{\sigma_k}$ is the sales-weighted average (inverse) trade elasticity. Evidently, the optimal 2nd-best trade policy consists of a uniform import tariff or export subsidy, which is blind to inter-industry misallocation and industry-level export market power. The logic is that any attempt at exploiting industry-level export market power exacerbates inter-industry misallocation and *vice versa*—leaving the government with no choice but to abandon these targeted policy aspirations. This consideration leads to *industry-blind* optimal trade taxes that solely manipulate the relative aggregate wage (w_i/w_{-i}) in country i's favor, with minimal reshuffling of resources across industries.

Proof of Claim #2. Suppose country i is initially operating under Laissez-Faire. The government has, moreover, agreed (under a shallow treaty) to limit itself to the efficient or cooperative policy choice specified in Section II. Our goal is to show that a unilateral implementation of efficient industrial subsidies by country i (without reciprocity by partners) causes a deterioration of country i's terms of trade (ToT) and even immiserizing growth. To this end, we build on the welfare accounting formulas derived earlier. In particular, the welfare impacts of unilateral markup-correcting subsidies by country i (i.e., $\mathbf{s}_i=\boldsymbol{\mu}$) can be written as

$$\Delta W_i=\int_{\mathbf{s}_i=\mathbf{0}}^{\boldsymbol{\mu}} \mathrm{d}\ln W_i,$$

where the change in welfare in response to modest industrial policy adjustments can, as earlier, be decomposed as

$$\mathrm{d}\ln W_i=\underbrace{\left(1-\pi_i^{\mathscr{S}}\right)Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i},\mathrm{d}\ln\rho_{i,k}\right)}_{\text{Allocative Efficiency}}+\overbrace{\left[-\pi_i^{\mathscr{S}}Cov_\chi\left(\sigma_k,\mathrm{d}\ln\left(1+s_{i,k}\right)\right)\;+\;\left(1-\lambda_{ii}\right)\mathrm{d}\ln w_i\right]}^{\text{Terms of Trade}}.$$

The above equation can be derived analogously to Equation L.3, with $\pi_i^{\mathscr{S}}$ denoting the share of production tax (or subsidy) revenues that are collected from foreign consumers. Intuitively, the fraction of tax revenues collected from domestic consumers deliver income gains that are exactly offset by the corresponding loss from price increases. Hence, the domestically-borne fraction of the tax revenue does not contribute to welfare changes, $\mathrm{d}\ln W_i$, beyond general equilibrium impacts on inter-industry labor allocation and wages.

Considering this background, we can specify the sign of the *Allocative Efficiency* and *Terms of Trade* effects (in the above equation) in response to markup/scale-correcting subsidies. The efficient policy $(s_{i,k}=0\to s'_{i,k}=\mu_k)$ subsidizes output in high-μ industries, relocates labor from the rest of the economy to these industries, and thereby improves allocative efficiency, i.e., $Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i},\mathrm{d}\ln\rho_{i,k}\right)$.

But following the logic presented earlier: If $Cov\left(\sigma_k, \mu_k\right) < 0$, it is the case that along the path of restoring marginal cost pricing, $Cov_\chi\left(\sigma_k, \mathrm{d}\ln\left(1+s_{i,k}\right)\right) > 0$ and $\mathrm{d}\ln w_i < 0$—both of which contribute to a deterioration of the ToT. In some cases, like the numerical example presented in Section **??**, the deterioration of the ToT is larger than the corresponding allocative efficiency gains—leading to immiserizing growth in country i.

Adverse Firm-Delocation Effects when Country i is Large—When country i is excessively large, a unilateral adoption of corrective industrial policies worsens its ToT through an additional channel: *firm-delocation* effects. To make this point, let $v_{i,k} \equiv \frac{(1-\lambda_{ii,k})e_{i,k}}{(1-\lambda_{ii})}$ denote the import share pertaining to industry k. Suppose without loss of generality that $\mathbb{E}_v\left[\mu_k\right] = 0$—since following Lemma 1, we can recalibrate the level of markups and wages in the rest of the world without changing welfare. Our welfare decomposition, in that case, takes the following form, internalizing the effect of country i's policy on entry and labor shares in the rest of the world:

$$\mathrm{d}\ln W_i = \left(1-\pi_i^{\mathscr{S}}\right) Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i}, \mathrm{d}\ln\rho_{i,k}\right) + \\ -\pi_i^{\mathscr{S}} Cov_\chi\left(\sigma_k, \mathrm{d}\ln\left(1+s_{i,k}\right)\right) + \left(1-\lambda_{ii}\right)\left[\mathrm{d}\ln w_i + Cov_v\left(\mu_k, \mathrm{d}\ln\rho_{-i,k}\right)\right].$$

Country i's corrective subsidies, by design, relocate labor to high-μ_k industries in the local economy, but have the opposite effect on labor allocation in the rest of the world. Put formally, $Cov_\rho\left(\frac{1+\mu_k}{1+\bar{\mu}_i}, \mathrm{d}\ln\rho_{i,k}\right) > 0$, while $Cov_v\left(\mu_k, \mathrm{d}\ln\rho_{-i,k}\right) < 0$. Improving allocative efficiency with corrective subsidies, therefore, worsens the ToT through an additional term $Cov_v\left(\mu_k, \mathrm{d}\ln\rho_{-i,k}\right)$, which represents *firm-delocation* effects. The logic is that promoting output and entry in high-μ domestic industries, diminishes output and entry in high-μ foreign industries—hence, the term *firm-delocation*. The reduction in foreign firm-level varieties inflates the price of imports $P_{-ii,k} \propto M_{-i,k}^{-\mu_k}$, thereby worsening country i's ToT.

M Optimal Import Tariff under IO Linkages: *Simple Case*

This appendix elucidates the result that optimal–import–tariff–formulas–are–IO-blind using a simpler case of our model without scale economies or markups (i.e., $\mu_k = 0$ for all k). To this end, we re-derive the first-order conditions associated with the (1st-best) optimal policy problem for this special case, focusing specifically on optimal import taxes. The optimal policy problem can be specified as

$$\max_{\tilde{\mathbb{P}}_i} \quad W_i\left(\tilde{\mathbb{P}}_i, \mathbf{w}\right) \equiv V_i\left(Y_i, \tilde{\mathbf{P}}_i\right),$$

where $\tilde{\mathbf{P}}_i \equiv \left\{\tilde{\mathbf{P}}_{ni}, \tilde{\mathbf{P}}_{ii}\right\}_{n\neq i}$, $\tilde{\mathbb{P}}_i \equiv \left\{\tilde{\mathbf{P}}_i, \tilde{\mathbf{P}}_{in}\right\}_{n\neq i}$, and income is the sum of wage income and taxes revenues:

$$Y_i = w_i L_i + \sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni}\right)\cdot\mathbf{Q}_{ni}\right] + \sum_{\ell}\left[\left(\tilde{\mathbf{P}}_{i\ell} - \mathbf{P}_{i\ell}\right)\cdot\mathbf{Q}_{i\ell}\right].$$

With IO linkages, producer prices in the rest of the world depend on the after-export-tax price of intermediate inputs supplied by country i. Likewise, producer prices in country i depend on the after-import-tax price of intermediate inputs sourced from abroad. Namely,

$$\mathbf{P}_{ni} \sim \mathbf{P}_{ni}\left(\mathbf{w}, \left\{\tilde{\mathbf{P}}_{i\ell}\right\}_{n\neq i}\right) \qquad\qquad \mathbf{P}_{i\ell} \sim \mathbf{P}_{i\ell}\left(w_i, \left\{\tilde{\mathbf{P}}_{ni}\right\}_{n}\right).$$

The solution to optimal policy problem identifies a vector of optimal after-tax prices, $\tilde{\mathbb{P}}_i^*$, which implicitly determine the optimal tax rates. For instance, the optimal choice w.r.t. tax-inclusive import prices ($\tilde{P}_{ji,k}^* \in \tilde{\mathbb{P}}_i^*$) implicitly determines the optimal import tariff as $1 + t_{ji,k}^* \equiv \frac{\tilde{P}_{ji,k}^*}{P_{ji,k}}$ for all (ji,k). Invoking our *wage neutrality* result, the optimal (after-tax) price of goods imported from country j ($\tilde{P}_{ji,k} \in \tilde{\mathbb{P}}_i$)

satisfies the following first-order condition:[34]

$$\frac{\partial W_i\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)}{\partial\ln\tilde{P}_{ji,k}}=\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial\ln\tilde{P}_{ji,k}}+\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial Y_i}\frac{\partial Y_i\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)}{\partial\ln\tilde{P}_{ji,k}}$$
$$=\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial\ln\tilde{P}_{ji,k}}+\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial Y_i}\left\{\tilde{P}_{ji,k}Q_{ji,k}+\sum_n\left[\left(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni}\right)\cdot\frac{\partial\mathbf{Q}_{ni}}{\partial\ln\tilde{P}_{ji,k}}\right]-\sum_\ell\left[\frac{\partial\mathbf{P}_{i\ell}}{\partial\ln\tilde{P}_{ji,k}}\cdot\mathbf{Q}_{i\ell}\right]\right\}=0$$

We can simplify the above equation by appealing to Roy's identity and Shephard's lemma, whereby

$$\text{[Roy's identity]}\qquad\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial\ln\tilde{P}_{ji,k}}=-\tilde{P}_{ji,k}Q^{\mathcal{C}}_{ji,k}$$
$$\text{[Shephard's lemma]}\qquad\sum_\ell\frac{\partial\mathbf{P}_{i\ell}}{\partial\ln\tilde{P}_{ji,k}}\cdot\mathbf{Q}_{i\ell}=\tilde{P}_{ji,k}Q^{\mathcal{I}}_{ji,k}.$$

Recall that $Q^{\mathcal{C}}_{ji,k}$ and $Q^{\mathcal{I}}_{ji,k}$ denote final good and intermediate input quantities which, together, determine total quantity, $Q_{ji,k}=Q^{\mathcal{C}}_{ji,k}+Q^{\mathcal{I}}_{ji,k}$. Plugging the above envelope identities into the first-order condition derived earlier, yields

$$\frac{\partial W_i\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)}{\partial\ln\tilde{P}_{ji,k}}=\sum_n\left[\left(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni}\right)\cdot\frac{\partial\mathbf{Q}_{ni}}{\partial\ln\tilde{P}_{ji,k}}\right]=0,$$

the solution to which implies $\tilde{\mathbf{P}}^*_{ni}=\mathbf{P}_{ni}$ $(\forall n)$, indicating that optimal tariffs,$1+t^*_{ni,k}\equiv\tilde{P}^*_{ni,k}/P_{ni,k}$, are either zero (or uniform based on the Lerner symmetry) and, thus, blind to input-output linkages. The logic is that, fixing export prices at their optimal level, $\tilde{\mathbf{P}}^*_{i\ell}\in\tilde{\mathbb{P}}_i$, import tariffs do not influence prices in the rest of the world via input output linkages. More formally,

$$\frac{\partial\mathbf{P}_{ni}\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)}{\partial\tilde{P}_{ji,k}}\sim\left(\frac{\partial\mathbf{P}_{ni}}{\partial\tilde{P}_{ji,k}}\right)_{\left\{\tilde{\mathbf{P}}_{i\ell}\right\}_{\ell\neq i},\mathbf{w}}=0\qquad\left(\forall n,j\neq i\right).$$

This logic, however, holds only *if* export tax-cum-subsidies are part of the government's policy artillery. To see this, consider the 2nd-best case where the government can only set import taxes and is unable to control the export prices via direct export policy measures. In that case, an import tariff on inputs is partially passed onto export prices via the IO network, influencing producer prices in the rest of the world. In particular,

$$\tilde{\mathbf{P}}_{in}\notin\tilde{\mathbb{P}}_i\qquad\Longrightarrow\qquad\frac{\partial\mathbf{P}_{ni}\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)}{\partial\tilde{P}_{ji,k}}\neq0.$$

Accordingly, the first-order condition *w.r.t.* $\tilde{P}_{ji,k}\in\tilde{\mathbb{P}}_i$ must be updated to account for tariff-reexportation, delivering

$$\frac{\partial W_i\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)}{\partial\ln\tilde{P}_{ji,k}}=\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial\ln\tilde{P}_{ji,k}}+\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial Y_i}\frac{\partial Y_i\left(\tilde{\mathbb{P}}_i,\mathbf{w}\right)}{\partial\ln\tilde{P}_{ji,k}}$$
$$=\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial\ln\tilde{P}_{ji,k}}+\frac{\partial V_i\left(Y_i,\tilde{\mathbf{P}}_i\right)}{\partial Y_i}\left\{\tilde{P}_{ji,k}Q_{ji,k}+\sum_{n\neq i}\left[\left(\tilde{\mathbf{P}}_{ni}-\mathbf{P}_{ni}\right)\cdot\frac{\partial\mathbf{Q}_{ni}}{\partial\ln\tilde{P}_{ji,k}}-\underbrace{\frac{\partial\mathbf{P}_{ni}}{\partial\ln\tilde{P}_{ji,k}}\cdot\mathbf{Q}_{ni}}_{\text{tariff re-exportation}}\right]-\sum_\ell\left[\frac{\partial\mathbf{P}_{i\ell}}{\partial\ln\tilde{P}_{ji,k}}\cdot\mathbf{Q}_{i\ell}\right]\right\}=0.$$

[34] One can verify wage neutrality in this particular setting as follows: First, note that when the government can control the after-tax price of all domestically-produced goods, wage perturbations are welfare-neutral conditional on domestic demand quantities. Namely, $\left(\frac{\partial W_i}{\partial\ln w_i}\right)_{\mathbf{Q}_i}=w_iL_i-\sum_{\ell,k}\frac{\partial\ln P_{i\ell,k}}{\ln w_i}P_{i\ell,k}Q_{i\ell,k}=0$, where the last equality follows from Shephard's lemma. Second, policy-led wage perturbations can influence welfare by raising income and modifying the domestic demand schedule,$\mathbf{Q}_i\equiv\left\{Q_{ni,k}\right\}_{n,k}$; but these effects are already encompassed by the term $\frac{\partial\mathbf{Q}_{ni}}{\partial\ln\tilde{P}_{ji,k}}$. And as our following derivation reveals, these wage-driven income effects turn our to be redundant in the neighborhood of the optimal policy solution, $\tilde{\mathbb{P}}^*_i$.

Simplifying the updated first-order condition using Roys identity and Shephard's lemma, yields

$$\sum_{n \neq i} \left[\left(\tilde{\mathbf{P}}_{ni} - \mathbf{P}_{ni} \right) \cdot \frac{\partial \mathbf{Q}_{ni}}{\partial \ln \tilde{P}_{ji,k}} - \frac{\partial \mathbf{P}_{ni}}{\partial \ln \tilde{P}_{ji,k}} \cdot \mathbf{Q}_{ni} \right] = 0,$$

indicating that 2nd-best optimal tariffs ($1 + t^*_{ni,k} \equiv \tilde{P}^*_{ni,k} / P_{ni,k}$) are non-uniform, with the optimal tariff formula explicitly depending on input-output shares in the rest of the world through the term $\partial \mathbf{P}_{ni} / \partial \ln \tilde{P}_{ji,k}$. The logic can be explained as follows echoing our previous claim: Absent the government's ability to set export taxes/subsidies, import tariffs on intermediate inputs are (partially) passed onto prices in the rest of world via input-output connections, acting as an indirect export tax. The optimal tariff formula, thus, internalizes the ToT benefits of these indirect export taxation effects. However, when the government can directly apply export taxes/subsidies, any such benefits are already internalize by the optimal export tax-cum-subsidy choice, rendering the optimal tariff formula IO-blind.

N Data Appendix

N.1 Transaction-level Colombian import data from DATAMYNE INC

DATAMYNE Inc provides transaction-level trade records from 43 countries including United States, Latin America, Asia and several European Union Member States (website: datamyne.com). Their data is obtained from customs authorities and government agencies. Access to the transaction-level Colombian trade data was purchased from DATAMYNE INC. in May 2014. The data were available for manual online download in segments of 5,000 observations per download. Our final dataset covers the universe of import transactions during 2007-2013 and contains more than 17 million observations. The data include detailed information about each transaction, such as the Harmonized System 10-digit product category (HS10), country of origin, importing and exporting firm IDs, quantity, f.o.b. (free on board), and c.i.f. (customs, insurance, and freight) transaction values, freight, insurance, and value-added tax in US dollars. As a unique feature, our data reports the identities of all foreign firms exporting to Colombia. This feature allows us to define import varieties as firm-product combinations rather than country-product combinations, which is the standard approach. HS product codes were updated by the Colombian Statistical Agency (DANE) during the 2007-2013 time period, and we use correlational tables by DANE to match product codes over time (DANE (Colombian Statistical Agency), 2017). Table N.1 reports a summary of basic trade statistics in the DATAMYNE data.

Table N.1: *Summary Statistics of the Colombian Import Data.*

	Year						
Statistic	2007	2008	2009	2010	2011	2012	2013
F.O.B. value (billion dollars)	30.69	37.25	31.17	38.47	52.00	55.74	56.90
$\frac{\text{C.I.F. value}}{\text{F.O.B. value}}$	1.08	1.07	1.05	1.06	1.05	1.05	1.05
$\frac{\text{C.I.F. + tax value}}{\text{F.O.B. value}}$	1.28	1.26	1.24	1.26	1.22	1.22	1.20
No. of exporting countries	210	219	213	216	213	221	224
No. of imported varieties	483,286	480,363	457,000	509,524	594,918	633,008	649,561

Notes: Tax value includes import tariff and value-added tax (VAT). The number of varieties corresponds to the number of country-firm-product combination imported by Colombia in a given year.

N.2 Cleaning data on the identity/name of exporting firms in DATAMYNE INC

Utilizing the information on the identity of the foreign exporting firm is a critical part of our empirical exercise. Unfortunately, the names of the exporting firms in our dataset are not standardized. As a result, there are instances when the same firm is recorded differently due to using or not using the abbreviations, capital and lower-case letters, spaces, dots, other special characters, etc. To standardize the names of the exporting firms, we used the following procedure.

i. We deleted all observations with the missing exporting names and/or zero trade values.

ii. We capitalized firms names and their contact information (which is either email or phone number of the firm).

iii. We eliminated abbreviation "LLC," spaces, parentheses, and other special characters (. , ; / @ ' } - & ") from the firms names.

iv. We eliminated all characters specified in 3. above and a few others (# : FAX) from the contact information.

v. We dropped observations without contact information (such as, "NOTIENE", "NOREPORTA", "NOREGISTRA," etc.), with non-existent phone numbers (e.g., "0000000000", "1234567890", "1"), and with six phone numbers which are used for multiple firms with different names (3218151311, 3218151297, 6676266, 44443866, 3058712865, 3055935515).

vi. Next, we kept only up to first 12 characters in the firm's name and up to first 12 characters in the firm's contact information (which is either email or phone number). In our empirics, we treat all transaction with the same updated name and contact information as coming from the same firm.

vii. We also analyzed all observations with the same contact information, but slightly different name spelling. We only focused on the cases in which there are up to three different variants of the firm name. For these cases, we calculated the Levenshtein distance in the names, which is the smallest number of edits required to match one name to another. We treat all export observations as coming from the same firm if the contact phone number (or email) is the same and the Levenshtein distance is four or less.

N.3 Daily exchange rate data

The database on the exchange rate between different international currencies with the Colombian Peso and the US dollar was complied by manually downloading historical daily exchange rate data from the BANK OF CANADA web portal (Bank of Canada (2017)). The underlying data is sourced from REFINITIV (formerly THOMSON REUTERS). Around 2017, the BANK OF CANADA discontinued publishing daily exchange rates for a wide set of countries. But the historical rates relevant to this project can be downloaded via Bank of Canada's legacy exchange rate portal.

N.4 Penn World Tables

We use PENN WORLD TABLES 9.1 (Feenstra, Inklaar, and Timmer (2021)) to obtains data on national accounts and in particular real GDP for year 2011 to construct Figure S.1. This data is in the public domain and can be downloaded via the Groningen Growth and Development Centre web portal (`https://www.rug.nl/ggdc/productivity/pwt/?lang=en`).

N.5 Exporter Dynamics Database

The Export Dynamics Database (EDD by World Bank, 2016) is based on firm-level customs data covering the universe of export transactions provided by customs agencies from 59 countries (Fernandes, Freund, and Pierola (2016)). For each country, the raw firm-level customs data contains annual export flows (in current values) disaggregated by firm, destination and Harmonized System (HS) 6-digit product. This data can be downloaded from the World Bank micro-data web portal (`https://doi.org/10.48529/agcr-yt74`) subject to a agreeing to the terms of use.

N.6 World Input-Output Database (WIOD)

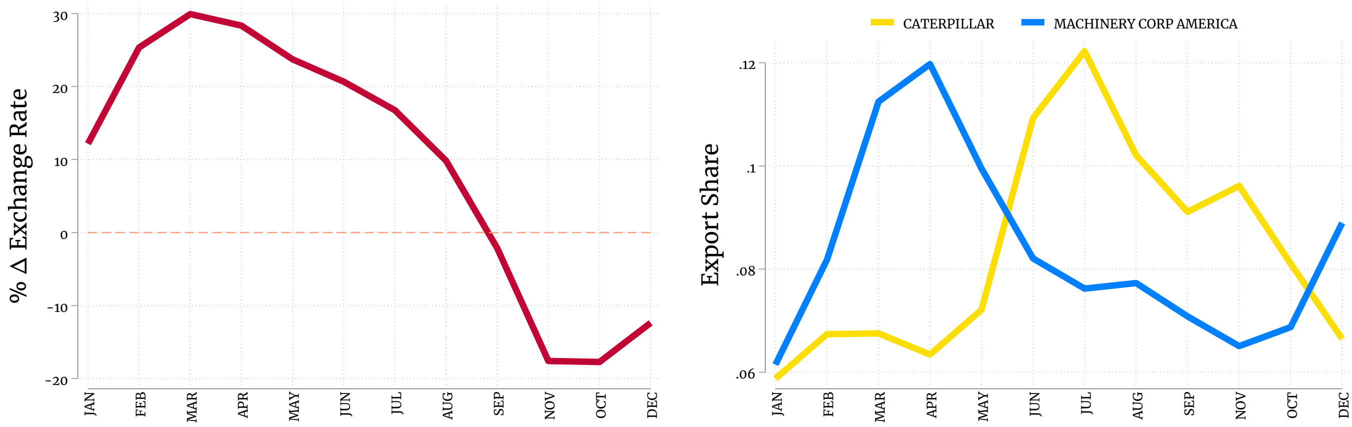
These World Input-Output Database contains data on sector-country flows of goods and services for the year 2014. Our analysis uses the World Input-Output Database 2016 Release (WIOD (2021); Timmer, Dietzenbacher, Los, Stehrer, and De Vries (2015)). The data are in the public domain and can be accessed via the Groningen Growth and Development Centre s web portal: `https://www.rug.nl/ggdc/valuechain/wiod/?lang=en`. We supplement the WIOD data with tariff data, which are compiled and matched to WIOD entries by closely following the guidelines in Kucheryavyy et al. (2023a; 2023b). The concordance between the HS product codes and ISIC rev 3. sector codes are from the World Integrated Trade Solutions (WITS) by the World Bank.

O Illustrative Example for our Instrumental Variable

This section presents an example to elucidate the logic behind our shift-share instrument, presented in Section V. The example compares two major U.S. firms that dominate exports to Colombia in product code HS8431490000 (PARTS AND ATTACHMENTS OTHER FOR DERRICKS). We have chosen this product code because it features two of the biggest exporters to Colombia: "CATERPILLAR" and "MACHINERY CORPORATION OF AMERICA."

The left panel of Figure O.1 shows the year-over-year change in the Peso-to-Dollar exchange rate for each month in 2008. The right panel plots monthly export shares for "CATERPILLAR" and "MACHINERY CORPORATION OF AMERICA" of HS8431490000. Notice that export patterns from "CATERPILLAR" and "MACHINERY CORPORATION OF AMERICA" are markedly different. The former exports primarily in the first half of the fiscal year, while the latter exports primarily in the second half. The prices charged by these two firms are, thus, differentially affected by aggregate exchange rate shocks.

Figure O.1: *Monthly export shares and exposure to aggregate exchange rate shocks*



Notes: The left panel reports the year-over-year change in the Peso-to-Dollar exchange rate for each month in 2009. The right panel reports monthly export sales shares for the two largest US firms serving product code HS8431490000—namely, Caterpillar and Machinery Corp. of America.

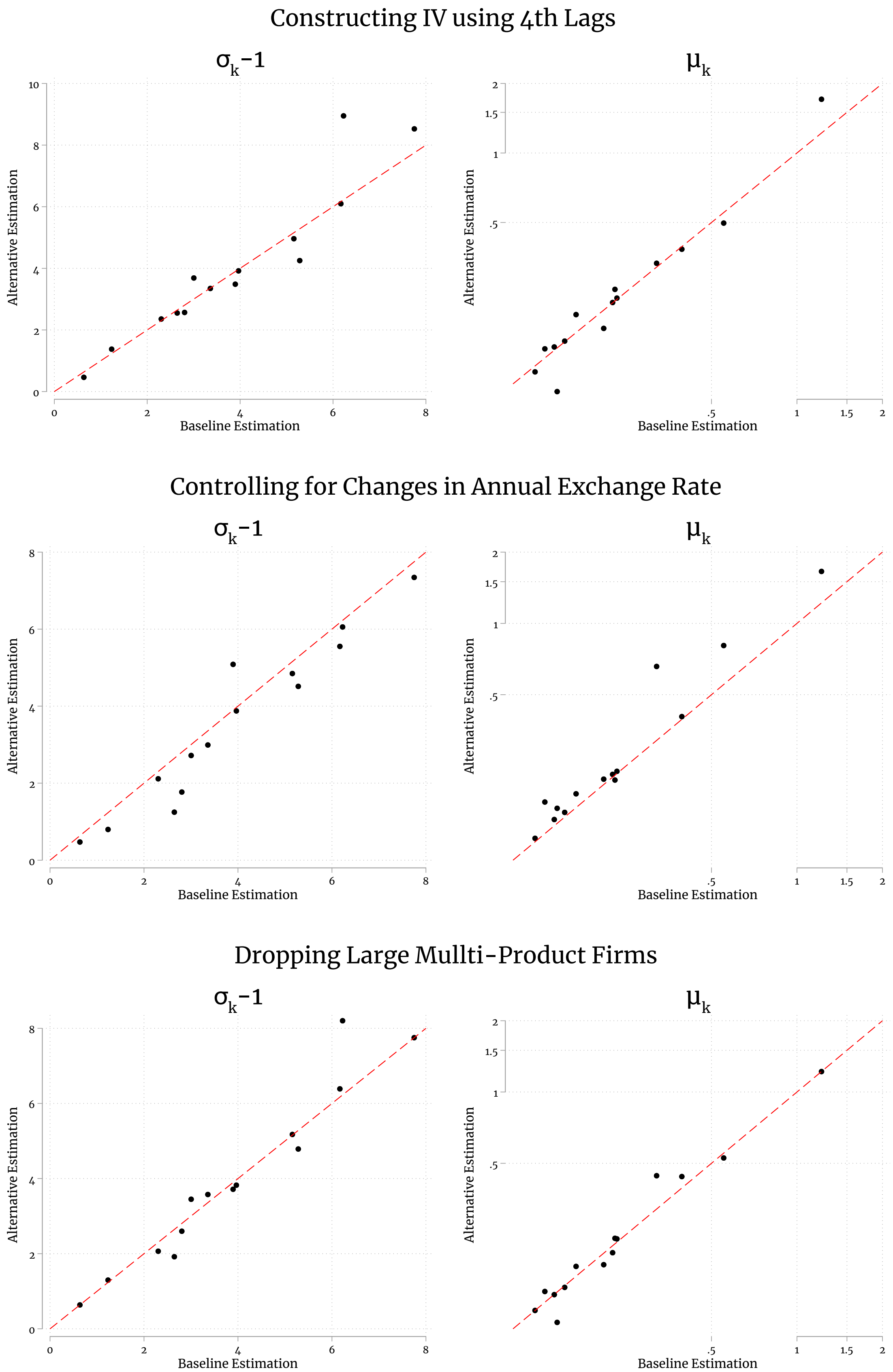
P Robustness Checks: *Import Demand Estimation*

This appendix reports three robustness checks that we described in Section V. The first check addresses the possibility that firms set prices in forward-looking manner. To restate the issue, when there are lags in inventory clearances, firms' optimal pricing decisions may be forward-looking. If true, such price-setting behaviors can violate assumption (a1). To address this concern, we reconstruct our shift-share instrument using 4 lags instead of 1. If inventories clear in at most 4 years, we can deduce that pricing decisions do not internalize expected demand shocks beyond the 4 year mark. As a result, $\mathbb{E}\left[\tilde{p}_{jkt-4}(\omega, m)\,\Delta\ln\varphi_{\omega jkt}\right] = 0$, and this more-stringent instrument will satisfy the exclusion restriction. The *top* panel in Figure P.1 compares the estimated σ_k and μ_k under the new and baseline estimations. Evidently, the ordering and magnitude of the estimated elasticities is rather preserved across industries. More importantly, the new estimation retains the negative correlation between σ_k and μ_k, which is the key assumption in Proposition 1.

The second check addresses the possibility that, in the presence of cross-inventory effects, $\Delta\ln\varphi_{\omega jkt}$ may encompass omitted variables that concern firms' dynamic inventory management decisions. These decisions internalize exchange rate movements, which may violate the identifying assumption (a2), i.e., $\mathbb{E}\left[\Delta\ln\mathcal{E}_{jt}(m)\,\Delta\ln\varphi_{\omega jkt}\right] \neq 0$. To address this concern, we reestimate the firm-level import demand function while directly controlling for changes on the annual exchange rate. In that case, $\mathbb{E}\left[z_{jk,t}(\omega)\,\Delta\ln\varphi_{\omega jkt} \mid \Delta\ln\mathcal{E}_{jt}\right]$, and the exclusion restriction will be satisfied insofar as dynamic demand optimization is a concern. The *middle* panel in Figure P.1 compares the estimated σ_k and $\mu_k$under the new and baseline estimations. Evidently, the ordering and magnitude of the estimated elasticities is rather preserved across industries. More importantly, the new estimation retains the negative correlation between σ_k and μ_k, which is a crucial statistic for the welfare impacts of policy.

The third check addresses large multi-product firms that export multiple product varieties to Colombia in a given year. Suppose a multi-product firm ω exports many products including products k and g to Colombia in year t. If demand shock are correlated across the varieties supplied by this firm (i.e., $\mathbb{E}\left[\Delta\ln\varphi_{\omega jkt}\,\Delta\ln\varphi_{\omega jgt}\right] \neq 0$), Assumption (a2) may be violated despite each variety's market share being infinitesimally small. To address this issue, we reestimate the firm-level import demand function on a restricted sample that drops excessively large firms with a within-national market share that exceeds 0.1%. The *bottom* panel in Figure P.1 compares the estimated σ_k and $\mu_k$under the new and baseline estimations. Evidently, the ordering and magnitude of the estimated elasticities is rather preserved across industries. More importantly, the new estimation retains the negative correlation between σ_k and μ_k, which is a crucial statistic for the welfare impacts of policy.

Figure P.1: *Robustness checks to address challenges to the identification of σ_k and μ_k*



Q Estimating the Import Demand Function in Levels

Our preferred estimates for μ_k and σ_k are obtained by estimating a firm-level import demand function in first-differences—see Section V. The first-difference approach for estimating elasticities in this context can be traced back to the seminal work of Feenstra (1994) and Broda and Weinstein (2006)—although both studies rely on *country-level* rather than *firm-level* data. Another body of literature estimates the trade elasticity by fitting a country-level import demand function in log-levels, while controlling for appropriate fixed effects (e.g., Hummels, Lugovskyy, and Skiba (2009); Caliendo and Parro; Shapiro (2016)).

Recently, Boehm, Levchenko, and Pandalai-Nayar (2020) have outlined the advantages and disadvantages of each approach: On the one hand, the first-difference approach performs better at handling the identification challenge poised by endogenous policy choices and omitted variable bias. On the other hand, the first difference estimator—at least when applied to country-level data—may not necessarily identify the long-run elasticity, which is the desired target for static trade models.

These issues pose a lesser problem for our firm-level estimation. We articulate this claim in two steps. First, we detail the *long-* versus *short-run* dilemma identified by Boehm et al. (2020), and explain why the same dilemma does not necessarily plague our firm-level estimation. Second, we establish our claim empirically by re-estimating our firm-level import demand function in levels. This exercise encouragingly confirms that our estimation in levels yields very similar results to our baseline estimation in differences.

The dilemma facing country-level estimations. Country level trade flows—which are traditionally used to estimate the trade elasticity—can be decomposed as follows:

$$\tilde{X}_{ji,k} = N_{ji,k}\tilde{p}_{ji,k}q_{ji,k}$$

where $\tilde{X}_{ji,k}$ denotes gross sales corresponding to *origin j–destination i–industry k*; $\tilde{p}_{ji,k}q_{ji,k}$ denotes average sales per firm (i.e., the intensive margin) and $N_{ji,k}$ denotes the total mass of firms associated with transaction ji, k (i.e., the extensive margin). Accordingly, the *long-run* trade elasticity is composed of an extensive and an intensive margin component:

$$\text{trade elasticity} \sim \frac{\partial \ln \tilde{X}_{ji,k}}{\partial \ln(1+t_{ji,k})} = \underbrace{\frac{\partial \ln N_{ji,k}}{\partial \ln(1+t_{ji,k})}}_{\varepsilon_n} + \underbrace{\frac{\partial \ln \tilde{p}_{ji,k}q_{ji}}{\partial \ln(1+t_{ji,k})}}_{\varepsilon_x}.$$

The issue raised by Boehm et al. (2020) concerns the fact that researchers do not separately observe $N_{ji,k}$ and $\tilde{p}_{ji,k}q_{ji,k}$ in country-level datasets. A standard solution to this limitation is to assume away firm-selection (i.e., set $N_{ji,k} = N_{j,k}$). Under this assumption, one can recover the trade elasticity by estimating an import demand function that controls for $N_{j,k}$ with *origin-industry* fixed effects. Crudely speaking, this solution is analogous to omitting the extensive margin component, i.e., setting $\varepsilon_n = 0$.

In practice, however, $N_{ji,k}$ may feature a bilateral element that accounts for firm-selection and which varies with the bilateral tariff rate—even after we control for a full set of origin and destination fixed effects. As noted above, traditional techniques that estimate the import demand function in *levels* with origin/destination fixed effects, are unable to account for the bilateral nature of $N_{ji,k}$. As such, traditional log-level estimators often suffer from an omitted variable bias.

Boehm et al. (2020) argue that we can overcome the omitted variable bias by estimating the country-level import demand function in differences rather than levels. Under this approach, however, one must employ long differences (over a sufficiently long time horizon) to credibly estimate the extensive margin component, ε_n. Nonetheless, the long-difference estimator may still fall short if tariff changes occur unevenly over the time-differencing horizon. In such cases, a correction must be applied to the estimated trade elasticity to account for lumpy longitudinal tariff changes.

Importantly, these limitations do not plague our firm-level estimation. We directly observe firm-level sales and need not to *infer* changes in $N_{ji,k}$ from changes in country level trade flows. Our data explicitly encompasses information on $N_{ji,k}$ and our identification strategy relies on the cross-sectional variation in firm-level variables within *importer–HS10 product–year* cells. With this level of disaggregation, our estimation is closer in spirit to the *Industrial Organization* literature on markup estimation. This literature routinely uses first difference estimators to recover markups (see, for example, equations 17-19 and related discussion in De Loecker and Warzynski (2012)). These markups estimates have been routinely used to discipline steady state models in the Macroeconomics literature (e.g., Baqaee and Farhi (2020b)).

Re-estimating our firm-level import demand function in levels. Above, we presented a conceptual argument that (compared to traditional country-level estimations) firm-level estimations should yield relatively similar results whether the import demand is estimated in *levels* or in *first differences*—provided that appropriate instruments are employed to adequately handle reverse causality. To illustrate the same point empirically, we re-estimate our firm-level import demand function in levels with *two-ways fixed effects*. We then compare the *two-ways-fixed-effects* estimates for $\mu_k = \frac{1}{\gamma_k - 1}$ and σ_k with our baseline estimates. The estimating equation in log-levels can be expressed as follows:

$$\ln \tilde{x}_{j,kt}(\omega) = (1-\sigma_k)\ln \tilde{p}_{j,kt}(\omega) + \left[1 - \frac{\sigma_k - 1}{\gamma_k - 1}\right] \ln \lambda_{j,kt}(\omega) + \underbrace{D_{kt}}_{\text{HS10-year FE}} + \underbrace{\varphi_{jk}(\omega)}_{\text{HS10-firm FE}} + \varphi_{\omega jkt}. \quad \text{(Q.1)}$$

Recall that $\tilde{x} \equiv \tilde{p}q$ denotes gross firm-level sales value; $\tilde{p}$ denotes the consumer price which includes taxes and tariffs; $\lambda_{j,kt}(\omega)$ denotes the *within-origin* $j \times$*product* k expenditure share on firm-level variety ω; D_{kt} accounts for product–year fixed effects, while $\varphi_{jk}(\omega)$ accounts for product-firm-origin fixed effects. The above equation differs from our baseline estimating equation in that the firm-product fixed effect, $\varphi_{jk}(\omega)$, is not differenced out. Instead the equation is estimated in levels.

As in the baseline case, we estimate the Equation (Q.1) using a 2SLS estimator. To this end, we modify our original shift-share instrument to make it consistent with the fixed-effects estimation, which is conducted in levels. The new instrument is calculated as follows

$$\acute{z}_{j,kt}(\omega) = \sum_{m=1}^{12} s_{j,kt-1}(\omega, m) \ln \mathcal{E}_{jt}(m)$$

where $s_{j,kt-1}(\omega, m)$ denotes the lagged share of Month m sales in firm ω's total annual export sales. $\mathcal{E}_{jt}(m)$, as before, denotes the exchange rate (between Origin j's currency and the Colombian Peso) in Month m of the current year. The other instrumental variables are adjusted accordingly, to be consistent with our estimation that is conducted in levels rather than in differences.

The estimation results are reported in Table Q.1. The estimated values for σ_k and $\mu_k = \frac{1}{\gamma_k} - 1$ are encouragingly similar to the baseline (first-differences) estimates. Most importantly, the new estimation quasi-maintains the ranking of industries in terms of the underlying degree of national-level market power (σ_k) and firm-level market power. Later, in Appendix Y, we recalculate the gains from optimal policy using the newly-estimated μ_k's and σ_k's. Encouragingly, the implied gains are starkly similar to those implied by our baseline estimates.

R Comparison of Scale Elasticity Estimation Techniques

This appendix overviews the various approaches to scale elasticity estimation, offering some perspective on the advantages of our indirect estimation technique. To provide a fair description of the existing techniques, we use an extended theoretical framework that accommodates (i) scale economies due to love-for-variety à la Krugman (1980), (ii) scale economies due to Marshallian externalities, and (iii) diseconomies of scale due to quasi-fixed factors of production. To this end, we begin this appendix by introducing a richer firm-level production function that accommodates Marshallian externalities and quasi-fixed inputs.

General Production Function—Firm ω located in origin i–industry k employs labor (L) and quasi-fixed inputs (F) using the following production function:

$$q_{i,k}(\omega) = \varphi_{i,k}(\omega)\left(L_{i,k}(\omega)^{1-\beta_{i,k}} F_{i,k}(\omega)^{\beta_{i,k}}\right) \times L_{i,k}^{\psi_k}.$$

Quasi-fixed inputs ($F_{i,k}(\omega)$) correspond to land, physical capital, or industry-specific human capital, the supply of which is fixed at the industry-level, i.e., $\sum_\omega F_{i,k}(\omega) = \overline{F}_{i,k}$. The last term in the production function accounts for Marshallian externalities, whereby the TFP of firm ω increases with industry-wide employment, $L_{i,k}$, at a constant elasticity ψ_k.

Table Q.1: *Two-ways fixed effects estimation results*

Sector	ISIC4 codes	Estimated Parameter $\sigma_k - 1$	$\frac{\sigma_k-1}{\gamma_k-1}$	μ_k	Obs.	Weak Ident. Test
Agriculture & Mining	100-1499	4.563	0.698	0.153	10,762	3.07
		(1.739)	(0.132)	(0.089)		
Food	1500-1699	2.476	0.927	0.374	17,594	5.01
		(0.818)	(0.050)	(0.284)		
Textiles, Leather & Footwear	1700-1999	3.256	0.685	0.210	110,925	59.94
		(0.297)	(0.023)	(0.024)		
Wood	2000-2099	2.093	0.893	0.427	5,282	2.12
		(1.812)	(0.200)	(3.541)		
Paper	2100-2299	7.858	0.895	0.114	35,058	2.00
		(3.953)	(0.154)	(0.177)		
Petroleum	2300-2399	2.080	1.028	0.494	3,675	1.53
		(0.342)	(0.081)	(1.584)		
Chemicals	2400-2499	4.738	0.913	0.193	127,946	29.71
		(0.496)	(0.031)	(0.071)		
Rubber & Plastic	2500-2599	4.025	0.664	0.165	101,730	9.95
		(0.791)	(0.062)	(0.045)		
Minerals	2600-2699	3.390	0.681	0.201	173,432	20.03
Basic & Fabricated Metals	2700-2899	(0.453)	(0.036)	(0.035)		
Machinery	2900-3099	4.402	0.710	0.161	257,788	19.88
		(0.662)	(0.044)	(0.034)		
Electrical & Optical Equipment	3100-3399	0.756	0.609	0.806	246,597	19.25
		(0.221)	(0.015)	(0.238)		
Transport Equipment	3400-3599	2.156	0.514	0.238	147,772	11.37
N.E.C. & Recycling	3600-3800	(0.462)	(0.032)	(0.053)		

Notes. Estimation results of Equation (16). Standard errors in parentheses. The estimation is conducted with HS10 product-year fixed effects. All standard errors are simultaneously clustered by product-year and by origin-product, which is akin to the correction proposed by Adao, Kolesár, and Morales (2019). The weak identification test statistics is the F statistics from the Kleibergen-Paap Wald test for weak identification of all instrumented variables. The test for over-identification is not reported due to the pitfalls of the standard over-identification Sargan-Hansen J test in the multi-dimensional large datasets pointed by Angrist, Imbens, and Rubin (1996).

Aggregation of Firm-Level Prices into Industry-Level Prices Indexes—Cost-minimization and profit-maximization imply that firm ω sets a price equal to $p_{in,k}(\omega) = \left(\frac{\gamma_k}{\gamma_k - 1}\right) \frac{d_{ij,k}}{\varphi_{i,k}(\omega)} w_i^{1-\beta_k} v_{i,k}^{\beta_k} L_{i,k}^{\psi_k}$. Variable $v_{i,k}$ denotes the unit price of the quasi-fixed input in industry k, which per cost minimization satisfies $v_{i,k} = \frac{1-\beta_k}{\beta_k} \frac{w_i L_{i,k}}{\overline{F}_{i,k}}$. Supposing that preferences have a nested-CES parameterization (per Assumption A1), we can use the logic in Section I to aggregate firm-level prices into industry-level price indexes subject to free entry ($M_{i,k} = L_{i,k}/\gamma_k f_k^e$). Doing so yields the following producer price index for goods associated with origin i–destination n–industry k,

$$P_{in,k} = \left(\frac{\gamma_k}{\gamma_k - 1}\right) \frac{\tau_{in,k}}{\overline{\varphi}_{i,k}} \, w_i \, L_{i,k}^{-\left(\frac{1}{\gamma_k - 1} + \psi_k\right) + \beta_k}, \tag{R.1}$$

where $\overline{\varphi}_i$ encompasses constant parameters including (average) firm productivity. Our baseline Krugman model is a special case of this equation, in which $\psi_k = \beta_k = 0$. Though, as we argue shortly, our estimation of the scale elasticity is insensitive to $\beta_k = 0$. Based on Equation R.1, the scale of employment, $L_{i,k}$, affects producer prices through increasing-returns to scale (i.e., *Jacobian* (love-for-variety) + *Marshallian* externalities) and decreasing returns to scale due to quasi-fixed inputs. More specifically, $\partial \ln P_{in,k} / \partial \ln L_{i,k} = -\left(\frac{1}{\gamma_k - 1} + \psi_k\right) + \beta_k$, where the sub-elasticity $\left(\frac{1}{\gamma_k - 1} + \psi_k\right)$ accounts for increasing-returns to scale that disrupt allocative efficiency, while β_k accounts for decreasing-returns to scale that do *not* undermine allocative efficiency given equilibrium constraints. Hence, for policy analysis, it is crucial to separately identify the former sub-elasticity from the latter—as the degree of allocative inefficiency depends solely on the sub-elasticity $\left(\frac{1}{\gamma_k - 1} + \psi_k\right)$. The following remark formalizes this point.

Remark 1. *For policy evaluation, it is useful to separately identify* $\frac{1}{\gamma_k - 1} + \psi_k$ *from* β_k*. The logic is that if* $\frac{1}{\gamma_k - 1} + \psi_k = 0$*, the market equilibrium is constrained-efficient irrespective of* β_k*, and there is no scope for improving allocative efficiency with policy. Correspondingly, the corrective gains from policy depend on the following notion of scale elasticity that differs from the reduced-form elasticity,* $\frac{\partial \ln P_{in,k}}{\partial \ln L_{i,k}}$*:*

$$\mu_k = \underbrace{(\gamma_k - 1)^{-1}}_{\text{Jacobian (love-for-variety)}} + \underbrace{\psi_k}_{\text{Marshallian}} \sim \text{scale elasticity}$$

The above remark is an immediate corollary of the First Welfare Theorem. In particular, letting $\gamma_k \to \infty$ and $\psi_k = 0$, our theoretical model reduces to a simple Arrow-Debreu model to which the fundamental welfare theorems apply. Our emphasis on $\mu_k \neq \partial \ln P_{in,k} / \partial \ln L_{i,k}$, as we elaborate shortly, speaks to one of the possible techniques for scale elasticity estimation, which infers μ_k from the reduced-form elasticity, $\ln P_{in,k} / \partial \ln L_{i,k}$.

Section III unveiled another consideration when estimating scale elasticities. We, in particular, argued that the cross-industry covariance between the scale elasticity (μ_k) and the trade elasticity (σ_k) is a crucial determinant of policy outcomes in open economies. Hence, it is advantageous to estimate these elasticities in a manner that ascertains mutual consistency.

Remark 2. *Policy outcomes in open economies depend on the cross-industry covariance between the scale and trade elasticities, i.e.,* $Cov(\mu_k, \sigma_k)$*. So, for policy evaluation, it is advantageous to jointly estimate* μ_k *and* σ_k *in a way that ascertains mutual consistency.*

Taking these remarks into consideration, we describe three techniques for estimating scale elasticities and identify their advantages and disadvantages. We begin with the indirect (demand-based) estimation technique developed in Section V of this paper.

Technique 1: Firm-Level Demand Estimation (Indirect)

Firm-level demand estimation can identify the scale elasticity, insofar as scale effects are driven by love-for-variety à la Krugman (1980). Demand estimation can also simultaneously identify the trade elasticity, σ_k, which is advantageous considering Remark 2. To unpack these points, let us revert to our generalized Krugman model for a moment while retaining the assumption that production exhibits diseconomies of scale (i.e., $\beta_{i,k} > 0$). In this setting, which widely used for trade policy analysis, the scale and trade elasticities become

$$\mu_k = \frac{1}{\gamma_k - 1} \sim \text{scale elasticity}; \qquad \sigma_k \sim \text{trade elasticity}.$$

The reason we can infer μ_k from demand parameters is that the scale elasticity in Krugman (1980) reflects the extent of *love-for-variety*—the social benefits of which are not internalized by firms' entry decisions. Recall from Section I (Assumption A1) that the nested-CES demand function facing firm ω

can be written in terms of sales ($\tilde{x} \equiv \tilde{p} \times q$) as follows

$$\tilde{x}_{ni,k}(\omega) = \xi_{ni,k}(\omega)\left(\frac{\tilde{p}_{ni,k}(\omega)}{\tilde{P}_{ni,k}}\right)^{1-\gamma_k}\left(\frac{\tilde{P}_{ni,k}}{P_{i,k}}\right)^{1-\sigma_k} Y_{i,k}.$$

One immediately notices that estimating the above function simultaneously determines the scale elasticity ($\mu_k = 1/(\gamma_k - 1)$) and the trade elasticity (σ_k). To perform the estimation, we take inspiration from Berry (1994) and rearrange and log-linearize the above demand function to obtain our estimating equation:

$$\ln \tilde{x}_{ni,k}(\omega) = (1-\sigma_k)\ln \tilde{p}_{i,k}(\omega) + \left(1 - \frac{\sigma_k - 1}{\gamma_k - 1}\right)\ln\lambda_{ni,k}(\omega) + D_{i,k} + \varepsilon_{in,k}(\omega),$$

where $\lambda_{ni,k}(\omega) = \frac{\tilde{x}_{in,k}(\omega)}{\sum_{\omega'}\tilde{x}_{in,k}(\omega')}$ is firm ω's observed conditional market share within nest (in, k); $D_{i,k}$ account for importer-industry fixed effects; and the demand residual $\varepsilon_{in,k}(\omega)$ encompasses good-specific demand-shifters and measurement error. Our identification of the demand function relies on a shift-share instrument that constitutes a firm-specific cost- or supply shifter—see Section V for specific details.

Advantages and Disadvantages of Technique 1—One advantage of our the indirect (demand-based) estimation technique is that it simultaneously identifies σ_k and μ_k—which is useful for policy evaluation following Remark 2. Moreover, our indirect estimation technique is robust to the presence of diseconomies of scale, as measured by β_k. Our estimation technique is not merely limited to the Krugman (1980) model either, and can also identify the scale elasticity in the more general Melitz (2003)-Pareto setting (see Appendixes D and R.1). A clear limitation of our approach is that it does not directly leverage scale-related moments, making it unable to identify Marshallian externalities, as measured by ψ_k.

Technique 2: National Labor Content Supply Estimation (direct)

The genesis of this technique is the observation that the producer price index, $P_{in,k}$, can be regarded as the of price of country i's labor services in destination n. Under this interpretation, the product of the trade and scale elasticity can be recovered from the labor content supply elasticity, $\frac{\partial \ln P_{in,k}}{\partial \ln L_{i,k}}$, *insofar* as production involves no quasi-fixed inputs ($\beta_k = 0$). As for the actual estimation, the trick is that even though aggregate price indexes ($P_{in,k}$) are unobserved, they can be proxied by aggregate sales. To sketch out the logic, let $\tilde{X}_{in,k} = \tilde{P}_{in,k}Q_{in,k}$ denote gross sales which satisfy the gravity equation in our framework. In particular,

$$\frac{1}{1-\sigma_k}\ln \tilde{X}_{in,k} = \left(\frac{1}{\gamma_k - 1} + \psi_k - \beta_k\right)\ln L_{i,k} + D_{i,k} + D_{n,k} + \phi_{ij,k},$$

where $D_{i,k}$ and $D_{n,k}$ are labor-size-adjusted exporter and importer fixed effects and $\phi_{in,k}$ is the bilateral resistance term. Suppose $D_{i,k} = \tilde{D}_i + \tilde{D}_k + \varepsilon^A_{i,k}$ and $\phi_{in,k} = \phi_{in} + \phi_{nk} + \varepsilon^B_{in,k}$, where ε^A and ε^B are mean-zero. Appealing to the above equation and noting that $\frac{1}{\gamma_k - 1} + \psi_k = \mu_k$, we can produce the following equation relating gross industry-level sales to employment size,

$$\frac{1}{1-\sigma_k}\ln\sum_n \tilde{X}_{in,k} = (\mu_k - \beta_k)\ln L_{i,k} + D_k + D_i + \varepsilon_{i,k}, \tag{R.2}$$

where $D_i \equiv \tilde{D}_i + \ln\sum_n(\exp(\phi_{in}))$ and $D_k \equiv \tilde{D}_k + \ln\sum_n[\exp(\phi_{nk})]$ are country and industry fixed effects, while the error term, $\varepsilon_{i,k}$, collects $\varepsilon^A_{i,k}$ (production cost shifters) and $\varepsilon^B_{in,k}$'s (trade cost shifters). Notice that $\varepsilon_{i,k}$ is akin to a supply shock here, but this interpretation rests on the implicit assumption that bilateral resistance terms have no demand-driven component. Importantly, the left-hand side variable in Equation R.2 can be regarded a proxy for the price of country i's labor services in industry k. To elucidate this connection, note that $\sum_n \tilde{X}_{in,k} = P^{1-\sigma_k}_{ii,k}\sum_n\left[\tau_{ni,k}d_{ni,k}\left(\tilde{P}^{\sigma_k - 1}_{n,k}\right)Y_{n,k}\right]$ where $\tau_{ni,k}$ collects all tax instruments associated with triplet in, k. It then follows that $\frac{1}{1-\sigma_k}\ln\sum_n \tilde{X}_{in,k} \sim P_{ii,k} + \delta_{n,k}$, where $\delta_{n,k} \equiv \sum_n\left[\tau_{ni,k}d_{ni,k}\left(\tilde{P}^{\sigma_k-1}_{n,k}\right)Y_{n,k}\right]$ can be broken down into components that are absorbed by D_k, D_i, and $\varepsilon_{i,k}$. Putting the pieces together, Equation R.2 can be regarded as a supply function for country i's labor services, with $\varepsilon_{i,k}$ representing idiosyncratic supply shocks.

One can utilize macro-level sales and employment data to estimate the following combination of parameters based on Equation R.2:

$$\frac{\partial \ln \sum_n \tilde{X}_{in,k}}{\partial \ln L_{i,k}} \sim (\mu_k - \beta_k)(\sigma_k - 1)$$

Identification in this case relies on plausibly exogenous demand-shifters that are orthogonal to $\varepsilon_{i,k}$—see Bartelme, Costinot, Donaldson, and Rodriguez-Clare (2019) for one such application. The product of the trade and scale elasticity can be, subsequently, recovered from $(\mu_k - \beta_k)(\sigma_k - 1)$ *insofar* as production involves no quasi-fixed inputs, i.e., $\beta_k = 0$. To isolate the scale elasticity (μ_k) from the trade elasticity ($\sigma_k - 1$), one must additionally rely on externally-estimated values for the trade elasticities.

Advantages and Disadvantages of Technique 2—This technique can detect Marshallian externalities as it directly leverages scale-related moments. This is a notable advantage, but comes with certain limitations, at least in the context of policy evaluation. Technique 2 cannot separately identify the scale elasticity, μ_k, from β_{ik}—which, following Remark 1, can be problematic for certain applications. Another possible drawback is Technique 2's inability to isolate the scale elasticity from the trade elasticity. This technique instead recovers the product of the two elasticities, and is mute about the sign of $Cov(\mu_k \sigma_k)$, which Following Remark 2 is an important statistic for trade policy evaluation in open economies.

Technique 3: Production Function Estimation

This technique is an augmentation of the standard production function estimation technique. Suppose we posses firm-level data on real output, $q_{i,k}(\omega)$, and input quantities, $\mathbf{X}_{i,k}(\omega) = \{L_{i,k}(\omega), F_{i,k}(\omega), ...\}$. We can, then, estimate the following log-linear production function, which includes industry-level employment as an additional covariate to identify the Marshallian component of the scale elasticity, ψ_k:

$$\ln q_{i,k}(\omega) = \boldsymbol{\beta}_k \cdot \ln \mathbf{X}_{i,k}(\omega) + \psi_k \ln L_{i,k} + \varepsilon_{i,k}(\omega).$$

The residual term $\varepsilon_{i,k}(\omega)$, in this specification, encompasses idiosyncratic firm productivity shifters and measurement error. The above function can be all but impossible to estimate at scale given the scarcity of firm-level data on real input and output quantities. To bypass this challenge, existing applications of the production function technique often estimate an aggregate version of the above equation that regresses industry-wide output, $Q_{i,k} = \sum_\omega q_{i,k}(\omega)$, on input quantities, $\mathbf{X}_{i,k} = \sum_\omega \mathbf{X}_{i,k}(\omega)$—see e.g., Basu and Fernald (1997). The scale elasticity is then recovered as $\mu_k = \sum_f \left(\beta_{f,k}\right) - 1$, where f indexes production inputs. Under this approach, $Q_{i,k}$ and $\mathbf{X}_{i,k}$ are calculated by deflating nominal sales and cost data using price indexes calculated by statistical agencies. The identification challenges relating to production function estimation of this sort are well-documented in the literature, so we refer readers to Ackerberg, Caves, and Frazer (2015) for a comprehensive synthesis of these issues.

Advantages and Disadvantages of Technique 3—The production function technique can detect Marshallian externalities, similar to Technique 2. It is also robust to the presence of quasi-fixed inputs, like Technique 1. Despite these appealing properties, the production function technique exhibits crucial limitations given its reliance on externally-constructed price indexes. This approach can credibly identify the scale elasticity, μ_k, only if the price indexes constructed by statistical agencies have adequately accounted for product quality and love-for-variety—which is often not the case. Another disadvantage of this approach is that it relies on domestic production data, meaning that the same data cannot be used to identify the trade elasticity ($\sigma_k - 1$). Instead, one must rely on completely different data to estimate ($\sigma_k - 1$), which can compromise mutual consistency as emphasized by Remark 2.

R.1 Love-for-Variety: *Krugman vs. Melitz*

As discussed earlier, the demand parameters, γ_k, fully determine the markups and scale elasticities in our baseline Krugman model (i.e., $\mu_k = \frac{1}{\gamma_k - 1}$). Section I also noted that the relationship between demand parameters and markups/scale elasticities is amended in richer environments. One such canonical case is the Melitz-Pareto model where firms incur a fixed overhead cost to serve individual markets. Following Appendix D, the markup and scale elasticities in this environment depend on the shape of the Pareto firm-level productivity distribution, ϑ_k, in addition to demand parameter,

γ_k. In particular,

$$\text{[Melitz-Pareto Model]} \qquad \mu_k^{\text{RE}} = \frac{\gamma_k \vartheta_k}{(\gamma_k - 1)(\vartheta_k + 1) - \vartheta_k} - 1 \sim \text{markup} \qquad \mu_k^{\text{FE}} = \frac{1}{\vartheta_k} \sim \text{scale elasticity}$$

To provide some intuition, the adjusted markup, $\frac{\gamma_k \vartheta_k}{(\gamma_k-1)(\vartheta_k+1)-\vartheta_k}$, corresponds to the gross markup, $\frac{\gamma_k}{\gamma_k-1}$, net of fixed cost payments. From a policy standpoint, the fraction of the markup paid to cover fixed costs is not a source of misallocation. In fact, if fixed cost payments consume the entire gross markup, the market equilibrium will be *constrained* efficient. We present a quantitative analysis of the Melitz-Pareto model in Appendix Y, elaborating more on the model's implications. Below, we discuss other settings in which the markup and scale elasticity values depend on factors other than the demand parameter γ_k.

R.2 Markups under Alternative Market Conducts

Our analysis thus far assumed that firms compete under monopolistic competition. Beyond this case, markups depend not only on demand parameters but also the conduct parameter. The markup associated with goods from origin i–industry k is, in particular, given by

$$\mu_{i,k} = v_{i,k} \times \frac{\gamma_k}{\gamma_k - 1},$$

where $v_{i,k}$ denotes the conduct parameter. Following Weyl and Fabinger (2013),(1) $v_{i,k} = 1$ under monopolistic or Bertrand competition, (2) $v_{i,k} = 0$ under perfect competition,[35] and (3) $v_{i,k} = 1/N_{i,k}$ under Cournot competition. In the spirit of Berry, Levinsohn, and Pakes (1995) and Berry (1994), our main analysis recovered markups from demand parameters by setting $v_{i,k} = 1$. We also discussed, in detail, how our optimal policy results change if we were to assume perfect or Bertrand competition instead.

Below, we discuss how our quantitative results may be be impacted by Cournot competition. The crucial takeaway from our baseline markup estimation was that trade elasticities and markups are negatively correlated across industries. This pattern may weaken or even reverse if the number of firms,$N_{i,k}$, is systematically correlated with γ_k. We investigate this possibility, using the World Bank's Exporter Dynamics Database (EDD) described in Fernandes et al. (2016). The publicly-available version of the EDD features data on firm-level exports provided by customs agencies from 60 countries for the 1997–2013 period. One of these datapoints is the number of exporters per origin and HS6 product code, from which we can infer $N_{i,k}$. Using this information, we update our baseline markup estimates as, $\mu_{i,k} = \frac{1}{N_{i,k}} \times \frac{\gamma_k}{\gamma_k-1}$ to make them compatible with Cournot competition. We then regress the our estimated trade elasticity $(\sigma_k - 1)$ on the Cournot-compatible markup values to investigate wether the negative correlation is persevered.

The results reported in Table R.1 indicate that the negative relationship between the trade elasticities and firm-level markups are robust to relaxing the monopolistic competition assumption with Cournot competition. The negative relationship becomes slightly weaker but remains significant and strong, nonetheless. Note once more, this structural relationship is the crucial driving force behind our quantitative findings that center around immiserizing growth. As detailed in Section III, if $\text{Cov}_k(\sigma_k, \mu_{i,k}) < 0$, non-cooperative second-best trade policies are ineffective at correcting misallocation in domestic industries and cooperative domestic policies trigger immiserizing growth unless they are internationally coordinated.

R.3 Scale Elasticities under Arbitrary Love-for-Variety

In our baseline Krugman model, there is a one-to-one link between the degree of firm-level market power and the love-for-variety in each industry. Baqaee and Farhi (2020a) demonstrate that this link has deep root, beyond CES models. This tight link, however, can be broken by introducing arbitrary love-for-variety into the CES demand aggregator à la Benassy (1996). In particular, suppose the sub-national CES aggregator in industry k is adjusted as follows:

$$Q_{ji,k} = \left(N_{j,k}^{\varsigma_k} \times \int_{\omega \in \Omega_{j,k}} \varphi_{ji,k}(\omega)^{\frac{1}{\gamma_k}} q_{ji,k}(\omega)^{\frac{\gamma_k - 1}{\gamma_k}} d\omega \right)^{\frac{\gamma_k}{\gamma_k - 1}}.$$

[35] Likewise, under Bertrand competition with homogeneous sub-products, $v_{i,k} = 0$.

Table R.1: The tension between ToT and sectoral misallocation under Cournot competition

	dependent: trade elasticity $(\sigma_k - 1)$		
$\mu_{i,k} = \frac{1}{N_{i,k}} \times \frac{\gamma_k}{\gamma_k - 1}$	-0.138***	-0.138***	-0.279***
	(0.0111)	(0.0111)	(0.0157)
Year fixed effects	No	Yes	Yes
Origin fixed effects	No	No	Yes
Observations	3,221	3,221	3,221

Note: This paper relationship between the firm-level markup under Cournot competition, $\frac{1}{N_{i,k}} \times \frac{\gamma_k}{\gamma_k-1}$, and the trade elasticity, $\sigma_k - 1$. Data for $N_{i,k}$ are from the World Bank's Exporter Dynamics Database. Parameters σ_k and γ_k are from the demand estimation conducted in Section V, where k denotes a WIOD industry. *** denotes significant at the 1\% level.

Parameter ς_k regulates the love-for-variety as measured by the number of firm-level varieties, $N_{j,k}$. The above CES aggregator coincides with our baseline CES aggregator if $\varsigma_k = 0$. It is straightforward to check that firm-level markups are unaffected by ς_k, as individual firms treat $N_{j,k}$ as given when setting their prices. The scale elasticity, however, should be adjusted as follows:

$$\text{[Krugman+Benassy]} \qquad \mu_k = \frac{1}{\gamma_k - 1} \sim \text{markup} \qquad 1 + \psi_k = \left(1 + \frac{1}{\gamma_k - 1}\right)(1 + \varsigma_k) \sim \text{scale leasticity}$$

The optimal domestic subsidy in this case is $1 + s_k^\star = \left(1 + \frac{1}{\gamma_k - 1}\right)(1 + \varsigma_k)$, and the gains from restoring efficiency are, accordingly, amplified. Notice, markup heterogeneity is no longer necessary to justify policy intervention. The heterogeneity in ς_k is sufficient, which echos Epifani and Gancia's (2011) findings in a single-sector economy.

Estimating the love-for-variety parameter, ς_k, with sales and price data is, however, challenging. In our *firm-level* estimation, $N_{j,k}^{\varsigma_k}$ will appear as an origin-and-industry-specific demand shifter and will be absorbed by our extensive set of fixed effects. Estimating ς_k with *national-level* sales and price data faces the same complications as external economies of scale. In particular, one cannot purge elasticity, ς_k, from the diseconomies of scale elasticity without explicit data on quasi-fixed factors of production. As detailed in Section V.E, the latter elasticity does not contribute to inefficiency and must be excluded from the optimal Pigouvian subsidy.

S Examining the Plausibility of Estimates

In this Appendix we examine the plausibility of our estimated parameters from a different angle. We show that when our estimated parameters are plugged into a workhorse trade model, they resolve the *income-size* elasticity puzzle. This puzzle, as noted by Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), concerns the fact that a large class of quantitative trade models—including Krugman (1980), Eaton and Kortum (2001), and Melitz (2003)—predict a counterfactually high income-size elasticity (i.e., the elasticity at which real per capita income increases with population size). One straightforward remedy for this counterfactual prediction is introducing domestic trade frictions into the aforementioned models. This treatment, however, is only a partial remedy. As shown by Ramondo et al. (2016), even after controlling for direct measures of internal trade frictions, the predicted income-size elasticity remains counterfactually strong.

To test macro-level predictions, we first produce economically-representative estimates for σ_k and μ_k. We do so by pooling data across all manufacturing and non-manufacturing industries and estimating Equation 16 on theses two pooled samples.The estimation results are reported in Table S.1, and imply that $\sigma \approx 3.8$ and $\frac{\sigma-1}{\gamma-1} \approx 0.66$ across manufacturing industries. For the sake of comparison, the same table also reports estimates produced using the standard OLS estimator.

To understand the income-size elasticity puzzle, consider a single-industry version of the model presented in Section I. Such a model implies the following expression relating country i's real income

Table S.1: *Pooled estimation results*

	Manufacturing		Non-Manufacturing	
Variable (log)	2SLS	OLS	2SLS	OLS
Price, $1-\sigma$	-2.766***	0.203***	-5.540***	0.102***
	(0.186)	(0.004)	(0.706)	(0.007)
Within-national share, $1-\mu(\sigma-1)$	0.340***	0.816***	0.167***	0.804***
	(0.010)	(0.002)	(0.033)	(0.010)
Weak Identification Test	259.91	...	28.83	...
Under-Identification P-value	0.00	...	0.00	...
Within-R^2	...	0.78	...	0.73
N of Product-Year Groups	21,416		8,903	
Observations	1,130,742		204,828	

Notes: *** denotes significant at the 1% level. The Estimating Equation is (16). Standard errors in brackets are robust to clustering within product-year. The estimation is conducted with HS10 product-year fixed effects. The reported R^2 in the OLS specifications correspond to within-group goodness of fit. Weak identification test statistics is the F statistics from the Kleibergen-Paap Wald test for weak identification of all instrumented variables. The p-value of the under-identification test of instrumented variables is based on the Kleibergen-Paap LM test. The test for over-identification is not reported due to the pitfalls of the standard over-identification Sargan-Hansen J test in the multi-dimensional large datasets pointed by Angrist et al. (1996).

per worker or TFP ($W_i = w_i/P_i$) to its structural efficiency, A_i, population size, L_i, trade-to-GDP ratio, λ_{ii}, and a measure of internal trade frictions, τ_{ii}:

$$W_i = \gamma\ A_i\ L_i^{\mu}\ \lambda_{ii}^{-\frac{1}{\sigma-1}}\ \tau_{ii}^{-1}. \tag{S.1}$$

The standard Krugman model assumes extreme love-of-variety (or extreme scale economies), which implies $\mu = 1/(\sigma-1)$ and precludes internal trade frictions, which results in $\tau_{ii} = 1$. Given these two assumptions, we can compute the real income per worker predicted by the standard Krugman model and contrast it to actual data for a cross-section of countries.

For this exercise, we use data on the trade-to-GDP ratio, real GDP per worker, and population size for 116 countries from the Penn World Tables in the year 2011. Given our micro-estimated trade elasticity, $\sigma-1$, and plugging $\tau_{ii} = 1$ as well as $\mu = 1/(\sigma-1)$ into Equation S.1, we can compute the real income per worker predicted by the Krugman model. Figure S.1 (top panel) reports these predicted values and contrasts them to factual values. Clearly, there is a sizable discrepancy between the income-size elasticity predicted by the standard Krugman model (0.36, standard error 0.03) and the factual elasticity (-0.04, standard error 0.06). To gain intuition, note that small countries import a higher share of their GDP (i.e., posses a lower λ_{ii}), which partially mitigates their size disadvantage. However, even after accounting for observable levels of trade openness, the scale economies underlying the Krugman model are so strong that they lead to a counterfactually high income-size elasticity.

One solution to the income-size elasticity puzzle is introducing internal trade frictions into the Krugman model (i.e., relaxing the $\tau_{ii} = 1$ assumption). Ramondo et al. (2016) perform this task using direct measures of domestic trade frictions. Their calibration is suggestive of $\tau_{ii} \propto L_i^{0.17}$. Plugging this implicit relationship into Equation S.1 and using data on population size and trade openness, we compute the model-predicted real income per worker and contrast it with actual data in Figure S.1 (middle panel). Expectedly, accounting for internal frictions shrinks the income-size elasticity. However, as pointed out by Ramondo et al. (2016), the income-size elasticity remains puzzlingly large.

We ask if our micro-estimated scale elasticity can help resolve the remaining income-size elasticity puzzle. To this end, in Equation S.1, we set the scale elasticity to $\mu = \alpha/(\sigma-1)$ where α is set to 0.65 as implied by our micro-level estimation. Then, using data on population size and trade-to-GDP ratios, we compute the real income per capita predicted by a model that features both domestic trade frictions *and* adjusted scale economies. Figure S.1 plots these predicted values, indicating that this adjustment indeed resolves the income-size elasticity puzzle. In particular, the income-size elasticity predicted by the Krugman model with adjusted scale economies is statistically insignificant (0.02, standard error 0.03), aligning very closely with the factual elasticity.

T Mapping Second-Best Tax Formulas to Data

In this appendix, we present an analog to Proposition 1, but for second-best trade taxes under restricted entry (as specified by Theorem 2). As in Section VI, we assume that preferences have a CES-Cobb-Douglas parametrization. We use the "$**$" superscript indicates that a variable is being evaluated in the counterfactual *second-best* optimal policy equilibrium. We assume hereafter that countries do not apply domestic subsidies in the factual equilibrium, i.e., $s_{n,k} = 0$ for all $n \in \mathbb{C}$. Using the hat-algebra notation and the expression of the good-specific supply elasticity, $\omega_{ji,k}$ (Equation 10), we can write the second-best tax formulas in changes as follows:

$$
\begin{aligned}
&\text{[optimal import tax]} \quad 1+t^{**}_{ji,k} = \frac{1+(\sigma_k-1)\hat{\lambda}_{ii,k}\lambda_{ii,k}}{1+\frac{1+\overline{\mu}^*_i}{1+\mu_k}(\sigma_k-1)\hat{\lambda}_{ii,k}\lambda_{ii,k}}\left(1+\omega^{**}_{ji,k}\right)\\
&\text{[optimal export subsidy]} \quad 1+x^{**}_{ij,k} = \frac{(\sigma_k-1)\sum_{n\neq i}\left[(1+\omega^{**}_{ni,g})\hat{\lambda}_{nj,k}\lambda_{nj,k}\right]}{1+(\sigma_k-1)(1-\hat{\lambda}_{ij,k}\lambda_{ij,k})}\left(\frac{1+\mu_k}{1+\overline{\mu}^{**}_i}\right),\\
&\text{[change in taxes]} \quad \widehat{1+s_{i,k}} = 1; \qquad \widehat{1+t_{ji,k}} = \frac{1+t^{**}_{ji,k}}{1+t_{ji,k}}; \qquad \widehat{1+x_{ij,k}} = \frac{1+x^{**}_{ij,k}}{1+x_{ij,k}}.
\end{aligned}
\tag{T.1}
$$

Since the rest of the world is passive in their use of taxes, $\widehat{1+s_{n,k}} = \widehat{1+t_{jn,k}} = \widehat{1+x_{nj,k}} = 1$ for all $n \neq i$. To determine the change in expenditure shares, $\hat{\lambda}_{ji,k}$, we need to determine the change in consumer price indexes. Invoking the CES structure of within-industry demand, we can express the change in *market i–industry k*'s consumer price index as

$$
\text{[price indexes]} \qquad \hat{\tilde{P}}_{i,k} = \sum_{n=1}^{N}\left(\lambda_{ni,k}\left[\frac{\widehat{1+t_{ni,k}}}{\widehat{1+x_{ni,k}}}\widehat{w}_n\right]^{1-\sigma_k}\right)^{\frac{1}{1-\sigma_k}}. \tag{T.2}
$$

Given $\hat{\tilde{P}}_{i,k}$, we can calculate the change in expenditure shares $\hat{\lambda}_{ji,k}$ and revenue shares $\hat{r}_{ji,k}$ as

$$
\begin{aligned}
&\text{[expenditure shares]} \qquad \hat{\lambda}_{ji,k} = \left[\frac{\widehat{1+t_{ji,k}}}{\widehat{1+x_{ji,k}}}\widehat{w}_j\right]^{1-\sigma_k}\hat{\tilde{P}}^{\sigma_k-1}_{i,k}\\
&\text{[revenue shares]} \qquad \hat{r}_{ji,k} = \left(\frac{\widehat{1+x_{ji,k}}}{\widehat{1+t_{ji,k}}}\hat{\lambda}_{ji,k}\hat{Y}_i\right)\left(\sum_{n=1}^{N}\frac{\widehat{1+x_{jn,k}}}{\widehat{1+t_{jn,k}}}\hat{\lambda}_{jn,k}\hat{Y}_n\right)^{-1}.
\end{aligned}
\tag{T.3}
$$

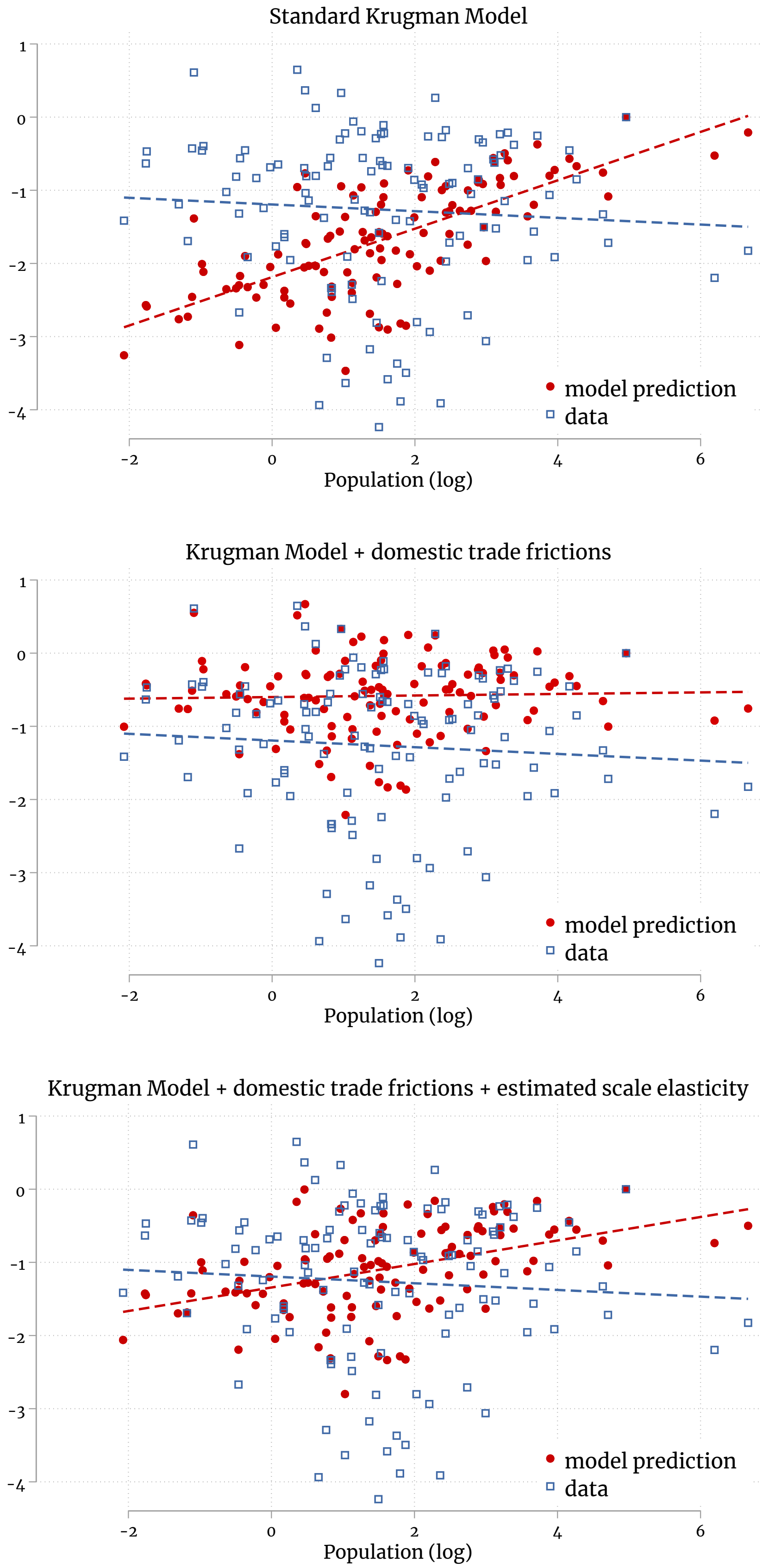
The change in wage rates, $\hat{w}_i$, and labor shares, $\hat{\rho}_{i,k}$, are dictated by the labor market clearing (LMC) condition, which ensures that industry-level sales match wage payments:

$$
\text{[LMC]} \qquad (1+\overline{\mu}^{**}_i)\hat{w}_i w_i L_i = \sum_{j\in\mathbb{C}}\sum_{k\in\mathbb{K}}\left[\frac{1+x^{**}_{ji,k}}{1+t^{**}_{ji,k}}\hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j\right]. \tag{T.4}
$$

where the output-weighted average markup in the counterfactual equilibrium is given by

$$
1+\overline{\mu}^{**}_i = \frac{\sum_{j\in\mathbb{C}}\sum_{k\in\mathbb{K}}\left[\frac{1+x^{**}_{ji,k}}{1+t^{**}_{ji,k}}\hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j\right]}{\sum_{j\in\mathbb{C}}\sum_{k\in\mathbb{K}}\left[\frac{1+x^{**}_{ji,k}}{(1+\mu_k)(1+t^{**}_{ji,k})}\hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j\right]}. \tag{T.5}
$$

Figure S.1: *Resolving the income-size elasticity puzzle*



The change in national expenditure, $\hat{Y}_i$, is governed by the balanced budget (BB) condition, which ensures that total expenditure matches total income from wage payments and tax revenues:

$$[\text{BB}] \quad \hat{Y}_i Y_i = +(1+\overline{\mu}_i^{**})\hat{w}_i w_i L_i + \sum_{j\neq i}\sum_k \left(\frac{t_{ji,k}^{**}}{1+t_{ji,k}^{**}} \lambda_{ji,k}\hat{\lambda}_{ji,k} e_{i,k}\hat{Y}_i Y_i + \frac{1-(1+x_{ij,k}^{**})}{1+t_{ij,k}^{**}} \lambda_{ij,k}\hat{\lambda}_{ij,k} e_{j,k}\hat{Y}_j Y_j \right). \tag{T.6}$$

Equations T.1-T.6 represent a system of $2N + NK + 2(N-1)K$ independent equations and unknowns. The independent unknowns are, namely, $\hat{w}_i$ (N unknowns), $\hat{Y}_i$ (N unknowns), $\hat{\rho}_{i,k}$ (NK unknowns), $\widehat{1+t_{ji,k}}$ ($(N-1)K$ unknowns), and $\widehat{1+x_{ij,k}}$ ($(N-1)K$ unknowns). Solving the aforementioned system is possible with information on observable data points, $\mathbb{D}$, and estimable parameters, $\Theta \equiv \{\mu_k, \sigma_k\}$. Once we solve this system, the welfare consequences of country i's optimal policy are also fully determined. The following proposition outlines this result.

Proposition 1. *Suppose we have data on observable shares, national accounts, and applied taxes,* $\mathbb{D}=\left\{\lambda_{ji,k}, r_{ji,k}, e_{i,k}, Y_i, w_i L\right.$ *and information on structural parameters,* $\Theta \equiv \{\mu_k, \sigma_k\}$*. We can determine the economic consequences of country i's second-best optimal policy by calculating* $\mathbb{X} = \left\{\hat{Y}_i, \hat{w}_i, \hat{\rho}_{i,k}, \widehat{1+t_{ji,k}}, \widehat{1+x_{ij,k}}\right\}$ *as the solution to the system of Equations T.1-T.6. After solving for* $\mathbb{X}$*, we can fully determine the welfare consequence of country i's optimal policy as*

$$\hat{W}_n = \hat{Y}_n / \prod_{k\in\mathbb{K}} \hat{\tilde{P}}_{n,k}^{e_{n,k}}, \qquad (\forall n \in \mathbb{C})$$

where $\hat{\tilde{P}}_{n,k}$ *can be computed as function of* $\mathbb{X}$ *and data,* $\mathbb{D}$*, using Equation T.2.*

U Additional Details about the World Input-Output Database

This appendix presents additional details about the World Input-Output Database analyzed in Section VI. Table U.1 describes our aggregation of WIOD industries into 16 industries. To summarize the information in this table, we aggregate the 'Agriculture' and Mining' industries into one non-manufacturing industry. We also follow Costinot and Rodríguez-Clare (2014) in two details: First, we aggregate the 'Textile' and 'Leather' industries into one industry. Second, we lump all service-related industries together treating them as one semi-non-tradable sector.

Following Proposition 1 in Section VI, we need data on observable shares, national accounts, and applied taxes ($\mathbb{D} = \left\{\lambda_{ji,k}, r_{ji,k}, e_{i,k}, \rho_{i,k}, Y_i, w_i L_i, x_{ij,k}, t_{ji,k}, s_{i,k}\right\}_{j,i,k}$) to compute the gains from policy. The WIOD reports data on trade values, $X_{ji,k} \equiv P_{ji,k}Q_{ji,k}$, for each origin j–destination i–industry k. The aggregated version of the data covers $N = 33$ countries (including the rest of the world) and $K = 16$ industries. Below, we describe how each element in $\mathbb{D}$ is computed based on $X_{ji,k}$ and our estimated values for μ_k. Assuming that countries impose no taxes under the status-quo, we can compute national income and the wage bill in each country i as follows:

$$Y_i = \sum_{k=1}^{K}\sum_{n=1}^{N} X_{ni,k}; \qquad w_i L_i = \begin{cases} \sum_{k=1}^{K}\sum_{n=1}^{N} X_{in,k} & \text{if entry is free} \\ \sum_{k=1}^{K}\sum_{n=1}^{N} \frac{1}{1+\mu_k} X_{in,k} & \text{if entry is restricted} \end{cases}$$

Next, we can compute the within-industry and industry-level expenditure shares for each market i based on the following calculations:

$$\lambda_{ji,k} = \frac{X_{ji,k}}{\sum_{n=1}^{N} X_{ni,k}}; \qquad e_{i,k} = \frac{\sum_{n=1}^{N} X_{ni,k}}{\sum_g \sum_{n=1}^{N} X_{ni,g}} = \frac{\sum_{n=1}^{N} X_{ni,k}}{Y_i}.$$

Lastly, we can compute the within-industry revenue share and the industry-level labor share in each country using the following equations:

$$r_{in,k} = \frac{X_{in,k}}{\sum_{n=1}^{N} X_{in,g}}; \qquad \rho_{i,k} = \frac{\sum_{n=1}^{N} X_{in,k}}{\sum_{g=1}^{K}\sum_{n=1}^{N} X_{in,g}}.$$

Table U.1: *List of industries in the World Input-Output Database*

WIOD Sector	Sector's Description	Trade Ealsticity	Scale Ealsticity
1	Agriculture, Hunting, Forestry and Fishing	6.227	0.143
2	Mining and Quarrying	6.227	0.143
3	Food, Beverages and Tobacco	2.303	0.393
4	Textiles and Textile Products Leather and Footwear	3.359	0.224
5	Wood and Products of Wood and Cork	3.896	0.229
6	Pulp, Paper, Paper , Printing and Publishing	2.646	0.320
7	Coke, Refined Petroleum and Nuclear Fuel	0.397	1.758
8	Chemicals and Chemical Products	3.966	0.232
9	Rubber and Plastics	5.157	0.140
10	Other Non-Metallic Mineral	5.283	0.167
11	Basic Metals and Fabricated Metal	3.004	0.209
12	Machinery, Nec	7.750	0.120
13	Electrical and Optical Equipment	1.235	0.552
14	Transport Equipment	2.805	0.129
15	Manufacturing, Nec; Recycling	6.169	0.152
16	Electricity, Gas and Water Supply Construction Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods Hotels and Restaurants Inland Transport Water Transport Air Transport Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies Post and Telecommunications Financial Intermediation Real Estate Activities Renting of M&Eq and Other Business Activities Education Health and Social Work Public Admin and Defence; Compulsory Social Security Other Community, Social and Personal Services Private Households with Employed Persons	11	0

V Quantitative Analysis with Exact Formulas

Our optimization-free quantitative approach in Section VI relied on approximate formulas for the export supply elasticity. The same analysis, however, can also be conducted with exact formulas. In this appendix, we demonstrate this point and show that both approaches deliver virtually identical output. Though, our suggested approximation saves computation time to a notable degree.

As a starting point, we appeal to our exact formula for the (general equilibrium) export supply elasticity,[36]

$$\omega_{ji,k} \equiv \frac{1}{r_{ji,k}\rho_{j,k}} \left[\frac{w_i L_i}{w_j L_j} \rho_{i,k} \left(\frac{\partial \ln P_{ii,k}}{\partial \ln Q_{ji,k}} \right)_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i} + \sum_{n \neq i} \frac{w_n L_n}{w_j L_j} r_{ni,} \rho_{n,k} \left(\frac{\partial \ln P_{ni,k}}{\partial \ln Q_{ji,k}} \right)_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i} \right].$$

As detailed in Appendix, $\left(\frac{\partial \ln P_{ii,k}}{\partial \ln Q_{ji,k}} \right)_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i}$ is a partial derivate holding constant the vector of wages, income, and "consumer" prices associated with economy i. The matrix consisting of these partial derivatives can be evaluated by inverting a system of equations as specified by Equation E.38. Namely,

$$\begin{bmatrix} \left(\frac{\partial \ln P_{11,k}}{\partial \ln Q_{1i,k}} \right)_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i} & \cdots & \left(\frac{\partial P_{11,k}}{\partial \ln Q_{Ni,k}} \right)_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i} \\ \vdots & \ddots & \vdots \\ \left(\frac{\partial \ln P_{NN,k}}{\partial \ln Q_{1i,k}} \right)_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i} & \cdots & \left(\frac{\partial P_{NN,k}}{\partial \ln Q_{Ni,k}} \right)_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i} \end{bmatrix} = - \underbrace{\begin{bmatrix} \frac{\partial F_{1i,k}(.)}{\partial \ln P_{11,k}} & \cdots & \frac{\partial F_{1i,k}(.)}{\partial \ln P_{NN,k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{Ni,k}(.)}{\partial \ln P_{11,k}} & \cdots & \frac{\partial F_{Ni,k}(.)}{\partial \ln P_{NN,k}} \end{bmatrix}}_{\mathbf{A}_i}^{-1} \begin{bmatrix} \frac{\partial F_{1i,k}(.)}{\partial \ln Q_{1i,k}} & \cdots & \frac{\partial F_{1i,k}(.)}{\partial \ln Q_{Ni,k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{Ni,k}(.)}{\partial \ln Q_{1i,k}} & \cdots & \frac{\partial F_{Ni,k}(.)}{\partial \ln Q_{Ni,k}} \end{bmatrix}.$$

where $F_{ni,k}(\mathbf{Q}_{i,k}, \mathbf{P}_k) \equiv P_{nn,k} - \bar{\varrho}_{nn,k} w_n \left[\tau_{ni,k} Q_{ni,k} + \sum_{\ell \neq i} \tau_{n\ell,k} Q_{n\ell,k} \right]^{-\frac{\mu_k}{1+\mu_k}} = 0$, with the corresponding derivatives specified in Appendix E.

With the above background, we now explain how the quantitative procedure explained in Section VI.A can be re-done without appealing to approximation or numerical optimization. In summary, one must now solve the exact optimal tax formulas in conjunction with the equilibrium condition in changes. The exact optimal tax/subsidy formulas can be expressed as

$$\text{[optimal import tax]} \quad 1 + t^*_{ji,k} = 1 + \frac{1}{\hat{r}_{ji,k}\hat{\rho}_{j,k} r_{ji,k}\rho_{j,k}} \left[\frac{\hat{w}_i w_i L_i}{\hat{w}_j w_j L_j} \hat{\rho}_{i,k}\rho_{i,k} \varkappa^*_{ij,k} + \sum_{n \neq i} \frac{\hat{w}_n w_n L_n}{\hat{w}_j w_j L_j} \hat{r}_{ni,k}\hat{\rho}_{n,k} r_{ni,k}\rho_{n,k} \varkappa^*_{nj,k} \right]$$

$$\text{[optimal export subsidy]} \quad 1 + x^*_{ij,k} = \frac{(\sigma_k - 1)\sum_{n \neq i} \left[(1 + t^*_{ni,k}) \hat{\lambda}_{nj,k}\lambda_{nj,k} \right]}{1 + (\sigma_k - 1)(1 - \hat{\lambda}_{ij,k}\lambda_{ij,k})},$$

$$\widehat{1 + s_{i,k}} = \frac{1 + \mu_k}{1 + s_{i,k}}; \qquad \widehat{1 + t_{ji,k}} = \frac{1 + t^*_{ji,k}}{1 + t_{ji,k}}; \qquad \widehat{1 + x_{ij,k}} = \frac{1 + x^*_{ij,k}}{1 + x_{ij,k}}. \tag{A}$$

The variable $\varkappa^*_{nj,k} \equiv \left(\frac{\partial \ln P_{nn,k}}{\partial \ln Q_{ji,k}} \right)^*_{\mathbf{w},\mathbf{Y},\tilde{\mathbb{P}}_i}$ refers to the partial price derivatives evaluated in the counterfactual optimal policy equilibrium. The entire matrix of $\varkappa^*_{nj,k}$'s can be recovered with information on structural parameters and the change to observable share variables. Namely,

$$\begin{bmatrix} \varkappa^*_{11,k} & \cdots & \varkappa^*_{1N,k} \\ \vdots & \ddots & \vdots \\ \varkappa^*_{N1,k} & \cdots & \varkappa^*_{NN,k} \end{bmatrix} = - \left(\mathbf{I}_N - \begin{bmatrix} a_{11,k} & \cdots & a_{1N,k} \\ \vdots & \ddots & \vdots \\ a_{N1,k} & \cdots & a_{NN,k} \end{bmatrix} \right)^{-1} \begin{bmatrix} \frac{\mu_k}{1+\mu_k} r_{1i,k}\hat{r}_{1i,k} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\mu_k}{1+\mu_k} r_{Ni,k}\hat{r}_{Ni,k} \end{bmatrix}, \tag{B}$$

where the elements of the first matrix on the right-hand side are

$$a_{nj,k} = \mathbb{1}_{j \neq i} \frac{\mu_k}{1 + \mu_k} \sum_{\ell \neq i} \left[\left(\mathbb{1}_{n=j}\sigma_k - (\sigma_k - 1)\lambda_{j\ell,k}\hat{\lambda}_{j\ell,k} \right) r_{n\ell,k}\hat{r}_{n\ell,k} \right].$$

Solving Equations (A) and (B) alongside the equilibrium conditions specified by Equations 18-21 in

[36] Notice, the above expression for $\omega_{ji,k}$ precludes cross-industry effects, given our Cobb-Douglas utility parameterization across industries.

the main text determines the entire vector of counterfactual outcomes after the imposition of optimal taxes/subsidies. The following proposition summarizes this point.

Proposition 2. *Suppose we have data on observables,* $\mathbb{D} = \left\{\lambda_{ji,k}, r_{ji,k}, e_{i,k}, \rho_{i,k}, Y_i, w_i L_i, x_{ij,k}, t_{ji,k}, s_{i,k}\right\}_{j,i,k}$, *and information on structural parameters,* $\Theta \equiv \{\mu_k, \sigma_k\}$. *We can determine the economic consequences of country i's optimal policy by calculating* $\mathbb{X} = \left\{\hat{Y}_i, \hat{w}_i, \hat{\rho}_{i,k}, \widehat{1+s_{i,k}}, \widehat{1+t_{ji,k}}, \widehat{1+x_{ij,k}}\right\}$ *as the solution to the system of Equations consisting of (A) and (B) plus equilibrium conditions 18-21. After solving for* $\mathbb{X}$, *we can fully determine the welfare consequence of country i's optimal policy as*

$$\hat{W}_i = \hat{Y}_i / \prod_{k \in \mathbb{K}} \hat{\tilde{P}}_{i,k}^{e_{i,k}}, \qquad (\forall n \in \mathbb{C})$$

where $\hat{\tilde{P}}_{i,k}$ *is determined by Equation 18 as a function of* $\mathbb{X}$ *and data,* $\mathbb{D}$.

Using Proposition 3, we recalculate the exact gains from optimal policy and compare them with the baseline gains implied by our approximate formulas. The results are displayed in Table V.1 for select countries. These are relatively large countries for which our approximation is more suspect. One immediately notices that our approximate formulas deliver identical numbers to the exact formulas. The intuition, as explained in Appendix E, is that the matrix $\mathbf{A}_k = \left[a_{nj,k}\right]$ is sufficiently sparse. To put these results in perspective, Table V.1 also reports the gains implied by the small open economy optimal policy formulas. These formulas are presented in Section II. The small open economy assumption is markedly more error-prone, as it attributes "zero" import market power to each country irrespective of market size and import composition.

Table V.1: *Gains from policy: exact vs. approximate optimal tax formulas*

	Exact Formula	Approximated Formula		Small Open Economy Formula	
Country	ΔW	ΔW	Error	ΔW	Error
BEL	1.3088%	1.3088%	0.00%	1.3007%	0.62%
DEU	1.7117%	1.7113%	0.02%	1.6885%	1.37%
NLD	1.3547%	1.3547%	0.00%	1.3450%	0.72%
NOR	1.1889%	1.1889%	0.00%	1.1757%	1.12%
USA	1.5283%	1.5278%	0.03%	1.5178%	0.69%

Note: The data source is the 2014 World Input-Output Database (WIOD, Timmer et al. (2015)). Policy outcomes in the small open economy case are calculated using the optimal policy specification under 12.

W Elucidating the Tension between ToT and Misallocation

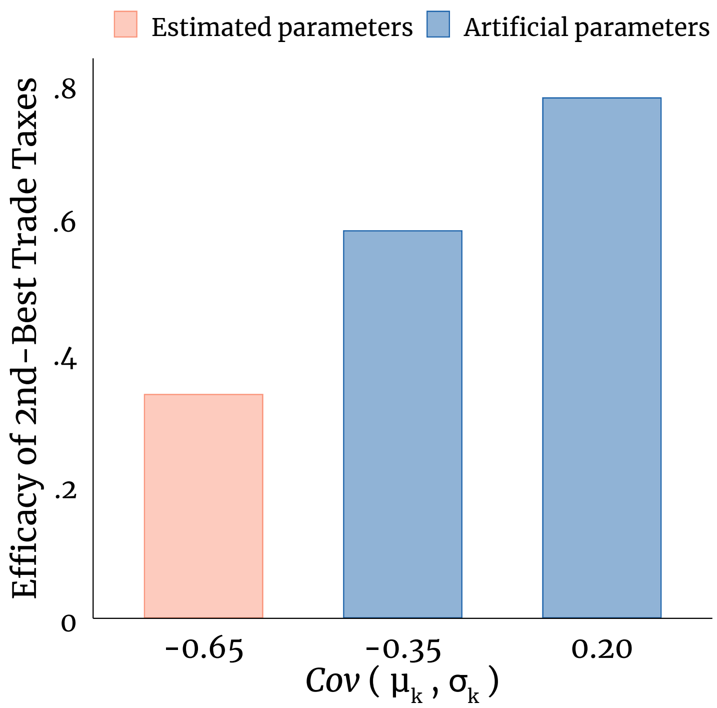
This appendix shows that the inefficacy of 2nd-best *non-cooperative* trade taxes stems from the tension between terms-of-trade (ToT) and misallocation. Our quantitative analysis, recall, indicated that 2nd-best trade taxes can replicate less than 40% of the gains from the 1st-best policy choice, which combines trade taxes with Pigouvian subsidies. In what follows, we argue that this apparent lack of efficacy is not a universal feature that merely reflects the targeting principle. Instead, it is an empirical result based on our estimated trade and scale elasticity values.

To establish this point, we artificially raise $Cov(\sigma_k, \mu_k)$ and recompute the gains from 2nd-best trade taxes. We then calculate the efficacy of 2nd-best trade taxes as the ratio of the corresponding gains relative to the 1st-best policy choice. Each iteration maintains the estimated vector of trade elasticities and adjusts the scale elasticities (or firm-level markups) to artificially inflate $Cov(\sigma_k, \mu_k)$ relative to its estimated value. Throughout this appendix, we report results for the case of *restricted entry*, noting that similar results hold under free entry.

The results reported in Figure W.1 confirm that 2nd-best trade taxes become increasingly more effective as $Cov(\sigma_k, \mu_k)$ is artificially inflated. Under our estimated parameters, $Cov(\sigma_k, \mu_k) \approx -0.60$ and 2nd-best trade taxes can replicate less than 40% of the 1st-best gains from policy. When $Cov(\sigma_k, \mu_k)$

is artificially raised to -0.35, 2nd-best trade taxes can replicate close to 60% of the gains from 1st-best gains from policy. When $Cov\,(\sigma_k, \mu_k)$ is raised further to 0.30, the efficacy of 2nd-best trade taxes improves to 80%. These results indicate that the inefficacy of 2nd-best trade taxes is not an exclusive reflection of the targeting principle. While one expects a less-than-100% efficacy based on the targeting principle, 2nd-best trade taxes become a remarkably weaker substitute for Pigouvian subsidies under lower values of $Cov\,(\sigma_k, \mu_k)$.

Figure W.1: *2nd-best trade taxes become more effective as $Cov\,(\mu_k, \sigma_k)$ is artificially inflated*



Note: The data source is the 2014 World Input-Output Database (WIOD, Timmer et al. (2015)). Each bar reports the average welfare gains when countries implement their 1st-best policy without retaliation by partners. The artificial parameters are constructed by fixing $\boldsymbol{\sigma}$ to its estimated values and adjusting $\boldsymbol{\mu}$to artificially inflate $Cov\,(\mu_k, \sigma_k)$.

We repeat the same exercise to elucidate the immiserizing growth effects of unilateral policy markup correction. In particular, we artificially raise $Cov\,(\sigma_k, \mu_k)$ and recompute the consequences of unilateral markup correction. Each iteration maintains the estimated vector of firm-level markups (or scale elasticities) and adjusts the trade elasticities to artificially inflate $Cov\,(\sigma_k, \mu_k)$ relative to its estimated value. This choice ensures that the degree of inter-industry misallocation remains approximately the same despite the change in $Cov\,(\sigma_k, \mu_k)$.

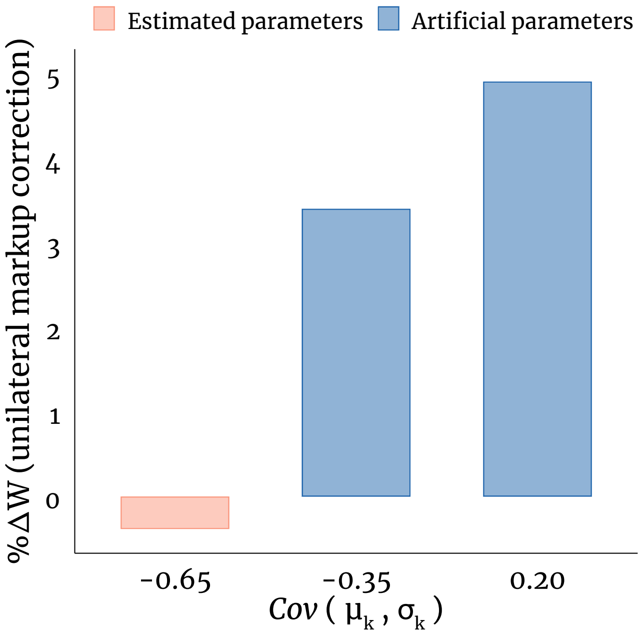
The results reported in Figure W.2 indicate that immiserizing growth effects fade and even reverse as $Cov\,(\sigma_k, \mu_k)$ is artificially inflated. Under our estimated parameters, where $Cov\,(\sigma_k, \mu_k) \approx -0.60$, unilateral markup correction prompts immiserizing growth and lowers welfare. When $Cov\,(\sigma_k, \mu_k)$ is artificially raised to -0.35, unilateral markup correction no longer triggers immiserizing growth. When $Cov\,(\sigma_k, \mu_k)$ is raised further to 0.30, unilateral markup correction becomes a promising policy choice as it restores allocative efficiency and improves the ToT at the same time.

X Country-Level Exposure to Immiserizing Growth

This appendix digs deeper into the immiserizing growth effects of unilateral industrial policy. Recall from section sec: Tension that immiserizing growth presents a grave challenge to industrial implementation in open economies. In Section III, we reported the extent of immiserizing growth for the average country. Here we unpack these numbers. First, by reporting immiserizing growth effects on a country-by-country basis. Second, by highlighting that trade-to-GDP is a crucial determinant of the extent to which countries experience immiserizing growth.

Figure X.1 displays welfare consequences when countries implement corrective policies without

Figure W.2: *Immiserizing growth effects diminish as* $Cov(\mu_k, \sigma_k)$ *is artificially inflated*



Note: The data source is the 2014 World Input-Output Database (WIOD, Timmer et al. (2015)). Each bar reports the average welfare change when countries undertake unilateral markup correction without reciprocity by partners. Thea artificial parameters are constructed by fixing $\boldsymbol{\mu}$ to its estimated values and adjusting $\boldsymbol{\sigma}$to artificially inflate $Cov(\mu_k, \sigma_k)$.

reciprocity by trading partners. The results in Figure X.1 highlight two rudimentary points: First, while most countries experience a deterioration of welfare, a few do not. But even for those few countries dampened gains from correcting misallocation than if they were operating as closed economies. Second, trade-to-GDP (measured as the value of imports divided by GDP) is strongly associated with the intensity at which countries experience immiserizing growth. Figure X.1, moreover, reveals that countries not experiencing immiserizing growth tend to trade relatively more with each other. Hence, even if these countries adopt corrective industrial policies, it does not spare others from immiserizing growth—hence the importance of multilateral coordination of corrective policies via deep agreements.

Y Gains from Policy Under Alternative Assumptions

In this appendix we quantify the gains from optimal policy under three alternative scenarios, comparing them to the baseline gains reported in Section VI. In each case, we contrast the new policy gains with the baseline gains along the two dimensions: First, in terms of the gains from first-best trade and industrial policies. Second, in terms of the effectiveness of second-best trade taxes at replicating the first-best outcome.

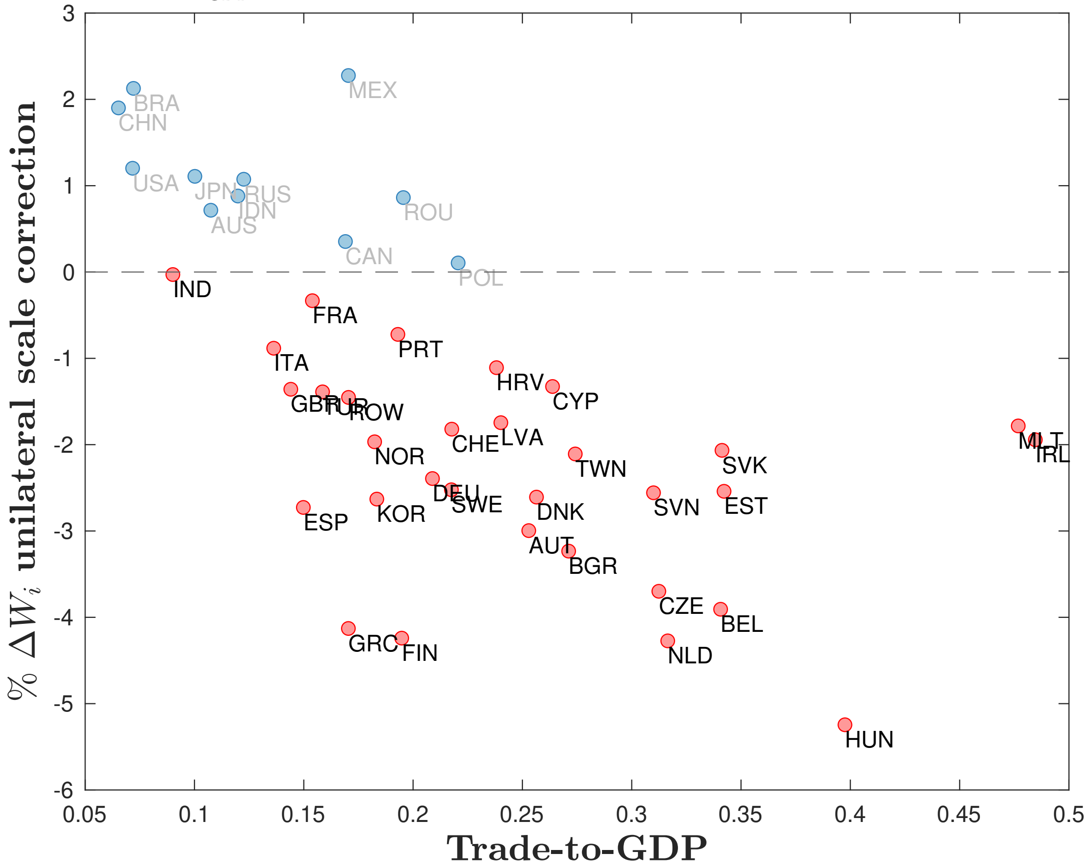
Y.1 Gains Implied by the Melitz-Pareto Model

Suppose the data generating process is consistent with a Melitz-Pareto model that accommodates firm-selection effects. In that case, Theorem 1 characterizes the optimal policy under the following reinterpretation of parameters—see Appendix D:

$$1+\mu_k^{\text{Melitz}} = \begin{cases} 1+\frac{1}{\vartheta_k} & \text{if entry is free} \\ \frac{\gamma_k\vartheta_k}{(\gamma_k-1)(\vartheta_k+1)-\vartheta_k} & \text{if entry is restricted} \end{cases}; \qquad \sigma_k^{\text{Melitz}} = 1+\frac{\vartheta_k}{1+\vartheta_k\left(\frac{1}{\sigma_k-1}-\frac{1}{\gamma_k-1}\right)}.$$

To compute the gains from policy we, therefore, need estimates for σ_k, γ_k, and ϑ_k. We have already produced estimates for the former two parameters. To estimate ϑ_k, we can first recover σ_k^{Melitz} using a standard gravity estimation à la Caliendo and Parro (2015). To explain the estimation procedure, suppose tariffs are applied before markups and industrial and export subsidies are zero ($x_{ji,k} = s_{j,k} = 0$ for all i, j, k). In that case, the national-level import demand function transforms into the following

Figure X.1: *Higher $\frac{Trade}{GDP}$ is associated with stronger immiserizing growth from unilateral corrective policies*



Note: The data source is the 2014 World Input-Output Database (WIOD, Timmer et al. (2015)). *The y-axis* corresponds to welfare gains when a country undertakes unilateral markup scale without reciprocity by partners.

industry-level gravity equation:[37]

$$\tilde{X}_{ji,k} \equiv \tilde{P}_{ji,k} Q_{ji,k} = \Phi_{j,k} \Omega_{i,k} \tau_{ji,k}^{1-\sigma_k^{\text{Melitz}}} (1 + t_{ji,k})^{1-\sigma_k^{\text{Melitz}}},$$

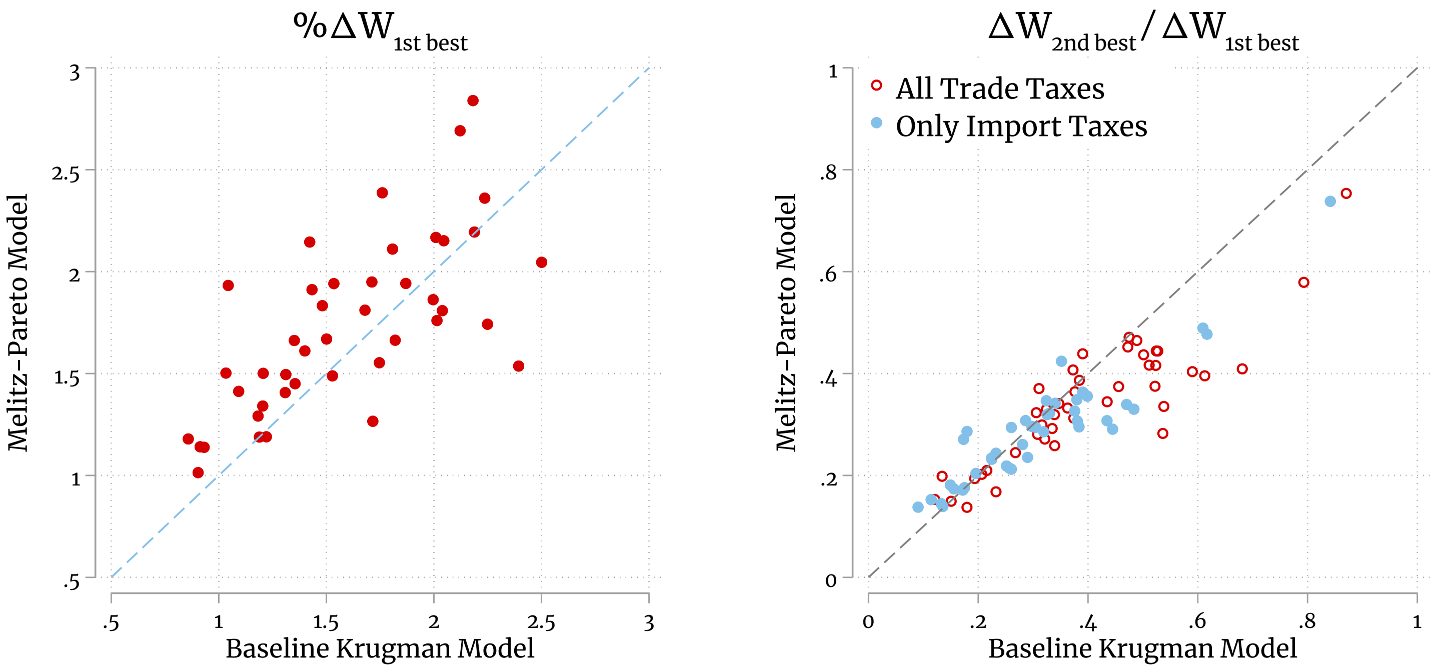
where $\Phi_{j,k} \equiv L_{j,k}^{\mu_k^{\text{Melitz}}} \bar{a}_{j,k}^{1-\sigma_k^{\text{Melitz}}} w_{j,k}^{1-\sigma_k^{\text{Melitz}}}$ and $\Omega_{i,k} \equiv \sum_n \left[\bar{a}_{n,k} w_{n,k}^{1-\sigma_k^{\text{Melitz}}} \tau_{ni,k}^{1-\sigma_k^{\text{Melitz}}} (1 + t_{ni,k})^{1-\sigma_k^{\text{Melitz}}} \right] e_{i,k} Y_{i,k}$ can be viewed as the exporter and importer fixed effects in the standard gravity estimation sense. To produce our final estimating equation, we assume that iceberg trade costs are given by $\ln \tau_{ji,k} = \ln d_{ji,k} + \varepsilon_{ji,k}$, where (i) $d_{ji,k} = d_{ij,k}$ is a systematic and symmetric cost component that accounts for the effect of distance, common language, and common border, while (ii) $\varepsilon_{ji,k}$ is a random disturbance term that represents any deviation from symmetry. Invoking this decomposition, we can produce the following estimating equation for any triplet (j, i, n):

$$\ln \frac{\tilde{X}_{ji,k} \tilde{X}_{in,k} \tilde{X}_{nj,k}}{\tilde{X}_{ij,k} \tilde{X}_{ni,k} \tilde{X}_{jn,k}} = -\left(\sigma_k^{\text{Melitz}} - 1\right) \ln \frac{(1 + t_{ji,k})(1 + t_{in,k})(1 + t_{nj,k})}{(1 + t_{ij,k})(1 + t_{ni,k})(1 + t_{jn,k})} + \varepsilon_{jin,k}.$$

The left-hand side variable, in the above equation, is composed of observable national-level trade values in industry k. The right-hand side variable is composed of observable industry-level tariff rates. The error term $\varepsilon_{jin,k} \equiv \theta_k(\varepsilon_{ij,k} - \varepsilon_{ji,k} + \varepsilon_{in,k} - \varepsilon_{ni,k} + \varepsilon_{nj,k} - \varepsilon_{jn,k})$ encompasses any idiosyncratic variation in non-tariff barriers. Under the identifying assumption that applied tariff rates are orthogonal to $\varepsilon_{jin,k}$, *i.e.*, $\mathbb{E}\left[t_{ji,k}\, \varepsilon_{ji,k}\right] = 0$, we can estimate σ_k^{Melitz} by estimation the above equation with data on

[37]The assumption that tariffs are applied before markups, amounts to saying that tariffs act as a cost-shifter. Alternatively, if tariffs are applied after markups, they act as a demand shifter. In the latter case, the elasticity of trade with respect to tariffs diverges from the trade elasticity in its standard definition—see Costinot and Rodríguez-Clare (2014) for more details.

Figure Y.1: The gains from policy under the Melitz-Pareto model



trade values, $\tilde{X}_{ji,k}$, and applied tariffs, $t_{ji,k}$, from the WIOD and TRAINS-UNCTAD datasets. After estimating σ_k^{Melitz}, we can recover ϑ_k for our previously-estimated values for σ_k and μ_k (which are reported in Table 3):

$$\vartheta_k = \frac{\hat{\sigma}_k^{\text{Melitz}} - 1}{1 + \left(\hat{\sigma}_k^{\text{Melitz}} - 1\right)\left(\frac{1}{\gamma_k - 1} - \frac{1}{\sigma_k - 1}\right)}.$$

For the analysis that follows, we borrow the estimated values for σ_k^{Melitz} from Lashkaripour (2020a; 2020b), which is based on the 2014 WIOD and TRAINS-UNCTAD datasets. After pinning down all the necessary parameters, we simply evaluate and plug σ_k^{Melitz} and μ_k^{Melitz} into our optimal tax formulas to compute the gains from optimal policy. The process is akin to that outlined in Section VI. Importantly, one should note that without our micro-level estimates for σ_k and μ_k, it is impossible to recover both σ_k^{Melitz} and μ_k^{Melitz} from macro-level trade and tariff data.

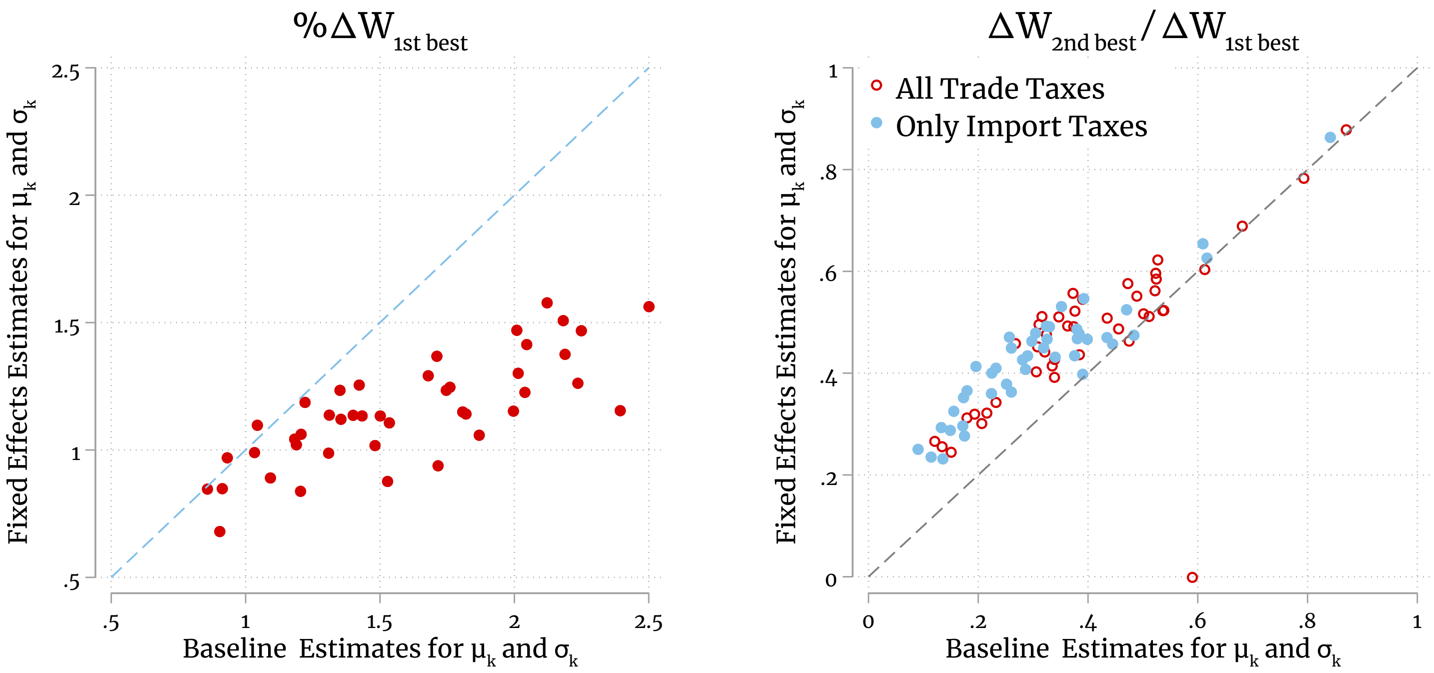
The optimal policy gains implied by the Melitz-Pareto model are reported under Figure Y.1. The results indicate that accounting for firm-selection (à la Melitz-Pareto) magnifies the gains from the first-best trade and industrial policy schedule. Moreover, accounting for firm-selection dampens the efficacy of second-best trade taxes at replicating the first-best policy gains. If anything, these results indicate that our baseline claim that trade taxes are an ineffective second-best substitute for industrial subsidies is strengthened once we account for firm-selection effects.

Y.2 Gains Implied by the Fixed-Effect Estimates for μ_k and σ_k

Our baseline estimation of the gains from policy in Section VI utilized the first-difference estimates for μ_k and σ_k—these estimates were reported under Table 3. In Appendix Q (under Table Q.1), we reported alternative estimates for μ_k and σ_k based on a two-ways fixed-effects estimation of the firm-level import demand function. In this appendix, we recompute the gains from policy using these alternative estimates for μ_k and σ_k.

The implied gains from optimal policy are reported under Figure Y.2. The fixed-effects estimates for σ_k and μ_k imply (on average) smaller gains from first-best trade and industrial policies. This outcome drives from two main factors: First, the fixed-effects estimates for μ_k exhibit smaller heterogeneity across industries. As such, they imply a small degree of misallocation in the economy compared to the baseline estimates. Second, the fixed-effects estimates for σ_k are generally smaller and imply larger unilateral gains from terms-of-trade manipulation.

Figure Y.2: The gains from policy under alternative estimates for σ_k and μ_k



Another takeaway from Figure Y.2 is that second-best trade taxes exhibit a greater degree of efficacy compared to the baseline case. This outcome reflects two issues: First, the corrective gains from policy are a smaller fraction of the overall first-best policy gains, once we plug the fixed-effects-estimated values for σ_k and μ_k. Second, the fixed-effects-estimated values for σ_k and μ_k exhibit a smaller negative correlation relative to the baseline estimates. As explained in Section IV, the less negative $Cov\left(\sigma_k, \mu_k\right)$, the smaller the implicit tensions between the terms-of-trade-improving and corrective gains from trade taxation—hence, the greater efficacy of second-best trade taxes.

Y.3 Assigning Alternative Values to μ_k and σ_k for the Service Sector

Our estimation of σ_k and μ_k in Section V relied on transaction-level trade data, which is scarce for (semi-non-traded) service industries. To address this issue, our baseline estimation of the gains from policy normalized the aforementioned parameters in service-related industries as follows:

$$\sigma_k = 11; \qquad \mu_k = 0 \qquad \text{if } k \in \text{Service}$$

The value assigned to σ_k for service-related industries is less consequential for our estimated welfare gains. The reason is that σ_k governs the gains from terms-of-trade manipulation. However, under the status quo, there is little-to-no trade occurring in service industries. With little-to-no service trade under the status quo equilibrium, the scope for terms-of-trade manipulation is limited in service industries—all irrespective of the value assigned to σ_k.[38]

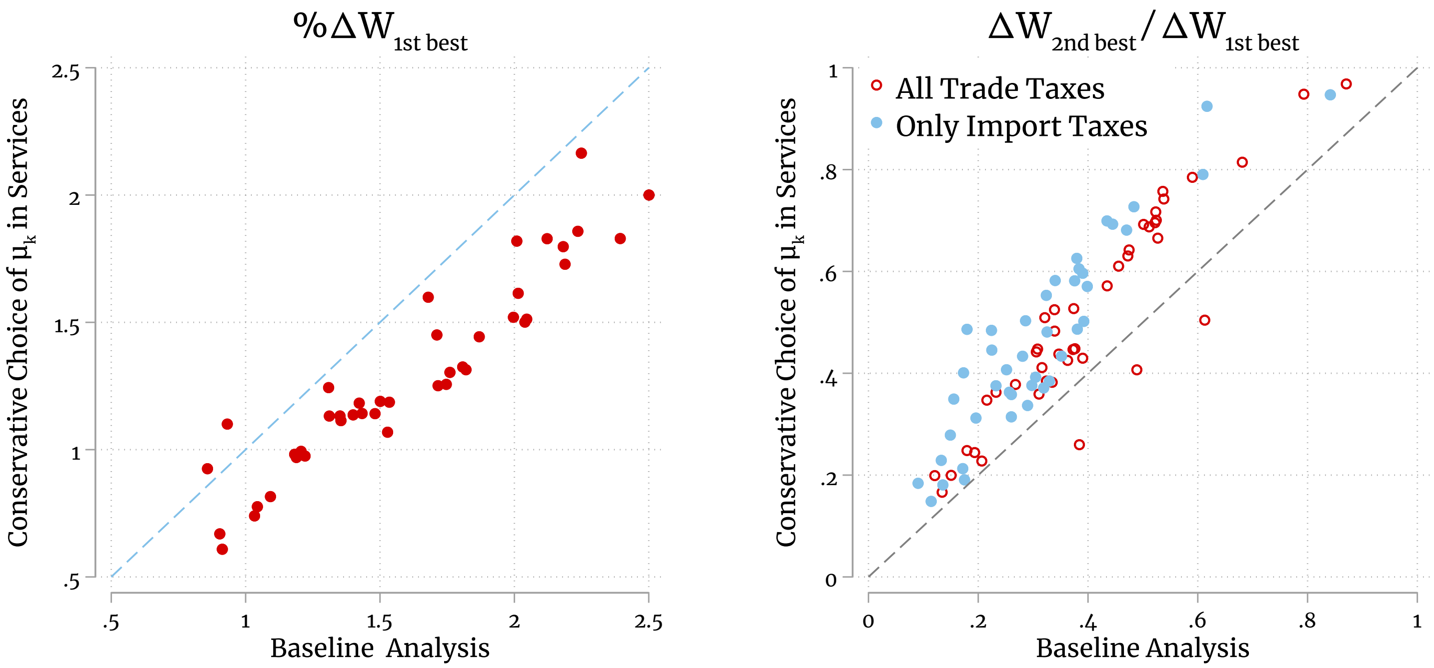
The value assigned to μ_k, however, can have a profound effect on the estimated gains from optimal policy. To elaborate on this point, recall that one function of optimal policy (in our framework) is to correct misallocation due to markup heterogeneity. The degree of misallocation can be crudely measured as the cross-industry variance in markups, i.e., $\text{Var}(\mu_k)$. Data indicate that the service

[38]This outcome is an artifact of the CES parametrization of import demand. Specifically, in response to a change, $\hat{\boldsymbol{\tau}}$, in trade taxes, the post-tax-change expenditure shares remain zero if start as zero in the initial equilibrium—all irrespective of the trade elasticity values. Stated in mathematical terms,

$$\lim_{\lambda_{ji,k}\to 0} \hat{\lambda}_{ji,k} = \frac{\lambda_{ji,k}\left(\hat{\tau}_{ji,k}\hat{w}_j\right)^{1-\sigma_k}}{\sum_n \lambda_{ni,k}\left(\hat{\tau}_{ni,k}\hat{w}_n\right)^{1-\sigma_k}} = 0 \qquad \forall \sigma_k \geq 1.$$

Since $\lambda_{ji,k} \approx 0$ in services, trade taxes have little-to-no ability at improving the terms-of-trade, as doing so requires policy to shrink exports/imports in the service sector away from their factual level.

Figure Y.3: The gains from policy when the service sector is modeled more conservatively



sector constitutes a non-trivial fraction of total output in each country. So, the value assigned to the service sector's μ_k is a non-trivial determinant of misallocation, as measured by $\mathrm{Var}(\mu_k)$.

As indicated above, our baseline analysis assumed that the service sector is perfectly competitive. This assumption, which is rather standard in the quantitative trade literature, amounts to setting $\mu_k = 0$ for any service-related industry, k. In this appendix, we contrast our baseline results with those obtained under the alternative but extremely conservative assumption that μ_k in services equals the average μ_k in traded (non-service) industries. This assumption is conservative because it artificially deflates $\mathrm{Var}(\mu_k)$ and, accordingly, dampens the corrective gains from optimal policy.

The gains computed under our conservative treatment of the service sector are reported under Figure Y.3. As expected, the gains from first-best policies are relatively lower under the conservative treatment—simply because the conservative value assigned to the service sector markup artificially lowers the degree of misallocation and the scope for policy intervention. Relatedly, second-best trade taxes are also more successful at replicating the gains obtainable under the first-best policy schedule. The intuition is that the *corrective* gains from policy constitute a smaller fraction of the first-best policy gains under the conservative model. Hence, the inability of trade taxes to replicate corrective gains becomes less consequential.

Z The Gains from Policy Under Artificial Parameter Values

Under what parameter values will our framework predict larger gains from policy? To answer this question, we simulate an artificial economy (with artificial values assigned to σ_k and μ_k) to examine the degree to which the gains from policy inflate under more extreme parameter values. Our theory indicates that the gains from optimal policy are regulate by two key statistics:

i. The variance in the industry-level scale elasticities, $\mathrm{Var}\left[\log \mu_k\right]$.

ii. The average level of the (inverse) industry-level trade elasticities, $\mathbb{E}\left[\frac{1}{\sigma_k - 1}\right]$.

The first statistic governs the extent to which countries can gain from restoring allocative efficiency. To explain this statistic, we can appeal to the Hsieh and Klenow (2009) exact formula for distance from the efficient frontier. Considering that preferences are Cobb-Douglas across industries, the distance from the efficient frontier in each country (net of trade effects) can be approximated to a first-order as

$$\text{Distance from efficient frontier} \approx \frac{1}{2}\mathrm{Var}\left[\log \mu_k\right].$$

The average level of μ_k is, however, inconsequential. To convey this point, suppose we multiply all the markups by some number $a \in \mathbb{R}_+$. Since this change is akin to offering a uniform industrial subsidy a to all industries, then it preserves real welfare based on Lemma 1.

The second statistic determines the degree of national-level market power and, thus, governs the degree to which countries can gain from ToT manipulation. To explain this statistic succinctly, consider a country that is sufficiently small in relation to the rest of the world. Following Theorem 1, the average optimal trade tax for this country is given by

$$\text{Avg. optimal trade tax} \approx \mathbb{E}\left[\frac{1}{\sigma_k - 1}\right].$$

If $\sigma_k \to \infty$ for all k, the average optimal trade tax approaches zero, leaving no room for unilateral ToT improvements. Conversely, as σ_k approaches 1 the average optimal trade tax increases and so do the implicit gains from unilateral trade restrictions.[39]

Noting the above background, we recompute the gains from policy by artificially increasing $\text{Var}\left[\log \mu_k\right]$ and decreasing $\mathbb{E}\left[\frac{1}{\sigma_k-1}\right]$, starting from our estimated vectors of $\{\sigma_k\}$ and $\{\mu_k\}$. The results are reported in Z.1 for a select set of countries—namely, the United States, China, Indonesia, and Korea. The graph indicates that the gains from policy nearly double if we artificially raise $\text{Var}\left[\log \mu_k\right]$ by a factor of two. A similar effect is borne out if we artificially raise $\mathbb{E}\left[\frac{1}{\sigma_k-1}\right]$ by a factor of about two. An apparent pattern, here, is that the gains from policy exhibit similar sensitivity levels to $\text{Var}\left[\log \mu_k\right]$ across all countries, but different sensitivity levels to $\mathbb{E}\left[\frac{1}{\sigma_k-1}\right]$. This pattern is expected, because $\mathbb{E}\left[\frac{1}{\sigma_k-1}\right]$ governs the gains from ToT-improvement which are smaller (by design) for larger economies like the United States or China. The gains for restoring allocative efficiency, however, depend less on size and more on a country's industrial pattern of specialization under the status quo—see Kucheryavyy et al. (2023a) for the role of specialization patterns.

These findings provide a platform to compare our estimated gains with alternatives in the literature. Our finding that the gains from restoring allocative efficiency are large sits well with the findings in Baqaee and Farhi (2017) that eliminating sectoral markup-heterogeneity in the U.S. economy can raise real GDP by 2.3%.[40] Bartelme et al. (2019), however, estimate smaller gains from similar policies. To understand these differences, note the formula for distance from the efficient frontier. Also note that *true* value for the scale elasticity, $\mu_k^{\text{True}} = \mu_k + \psi_k$, where ψ_k denotes the elasticity of Marshallian externalities. Accordingly, the *true* distance from the frontier can be approximated as follows:

$$\mathcal{L}_{\text{True}} \approx \frac{1}{2}\text{Var}\left[\log\left(\psi_k + \mu_k\right)\right]$$

Our analysis like Baqaee and Farhi (2017) sets $\psi_k = 0$, and measures the degree of allocative inefficiency as $\mathcal{L}_{LL} \approx \frac{1}{2}\text{Var}\left[\log\left(\psi_k + \mu_k\right)\right]$. This approach can lead to an overstatement of $\mathcal{L}$ if ψ_k is negatively correlated with firm-level market power, μ_k.[41] In comparison and as noted in Section V.D, the degree of misallocation in BDCR's analysis is measured as $\mathcal{L}_{BCDR} \approx \frac{1}{2}\text{Var}\left[\log\left(\mu_k + \psi_k - \frac{\beta_k}{\sigma-1}\right)\right]$, where β_k is the share of industry-specific factors in production. This approach can understate $\mathcal{L}$ when there are significant diseconomies of scale due to a high β_k.

References

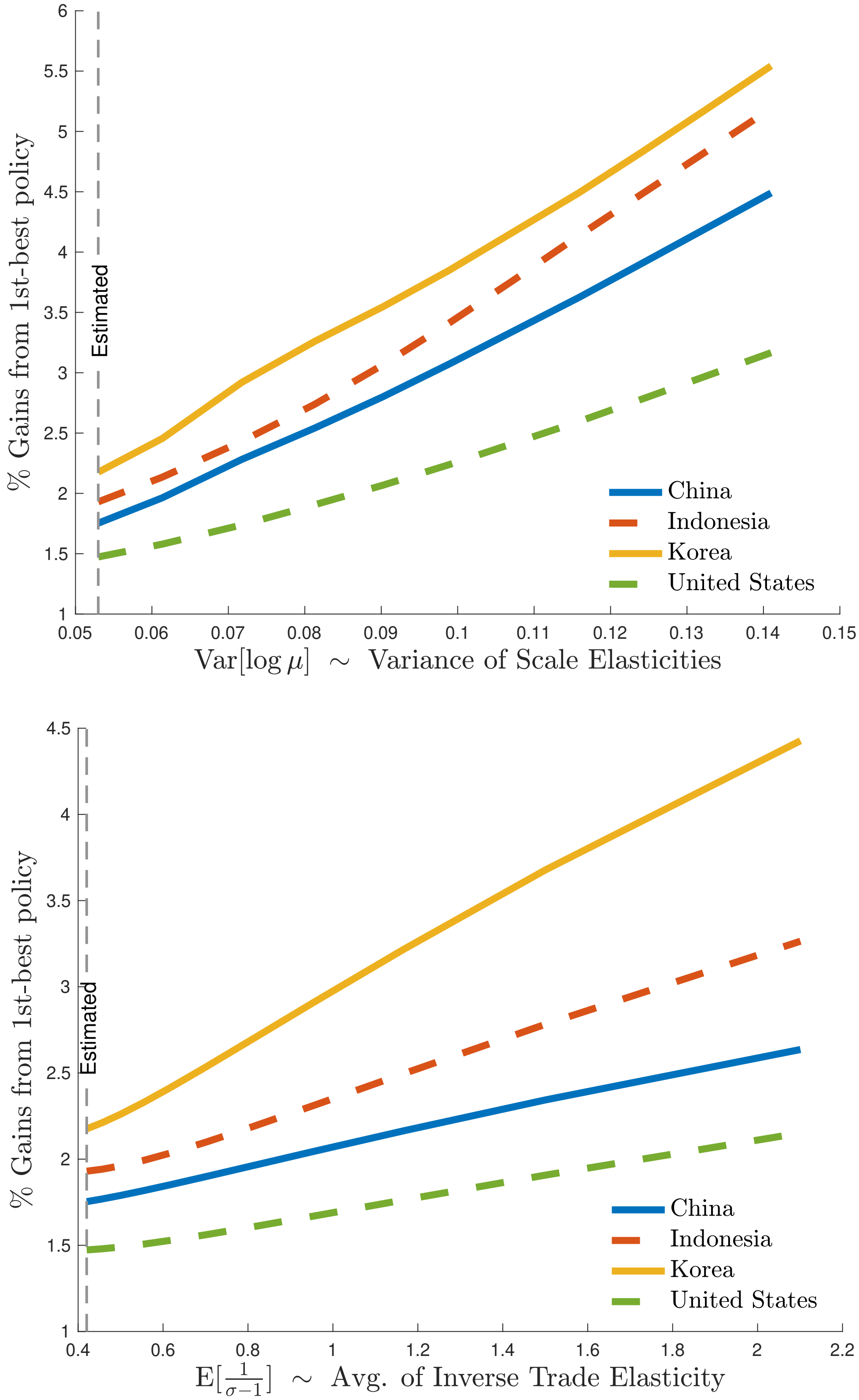
Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

[39] Since there is no choke price in our setup, the optimal export tax can approach infinity in the limit where $\sigma_k \to 1$. Introduce a choke price, then the optimal export tax will exhibit a limit-pricing formulation—see Costinot, Donaldson, Vogel, and Werning (2015).

[40] This number corresponds to the average of the numbers reported in the last column of Table 2 in Baqaee and Farhi (2017).

[41] Another issue is that we are assuming away selection effects in our quantitative analysis. In the presence of selection effects, we can still use our estimates for σ_k and μ_k to identify the scale elasticity up-to an externally chosen trade elasticity. Doing so, however, may lead to a lower or higher $\mathcal{L}$.

Figure Z.1: *The gains from policy under artificially higher Var* $[\log \mu_k]$ *and* $\mathbb{E}\left[\frac{1}{\sigma_k - 1}\right]$



Note: The data source is the 2014 World Input-Output Database (WIOD, Timmer et al. (2015)). The *1st best* policy is characterized by Theorem 1 for the case of restricted entry.

Adao, R., M. Kolesár, and E. Morales (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics 134*(4), 1949–2010.

Alvarez, F. and R. E. Lucas (2007). General equilibrium analysis of the Eaton–Kortum model of international trade. *Journal of Monetary Economics 54*(6), 1726–1768.

Angrist, J., G. Imbens, and D. Rubin (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association 91*, 444–455.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). Clare, 2012, new trade models, same old gains. *The American Economic Review* (1), 94.

Atkin, D. and D. Donaldson (2021). The role of trade in economic development. Technical report, National Bureau of Economic Research.

Bagwell, K. and R. W. Staiger (2001). Domestic policies, national sovereignty, and international economic institutions. *The Quarterly Journal of Economics 116*(2), 519–562.

Bagwell, K. and R. W. Staiger (2004). *The economics of the world trading system*. MIT press.

Bank of Canada (2017). Exchange rates: Historical noon and closing rates. https://www.bankofcanada.ca/rates/exchange/legacy-noon-and-closing-rates/. (accessed between 2015 and 2016).

Baqaee, D. and E. Farhi (2019). Networks, barriers, and trade. Technical report, National Bureau of Economic Research.

Baqaee, D. and E. Farhi (2020a). Entry vs. rents. Technical report, National Bureau of Economic Research.

Baqaee, D. R. and E. Farhi (2017). Productivity and misallocation in general equilibrium. Technical report, National Bureau of Economic Research.

Baqaee, D. R. and E. Farhi (2020b). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Bartelme, D., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2019). The Textbook Case for Industrial Policy: Theory Meets Data. *Working Paper*.

Basu, S. and J. G. Fernald (1997). Returns to Scale in US Production: Estimates and Implications. *Journal of Political Economy 105*(2), 249–283.

Benassy, J.-P. (1996). Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters 52*(1), 41–47.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Boehm, C. E., A. A. Levchenko, and N. Pandalai-Nayar (2020). The long and short (run) of trade elasticities. Technical report, National Bureau of Economic Research.

Broda, C. and D. E. Weinstein (2006). Globalization and the Gains from Variety. *The Quarterly Journal of Economics 121*(2), 541–585.

Caliendo, L. and F. Parro (2015). Estimates of the Trade and Welfare Effects of NAFTA. *Review of Economic Studies 82*(1), 1–44.

Costinot, A., D. Donaldson, J. Vogel, and I. Werning (2015). Comparative Advantage and Optimal Trade Policy. *The Quarterly Journal of Economics 130*(2), 659–702.

Costinot, A. and A. Rodríguez-Clare (2014). Trade theory with numbers: Quantifying the consequences of globalization. In *Handbook of international economics*, Volume 4, pp. 197–261. Elsevier.

DANE (Colombian Statistical Agency) (2017). Economics and social correlative tables. https://www.dane.gov.co/files/sen/nomenclatura/tablasCorrelativas/CORARAN-v44-oct-2017.xls.

De Loecker, J. and F. Warzynski (2012). Markups and firm-level export status. *American economic review 102*(6), 2437–71.

Demidova, S. and A. Rodriguez-Clare (2009). Trade policy under firm-level heterogeneity in a small economy. *Journal of International Economics 78*(1), 100–112.

Dixit, A. and V. Norman (1980). *Theory of international trade: A dual, general equilibrium approach*. Cambridge University Press.

Eaton, J. and S. Kortum (2001). Technology, trade, and growth: A unified framework. *European Economic Eeview 45*(4), 742–755.

Eaton, J. and S. Kortum (2002). Technology, Geography, and Trade. *Econometrica 70*(5), 1741–1779.

EDD by World Bank (2016). Exporter dynamics database version 2.0, 1997-2014,. https://doi.org/10.48529/agcr-yt74.

Epifani, P. and G. Gancia (2011). Trade, markup heterogeneity and misallocations. *Journal of International Economics 83*(1), 1–13.

Fajgelbaum, P., G. M. Grossman, and E. Helpman (2011). Income Distribution, Product Quality, and International Trade. *Journal of Political Economy 118*(4), 721.

Feenstra, R. C. (1994). New Product Varieties and the Measurement of International Prices. *The American Economic Review*, 157–177.

Feenstra, R. C. (2015). *Advanced international trade: theory and evidence*. Princeton university press.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2021). Penn world table version 9.1. Groningen Growth and Development Centre [distributor], https://www.rug.nl/ggdc/productivity/pwt/?lang=en (accessed October 6th of 2022).

Fernandes, A. M., C. Freund, and M. D. Pierola (2016). Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database. *Journal of Development Economics 119*, 121–137.

Gros, D. (1987). A note on the optimal tariff, retaliation and the welfare loss from tariff wars in a framework with intra-industry trade. *Journal of International Economics 23*(3-4), 357–367.

Hicks, J. R. (1939). The foundations of welfare economics. *The economic journal 49*(196), 696–712.

Horn, R. A. and C. R. Johnson (2012). *Matrix analysis*. Cambridge university press.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies 45*(3), 511–518.

Hummels, D., V. Lugovskyy, and A. Skiba (2009). The trade reducing effects of market power in international shipping. *Journal of Development Economics 89*(1), 84–97.

Kaldor, N. (1939). Welfare propositions of economics and interpersonal comparisons of utility. *The economic journal*, 549–552.

Kortum, S. S. (1997). Research, Patenting, and Technological Change. *Econometrica 65*(6), 1389–1420.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review 70*(5), 950–959.

Kucheryavyy, K., G. Lyn, and A. Rodríguez-Clare (2023a). Grounded by gravity: A well-behaved trade model with industry-level economies of scale. *American Economic Journal: Macroeconomics 15*(2), 372–412.

Kucheryavyy, K., G. Lyn, and A. Rodríguez-Clare (2023b). Replication data for: Grounded by gravity: A well-behaved trade model with industry-level economies of scale. American Economic Journal: Macroeconomics [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.1257/mac.20190156.

Lashkaripour, A. (2020a). Can Trade Taxes be a Major Source of Government Revenue? *Journal of the European Economic Association 19*(5), 2399–2428.

Lashkaripour, A. (2020b). Data for: Can Trade Taxes be a Major Source of Government Revenue? Journal of the European Economic Association [publisher], https://doi.org/10.1093/jeea/jvaa058.

Lashkaripour, A. (2021). The cost of a global tariff war: A sufficient statistics approach. *Journal of International Economics 131*, 103419.

Manski, C. F. and D. McFadden (1981). *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press Cambridge, MA.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995). *Microeconomic theory*, Volume 1. Oxford university press New York.

Melitz, M. (2003). The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations. *Econometrica 71*(6), 1695–1725.

Ossa, R. (2011). A "New Trade" Theory of GATT/WTO Negotiations. *Journal of Political Economy 119*(1), 122–152.

Ossa, R. (2014). Trade Wars and Trade Talks with Data. *The American Economic Review 104*(12), 4104–46.

Ossa, R. (2016). Quantitative Models of Commercial Policy. In *Handbook of Commercial Policy*, Volume 1, pp. 207–259. Elsevier.

Ostrowski, A. M. (1952). Note on bounds for determinants with dominant principal diagonal. *Proceedings of the American Mathematical Society 3*(1), 26–30.

Ramondo, N., A. Rodríguez-Clare, and M. Saborío-Rodríguez (2016, October). Trade, Domestic Frictions, and Scale Effects. *The American Economic Review 106*(10), 3159–84.

Shapiro, J. S. (2016, November). Trade costs, co2, and the environment. *American Economic Journal: Economic Policy 8*(4), 220–54.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. De Vries (2015). An illustrated user guide to the world input–output database: the case of global automotive production. *Review of International Economics 23*(3), 575–605.

Weyl, E. G. and M. Fabinger (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy 121*(3), 528–583.

WIOD (2021). World input-output database 2016 release, 2000-2014. Groningen Growth and Development Centre [distributor], https://doi.org/10.34894/PJ2M1C.

World Integrated Trade Solutions (WITS) (2010). Product concordance. https://wits.worldbank.org/product_concordance.html.

Wu, M., B. Yin, A. Vosoughi, C. Studer, J. R. Cavallaro, and C. Dick (2013). Approximate matrix inversion for high-throughput data detection in the large-scale mimo uplink. In *2013 IEEE international symposium on circuits and systems (ISCAS)*, pp. 2155–2158. IEEE.