Trade, Firm-Delocation, and Optimal Environmental Policy

Farid Farrokhi
Purdue University

Ahmad Lashkaripour
Indiana University

August 26, 2020

Abstract

To what extent can trade policy serve as a remedy for environmental pollution? We examine this question using a multi-country multi-industry general equilibrium trade model with transboundary pollution externalities. Our framework features important margins such as firm-delocation in response to policy, multilateral carbon leakage, and returns to scale in abatement. Our central result is a set of simple formulas for unilaterally optimal trade and emission taxes in an open economy. The optimal policy consists of (i) a uniform emission tax across all industries; (ii) industry-level production subsidies that restore marginal-cost-pricing independent of the industry's emission intensity; (iii) industry-level import taxes that penalize high-emission imports but less so in high-returns-to-scale industries; and (vi) industry-level export taxes that (in addition to improving the terms of trade) promote clean exports against high-emission foreign competition. These formulas reveal a tension between emission reduction and scale economies that limits the efficacy of trade taxes at correcting transboundary emission. Our formulas parsimoniously map to data, enabling us to uncover the full potential of trade policy at tackling global emission.

1 Introduction

Despite growing concerns over greenhouse gas emissions, international agreements such as the 1997 KYOTO PROTOCOL and the 2015 PARIS CLIMATE ACCORD have failed to deliver desirable outcomes. This failure has prompted experts to propose agreements that incentivize global cooperation with a mixture of environmental taxes and trade policies (Nordhaus (2015)). Relatedly, some experts have advocated for sub-global agreements wherein a bloc of nations use trade policy to achieve extraterritorial environmental objectives.[1]

[1] See Copeland et Taylor (2004) for a survey of the literature on trade and environment.

The basic idea behind these proposals is that governments can tackle cross-border carbon leakage or penalize non-cooperative governments with trade taxes. Evaluating these proposals is challenging since it requires a full characterization of each government's optimal environmental and trade policy in a multilateral, general equilibrium setting.

Our understanding of optimal trade taxes under environmental concerns is, meanwhile, limited. On the theoretical side, the literature has analyzed the use of trade taxes for environmental objectives in partial equilibrium or two-country settings. These analyses typically abstract from product differentiation, abatement decisions, or scale economies, and are difficult to map to data for quantitative applications (see e.g., Markusen (1975)).[2] On the quantitative side, most research has circumvented the task of computing *optimal* trade taxes for environmental objectives. Instead, trade taxes have been chosen sub-optimally based on easy-to-implement criteria. As such, little is known about the full potential of trade taxes at tackling carbon leakage or enforcing global environmental cooperation.

In this paper, we characterize the unilaterally optimal trade and environmental policy in a multi-country multi-industry general equilibrium framework that incorporates international trade costs, product differentiation, endogenous entry, scale economies, and firm-level abatement decisions. We also characterize the optimal policy under second-best scenarios, in which governments are banned from using certain policy instruments.

To this goal, we produce an intermediate *envelop result* that helps us overcome the main challenges in characterizing optimal policy in general equilibrium. Our result concerns the welfare-neutrality of general equilibrium wage and income effects at the optimum. This envelope result holds under general conditions, and can be employed for policy analysis in many alternative setups.

We use our theory to cast light on several unresolved questions: First, to what extent could trade taxes correct transboundary emission externalities and cross-border pollution leakages? Second, can trade taxes serve as an effective remedy for the free-riding problem in environmental agreements? Third, how much can unilateralism achieve relative to global cooperation, insofar as environmental objectives are concerned?

Our framework exhibits two properties that make it suitable for addressing these questions. First, our model accommodates endogenous entry and tracks the relocation of firms across space and industries in response to policy. These responses generate scale economies in both production and abatement. To put this property in perspective, the prior literature on trade and environmental pol-

[2]Several papers have advanced the pioneering work of Markusen (1975). See Sections 2 and 3 of Sturm (2003) for a detailed survey of subsequent works.

icy has often overlooked issues involving endogenous entry, economies of scale, or firm-delocation. However, recent evidence suggest that these features are crucial to how firms respond to environmental and trade policy changes (Shapiro et Walker (2018)).[3]

Second, our theory delivers sufficient statistics formulas for optimal taxes on emission, production, and trade, which can be readily mapped to data. This feature simplifies and enriches our quantitative analysis in several ways. Above all, our analytic formulas simplify the task of computing *optimal* trade taxes under environmental concerns —a task that has often eluded the past literature. Our analytical characterization of optimal policy delivers the minimum set of elasticities and baseline data required for quantitative analysis. Invoking this feature, we can uncover the full potential of trade taxes at delivering environmental objectives, and tractably account for previously-neglected interactions between firm entry, abatement, and scale economies.

At another level, our theory uncovers a set of new environmental trade-offs that trade taxes face. Most importantly, we find that endogenous entry creates scale economies in abatement, which dampen the effectiveness of optimal trade taxes at correcting cross-border carbon leakage. Specifically, governments may compromise on tackling cross-border carbon leakage by lowering taxes on dirty imports in order to exploit scale effects in abatement. Furthermore, we find that (absent environmental concerns) non-cooperative tariffs subsidize high-returns to scale industries. Since high returns-to-scale industries are also carbon-intensive, our finding highlights a new terms-of-trade-driven rationale for the environmental bias of trade policy—a bias that has been recently documented by Shapiro (2020).

Related Literature

Our work is related to several strands of literature. We integrate efforts to characterize optimal policies in modern trade theories with the literature on trade and environment, in a manner that can be connected to data for quantitative analyses.

First, we contribute to the theoretical literature on optimal trade and emission taxes in open economy. A central insight from this literature is that optimal unilateral tariffs include a tax on transboundary emission (e.g., Markusen (1975); Copeland (1996)). For all its merits, this body of literature has generally relied on partial equilibrium or two-country models that abstract from product dif-

[3] A special case of our framework is a multi-industry Ricardian model that itself nests Dornbusch *et al.* (1977) and Eaton et Kortum (2002). We use this special case to highlight the policy effects derived from non-Ricardian forces such as scale economies.

ferentiation, endogenous abatement, or firm-delocation. As a result, the results from this literature have been rarely used to guide the quantitative analysis of trade and environmental policy. We complement this literature by characterizing optimal policy in a multi-country general equilibrium trade model that accommodates several previously-overlooked features of the global economy and is straightforward to calibrate to data.

Second, our analysis is related to an emerging body of quantitative work that analyzes the efficacy of trade policy at tackling environmental pollution (e.g., Babiker (2005); Elliott *et al.* (2010); Nordhaus (2015); Böhringer *et al.* (2016)). Despite their rich structure, existing analyses have mostly quantified the efficacy of easy-to-implement but sub-optimal trade policy initiatives. This approach allows researchers to circumvent the computational difficulties associated with optimal policy analysis. However, it does not uncover the full potential of trade taxes at tackling environmental pollution. In comparison, we derive analytic formulas for optimal policy, which help us bypass computational difficulties, making us able to uncover the full potential of trade taxes at tackling environmental pollution.

Third, our intermediate envelope result speaks to an emerging literature that studies optimal policy in modern quantitative trade models (Costinot *et al.* (2015, 2016); Lashkaripour et Lugovskyy (2016); Bartelme *et al.* (2019); Beshkar et Lashkaripour (2020)). These studies have bridged a long-standing divide between classic partial equilibrium trade policy models and modern general equilibrium trade theories. This divide is partly driven by classic trade policy models assuming away general equilibrium wage and income effects. Our envelope result is a step forward in filling this divide. Specifically, it shows that the simplifying assumptions in dealing with wage and income effects can be relaxed without sacrificing richness of analysis.

Finally, we contribute to the ongoing revival and enhancement of quantitative trade theories. Over the past two decades, quantitative trade models have been enriched to account for firm-selection, scale economies, input-output linkages, multinational production, and more (Costinot et Rodríguez-Clare (2014)). But less attention has been paid to embedding pollution externalities into the state-of-the-art quantitative trade models (Cherniwchan *et al.* (2017)). Our conceptual framework and optimal policy formulas can help bridge this gap. They can enable future analyses of trade liberalization to formally account for environmental costs and benefits.

This paper is organized as follows: In Section 2 we present our theoretical framework. In Section 3 we present our intermediate envelope result which we use to derive simple formulas for optimal unilateral policy. In Section 4 we discuss second-best scenarios as well as international cooperative or non-cooperative policies. In Section 5 we map our theory with our optimal policy formulas to data in order to quantify the efficacy of trade policy at tackling environmental pollution.

2 Theoretical Setup

The global economy consists of multiple countries indexed by $i, j, n \in \mathbb{C}$ and multiple industries indexed by $k, g \in \mathbb{K}$. Each country i is endowed by $\bar{L}_i$ units of workers who are perfectly mobile across industries but immobile across countries.

2.1 Demand

We denote by subscript ji, k the composite variety that corresponds to supplying country j, purchasing country i, industry k. The representative consumer in country i maximizes a non-parametric utility function $U_i(\boldsymbol{Q}_i)$ by choosing the vector of quantities, $\boldsymbol{Q}_i = \{Q_{ji,k}\}_{j\in\mathbb{C},, k\in\mathbb{K}}$ subject to the budget constraint,

$$Y_i = \sum_j \sum_k \tilde{P}_{ji,k} Q_{ji,k} + \bar{D}_i, \tag{1}$$

where Y_i is national income, $\bar{D}_i$ is exogenous national debt (trade deficits), and $\tilde{P}_{ji,k}$ denotes the consumer price index of composite variety ji, k. The tilde notation on price distinguishes between after-tax consumer prices ($\tilde{P}_{ji,k}$) and before-tax producer prices ($P_{ji,k}$). Let $\tilde{\mathbf{P}}_i = \{\tilde{P}_{ji,k}\}$ be the vector of after-tax consumer prices in country i. The consumer problem implies $V_i(Y_i, \tilde{\mathbf{P}}_i)$ as indirect utility function, and $Q_{ji,k} = \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$ as quantity of demand for variety ji, k. The behavior of the demand function is characterized by a set of demand elasticities. First, the elasticity of demand for (ji, k) relative to price of every variety (ni, g),

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

with $\varepsilon_{ji,k} \equiv \varepsilon^{(ji,k)}_{ji,k}$ denoting the own-price elasticity. Second, the elasticity of demand for ji, k relative to income is:

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

2.2 Supply

Firms and Market Structure. Production in every country-industry j, k takes place by monopolistically competitive firms indexed by $\omega \in \Omega_{j,k}$. Firms employ labor for entry and production. A large pool of ex-ante identical firms can pay entry costs $w_j \bar{f}_{j,k}$ to serve markets with their firm-level variety. w_j is labor wage and $\bar{f}_{j,k}$ is the labor requirement for entry. Upon entry, every firm ω may devote a fraction $a_{j,k}(\omega) \in [0,1]$ of its labor input to abatement activities, and the rest to production. The choice of $a_{j,k}(\omega)$ is regulated by country-industry specific pollution tax, $\tau_{j,k}$. Firms choose the abatement $a_{j,k}(\omega) \equiv a_{j,k}$, deliver the output quantity $q_{ji,k}(\omega) \equiv q_{jik}$ to every market i where they incur iceberg trade costs $\bar{d}_{ji,k} \geq 1$ with $\bar{d}_{jj,k} = 1$,[4] and as an externality they generate pollution $z_{ji,k}(\omega) \equiv z_{ji,k}$.

Following Copeland et Taylor (2004), the pollution per unit of output in the location of firm, $z_{ji,k}/(d_{ji,k} q_{ji,k})$, equals $(1 - a_{j,k})^{1/\alpha_k - 1}$. Here, $\alpha_k > 0$ is the "pollution elasticity" which varies across industries $k \in \mathbb{K}$. Given a choice of abatement $a_{j,k}$, marginal cost of production equals $c_{ji,k} = \bar{d}_{ji,k}(1-\alpha_k)^{-1}(1-a_{j,k})^{-1}(w_j/\bar{\varphi}_{j,k})$, where $\bar{\varphi}_{j,k}$ is labor productivity. A higher level of abatement means less pollution, and so, less pollution taxes paid by the firm, whereas it raises the marginal cost of production.

Industry-Level Aggregates. The composite output of ji, k, $Q_{ji,k}$, aggregates over firm-level quantities $q_{ji,k}(\omega)$,

$$Q_{ji,k} = \left(\int_{\omega \in \Omega_{j,k}} q_{ji,k}(\omega)^{\frac{\gamma_k - 1}{\gamma_k}} d\omega \right)^{\frac{\gamma_k}{\gamma_k - 1}}$$

where $\gamma_k > 1$ is the elasticity of substitution across firm-level varieties. Facing this substitution elasticity, firms charge a constant markup over their marginal cost, implying the following industry-level producer price index:

$$P_{ji,k} = M_{j,k}^{\frac{1}{1-\gamma_k}} \frac{\gamma_k}{\gamma_k - 1} \frac{\bar{d}_{ji,k} w_j}{\bar{\varphi}_{j,k}(1-\alpha_k)(1-a_{j,k})}, \qquad \text{(Price)}$$

where $M_{j,k} \equiv |\Omega_{j,k}|$ denotes the mass of firms. $M_{j,k}$ is pinned down by the free entry condition, that requires entry costs, $M_{j,k} w_j \bar{f}_{j,k}$, to equal gross profits, $\sum_i \frac{1}{\gamma_k} P_{ji,k} Q_{ji,k}$. Putting these together with $P_{ji,k} = \bar{d}_{ji,k} P_{jj,k}$ and $Q_{j,k} = \sum_i \bar{d}_{ji,k} Q_{ji,k}$, implies the mass of firms:

$$M_{j,k} = \frac{P_{jj,k} Q_{j,k}}{\gamma_k \bar{f}_{j,k} w_j} \qquad \text{(Entry)}$$

[4] Our model allows for nontradeables: Products in industry k are nontradeable if $\bar{d}_{ji,k} \to \infty$ for all i, $j \neq i$.

Using the equations of (Price) and (Entry), we rewrite the following industry-level variables as functions of abatement and quantities:

$$P_{ji,k}(w_j, a_{j,k}; Q_{j,k}) = \bar{d}_{ji,k}\bar{p}_{jj,k}w_j(1-a_{j,k})^{\frac{1}{\gamma_k}-1}Q_{j,k}^{-\frac{1}{\gamma_k}} \tag{2}$$

$$Z_{j,k}(a_{jk}; Q_{j,k}) = \bar{z}_{j,k}(1-a_{j,k})^{\frac{1}{\alpha_k}+\frac{1}{\gamma_k}-1}Q_{j,k}^{1-\frac{1}{\gamma_k}} \tag{3}$$

$$M_{j,k}(a_{jk}; Q_{j,k}) = \bar{m}_{j,k}(1-a_{j,k})^{-1+\frac{1}{\gamma_k}}Q_{j,k}^{1-\frac{1}{\gamma_k}} \tag{4}$$

where $\bar{p}_{jj,k}$, $\bar{z}_{j,k}$, $\bar{m}_{j,k}$ are exogenous shifters of price, pollution, and mass of firms.[5] Internal economies of scale operate through endogenous mass of firms, given by equation (4). The resulting scale effects on price and pollution are reflected by $\left(Q_{j,k}/(1-a_{j,k})\right)^{-1/\gamma_k}$ in equations (2) and (3), which highlights that scale economies are operative through both production and abatement, to a common extent governed by γ_k.

Lastly, the optimal choice of abatement is given by:[6]

$$(1-a_{j,k}) = \left(\frac{\alpha_k}{1-\alpha_k}\right)^{\alpha_k}\left(\frac{w_j/\bar{\varphi}_{j,k}}{\tau_{j,k}}\right)^{\alpha_k}. \tag{5}$$

Optimal abatement is a function of wage relative to pollution tax, with the extent of the relationship controlled by the pollution elasticity α_k.[7]

2.3 Policy Instruments

The government in country i has access to the following set of tax instruments:[8]

[5] Specifically, $\bar{p}_{jj,k} \equiv \left(\gamma_k \bar{f}_{j,k}\right)^{1/\gamma_k}\left(\frac{\gamma_k}{\gamma_k-1}\frac{1}{\bar{\varphi}_{j,k}(1-\alpha_k)}\right)^{(\gamma_k-1)/\gamma_k}$, $\bar{z}_{j,k} \equiv \sum_i \left(\gamma_k \bar{f}_{j,k}/\bar{p}_{jj,k}\right)^{1/(\gamma_k-1)}$, $\bar{m}_{j,k} \equiv \bar{p}_{jj,k}/(\gamma_k \bar{f}_{j,k})$.

[6] Our specification can be alternatively interpreted as one in which production employs a pollutant input (whose unit cost is the pollution tax) and labor, $\bar{d}_{ji,k}q_{ji,k}(\omega) = \left(z_{ji,k}(\omega)\right)^{\alpha_k}\left(\varphi_{j,k}l_{ji,k}(\omega)\right)^{1-\alpha_k}$ where $\bar{d}_{ji,k}q_{ji,k}(\omega) \equiv \bar{d}_{ji,k}q_{ji,k}$, $z_{ji,k}(\omega) \equiv z_{ji,k}$ and $l_{jik}(\omega) \equiv l_{ji,k}$ are firm-level output, pollution, and labor for production of ji,k. In this alternative formulation, the unit cost of ji,k is given by $c_{ji,k} = \bar{d}_{ji,k}\alpha_k^{-\alpha_k}(1-\alpha_k)^{-(1-\alpha_k)}\tau_{j,k}^{\alpha_k}(w_j/\bar{\varphi}_{j,k})^{1-\alpha_k}$, pollution by $z_{ji,k} = \alpha_k c_{ji,k}\bar{d}_{ji,k}q_{ji,k}/\tau_{j,k}$, and labor by $l_{ji,k} = (1-\alpha_k)c_{ji,k}\bar{d}_{ji,k}q_{ji,k}/w_j$. Replacing these in the relation between abatement and pollution, $(1-a_{j,k}) = \left(z_{ji,k}\big/\bar{\varphi}_{j,k}l_{ji,k}\right)^{\alpha_k}$, delivers equation (5). In addition, note that our framework nests a model with exogenous emission intensities (no abatement) if $\alpha_k \to 0$. In this case, $\bar{d}_{ji,k}q_{ji,k}(\omega) = \varphi_{j,k}l_{ji,k}(\omega)$, and $a_{j,k} = 0$.

[7] For completeness, (for $0 < \alpha_k < 1$) we specify that $a_{j,k} = 0$ if $\tau_{j,k} \leq \tau_{j,k}^{\min} \equiv \frac{\alpha_k}{1-\alpha_k}(w_j/\bar{\varphi}_{j,k})$.

[8] Adding consumption and abatement taxes does not bring any new potential in policy since the entire effect from these two taxes can be replicated by an appropriate choice of the current instruments.

1. An import tax, $t_{ji,k}$, applied to each imported variety ji,k ($t_{ii,k}=0$ by design)

2. An export tax, $x_{ij,k}$, applied to each exported variety ij,k ($x_{ii,k}=0$ by design)

3. A production tax, $s_{i,k}$, applied to all outputs in country-industry i,k irrespective of the location of final sales.

4. A pollution tax, $\tau_{i,k}$, applied to all outputs in country-industry i,k irrespective of the location of final sales.

The first three tax instruments create a wedge between the consumer and producer price of a given variety. The following one-to-one mapping holds between the set of instruments $\{t_{ji,k}, x_{ij,k}, s_{i,k}\}_{j,k}$ and the set of prices $\{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}, \tilde{P}_{ii,k}\}_{j\neq i,k}$,

$$(1+t_{ji,k}) = \frac{\tilde{P}_{ji,k}}{P_{ji,k}}, \quad (1+x_{ij,k}) = \frac{\tilde{P}_{ij,k}}{P_{ij,k}}\frac{\tilde{P}_{ii,k}}{P_{ii,k}}, \quad (1+s_{i,k}) = \frac{\tilde{P}_{ii,k}}{P_{ii,k}} \tag{6}$$

In addition to these price wedges, according to equation (5), the government can choose abatement levels $\{a_{i,k}\}_k$ to replicate pollution taxes $\{\tau_{i,k}\}$. These equivalences highlight which variables every tax instrument directly targets.

2.4 General Equilibrium

Revenues. Total income in country i, Y_i, which equals total expenditure in i given by equation (1), is the sum of wages, lump-sum tax revenues, T_i, which are collected from pollution and non-pollution taxes, and trade deficit, $\bar{D}_i$:

$$Y_i = w_i\bar{L}_i + \bar{D}_i + T_i \tag{7}$$

where T_i is the sum of payments for pollution taxes (equivalent to the compensation of pollution as a factor of production), and imports, exports, and production tax revenues:

$$\begin{aligned} T_i = & \overbrace{\sum_{k\in\mathbb{K}}\sum_{j\in\mathbb{C}}\left(\alpha_k\frac{\gamma_k-1}{\gamma_k}P_{ij,k}Q_{ij,k}\right)}^{\text{pollution taxes}} + \overbrace{\sum_{k\in\mathbb{K}}\left[\left(\tilde{P}_{ii,k}-P_{ii,k}\right)Q_{ii,k}\right]}^{\text{production taxes}} \\ & + \underbrace{\sum_{k\in\mathbb{K}}\sum_{j\in\mathbb{C},j\neq i}\left[\left(\tilde{P}_{ji,k}-P_{ji,k}\right)Q_{ji,k}\right]}_{\text{imports taxes}} + \underbrace{\sum_{k\in\mathbb{K}}\sum_{j\in\mathbb{C},j\neq i}\left[\left(\tilde{P}_{ij,k}-P_{ij,k}\right)Q_{ij,k}\right]}_{\text{exports taxes}} \end{aligned} \tag{8}$$

We treat the trade deficit as exogenous in our model. By construction, trade deficits satisfy an adding-up constraint, $\sum_i \bar{D}_i = 0$.

Definition. For every country $i, j \in \mathbb{C}$ and industry $k \in \mathbb{K}$, given taxes $\{t_{ji,k}, x_{ij,k}, s_{i,k}, \tau_{i,k}\}_{j,k}$, an *equilibrium* is a vector of wages $\{w_j\}$ such that before-tax prices $\{P_{ji,k}\}$ are given by (2), pollution $\{Z_{j,k}\}$ by (3), mass of firms $\{M_{j,k}\}$ by (4), abatement $\{a_{j,k}\}$ by (5), demand quantities by $Q_{ji,k} = \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$ in which after-tax prices $\{\tilde{P}_{ji,k}\}$ are given by (6) and Y_i is national expenditure according to (1) that equals national income according to (7) where tax revenues are given by (8)" and labor markets clear:[9]

$$w_i \bar{L}_i - \sum_{k \in \mathbb{K}} \sum_{i \in \mathbb{C}} (1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}) P_{ij,k} Q_{ij,k} = 0 \tag{9}$$

Expenditure/Revenue Shares and Pollution Intensity. To streamline the presentation of our theory, we define the following variables. The share of country i's expenditure on variety ji, k is denoted by $e_{ji,k}$,

$$e_{ji,k} \equiv \frac{\tilde{P}_{ji,k} Q_{ji,k}}{\sum_{\hat{j} \in \mathbb{C}} \sum_{\hat{k} \in \mathbb{K}} \tilde{P}_{\hat{j}i,\hat{k}} Q_{\hat{j}i,\hat{k}}} = \frac{\tilde{P}_{ji,k} Q_{ji,k}}{Y_i} \tag{10}$$

The within-industry share of country j's revenues collected from sales of variety ji, k is denoted by $r_{ji,k}$,

$$r_{ji,k} \equiv \frac{P_{ji,k} Q_{ji,k}}{\sum_{\hat{k} \in \mathbb{K}} P_{j\hat{i},\hat{k}} Q_{j\hat{i},\hat{k}}}, \tag{11}$$

In addition, we use $v_{j,k}$ to denote the pollution intensity per unit value of output in country-industry j, k.

$$v_{j,k} \equiv \frac{Z_{j,k}}{P_{jj,k} Q_{j,k}} = \frac{\gamma_k - 1}{\gamma_k} \frac{\alpha_k}{\tau_{j,k}} \tag{12}$$

Lastly, we denote within-industry expenditure share on ji, k by $\lambda_{ji,k}$,

$$\lambda_{ji,k} \equiv \frac{\tilde{P}_{ji,k} Q_{ji,k}}{\sum_{\hat{j} \in \mathbb{C}} \tilde{P}_{\hat{j}i,k} Q_{\hat{j}i,k}} \tag{13}$$

2.5 Governments and Their Objectives

In this section, we define the objective function that the government in a country aims to maximize. Let $\mathbb{I}_i$ stack the instruments of policy for the government in country i, $\mathbb{I}_i \equiv \{t_{ji,k}, x_{ij,k}, s_{i,k}, \tau_{i,k}\}_{j,k}$. The objective function of the government in country i is given by:

[9] The labor market clearing condition (LMC) is equivalent to trade deficit condition (TDC), $\sum_{k \in \mathbb{K}} \sum_{j \neq i \in \mathbb{C}} \left(P_{ji,k} Q_{ji,k} - \tilde{P}_{ij,k} Q_{ij,k} \right) = \bar{D}_i$, where exports and imports of every country i are measured in values outside the border of i (that are, exports are after-tax, but imports are before-tax). In our policy analysis, we sometimes use (TDC) instead of (LMC).

$$\mathcal{W}_i = V_i(Y_i(\mathbb{I}_i, \mathbf{w}), \tilde{\mathbf{P}}_i) - \sum_n \sum_k (\delta_{ni} Z_{n,k}) \quad (14)$$

The first term in this objective function reproduces indirect utility, taking into account that income Y_i depends on the vector of wages $\mathbf{w} = \{w_i\}$ as well as policy instruments $\mathbb{I}_i$. The second term sums over all pollution externalities from global production. δ_{ni} is the disutility to residents of country i from every unit of pollution generated in every country n. For instance, a unit of pollution generated in a country may have a greater negative effect on that country or its nearby countries than those that are faraway. One particular case in which $\delta_{ni} = \delta_A + \delta_B$ if $n = i$, and $\delta_{ni} = \delta_B$ if $n \neq i$, sets the disutility at ($\delta_A \times$ local pollution $+ \delta_B \times$ global pollution). That is, the government may assign an additional weight to local pollution beyond its care for global pollution. One can interpret the disutility from global pollution as the present discounted externality from global climate change—see Shapiro (2016).

We can write $\delta_{ni} = \bar{L}_i \bar{\delta}_{ni}$ to reflect that the disutility to every nation i is scaled by its population $\bar{L}_i$. In addition, we will use $\tilde{\delta}_{ni} \equiv \tilde{P}_i \delta_{ni}$ as the *CPI-adjusted* welfare cost per unit of pollution, where $\tilde{P}_i \equiv (\partial V_i(.)/\partial Y_i)^{-1}$.

Definition. The *Optimal Unilateral Policy* for country i is achieved by choosing policy instruments, $\mathbb{I}_i$, that maximize country i's welfare, $\mathcal{W}_i$ (equation 14), subject to equilibrium conditions (1)-(9).

Following the above definition, we present in Appendix A.1 a minimal set of equations that describe the unilateral policy problem.

3 Optimal Unilateral Policy

In this section, we characterize the unilaterally optimal tax schedule. The unilateral policy incorporates a number of non-cooperative motives. First, non-cooperative governments only care about the domestic disutility of (local or global) emission and neglect their country's transboundary emission externality on the rest of the world. Second, non-cooperative governments resort to trade taxes to correct the emission externality imposed on them by the rest of the world. Third, a non-cooperative government may use trade taxes to improve its country's terms-of-trade at the expense of trading partners.

We currently have a limited understanding of how these distinct policy motives interact. To shed light on their interaction, we analytically characterize the optimal unilateral tax schedule. This is a challenging task, which explains why previous characterizations of optimal trade and environmental policies have typically restricted their attention to two-country or partial equilibrium setups, with all or some of these simplifying assumptions: perfect competition, fixed location of firms, fixed set of products, exogenous emission intensities, and constant-returns-to-scale production technologies.

Before turning to present our results, we discuss our methodological approach. Here, our goal is to demonstrate that we have a systematic way of deriving the optimal policy schedules with applications beyond this work. The analysis of optimal policy in open economy introduces a number of challenges, such as the way one has to address general equilibrium wage and income effects of policy. These challenges are in turn responsible for some of the limitations in the previous literature. To fully address these issues, we establish an intermediate envelope result which we continue to explain. Throughout the paper, if not reported in the main text, we will report our derivations and proofs in the appendix.

3.1 Intermediate Envelope Result

In this section, we present an intermediate envelope result that greatly facilitates our optimal policy analysis. In summary, this result allows us to convert our general equilibrium optimization problem into a simpler problem characterized by a set of partial equilibrium derivatives. We establish this result in three steps.

Step 1: Reformulate the optimal policy problem in terms of consumer prices and abatement

The government in i can choose consumer prices $\{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}, \tilde{P}_{ii,k}\}_{j\neq i,k}$ according to equation (6) to replicate any set of $\{t_{ji,k}, x_{ij,k}, s_{i,k}\}_{j,k}$, and abatement levels $\{a_{i,k}\}_k$ to replicate any set of pollution taxes $\{\tau_{i,k}\}$ according to equation (5). These equivalences highlight which variables every tax directly targets. Shifting the focus from the vector of taxes $\mathbb{I}_i \equiv \{t_{ji,k}, x_{ij,k}, s_{i,k}, \tau_{i,k}\}_{j,k}$ to their target variables $\mathbb{P}_i \equiv \{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}, \tilde{P}_{ii,k}, a_{i,k}\}_{j\neq i,k}$ proves useful in our analysis. As a point of reference, we define $\mathbb{P}_i$ formally.

Definition 1. $\mathbb{P}_i \equiv \left\{\tilde{\mathbf{P}}_{ij}, \tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ii}, \boldsymbol{a}_i\right\}$ denotes the vector of policy instruments for country i in the reformulated optimal policy problem, where $\tilde{\mathbf{P}}_{ji} = \left\{P_{ji,k}\right\}_{j\neq i,k}$, $\tilde{\mathbf{P}}_{ij} = \left\{P_{ij,k}\right\}_{j\neq i,k}$, $\tilde{\mathbf{P}}_{ii} = \{P_{ii,k}\}_k$, and $\mathbf{a}_i = \{a_{i,k}\}_k$.

Appealing to the above definition, our optimal policy problem can be simplified when it is cast as the problem of choosing consumers prices and abatement instead of taxes. Once we solve that problem, then we can recover optimal taxes using equation (6) and (12).

Lemma 1. *Given optimal prices and abatement levels, $\mathbb{P}_i^\star = \{\tilde{P}_{ji,k}^\star, \tilde{P}_{ij,k}^\star, \tilde{P}_{ii,k}^\star, a_{i,k}^\star\}_{j\neq i,k}$, optimal taxes $\mathbb{I}_i^\star = \{t_{ji,k}^\star, x_{ij,k}^\star, s_{i,k}^\star, \tau_{i,k}^\star\}_{j,k}$ can be recovered according to the following one-to-one mapping:*

$$1 + t_{ji,k}^\star = \frac{\tilde{P}_{ji,k}^\star}{P_{ji,k}}, \quad 1 + x_{ij,k}^\star = \frac{\tilde{P}_{ij,k}^\star}{P_{ij,k}}\frac{\tilde{P}_{ii,k}^\star}{P_{ii,k}}, \quad 1 + s_{i,k}^\star = \frac{\tilde{P}_{ii,k}^\star}{P_{ii,k}}, \quad \tau_{i,k}^\star = \frac{\gamma_k - 1}{\gamma_k}\frac{\alpha_k}{v_{i,k}(a_{i,k}^\star)}$$

This lemma states that once we find the vector of optimal consumers prices and abatement levels, we can easily recover the vector of optimal taxes.

Step 2: Conditional welfare-neutrality of wage effects

The choice of $\mathbb{P}_i$ affects the vector of wages whose subsequent effect on welfare complicates the analysis. We show that conditional on holding policy $\mathbb{P}_i$ fixed, the wage effects are welfare-neutral. To make this point, we formulate all variable outcomes as a function of $\mathbb{P}_i$ and wage vector $\mathbf{w}$ in a system where all equilibrium relationships hold with the exception of labor market clearing condition. Appendix A.2 details this formulation. This characterizes welfare in country i as $W_i(\mathbb{P}_i; \mathbf{w})$, prices as $P_{ij,k}(\mathbb{P}_i; \mathbf{w})$, quantities as $Q_{ij,k}(\mathbb{P}_i; \mathbf{w})$, etc. Note, however, that all $(\mathbb{P}_i; \mathbf{w})$ pairs are not feasible. Given $\mathbb{P}_i$, a feasible vector of wages must satisfy labor market clearing conditions.

Definition 2. A policy-wage pair, $(\mathbb{P}_i; \mathbf{w})$ is feasible iff, the vector of wages $\mathbf{w} \equiv \{w_n\}_{n\in\mathbb{C}}$ satisfy the labor-market clearing conditions, given a policy vector $\mathbb{P}_i$,

$$(\mathbb{P}_i; \mathbf{w}) \in \mathbb{F}_i^w \iff \sum_{j,k}\left[\left(1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}\right) P_{nj,k}(\mathbb{P}_i; \mathbf{w}) Q_{nj,k}(\mathbb{P}_i; \mathbf{w})\right] = w_n \bar{L}_n, \qquad \text{for all } n \in \mathbb{C}. \quad (15)$$

Using this definition, we express the government's problem (P1) as:

$$\max_{\mathbb{P}_i} W_i(\mathbb{P}_i; \mathbf{w}), \quad \text{subject to } (\mathbb{P}_i; \mathbf{w}) \in \mathbb{F}_i^w \qquad (P1)$$

where $W_i(\mathbb{P}_i; \mathbf{w}) = V_i(Y_i(\mathbb{P}_i; \mathbf{w}), \tilde{\boldsymbol{P}}_i) - \boldsymbol{\delta}_i \cdot \mathbf{Z}(\mathbb{P}_i; \mathbf{w})$. Here, the inner product $\boldsymbol{\delta}_i \cdot \mathbf{Z}(\mathbb{P}_i; \mathbf{w}) = \sum_{j,k} \delta_{ji} Z_{j,k}(\mathbb{P}_i; \mathbf{w})$ summarizes the disutility from global pollution to country i. The necessary condition for the optimality of a policy instrument $\mathcal{P} \in \mathbb{P}_i$ is then given by:

$$\frac{\mathrm{d}W_i(\mathbb{P}_i; \mathbf{w})}{\mathrm{d}\ln\mathcal{P}} = \frac{\partial V_i(.)}{\partial \ln \mathcal{P}} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial Y_i(\mathbb{P}_i; \mathbf{w})}{\partial \ln \mathcal{P}}\right)_{\mathbf{w}} - \left(\frac{\partial \boldsymbol{\delta}_i \cdot \mathbf{Z}(\mathbb{P}_i; \mathbf{w})}{\partial \ln \mathcal{P}}\right)_{\mathbf{w}} + \underbrace{\frac{\partial W_i(\mathbb{P}_i; \mathbf{w})}{\partial \mathbf{w}}\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\ln\mathcal{P}}}_{\text{wage effects}} = 0,$$

Recall that $V_i(.) \equiv V_i(Y_i, \tilde{\mathbf{P}}_i)$ denotes the indirect utility function from consumption, and $\frac{\partial V_i(.)}{\partial \ln \mathcal{P}}$ is nonzero only if $\mathcal{P}$ is one of prices faced by home consumers, $\mathcal{P} \in \tilde{\mathbf{P}}_i \equiv \{\tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ii}\}$. In the above

FOC, the first three terms correspond to the effects of policy $\mathcal{P} \in \mathbb{P}_i$ on welfare holding $\mathbf{w} = \{w_n\}_{n\in\mathbb{C}}$ fixed.[10] The last term accounts for the general equilibrium wage effects from every country. By choice of numeraire, we normalize wage in one of foreign countries, say n, to unity. That implies $\frac{\mathrm{d}w_n}{\mathrm{d}\ln\mathcal{P}} = 0$. We show in Appendix A.4 that

$$r_{ji} \times \lambda_{\ell i,k} \approx 0 \quad \text{if } (j \neq i) \wedge (\ell \neq i) \implies \frac{\mathrm{d}\mathbf{w}_{-\{i,n\}}}{\mathrm{d}\ln\mathcal{P}} \approx 0.$$

Throughout this section we maintain the assumption that $r_{ji} \times \lambda_{\ell i,k} \approx 0$ if j and $\ell \neq i$. Later, when mapping out theory to data, we show that this assumption is strikingly consistent with actual data. Regardless, the most important wage effect is the one from home itself. Finding this wage effect appears to be a major obstacle when solving problem (P1). The next lemma allows us to overcome this challenge. It states that for any $(\mathbb{P}_i;\mathbf{w}) \in \mathbb{F}_i^w$, if the government has access to all policy instruments, country i's own wage effects are also welfare-neutral.

Lemma 2. *Within the feasible policy-wage set $(\mathbb{P}_i;\boldsymbol{w}) \in \mathbb{F}_i^w$, conditional on a choice of policy vector $\mathbb{P}_i$, welfare in country i is invariant to wage w_i:*

$$\left(\frac{\partial W_i(\mathbb{P}_i;\mathbf{w})}{\partial w_i}\right)_{\mathbf{w}_{-i}} = 0, \qquad \forall(\mathbb{P}_i;\mathbf{w}) \in \mathbb{F}_i^w.$$

To take stock, the above result indicates that home's wage has no effect on home's welfare, provided that the labor market clearing condition holds and the government has access to all policy instruments. This result would hold even if the government did *not* choose the policy vector optimally. To provide intuition, note that as long as policy $\mathbb{P}_i$ is fixed, w_i affects welfare, W_i, only though its effect on income Y_i. Lemma 2 can, thus, be established by showing that $\partial Y_i/\partial w_i = 0$. To show that $\partial Y_i/\partial w_i = 0$, note that an increase in w_i has two opposite but equal-size effects on income Y_i, as long as the policy vector, $\mathbb{P}_i$, is held fixed. On one hand, an increase in wage w_i raises income Y_i directly through wage incomes $w_i\bar{L}_i$. On the other hand, it decreases income indirectly through raising producer prices, which reduce tax revenues. Importantly, this latter effect appears because the after-tax price of home-made varieties, $\{\tilde{\mathbf{P}}_{ij}, \tilde{\mathbf{P}}_{ii}\}$, are held fixed. The two opposing effects also sum up to zero, because the tax revenue effect is proportional to country i's total sales, and total (net) sales equal wage incomes in equilibrium.

[10] On our notation: (1) For any vector $\mathbf{y}$, $\mathbf{y}_{-n} \equiv \mathbf{y}/\{y_n\}$. (2) In cases where there might be ambiguity, we include endogenous variables that we hold fixed in the subscript of a derivative. For function $G(\mathbf{x};\mathbf{y})$ with $\mathbf{x}$ as the policy vector, and $\mathbf{y}$ as the vector of endogenous variables, $\left(\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial x_m}\right)_{\mathbf{y}}$ denotes the derivative of G wrt x_m, holding fixed $\mathbf{y}$ and $\mathbf{x}_{-m}$, and $\left(\frac{\partial G(\mathbf{x},\mathbf{y})}{\partial y_n}\right)_{\mathbf{y}_{-n}}$ denotes the derivative of G wrt y_n, holding fixed $\mathbf{y}_{-n}$ and $\mathbf{x}$.

Lemma 2 greatly facilitates our analysis, since it allows us to identify the optimal policy by treating the wage vector $\mathbf{w}$ as fixed. Once we fix $\mathbf{w}$, income in the rest of the world, $\mathbf{Y}_{-i} = \mathbf{w}_{-i} \odot \bar{\mathbf{L}}_{-i}$, is also fixed by construction. The next step shows that income in country i, Y_i, can also be treated as fixed since domestic income effects are welfare neutral at the optimum.

Step 3: Conditional welfare-neutrality of income effects at the optimum

Following Step 2, we treat wages as invariant to policy. This means that, for a given vector of policy $\mathbb{P}_i$, we can hold wages fixed at their values that satisfy market clearing conditions, $\mathbf{w} = \bar{\mathbf{w}}$. This intermediate result also implies that we can hold income in foreign countries fixed, $Y_n = \bar{Y}_n$ for $n \neq i$. With these considerations, we re-formulate all equilibrium variables as a function of policy vector $\mathbb{P}_i$ and income Y_i. Under this formulation, all equilibrium relationships hold except the budget constraint, $Y_i = \bar{w}_i \bar{L}_i + \bar{D}_i + T_i(\mathbb{P}_i; Y_i)$. This formulation is detailed in Appendix A.3. Notice, tax revenues T_i depend on income Y_i because home's demand schedule, whose position matters for the amount of tax revenues, depends on income. This brings us to define feasible pairs of policy-income.

Definition 3. *A policy-income pair, $(\mathbb{P}_i; Y_i)$ is feasible iff , income Y_i equals total wages plus tax revenues, for a given policy $\mathbb{P}_i$,*

$$(\mathbb{P}_i; Y_i) \in \mathbb{F}_i^Y \iff Y_i = \bar{w}_i \bar{L}_i + \bar{D}_i + T_i(\mathbb{P}_i; Y_i). \tag{16}$$

We continue with an observation that further facilitates our analysis. Restricting the system to the feasible policy-income pairs, we observe that income Y_i affects welfare *exclusively* through demand quantities. Behind this observation is that income affects producer prices, emissions, and taxes only though income effects in demand, meaning that we can express these variables as $P_{ni,g} = P_{ni,g}(\mathbb{P}_i, \mathbf{Q}_i(\mathbb{P}_i, Y_i))$, $Z_{n,g} = Z_{n,g}(\mathbb{P}_i, \mathbf{Q}_i(\mathbb{P}_i, Y_i))$, $T_i = T_i(\mathbb{P}_i, \mathbf{Q}_i(\mathbb{P}_i, Y_i))$, where $\mathbf{Q}_i \equiv \{Q_{ni,g}, Q_{in,g}\}_{n \in \mathbb{C}, g \in \mathbb{K}}$ is the vector of country i's output and consumption quantities. The equilibrium value for consumption quantities are given by $Q_{ni,g} = \mathcal{D}_{nig}(Y_i, \tilde{\mathbf{P}}_i)$. Export quantities are $Q_{in,g} = \mathcal{D}_{ing}(\bar{Y}_j = \bar{w}_j \bar{L}_j, \tilde{\mathbf{P}}_{in}, \tilde{\mathbf{P}}_{-in}(\bar{\mathbf{w}}_{-i}))$.

The optimal policy problem of country i can now be expressed as:

$$\max_{\mathbb{P}_i} \quad W_i(\mathbb{P}_i; \mathbf{Q}_i(\mathbb{P}_i, Y_i)) \quad \text{subject to } (\mathbb{P}_i; Y_i) \in \mathbb{F}_i^Y \tag{$P2$}$$

where:

$$W_i(\mathbb{P}_i, \mathbf{Q}_i(\mathbb{P}_i, Y_i)) = V_i(\underbrace{\bar{w}_i \bar{L}_i + \bar{D}_i + T_i(\mathbb{P}_i, \mathbf{Q}_i(\mathbb{P}_i, Y_i))}_{Y_i}, \tilde{\mathbf{P}}_i) - \delta_i \cdot \mathbf{Z}(\mathbb{P}_i, \mathbf{Q}_i(\mathbb{P}_i, Y_i))$$

Capitalizing on our reformulation with (P2), we explain the conditions for the welfare neutrality of income effects. The first order condition w.r.t. to policy $\mathcal{P} \in \mathbb{P}_i$ is given by $\left[\frac{\partial V_i}{\partial Y_i}\frac{\partial T_i(.)}{\partial \mathcal{P}} + \frac{\partial V_i}{\partial \mathcal{P}} - \frac{\partial \delta_i \cdot \mathbf{Z}(.)}{\partial \mathcal{P}}\right] = 0$. We expand the components of this equation using the following derivatives,

$$\begin{cases} \frac{\partial T_i(.)}{\partial \mathcal{P}} = \left(\frac{\partial T_i}{\partial \mathcal{P}}\right)_{\mathbf{Q}_i} + \frac{\partial T_i}{\partial \mathbf{Q}_i} \cdot \left[\left(\frac{\partial \mathbf{Q}_i}{\partial \mathcal{P}}\right)_{Y_i} + \frac{\partial \mathbf{Q}_i}{\partial Y_i}\frac{\mathrm{d}Y_i}{\mathrm{d}\mathcal{P}}\right] \\ \frac{\partial \delta_i \cdot \mathbf{Z}(.)}{\partial \mathcal{P}} = \left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathcal{P}}\right)_{\mathbf{Q}_i} + \frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathbf{Q}_i} \cdot \left[\left(\frac{\partial \mathbf{Q}_i}{\partial \mathcal{P}}\right)_{Y_i} + \frac{\partial \mathbf{Q}_i}{\partial Y_i}\frac{\mathrm{d}Y_i}{\mathrm{d}\mathcal{P}}\right] \end{cases},$$

where $\frac{\mathrm{d}Y_i}{\mathrm{d}\mathcal{P}}$ can be calculated by applying the Implicit Function Theorem to Equation 16 to ensure feasibility. To elaborate, tax revenues $T_i(.)$ and emission disutility $\delta_i \cdot \mathbf{Z}(.)$ react to policy $\mathcal{P}$ directly by fixing quantities, and indirectly through quantities. Note that, once we hold quantities fixed, we are also holding income fixed. Putting together, and recalling that $\tilde{P}_i \equiv \left(\frac{\partial V_i(.)}{\partial Y_i}\right)^{-1}$, the FOC collapses to:

$$\underbrace{\tilde{P}_i\frac{\partial V_i(.)}{\partial \mathcal{P}} + \left(\frac{\partial T_i}{\partial \mathcal{P}}\right)_{\mathbf{Q}_i} - \tilde{P}_i\left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathcal{P}}\right)_{\mathbf{Q}_i} + \left[\frac{\partial T_i}{\partial \mathbf{Q}_i} - \tilde{P}_i\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathbf{Q}_i}\right] \cdot \left(\frac{\partial \mathbf{Q}_i}{\partial \mathcal{P}}\right)_{Y_i}}_{\left(\frac{\partial W_i}{\partial \mathcal{P}}\right)_{Y_i}} + \underbrace{\left[\frac{\partial T_i}{\partial \mathbf{Q}_i} - \tilde{P}_i\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathbf{Q}_i}\right] \cdot \frac{\partial \mathbf{Q}_i}{\partial Y_i}}_{\frac{\partial W_i}{\partial Y_i}}\frac{\mathrm{d}Y_i}{\mathrm{d}\mathcal{P}} = 0 \tag{17}$$

In equation (17), the first four terms represent the direct welfare effect of policy $\mathcal{P}$ holding income fixed, $\left(\frac{\partial W_i}{\partial \mathcal{P}}\right)_{Y_i}$, and the last term represents the indirect general equilibrium effect of policy $\mathcal{P}$ on welfare through income (hence, the term "income effects"). We will rely on this system of FOCs to solve for the optimal policy schedule, but we pause that analysis for the moment to illustrate the conditions for the neutrality of income effects.

Suppose $\mathcal{P}$ be one of consumer prices in home $\tilde{P}_{ji,k} \in \tilde{\mathbf{P}}_i$. The price index $\tilde{P}_{ji,k}$ can be for one of domestic ($j = i$) or imported ($j \neq i$) varieties. In this case,

$$\left(\frac{\partial T_i}{\partial \tilde{P}_{ji,k}}\right)_{\mathbf{Q}_i} = Q_{ji,k}, \quad \tilde{P}_i\left(\frac{\partial V_i}{\partial \tilde{P}_{ji,k}}\right) = -Q_{ji,k}, \quad \left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \tilde{P}_{ji,k}}\right)_{\mathbf{Q}_i} = 0 \quad \Longrightarrow \quad \tilde{P}_i\frac{\partial V_i(.)}{\partial \tilde{P}_{ji,k}} + \left(\frac{\partial T_i}{\partial \tilde{P}_{ji,k}}\right)_{\mathbf{Q}_i} - \tilde{P}_i\left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \tilde{P}_{ji,k}}\right)_{\mathbf{Q}_i} = 0$$

where the first equality reflects the direct effect of consumer price $\tilde{P}_{ji,k}$ on tax revenues, the second equality is Roy's identity, and the third equality holds because emission depends on abatement and quantities. From setting $\tilde{P}_i\frac{\partial V_i(.)}{\partial \mathcal{P}} + \left(\frac{\partial T_i}{\partial \mathcal{P}}\right)_{\mathbf{Q}_i} - \tilde{P}_i\left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathcal{P}}\right)_{\mathbf{Q}_i} = 0$ in equation (17) and noting that $\frac{\partial \mathbf{Q}_{in}}{\partial \tilde{P}_{ji,k}} = \frac{\partial \mathbf{Q}_{in}}{\partial \tilde{P}_{ii,k}} = \frac{\partial \mathbf{Q}_{in}}{\partial Y_i} = 0$, we can conclude that a trivial solution in case of $\mathcal{P} = \tilde{P}_{ji,k}$ or $\tilde{P}_{ii,k} \in \tilde{\mathbf{P}}_i \subset \mathbb{P}_i$ is achieved where

$$\frac{\partial T_i}{\partial \mathbf{Q}_{ni}} - \tilde{P}_i\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathbf{Q}_{ni}} = \sum_{n=1}^{N}\sum_{k=1}^{K}\left[\frac{\partial T_i}{\partial Q_{ni,k}} - \tilde{P}_i\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial Q_{ni,k}}\right] = 0$$

We show in the next section that this trivial solution is also the unique solution to the system of FOCs. It, therefore, follows that for the choices of $\tilde{P}_{ii,k}$ and $\tilde{P}_{ji,k} \in \tilde{\mathbf{P}}_i$ to be optimal, it is necessary that income effects be welfare neutral: $\frac{\partial W_i}{\partial Y_i} = 0$. We summarize our discussion in the following lemma.

Lemma 3. *Within the feasible policy-income set, $(\mathbb{P}_i; Y_i) \in \mathbb{F}_i^Y$, if $\tilde{\mathbf{P}}_i \subset \mathbb{P}_i$ is chosen optimally, then the income effect is welfare-neutral, $\frac{\partial W_i}{\partial Y_i} = 0$.*

Putting the Three Steps Together

We outline the results from Lemmas 1,2,3 in the following proposition.

Proposition 1. [***Intermediate Envelope Result***] *Country i's optimal policy, $\mathbb{P}_i^\star$, is the solution to a system of equations that asserts optimality w.r.t. all $\mathcal{P}_i \in \mathbb{P}_i$ by holding fixed wages and income,*

$$\tilde{P}_i \frac{\partial V_i(.)}{\partial \ln \mathcal{P}} + \left(\frac{\partial T_i}{\partial \ln \mathcal{P}}\right)_{\mathbf{w},\mathbf{Q}_i} - \tilde{P}_i \left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \ln \mathcal{P}}\right)_{\mathbf{w},\mathbf{Q}_i} + \left[\left(\frac{\partial T_i}{\partial \mathbf{Q}_i}\right)_{\mathbf{w}} - \tilde{P}_i \left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \mathbf{Q}_i}\right)_{\mathbf{w}}\right] \cdot \left(\frac{\partial \mathbf{Q}_i}{\partial \mathcal{P}}\right)_{Y_i} = 0 \quad (\star).$$

We refer to Proposition 1 as an intermediate envelop result, because it reduces our general equilibrium optimal policy problem into one in which wage and income effects can be ignored. In other words, we can derive the optimal policy schedule while treating $\mathbf{w}$ as constant and ignoring Y_i's impact on country i's demand schedule. Below, we discuss several aspects of this intermediate envelope result.

As noted in the build up to Lemma 3, the first three terms in Equation ($\star$) collapse to zero when $\mathcal{P} = \tilde{P}_{ji,k}$ or $\tilde{P}_{ii,k} \in \mathbb{P}_i$. Relatedly, $\frac{\partial V_i(.)}{\partial \ln(1-a_{i,k})} = \frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ij,k}} = 0$ since neither $a_{i,k}$ or $\tilde{P}_{ij,k}$ explicitly enter the indirect utility function. Furthermore, $\left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},\mathbf{Q}_i} = 0$ since $\tilde{P}_{ij,k}$ affects emission only through its effect on output quantities, $\mathbf{Q}_i$; and $\left(\frac{\partial \mathbf{Q}_i}{\partial \ln(1-a_{i,k})}\right)_{Y_i} = 0$ since holding prices and income fixed abatement has not effect on the demand schedule. Accounting for these equal-to-zero terms, $\mathbb{P}_i^\star$ solve the following system according to Proposition 1:

$$\begin{cases} \left(\frac{\partial T_i}{\partial \ln(1-a_{i,k})}\right)_{\mathbf{w},\mathbf{Q}_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}}{\partial \ln(1-a_{i,k})}\right)_{\mathbf{w},\mathbf{Q}_i} = 0 & [a_{i,k}] \\ \tilde{P}_{ij,k} Q_{ij,k} + \sum_{n \in \mathbb{C}} \sum_{k \in \mathbb{K}} \left[\left(\frac{\partial T_i}{\partial Q_{nj,k}} - \tilde{P}_i \frac{\partial \delta_i \cdot \mathbf{Z}}{\partial Q_{nj,k}}\right) \frac{\partial \mathcal{D}_{nj,k}(.)}{\partial \ln \tilde{P}_{nj,k}}\right] = 0 & [\tilde{P}_{ij,k}] \\ \sum_{n \in \mathbb{C}} \sum_{k \in \mathbb{K}} \left[\frac{\partial T_i}{\partial Q_{ni,k}} - \tilde{P}_i \frac{\partial \delta_i \cdot \mathbf{Z}}{\partial Q_{ni,k}}\right] = 0 & [\tilde{P}_{ji,k},\ \tilde{P}_{ii,k}] \end{cases} \tag{18}$$

Before solving the above system, a few details about Proposition 1 are in order. Above all, Proposition 1 holds when country i's government has access to all *price-related* policy instruments. As for wage effects, if the government is banned from manipulating any of instruments belonging to after-tax prices of varieties originated from home, $\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ij}\}$, then Lemma 2 fails. The intuition is the following. The government in home country i can improve its terms-of-trade by manipulating home's relative wage , w_i (relative to value of labor in one of foreign countries that is chosen as numeraire). The gains from manipulating wage w_i can be perfectly mimicked with an appropriate adjustment in production and export taxes, $\{s_{i,k}, x_{ij,k}\}_{j\neq i,k}$, that are associated with committing to price vectors, $\tilde{\mathbf{P}}_{ii}$ and $\tilde{\mathbf{P}}_{ij}$. More specifically, with a proper adjustment in production and export taxes the government can achieve any level of national sales; and, as long as the labor market clears, national sales pin down home's wage. This argument holds even if choices of $\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ij}\}$ are not optimal, but it fails if the government is banned from manipulating one element of these price vectors. In that case, wage effects should be properly tracked when solving the optimal policy problem

A similar argument applies to Lemma 3. If the government can set all price variables associated with the local consumption market optimally, then income effects are redundant. Because any gains from raising Y_i are already internalized by the vector of consumer prices in home. But if the government is banned from manipulating $\tilde{\mathbf{P}}_i \equiv \{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}\}$ the argument no longer holds. Also notice, the welfare-neutrality of income effects explain why income elasticities of demand play no role in the optimal policy schedule.

Finally, note that the ability to set prices in foreign markets, $\tilde{\mathbf{P}}_{ij}$, is only relevant to Lemma 2 but irrelevant to 3. So even if the government cannot set $\tilde{\mathbf{P}}_{ij}$, we can still invoke Lemma 3 to simplify the optimal policy problem. . In addition, if a_i is set sub-optimally, Lemmas 2 and 3 continue to hold. Hence, Proposition 1 applies to scenarios where governments cannot tax emission but can manipulate the entire vector of after-tax prices $\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij}\}$.

3.2 Characterizing the Optimal Tax Schedule

We show in Proposition 1 the system of F.O.C.s that characterize the optimal policy schedule. Since we assume a non-parametric demand function, we present this system using the own- and cross-price demand elasticities defined in Section 2.1. The following lemma summarizes this step—see Appendixes B.2 and B.4 for details.

Lemma 4. Country i's optimal policy, $\mathbb{P}_i^\star = \left\{\tilde{\mathbf{P}}_{ij}^\star, \tilde{\mathbf{P}}_{ji}^\star, \tilde{\mathbf{P}}_{ii}^\star, \mathbf{a}_i^\star\right\}$, solves the following system of F.O.C.s:

$$
\begin{aligned}
&[a_{i,k}] && \tilde{\delta}_{ii} v_{i,k}(a_{i,k}^\star) - \alpha_k \frac{\gamma_k - 1}{\gamma_k} = 0; \\
&[\tilde{P}_{ni,k} = \tilde{P}_{ji,k},\ \tilde{P}_{ii,k}] && \sum_{n \neq i} \sum_g \left[\left(\frac{\tilde{P}_{ji,g}^\star}{P_{ji,g}} - \left(1 - \omega_{ji,g} - \tilde{\delta}_{ji} v_{j,g} \frac{\gamma_g - 1}{\gamma_g}\right)\right) e_{ji,g} \varepsilon_{ji,g}^{(ni,k)} \right] \\
& && + \sum_g \left[\left(\frac{\tilde{P}_{ii,g}^\star}{P_{ii,g}} - \left(1 - \alpha_g \frac{\gamma_g - 1}{\gamma_g} + \tilde{\delta}_{ii} v_{i,g}\right) \frac{\gamma_g - 1}{\gamma_g}\right) \right] e_{ii,g} \varepsilon_{ii,g}^{(ji,k)} = 0 \\
&[\tilde{P}_{ij,k}] && 1 - \sum_{\ell \neq i} \sum_g \left[\left(\omega_{\ell i,g} + \tilde{\delta}_{\ell i} v_{\ell,g} \frac{\gamma_g - 1}{\gamma_g}\right) \frac{e_{\ell j,g}}{e_{ij,k}} \varepsilon_{\ell j,g}^{(ij,k)} \right] \\
& && + + \sum_g \left[\left(1 - \left(1 - \alpha_g \frac{\gamma_g - 1}{\gamma_g} + \tilde{\delta}_{ii} v_{i,g}\right) \frac{\gamma_g - 1}{\gamma_g} \frac{P_{ij,g}^\star}{\tilde{P}_{ij,g}} \right) \frac{e_{ij,g}}{e_{ij,k}} \varepsilon_{ij,g}^{(ij,k)} \right] = 0
\end{aligned}
$$

where $\omega_{ji,k} \equiv \left(\frac{\partial \ln P_{jj,k}}{\partial \ln Q_{ji,k}}\right)_{\mathbb{P}_i, \mathbf{w}, Y_i} = \frac{-r_{ji,k}}{\gamma_k + \sum_{n \neq i} r_{jn,k} \varepsilon_{jn,k}}$ denotes the inverse of (backward-falling) export supply elasticity.

The optimality condition w.r.t. $a_{i,k}$ equalizes the marginal utility loss that stems from raising marginal cost of production and the marginal utility gain that is associated with a cleaner environment. Using the F.O.C w.r.t $a_{i,k}$ together with equation (12) that relates emission intensity to emission tax, and noting that $\tilde{\delta}_{ii} \equiv \tilde{P}_i \delta_{ii}$, the optimal emission tax, $\tau_{i,k}^*$, equals:

$$\tau_{i,k}^\star = \tilde{\delta}_{ii}. \tag{19}$$

The F.O.C.s w.r.t. $\tilde{P}_{ji,k}$ and $\tilde{P}_{ii,k}$ are inter-dependent, and contain price ratios in the form of $\frac{\tilde{P}_{ji,g}^\star}{P_{ji,g}}$ that do not show up in the F.O.C.s w.r.t. $\tilde{P}_{ij,k}$. Setting $\tau_{i,k}^\star = \tilde{\delta}_{ii}$, these F.O.C.s amount to NK equations in NK unknowns, that can be put in the following matrix equation:

$$
\begin{bmatrix}
e_{1i,1}\varepsilon_{1i,1}^{(1i,1)} & \cdots & e_{Ni,}\varepsilon_{Ni,1}^{(1i,1)} & \cdots & e_{1i,K}\varepsilon_{1i,K}^{(1i,1)} & \cdots & e_{Ni,K}\varepsilon_{Ni,K}^{(1i,1)} \\
\vdots & \ddots & & \ddots & & \ddots & \vdots \\
e_{1i,1}\varepsilon_{1i,1}^{(Ni,K)} & \cdots & e_{Ni,}\varepsilon_{Ni,1}^{(Ni,K)} & \cdots & e_{1i,K}\varepsilon_{1i,K}^{(Ni,K)} & \cdots & e_{Ni,K}\varepsilon_{Ni,K}^{(Ni,K)}
\end{bmatrix}
\begin{bmatrix}
\frac{\tilde{P}_{1i,k}^\star}{P_{1i,1}} - \left(1 - \omega_{1i,k} - \tilde{\delta}_{1i} \nu_{1,k} \frac{\gamma_k - 1}{\gamma_k}\right) \\
\vdots \\
\frac{\tilde{P}_{ii,k}^\star}{P_{ii,k}} - \frac{\gamma_k - 1}{\gamma_k} \\
\vdots \\
\frac{\tilde{P}_{Ni,k}^\star}{P_{Ni,k}} - \left(1 - \omega_{Ni,k} - \tilde{\delta}_{Ni} \nu_{N,k} \frac{\gamma_k - 1}{\gamma_k}\right)
\end{bmatrix}
= \mathbf{0}. \tag{20}
$$

where the first matrix is $NK \times NK$ and the second is $NK \times 1$. Importantly, the above equation identifies the optimal tariff, $1 + t_{ji,k}^\star = \tilde{P}_{ji,k}^\star / P_{ji,k}$, and production tax, $1 + s_{i,k}^\star = \tilde{P}_{ii,k}^\star / P_{ii,k}$ independently from choices of export taxes, $1 + x_{ij,k}^\star = \tilde{P}_{ij,k}^\star / P_{ij,k}$. To solve the above matrix equation we invoke on another intermediate result, which ensures the invertibility of the system.

Lemma 5. *The $KN \times KN$ matrix, $\Xi = \left[e_{ji,k}\varepsilon_{ji,k}^{(ni,g)}\right]_{ng,jk}$, is non-singular, with $\mid \det(\Xi) \mid > \prod_{n,k} e_{ni,k} > 0$.*

Given Lemma 4, the unique solution to Equation 20 is the trivial solution, which indicates that:

$$1 + t_{ji,k}^{\star} = \frac{\tilde{P}_{ji,k}^{\star}}{P_{ji,k}} = 1 - \omega_{ji,k} - \tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k}; \qquad 1 + s_{i,k}^{\star} = \frac{\tilde{P}_{ii,k}^{\star}}{P_{ii,k}} = \frac{\gamma_k - 1}{\gamma_k}.$$

Lastly, we can plug the already-derived values of $\{\tau_{i,k}^{\star}, t_{ji,k}^{\star}, s_{i,k}^{\star}\}_{j \neq i,k}$ (or equivalently, $\{a_{i,k}^{\star}, \frac{\tilde{P}_{ji,k}^{\star}}{P_{ji,k}}, \frac{\tilde{P}_{ii,k}^{\star}}{P_{ii,k}}\}_{j \neq i,k}$) into the first-order conditions w.r.t. $\{\tilde{P}_{ij,k}\}_{j \neq i,k}$. This final step, which solves for $x_{ij,k}^{\star}$, is outlined in Appendix B.4. The following proposition summarizes the optimal policy schedule in its final form.

Theorem 1. *The optimal unilateral tax schedule for country i is given by*

$$\begin{cases} 1 + t_{ji,k}^{\star} = 1 + \omega_{ji,k} + \tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k} & \forall j, k \\ 1 + x_{ij,k}^{\star} = \frac{\varepsilon_{ij,k}}{1 + \varepsilon_{ij,k}} \chi_{ij,k}^{-1} & \forall j, k \\ 1 + s_{i,k}^{\star} = \frac{\gamma_k - 1}{\gamma_k} & \forall k \\ \tau_{i,k}^{\star} = \tau_i^{\star} = \tilde{\delta}_{ii} & \forall k \end{cases} \tag{21}$$

where $\chi_{ij,k}$ is an export subsidy intended at lowering the emission of product varieties competing with ij,k, and is given by $\chi_{ij} = \left[\frac{\tilde{e}_{ij,g}\varepsilon_{ij,g}^{(ij,k)}}{\tilde{e}_{ij,k}\varepsilon_{ij,k}}\right]_{k,g}^{-1} \left[\frac{\sum_{n \neq i} t_{ni,g}^{\star} \tilde{e}_{nj,g} \varepsilon_{nj,g}^{(ij,k)}}{\sum_{\hat{n} \neq i,\hat{g}} \tilde{e}_{\hat{n}j,\hat{g}} \varepsilon_{\hat{n}j,\hat{g}}^{(ij,k)}}\right]_{k,g} \mathbf{1}_K$.

To put in words, the optimal unilateral policy for country i includes (i) a uniform Pigouvian tax on emission, $\tau_i^{\star}$, (ii) an industry-specific Pigouvian production subsidy, $s_{i,k}^{\star}$, that eliminates the cross-industry markup heterogeneity, (iii) import taxes, $t_{ji,k}^{\star}$ that penalize high-emission imports and take advantage of unexploited import market power, and (iv) export taxes, $x_{ij,k}^{\star}$, that promote low-emission exports and take advantage of unexploited export market power.

Optimal trade taxes are designed to both improve the terms-of-trade (ToT) and correct trans-boundary emission. So, a decomposition of these taxes is in order. First, consider the import tax on variety ji, k. The optimal rate, as implied by Proposition 1, can be decomposed as follows:

$$1 + t_{ji,k}^{\star} = \underbrace{1 + \omega_{ji,k}}_{\text{ToT driven}} \quad + \quad \underbrace{\tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k}}_{\text{Emission correcting}} \; .$$

The ToT-driven component is motivated by country i's import market power. It corresponds to an optimal mark-down on the producer price of goods imported from country j. This mark-down equals the inverse of the *export supply elasticity*, $\omega_{ji,k} < 0$. The environmentally-driven component is intended to correct the transboundary emission externality of goods imported from country j. Intuitively, this component is higher on high-emission (high-v) partners.

Likewise, export taxes are also designed to both improve the terms-of-trade (ToT) and correct transboundary emission. The export tax on good ij, k, therefore, exhibit two distinct components:

$$1 + x^*_{ij,k} = \underbrace{\frac{\varepsilon_{ij,k}}{1 + \varepsilon_{ij,k}}}_{\text{ToT driven}} \times \underbrace{\chi^{-1}_{ij,k}}_{\text{Emission correcting}} .$$

The ToT-driven term equals an optimal markup, if country i was pricing its composite export goods as a single representative monopolist. The emission-correcting term subsidizes exports of varieties that compete with dirty (high-v) varieties in foreign markets. This correcting term itself incorporates import taxes of home across its trade partners. The intuition can be put as follows. If home country i exported more to country j, then a third country n would reallocate its export away from j and possibly back to home country i. This gives home some leverage to curb emission in every third-country n using its import tax. We provide a more detailed discussion when we consider CES-Cobb-Douglas preferences.

Special Case: *Ricardian Model.* In the limit where $\gamma_k \to \infty$ and $f^e_k \to 0$, firms can be viewed as perfectly competitive and our framework reduces to Ricardian trade model. The Ricardian special case of our framework is isomorphic to the multi-industry Eaton et Kortum (2002) model. The optimal tax formulas in the Ricardian case can be attained by plugging the following values into Theorem 1:

$$\frac{\gamma_k}{\gamma_k - 1} \to 1; \quad \omega_{ji,k} \to 0 \qquad \text{(Ricardian Model)}$$

Note that, in principle, Proposition 1 applies equally to a model with a continuum of industries. As a result, in the limit where $\varepsilon_{ij,k} \to \infty$, our optimal tax formulas characterize the optimal policy in the Dornbusch *et al.* (1977) model.

Special Case: *Cobb-Douglas-CES preferences.* To gain further intuition about the optimal policy schedule, consider the special case where preferences have a Cobb-Douglas-CES formulation,

$$U_i(\boldsymbol{Q}_i) = \prod_k \left(\sum_j b^{1/\sigma_k}_{ji,k} Q^{\frac{\sigma_k - 1}{\sigma_k}}_{ji,k} \right)^{e_{i,k} \frac{\sigma_k}{\sigma_k - 1}}, \tag{22}$$

where $e_{i,k}$ is expenditure share of country i on industry k, and σ_k is the (Armington) elasticity of substitution between origin countries. In this case, $\varepsilon_{ji,k} = -1 - (\sigma_k - 1)(1 - \lambda_{ji,k})$, $\varepsilon^{(ji,k)}_{ni,g} = 0$ if $g \neq k$, and $\varepsilon^{(ji,k)}_{ni,k} = (\sigma_k - 1)\lambda_{ji,k}$. Also, for the sake of exposition, suppose each country is sufficiently small relative to the rest of the world, so that $\lambda_{ji,k},\ r_{ij,k} \approx 0$ if $j \neq i$. Plugging these values into Proposition 1, yields the following expression for the optimal trade taxes,

$$1 + t_{ji,k} = 1 + \tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k}$$
$$1 + x^*_{ij,k} = \left(\frac{\sigma_k}{\sigma_k - 1}\right) \left[1 + \frac{\gamma_k - 1}{\gamma_k} \sum_{n \neq i} \tilde{\delta}_{ni} v_{n,k} \lambda_{nj,k}\right]^{-1}. \tag{23}$$

The optimal import tax is the product of (1) $\tilde{\delta}_{ji} v_{j,k}$, which taxes high-emission imports, and (2) $\frac{\gamma_k - 1}{\gamma_k}$, which operates as a tax deflator for industries that exhibit high returns-to-scale in emission (low-γ_k). Considering this decomposition, the effectiveness of import taxes at reducing transboundary is dictated by $\mathrm{Cov}_k(v_{j,k}, \gamma_k)$. If $\mathrm{Cov}_k(v_{j,k}, \gamma_k) > 0$ import taxes are an ineffective emission-reducing instrument because the high-emission industries that have to be penalized are also the high-returns-to-scale industries whose production should not be contracted. Alternatively, if $\mathrm{Cov}_k(v_{j,k}, \gamma_k) < 0$ import taxes can be quite effective as they are hitting two birds with one stone.

Similarly, the optimal export tax includes an emission-correcting term (in brackets) that promotes country i's clean exports against its high-emission competition in market j. Specifically, suppose that good ij, k competes with high-emission (high-$\tilde{\delta}_{ni} v_{n,k}$) varieties in market j. In that case, country i's government will apply a relatively low export tax or even an export subsidy to good ij, k to increase its sales in market j against high-emission rivals there.[11]

To dig deeper, the magnitude of the emission-correcting term depends on the interaction between three terms. First, the lower γ_k, the larger the scope for scale economies in abatement. Hence, penalizing foreign varieties with export tax adjustments is less effective. Second, the smaller the perceived disutility from foreign emissions (lower $\delta_{ni} v_{ni}$ for $n \neq i$), the larger the incentive to use export policy to correct these emissions. Third, the greater the market share of high-emission international varieties in market j (higher $\lambda_{nj,k}$), the greater the incentive to promote exports of clean, locally-produced varieties to that market.

[11] Recall from Theorem 1, that emission-correcting term is governed by the emission externality of rival varieties ($\{\tilde{\delta}_{ni} v_{n,k}\}_{n \neq i}$) and the degree of cross-substitutability between ij, k and these rival varieties ($\varepsilon^{(ij,k)}_{nj,g}$). The latter effect in this special case is factored out in the term that depends on σ_k.

4 Policy Outcomes under Alternative Scenarios

In this section, we discuss policy outcomes in cases other than the unilateral first-best. In Section 4.1, we examine second-best scenarios in which governments act unilaterally in their self interest, but are afforded fewer tax instruments than is necessary to attain the first-best. We then change our perspective to the case of multi-lateral policies. In Section 4.2, we discuss global optimum that can be achieved via a deep international agreement among cooperative members. In Section 4.3, we characterize the non-cooperative Nash equilibrium where non-cooperative countries simultaneously chose their optimal unilateral policy.

4.1 Optimal Environmental Policy in Second-Best Scenarios

Theorem 1 concerns a unilaterally first-best scenario where government have access to a complete set of policy instruments. However, a government may face limitations in using all instruments of policy to achieve the first-best. The following definition puts it formally.

Definition. The *Second-best Unilateral Policy* for country i is achieved by choosing a *subset* of policy instruments to maximize $\mathcal{W}_i$ (equation 14) subject to equilibrium conditions (1)-(9).

In addition to prevalent political economy issues, second-best scenarios may arise from agreements on trade or environment that ties the hands of policymakers to flexibly exercise policy tools. For example, the WTO requires its members not to use export subsidies and set tariffs based on the principle of most-favored-nation among other principles. In that case, a country may manipulate its environmental policy not only for environmental objectives but also to manipulate its imports and exports. The opposite case would be a country that may use trade policy partly for environmental objectives. In line with these cases, we consider two second-best scenarios that are both of practical importance and have received considerable attention in the prior literature.

Case #1: Emission Taxes are Unavailable As the emission elasticity approaches zero, i.e., $\alpha_k \to 0$, our model collapses to a model with exogenous pollution intensity à la Markusen (1975) (See footnote (6) As such, emission taxes can be dropped from the model as firms do not undertake abatement. In this case, the optimal production tax will include the markup-correcting term $\frac{\gamma_k - 1}{\gamma_k}$ plus an extra term that taxes high-emission (high-v) industries. Namely,

$$1 + s_{i,k}^{\star\star} = \frac{\gamma_k - 1}{\gamma_k} \left(1 + \tilde{\delta}_{ii} v_{i,k}\right)$$

As before, the emission-correcting term depends on $\frac{\gamma_k - 1}{\gamma_k}$ because there are scale economies in emission. For instance, it may be optimal to subsidize a high-returns-to-scale industry that exhibits a high emission intensity. That is because subsidizing such an industry may lower emission through scale effects that dominate the higher firm-level emission intensity.

Alternatively, maintaining the assumption that $\alpha_k \in (0, 1)$, we could examine second-best production taxes within our general model. In this case, under non-optimal emission taxes, production taxes must correct the remaining emission externalities:

$$1 + s^{\star\star}_{i,k} = \frac{\gamma_k - 1}{\gamma_k}\left[1 + \tilde{\delta}_{ii}(v_{i,k} - v^{\star}_{i,k})\right]$$

where $v^{\star}_k$ is the emission intensity attainable under the first-best policy schedule.

Case #2: Emission Taxes are Used as Protection in Disguise There is widespread skepticism that environmental policies are occasionally used as protection in disguise. The argument is that when governments are banned from using trade taxes, they may turn to emission taxes as a second best trade-restricting instrument. Against this backdrop, we establish two important results. First, suppose export taxes are banned but governments can apply production and import taxes. In that case, the optimal emission tax remains uniform and coincides with the first-best optimal rate:

$$\tau^{*}_{i,k} = \tau^{*}_{i} = \tilde{\delta}^{*}_{ii}. \qquad (s, \tau, t \text{ available but } x \text{ banned})$$

Intuitively, import and production taxes are strictly more effective than emission taxes at mimicking export taxes. So, when import and production taxes are applicable, there is no rationale for using emission taxes to mimic export taxes.

Second, suppose that all tax instruments but emission taxes are banned. In that case, optimal emission taxes will be no longer uniform. Instead, it is optimal for country i to apply a higher emission tax on industries where it possesses more export market power. To make this point succinctly, consider a perfectly competitive economy ($f^{e}_{i,k} = 0, \gamma_k \to \infty$) in which $\alpha_k = \alpha$ is uniform across industries and preferences have a Cobb-Douglas-CES parmaterization given by equation (22). Then, as shown in Appendix C, the optimal emission tax is given by

$$\tau^{*}_{i,k} = \left(\frac{\alpha(1-\sigma_k)\left(1 - \lambda_{ii,k} r_{ii,k}\right) + 1}{\widetilde{\alpha}_i(1-\sigma_k)\left(1 - \lambda_{ii,k} r_{ii,k}\right) + r_{ii,k}}\right)\tilde{\delta}_{ii} \qquad (\text{only } \tau \text{ available})$$

where $\tilde{\alpha}_i > \alpha$ is a country-wide term that depends on the industry-composition of country i's production.[12] The above formula suggests that it is optimal to tax emission above the first-best level in low-σ industries. Doing so, enables country i to contract exports in high-market-power industries as an indirect means to extract a markup from the rest of the world.

Next, we present the optimal tax schedule in the case of international cooperation as well the non-cooperative Nash equilibrium case where countries simultaneously set their optimal unilateral policy. The former is akin to setting taxes in a closed-economy, and the latter relies on the optimal unilateral tax seclude presented in the previous section.

4.2 Global Optimum under International Cooperation

The globally optimal policy is the first-best from the global perspective. In other words, this is the best outcome expected from a deep international agreement on trade and environment between co-operative countries.

Definition. The *Global Optimum* is achieved by choosing policy instruments in each country to maximize global welfare, $\sum_i \mathcal{W}_i$, subject to equilibrium condition (1)-(9).

The globally optimal outcome involves zero trade taxes, as these taxes create inefficient distortions from a global perspective. The globally optimal production taxes solely correct the inefficiency from cross-industry markup heterogeneity. The globally optimal emission taxes are of Pigouvian nature correcting the local and transboundary emission externality. Stated formally, the optimal production subsidy and emission tax in country i is given by

$$\begin{aligned}
\boldsymbol{x}_i^* = \boldsymbol{t}_i^* &= \mathbf{0} \\
1 + s_{i,k}^* &= \frac{\gamma_k - 1}{\gamma_k} \\
\tau_{i,k}^* = \tau_i^* &= \sum_{j \in \mathbb{C}} \tilde{P}_j \delta_{ij}
\end{aligned}$$

To provide intuition, optimal emission taxes discriminate by country of origin because (i) the disutility δ_{ij} from emission in i is non-uniform across locations j, and (ii) converting a dollar loss in i (as a

[12] More specifically, $\widetilde{\alpha}_i$ exhibits the following formulation:

$$\widetilde{\alpha}_i = \alpha + \frac{\sum_g \left(\frac{\left[1-\alpha(1-r_{ii,g})\right]\left(1-r_{ii,g}\right)}{\left[\alpha\epsilon_g\left(1-\lambda_{ii,g} r_{ii,g}\right)-1\right](1-\lambda_{ii})} r_{i,g} \right)}{1 - \sum_g \left(\frac{\left[\epsilon_g\left(1-\lambda_{ii,g} r_{ii,g}\right)\right]\left(1-\alpha(1-r_{ii,g})\right)}{\left[\alpha\epsilon_g\left(1-\lambda_{ii,g} r_{ii,g}\right)-1\right](1-\lambda_{ii})} r_{i,g} \right)}$$

result of higher emission taxes) to a welfare loss requires an adjustment by the consumer price index in i. To gains more intuition, consider the special case where $\delta_{ij} = \bar{L}_j \bar{\delta}$, i.e., the disutility from a unit of emission produced in origin i is uniformly harmful to each individual irrespective of residence. In that case, the optimal emission tax relative to CPI is uniform across countries, $\tau_i^* / \tilde{P}_j = \bar{\delta} \bar{L}^{world}$.

4.3 Non-Cooperative Nash Equilibrium

As in the previous cases, we start with a formal definition of the non-cooperative Nash equilibrium.

Definition. The *Non-Cooperative Nash Equilibrium* corresponds to a case where non-cooperative countries simultaneously choose their optimal unilateral policy taking the policy choice in the rest of the world as given.[13]

In the Nash equilibrium, the unilaterally optimal emission and production tax formulas are still characterized by Theorem 1. However, the trade share, $\lambda_{nj,k}$, and emission intensities, $\nu_{j,k}$, in these formulas now depend on policy choices in the rest of the world. Specifically, Consider country i's optimal export and import taxes. They depend on transboundary emission intensities, $\{v_{j,k}\}_{j \neq i}$, which are regulated by optimal emission taxes adopted by other countries ($j \neq i$). Using equation (12) and given that $\tau_{j,k}^* = \tilde{\delta}_{jj}$ for all $j \in \mathbb{C}$,

$$v_{j,k}^{\star} = \alpha_k \frac{\gamma_k - 1}{\gamma_k} \tilde{\delta}_{jj}^{-1}$$

Supposing preferences are Cobb-Douglas-CES and each country is sufficiently small relative to the rest of the world, we can plug the above expression into Equation (23) to arrive at the following optimal trade tax schedule.

Proposition 2. *The non-cooperative Nash equilibrium is characterized by each country applying the following tax schedule:*

$$\begin{cases} 1 + t_{ji,k}^* = 1 + \alpha_k \left(\frac{\gamma_k - 1}{\gamma_k} \right)^2 \frac{\tilde{\delta}_{ji}}{\tilde{\delta}_{jj}} \\ 1 + x_{ij,k}^* = \frac{\sigma_k}{\sigma_k - 1} \left[1 + \alpha_k \left(\frac{\gamma_k - 1}{\gamma_k} \right)^2 \sum_{n \neq i} \frac{\tilde{\delta}_{ni}}{\tilde{\delta}_{nn}} \lambda_{nj,k} \right]^{-1} \\ 1 + s_{i,k}^* = \frac{\gamma_k - 1}{\gamma_k} \\ \tau_{i,k}^* = \tau_i^* = \tilde{\delta}_{ii} \end{cases}.$$

[13] This situation is akin to a one-shot non-cooperative Nash game.

The optimal emission and production taxes remain the same (as in Equation 21), even though all countries simultaneously apply trade, emission, and production taxes. That is because, as long as trade taxes are available, the unilaterally optimal tax rate for emission and production is independent of economic variables in the rest of the world—see Theorem 1.

As in the unilateral case, optimal trade taxes in country i correct transboundary emission externalities. But the extent of these externalities, here, depends on cross-national differences in the perceived cost of emission. For instance, suppose country i's government cares significantly more about pollution than its counterpart in country j. This situation corresponds to a high $\tilde{\delta}_{ji}/\tilde{\delta}_{jj}$, and asks for a large tariff on imports from country j. Correspondingly, if governments care significantly less about transboundary versus local emission, then $\tilde{\delta}_{ji}/\tilde{\delta}_{jj} \approx 0$ and tariffs may become redundant. In the case that countries have similar and symmetric preferences *vis-à-vis* local and transboundary emission, i.e., $\tilde{\delta}_{jj} = \tilde{\delta}_{ji}$, it is optimal for country i to charge an import tax that is proportional to the industry-level emission elasticity:

$$1 + t^*_{ji,k} = 1 + \alpha_k \left(\frac{\gamma_k - 1}{\gamma_k} \right)^2.$$

Intuitively, from country i's perspective, country j's emission tax on good ji, k is sub-optimal as it does not internalize the transboundary cost of emission. So, it is optimal for country i to tax imports originating from *high-*$\alpha_k \times$*high-*γ_k industries in country j to partially correct the transboundary pollution externality.

A similar logic explains why the square of the inverse markup, $\left(\frac{\gamma_k - 1}{\gamma_k}\right)^2$, appears in formulas specified under Proposition 2. According to equation (3), emission intensity per unit of production, $Z_{n,k}/Q_{n,k}$ is proportional to $\left(Q_{n,k}/(1-a_{n,k})\right)^{-1/\gamma_k}$. That is, the emission intensity is affected by scale economies in both production and abatement, governed by a common parameter γ_k. Now, consider the formula for optimal import taxes $t^*_{ji,k}$. The first $(\gamma_k - 1)/\gamma_k$ reflects the importing country i's desire to dampen the emission-correcting tariff given scale economies in "production". The second $(\gamma_k - 1)/\gamma_k$ is due to the origin country j's emission taxes interacting with scale economies in "emission".

5 Mapping Theory to Data

This section describes how our optimal tax formulas map to data. Our objective is to use this mapping to quantify the environmental gains from optimal trade taxation. To this end, we can run different exercises depending on which scenario we choose for the baseline emission taxes. Below, we adopt the conservative assumption that, under the status quo, home country i applies its optimal emission taxes and every foreign country $j \neq i$ has zero emission taxes. This scenario is not as restrictive as it may appear, since the optimal emission tax can assume any value depending on δ_{ii}. Perhaps more restrictive, we assume that the applied emission tax prompts abatement by the local firms.

To quantify the gains from taxation with our formulas, we focus on the Cobb-Douglas-CES case of our model and employ the exact hat-algebra technique. The basic idea behind our approach is to track the (taxation-induced) change in every variable z, using the hat notation: $\hat{z} = z'/z$. Invoking this notation, we jointly solve a system of equations consisting of our optimal tax formula and equilibrium conditions.

The system of equations that determine the gains from optimal policy can be expressed as follows:

$$1 + t^*_{ji,k} = 1 + \tilde{\delta}_{ji} v_{j,k} \hat{v}_{j,k} \hat{\tilde{P}}_i - \frac{r_{ji,k}\hat{r}_{ji,k}}{\gamma_k - \sum r_{jn,k}\hat{r}_{jn,k}\left(1 + \epsilon_k(1 - \lambda_{jn,k}\hat{\lambda}_{jn,k})\right)} \quad \text{(import tax)}$$

$$1 + x^*_{ij,k} = \left(1 + \frac{1}{\epsilon_k(1 - \lambda_{ij,k}\hat{\lambda}_{ij,k})}\right)\left[1 + \sum_{n \neq i} \tilde{\delta}_{ni} v_{n,k} \lambda_{nj,k} \hat{v}_{n,k} \hat{\lambda}_{nj,k}\ \hat{\tilde{P}}_i\right]^{-1} \quad \text{(export tax)}$$

$$1 + s^*_{i,k} = (\gamma_k - 1)/\gamma_k; \quad \hat{\tau}^*_{i,k} = \hat{\tilde{P}}_i; \qquad \text{(emission and production tax)}$$

$$\hat{v}_{i,k} = \alpha_k / \hat{\tau}^*_{i,k}; \quad \widehat{1 - a_{i,k}} = \left(\hat{w}_i / \hat{\tau}_{i,k}\right)^{\alpha_k}; \quad \widehat{1 - a_{j,k}} = 1,\ j \neq i$$

$$\hat{\tilde{P}}_{i,k} = \left[\sum_n \lambda_{ni,k} \tilde{P}^{-\epsilon_k}_{ni,k}\right]^{-1/\epsilon_k}; \quad \hat{\tilde{P}}_{ni,k} = \hat{w}_n (1 + t^*_{ni,k})(1 + x^*_{ni,k})(1 + s^*_{n,k})(\widehat{1 - a_{n,k}})^{\frac{1}{\gamma_k} - 1} \hat{Q}_{n,k}^{-\frac{1}{\gamma_k}}$$

$$\hat{\lambda}_{ji,k} = \left(\hat{w}_n (1 + t^*_{ji,k})(1 + x^*_{ji,k})(1 + s^*_{j,k})\right)^{-\epsilon_k} \hat{\tilde{P}}^{\epsilon_k}_{i,k}; \quad \hat{Q}_{i,k} = \sum_n r_{in,k} \frac{\hat{\lambda}_{in,k}\hat{Y}_n}{\hat{\tilde{P}}_{in,k}}$$

$$\hat{Y}_i Y_i = \hat{w}_i w_i \bar{L}_i + \hat{T}_i T_i + \hat{R}_i R_i; \qquad \text{(Income = Wage Bill + Emission Tax Rev. + Production/Trade Tax Rev.)}$$

$$\hat{w}_i w_i \bar{L}_i = \sum_{k \in \mathbb{K}} \sum_{i \in \mathbb{C}} \left[\frac{(1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k})\hat{\lambda}_{ij,k}\lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j}{(1 + s_{i,k})(1 + x_{ij,k})(1 + t_{ij,k})}\right]; \quad \hat{T}_i T_i = \sum_{k \in \mathbb{K}} \sum_{i \in \mathbb{C}} \left[\frac{\alpha_k \frac{\gamma_k - 1}{\gamma_k} \hat{\lambda}_{ij,k}\lambda_{ij,k} e_{j,k}\hat{Y}_j Y_j}{(1 + s_{i,k})(1 + x_{ij,k})(1 + t_{ij,k})}\right]$$

$$\hat{R}_i R_i = \sum_{k \in \mathbb{K}} \sum_{i \in \mathbb{C}} \left[\frac{\left[(1 + s_{i,k})(1 + x_{ij,k}) - 1\right]\hat{\lambda}_{ij,k}\lambda_{ij,k} e_{j,k}\hat{Y}_j Y_j}{(1 + s_{i,k})(1 + x_{ij,k})(1 + t_{ij,k})} + \frac{t_{ji,k}\hat{\lambda}_{ji,k}\lambda_{ji,k} e_{i,k}\hat{Y}_i Y_i}{(1 + s_{j,k})(1 + x_{ji,k})(1 + t_{ji,k})}\right]$$

The first three rows, in the above system, govern the optimal tax choice. The remaining rows govern the change in equilibrium variables, subject to optimal production/consumption choices and market clearing conditions.

To evaluate the above system we need to estimate following parameters per industry:

1. The trade elasticity, $\epsilon_k \equiv (\sigma_k - 1)$;

2. The emission elasticity, α_k; and

3. The degree of firm-level market power, γ_k, which is tied to the markup, $\mu_k \equiv \gamma_k/(\gamma_k - 1)$.

We also need data on total expenditure Y_i; expenditure shares, $\lambda_{ji,k}$, and $e_{i,k}$; revenue shares, $r_{ji,k}$; applied taxes, $t_{ji,k}$, $x_{ij,k}$, and $s_{i,k}$; emission intensities, $v_{i,k}$; and the unit cost of emission to welfare, $\tilde{\delta}_{ni}$.

5.1 Data Sources

[To be completed]

5.2 Estimation of γ_k, ϵ_k, and α_k

[To be completed]

5.3 Quantitative Exercises

[To be completed]

References

Mustafa H BABIKER : Climate change policy, market structure, and carbon leakage. *Journal of international Economics*, 65(2):421–445, 2005. 1

Dominick G BARTELME, Arnaud COSTINOT, Dave DONALDSON et Andrés RODRÍGUEZ-CLARE : The textbook case for industrial policy: Theory meets data. Rapport technique, National Bureau of Economic Research, 2019. 1

Mostafa BESHKAR et Ahmad LASHKARIPOUR : The cost of dissolving the wto: The role of global value chains. 2020. 1

Christoph BÖHRINGER, Jared C CARBONE et Thomas F RUTHERFORD : The strategic value of carbon tariffs. *American Economic Journal: Economic Policy*, 8(1):28–51, 2016. 1

Jevan CHERNIWCHAN, Brian R COPELAND et M Scott TAYLOR : Trade and the environment: New methods, measurements, and results. *Annual Review of Economics*, 9:59–85, 2017. 1

Brian R COPELAND : Pollution content tariffs, environmental rent shifting, and the control of cross-border pollution. *Journal of international Economics*, 40(3-4):459–476, 1996. 1

Brian R Copeland et M Scott Taylor : Trade, growth, and the environment. *Journal of Economic literature*, 42(1):7–71, 2004. 1, 2.2

Arnaud Costinot, Dave Donaldson, Jonathan Vogel et Iván Werning : Comparative advantage and optimal trade policy. *The Quarterly Journal of Economics*, 130(2):659–702, 2015. 1

Arnaud Costinot et Andrés Rodríguez-Clare : Trade theory with numbers: Quantifying the consequences of globalization. *In Handbook of international economics*, volume 4, pages 197–261. Elsevier, 2014. 1

Arnaud Costinot, Andrés Rodríguez-Clare et Iván Werning : Micro to macro: Optimal trade policy with firm heterogeneity. Rapport technique, National Bureau of Economic Research, 2016. 1

Rudiger Dornbusch, Stanley Fischer et Paul Anthony Samuelson : Comparative advantage, trade, and payments in a ricardian model with a continuum of goods. *The American Economic Review*, 67(5):823–839, 1977. 3, 3.2

Jonathan Eaton et Samuel Kortum : Technology, geography, and trade. *Econometrica*, 70(5):1741–1779, 2002. 3, 3.2

Joshua Elliott, Ian Foster, Samuel Kortum, Todd Munson, Fernando Perez Cervantes et David Weisbach : Trade and carbon taxes. *American Economic Review*, 100(2):465–69, 2010. 1

Roger A Horn et Charles R Johnson : *Matrix analysis*. Cambridge university press, 2012. B.3, B.4

Ahmad Lashkaripour et Volodymyr Lugovskyy : Profits, scale economies, and the gains from trade and industrial policy. 2016. 1

James R Markusen : International externalities and optimal tax structures. *Journal of international economics*, 5(1):15–29, 1975. 1, 2, 1, 4.1

Andreu Mas-Colell, Michael Dennis Whinston, Jerry R Green *et al.* : *Microeconomic theory*, volume 1. Oxford university press New York, 1995. B.3

William Nordhaus : Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, 105(4):1339–70, 2015. 1, 1

Alexander M Ostrowski : Note on bounds for determinants with dominant principal diagonal. *Proceedings of the American Mathematical Society*, 3(1):26–30, 1952. B.3

Joseph S SHAPIRO : Trade costs, co 2, and the environment. *American Economic Journal: Economic Policy*, 8(4):220–54, 2016. 2.5

Joseph S SHAPIRO : The environmental bias of trade policy. Rapport technique, National Bureau of Economic Research, 2020. 1

Joseph S SHAPIRO et Reed WALKER : Why is pollution from us manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review*, 108(12):3814–54, 2018. 1

Daniel M STURM : Trade and the environment: A survey of the literature. *In Environmental policy in an international perspective*, pages 119–149. Springer, 2003. 2

Appendix

A Theoretical Preliminaries

A.1 Detailed Statement of the Optimal Unilateral Policy Problem

We derive optimal unilateral policy for the government in country i, which here we refer to as the home country. We denote by $\mathbb{P}_i \equiv \{\tilde{P}_{ii,k}, \tilde{P}_{ji,k}, \tilde{P}_{ij,k}, a_{i,k}\}_{j\neq i, j\in\mathbb{C},\, k\in\mathbb{K}}$ the policy instruments in country i, by $\tilde{\mathbf{P}}_i \equiv \{\tilde{P}_{ji,k}\}_{j\in\mathbb{C},\, k\in\mathbb{K}}$ the vector of consumer prices in country i, and by $\mathbf{w} \equiv \{w_j\}_{j\in\mathbb{C}}$ the vector of wages. The problem of the government in country i is:

$$\max_{I_i} V_i(Y_i, \tilde{\mathbf{P}}_i) - \sum_{n\in\mathbb{C}} \sum_{g\in\mathbb{K}} \delta_{ni} Z_{n,g}(a_{n,g}; Q_{n,g})$$

subject to the following equilibrium relationships, for all $i, j \in \mathbb{C}$,and $k \in \mathbb{K}$,

$$\text{(Optimal Demand)} \qquad Q_{ji,k} = \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$$

$$\text{(Producer Price)} \qquad P_{ji,k}(w_j, a_{j,k}; Q_{j,k}) = \bar{d}_{ji,k}\bar{p}_{jj,k} w_j (1-a_{j,k})^{\frac{1}{\gamma_k}-1} Q_{j,k}^{-\frac{1}{\gamma_k}}$$

$$\text{(Pollution)} \qquad Z_{j,k}(a_{j,k}; Q_{j,k}) \equiv \bar{z}_{j,k}(1-a_{j,k})^{\frac{1}{\alpha_k}+\frac{1}{\gamma_k}-1} Q_{j,k}^{1-\frac{1}{\gamma_k}}$$

$$\text{(Income = Revneue)} \qquad Y_i = w_i\bar{L}_i + \sum_{k,\, j\neq i} \left[\left(\tilde{P}_{ji,k} - P_{ji,k}\right) Q_{ji,k}\right] + \sum_{k,\, j} \left[\left(\tilde{P}_{ij,k} - (1-\alpha_k\frac{\gamma_k - 1}{\gamma_k}) P_{ij,k}\right) Q_{ij,k}\right] + \bar{D}_i$$

$$\text{(Trade Deficit)} \qquad B_i \equiv \sum_{j\neq i}\sum_k P_{ji,k}Q_{ji,k} - \sum_{j\neq i}\sum_k \tilde{P}_{ij,k}Q_{ij,k} - \bar{D}_i = 0$$

Here, we have written every variable as a function of (1) wages, (2) all or a subset of policy instruments, and (3) quantities. Equations for producer price and pollution reproduce (2) and (3), in which $Q_{j,k} = \sum_i \bar{d}_{ji,k} Q_{ji,k}$. The equation for income reproduces (7) only in a more compact way by replacing for taxes from (8), and the trade deficit condition is equivalent to factor market clearing condition (See footnote 9). The demand function $\mathcal{D}_{ji,k}$ is characterized by the set demand elasticities defined in Section 2.1. Throughput our proof, we assign the factor in one foreign country as the numeraire.

A.2 Expressing Equilibrium Outcomes as a Function of $(\mathbb{P}_i; \mathbf{w})$

Consider system (S^w) that incorporates all equilibrium conditions excluding the labor-market clearing condition. For all $n, j \in \mathbb{C}$,and $k \in \mathbb{K}$,

$$
\begin{array}{rl}
\text{(Optimal Demand)} & Q_{nj,k}(\mathbb{P}_i; \mathbf{w}) = \begin{cases} D_{ni,k}(\tilde{\mathbf{P}}_i, Y_i(\mathbb{P}_i; \mathbf{w})) & \text{if } j = i \\ D_{nj,k}(\tilde{\mathbf{P}}_{ij}, \{\tilde{\mathbf{P}}_{nj}(\mathbb{P}_i; \mathbf{w})\}_{n \neq i}, Y_j(\mathbb{P}_i; \mathbf{w})) & \text{if } j \neq i \end{cases} \\
\text{(Indusry Output)} & Q_{n,k}(\mathbb{P}_i; \mathbf{w}) = \sum_{j \in \mathbb{C}} \bar{d}_{nj,k} Q_{nj,k}(\mathbb{P}_i; \mathbf{w}) \\
\text{(Producer Price)} & P_{nj,k}(\mathbb{P}_i; \mathbf{w}) = \bar{d}_{nj,k} \bar{p}_{nn,k} w_n (1 - a_{n,k})^{\frac{1}{\gamma_k} - 1} \left(Q_{n,k}(\mathbb{P}_i; \mathbf{w})\right)^{-\frac{1}{\gamma_k}} \\
\text{(Pollution)} & Z_{n,k}(\mathbb{P}_i; \mathbf{w}) = \bar{z}_{n,k} (1 - a_{n,k})^{\frac{1}{\alpha_k} + \frac{1}{\gamma_k} - 1} \left(Q_{n,k}(\mathbb{P}_i; \mathbf{w})\right)^{-\frac{1}{\gamma_k}} \\
\text{(Tax Revenues)} & T_n(\mathbb{P}_i; \mathbf{w}) = \begin{cases} \sum_{k,\, j \neq i} \left[\left(\tilde{P}_{ji,k} - P_{ji,k}(\mathbb{P}_i; \mathbf{w}) \right) Q_{ji,k}(\mathbb{P}_i; \mathbf{w}) \right] & \text{if } n = i \\ + \sum_{k,\, j} \left[\left(\tilde{P}_{ij,k} - (1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}) P_{ij,k}(\mathbb{P}_i; \mathbf{w}) \right) Q_{ij,k}(\mathbb{P}_i; \mathbf{w}) \right] & \\ 0 & \text{if } n \neq i \end{cases} \\
\text{(Income)} & Y_n(\mathbb{P}_i; \mathbf{w}) = w_n \bar{L}_n + \bar{D}_n + T_n(\mathbb{P}_i; \mathbf{w})
\end{array}
\qquad (S^w)
$$

Here, $\tilde{\mathbf{P}}_i \subset \mathbb{P}_i$ is the vector of consumer prices in home country i, $\tilde{\mathbf{P}}_{ij} \subset \mathbb{P}_i$ is the vector of consumer prices in foreign country j of varieties produced in home, and $a_{i,k} \in \mathbb{P}_i$ is the abatement in home. All these are instruments of policy to be chosen by the home government. In contrast, every foreign country $n \neq i$ has some fixed abatement level $a_{n,k} = \bar{a}_{n,k}$ and no tax revenues $T_n = 0$. System (S^w) characterizes quantities, producer prices, emissions, tax revenues, and income in all economies as a function of $(\mathbb{P}_i, \mathbf{w})$. Correspondingly, welfare in country i can be formulated as,

$$W_i(\mathbb{P}_i; \mathbf{w}) = V_i(Y_i(\mathbb{P}_i; \mathbf{w}), \tilde{\mathbf{P}}_i) - \sum_{n,k} \delta_{ni} Z_{i,k}(\mathbb{P}_i; \mathbf{w}).$$

By design, system (S^w) excludes the labor-market clearing condition, and it is understood that many wage vectors may satisfy (S^w). For a given choice of policy, $\mathbb{P}_i$, a wage vector, $\mathbf{w}$, is in the feasible set

$\mathbb{F}_i^w$ if and only if it satisfies the labor-market clearing conditions:

$$\sum_{j,k}\left[\left(1-\alpha_k\frac{\gamma_k-1}{\gamma_k}\right)P_{nj,k}(\mathbb{P}_i;\mathbf{w})Q_{nj,k}(\mathbb{P}_i;\mathbf{w})\right]=w_n\bar{L}_n,\qquad \forall n.$$

A.3 Expressing Equilibrium Outcomes as a Function of $(\mathbb{P}_i; Y_i)$

Following Lemma 2, we treat wages as fixed. Consider system (S^Y) that incorporates all equilibrium conditions excluding the budget constraint. For all $n, j \in \mathbb{C}$,and $k \in \mathbb{K}$,

$$\begin{aligned}
&\text{(Optimal Demand)} && Q_{nj,k}(\mathbb{P}_i;Y_i)=\begin{cases}D_{ni,k}(\tilde{\mathbf{P}}_i,Y_i) & \text{if } j=i\\ D_{nj,k}(\tilde{\mathbf{P}}_{ij},\{\tilde{\mathbf{P}}_{nj}(\mathbb{P}_i;Y_i)\}_{n\neq i},\bar{Y}_j) & \text{if } j\neq i\end{cases}\\
&\text{(Indusry Output)} && Q_{n,k}(\mathbb{P}_i;Y_i)=\textstyle\sum_{j\in\mathbb{C}}\bar{d}_{nj,k}Q_{nj,k}(\mathbb{P}_i;Y_i)\\
&\text{(Producer Price)} && P_{nj,k}(\mathbb{P}_i;Y_i)=\bar{d}_{nj,k}\bar{p}_{nn,k}\bar{w}_n(1-a_{n,k})^{\frac{1}{\gamma_k}-1}\left(Q_{n,k}(\mathbb{P}_i;Y_i)\right)^{-\frac{1}{\gamma_k}}\\
&\text{(Pollution)} && Z_{n,k}(\mathbb{P}_i;Y_i)=\bar{z}_{n,k}(1-a_{n,k})^{\frac{1}{\alpha_k}+\frac{1}{\gamma_k}-1}\left(Q_{n,k}(\mathbb{P}_i;Y_i)\right)^{1-\frac{1}{\gamma_k}}\\
&\text{(Taxes)} && T_n(\mathbb{P}_i;Y_i)=\begin{cases}\sum_{k,\,j\neq i}\left[\left(\tilde{P}_{ji,k}-P_{ji,k}(\mathbb{P}_i;Y_i)\right)Q_{ji,k}(\mathbb{P}_i;Y_i)\right] & \text{if } n=i\\ +\sum_{k,\,j}\left[\left(\tilde{P}_{ij,k}-(1-\alpha_k\frac{\gamma_k-1}{\gamma_k})P_{ij,k}(\mathbb{P}_i;Y_i)\right)Q_{ij,k}(\mathbb{P}_i;Y_i)\right] & \\ 0 & \text{if } n\neq i\end{cases}
\end{aligned}\qquad (S^Y)$$

System (S^Y) characterizes quantities, producer prices, emissions, and tax revenues in all economies as a function of $(\mathbb{P}_i, Y_i)$. Correspondingly, welfare in country i can be formulated as,

$$W_i(\mathbb{P}_i;Y_i)=V_i(\bar{w}_i\bar{L}_i+\bar{D}_i+T_i(\mathbb{P}_i;Y_i),\tilde{\mathbf{P}}_i)-\sum_{n,k}\delta_{ni}Z_{i,k}(\mathbb{P}_i;Y_i).$$

A policy-income pair is feasible, denoted by $(\mathbb{P}_i, Y_i) \in \mathbb{F}_i^Y$, if and only if $Y_i = \bar{w}_i\bar{L}_i + \bar{D}_i + T_i(\mathbb{P}_i; Y_i)$.

A.4 Characterizing Equilibrium Wage Effects

Suppose we formulate all equilibrium variables as a function of $\mathbb{P}_i$ and $\mathbf{w}$ (described in Appendix A.2). The feasible vector of wages, $\mathbf{w}$, solves the following system of labor market clearing conditions:

$$\begin{cases}f_1(\mathbb{P}_i;\mathbf{w})\equiv w_1L_1-\sum_{j\in\mathbb{C}}\sum_{k\in\mathbb{K}}(1-\alpha_k\frac{\gamma_k-1}{\gamma_k})P_{1j,k}(\mathbb{P}_i;\mathbf{w})Q_{1j,k}(\mathbb{P}_i;\mathbf{w})=0\\ \vdots\\ f_N(\mathbb{P}_i;\mathbf{w})\equiv w_NL_N-\sum_{j\in\mathbb{C}}\sum_{k\in\mathbb{K}}(1-\alpha_k\frac{\gamma_k-1}{\gamma_k})P_{Nj,k}(\mathbb{P}_i;\mathbf{w})Q_{Nj,k}(\mathbb{P}_i;\mathbf{w})=0\end{cases}\qquad (24)$$

Also, note that by Walras' law one equation is redundant so we can assign one element of $\mathbf{w}$ as the numeraire:

$$\sum_n f_n(\mathbb{P}_i; \mathbf{w}) = 0. \qquad \text{[Walras' Law]}$$

To characterize the term $\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}\mathcal{P}_i}$ in the F.O.C., we can apply the Implicit Function Theorem to the above system as follows:

$$\frac{\mathrm{d}\ln \mathbf{w}}{\mathrm{d}\ln \mathcal{P}_i} = -\left(\frac{\partial \boldsymbol{f}}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i}^{-1} \frac{\partial \boldsymbol{f}}{\partial \ln \mathcal{P}_i}.$$

To characterize the matrix $\frac{\partial f}{\partial \mathbf{w}}$, let us briefly abstract from scale economies and abatement, which amounts to setting $\alpha_k \frac{\gamma_k - 1}{\gamma_k} = 0$ in System 24. This simplification helps us convey our main point succinctly; but it does not imply it. As we argue shortly, our main claim goes through without this simplification. Taking partial derivatives from System 24 w.r.t. $\mathbf{w}$ holding $\mathbb{P}_i$ fixed, yields

$$\left(\frac{\partial \boldsymbol{f}}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i} = \begin{bmatrix} \frac{\partial f_1}{\partial \ln w_1} & \frac{\partial f_1}{\partial \ln w_2} & \cdots & \frac{\partial f_1}{\partial \ln w_N} \\ \frac{\partial f_2}{\partial \ln w_1} & \frac{\partial f_2}{\partial \ln w_2} & \cdots & \frac{\partial f_2}{\partial \ln w_N} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial f_N}{\partial \ln w_1} & \frac{\partial f_N}{\partial \ln w_2} & \cdots & \frac{\partial f_N}{\partial \ln w_N} \end{bmatrix} = \begin{bmatrix} 1 - \sum_{k,g} r_{11,k}\left(\eta_{11,k} + \varepsilon_{11,k}^{(11,g)}\right) & \cdots & -\sum_{k,g} r_{1N,k}\left(\eta_{1N,k} + \varepsilon_{1N,k}^{(NN,g)}\right) \\ 1 - \sum_{k,g} r_{21,k}\left(\eta_{21,k} + \varepsilon_{21,k}^{(11,g)}\right) & \cdots & -\sum_{k,g} r_{2N,k}\left(\eta_{2N,k} + \varepsilon_{2N,k}^{(NN,g)}\right) \\ \vdots & \ddots & \vdots \\ 1 - \sum_{k,g} r_{N1,k}\left(\eta_{N1,k} + \varepsilon_{N1,k}^{(11,g)}\right) & \cdots & -\sum_{k,g} r_{NN,k}\left(\eta_{NN,k} + \varepsilon_{NN,k}^{(NN,g)}\right) \end{bmatrix}.$$

Under Cobb-Douglas-CES preferences, the above matrix assumes the following parameterization:

$$\left(\frac{\partial \boldsymbol{f}}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i} = \mathbf{I} - \underbrace{\begin{bmatrix} -\sum_k [r_{11,k}\epsilon_k(1-\lambda_{11,k})] & \sum_k [r_{12,k}(1+\epsilon_k\lambda_{22,k})] & \cdots & \sum_k [r_{1N,k}(1+\epsilon_k\lambda_{NN,k})] \\ \sum_k [r_{21,k}(1+\epsilon_k\lambda_{11,k})] & -\sum_k [r_{22,k}\epsilon_k(1-\lambda_{22,k})] & \cdots & \sum_k [r_{2N,k}(1+\epsilon_k\lambda_{NN,k})] \\ \vdots & \ddots & \ddots & \vdots \\ \sum_k [r_{N1,k}(1+\epsilon_k\lambda_{11,k})] & \sum_k [r_{N2,k}(1+\epsilon_k\lambda_{22,k})] & \cdots & -\sum_k [r_{NN,k}\epsilon_k(1-\lambda_{NN,k})] \end{bmatrix}}_{\mathbf{A}}$$

Noting that $r_{ij,k}\epsilon_k(1-\lambda_{jj,k}) \ll 1$ if $j \neq i$, we can produce the following approximation:[14]

$$\left(\frac{\partial \boldsymbol{f}}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i}^{-1} = (\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots \approx$$

$$\mathbf{I} + \sum_{m=1}^{\infty} \operatorname{diag}\left(\left[-\sum_{k\in\mathbb{K}} r_{ii,k}\epsilon_k(1-\lambda_{ii,k})\right]^m\right) = \operatorname{diag}\left(\left[\sum_{k\in\mathbb{K}} 1 + r_{ii,k}\epsilon_k(1-\lambda_{ii,k})\right]^{-1}\right).$$

[14]The last line follows from the fact that for $a \in \mathbb{R}_+$,

$$\sum_{n=1}^{\infty} (-a)^n = -\frac{a}{1+a}.$$

The above equation indicates that $\left(\frac{\partial \boldsymbol{f}}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i}$ is *nearly* diagonal with smaller-than-unity diagonal elements. Now, consider the case where $\mathcal{P}_i = \tilde{P}_{ji,k}$ and assign w_j as the numeraire. The derivative of $\boldsymbol{f}_{-j}$ (i.e., $\boldsymbol{f}$ excluding row j) w.r.t. $\tilde{P}_{ji,k}$ holding $\mathbf{w}$ and $\mathbb{P}_i - \tilde{P}_{ji,k}$ fixed is given by:

$$\frac{\partial \boldsymbol{f}_{-j}}{\partial \ln \tilde{P}_{ji,k}} = \begin{bmatrix} \frac{\partial f_1}{\partial \ln \tilde{P}_{ji,k}} \\ \frac{\partial f_2}{\partial \ln \tilde{P}_{ji,k}} \\ \vdots \\ \frac{\partial f_N}{\partial \ln \tilde{P}_{ji,k}} \end{bmatrix} = \begin{bmatrix} \sum_g r_{1i,g}\varepsilon_{1i,g}^{(ji,k)} \\ \sum_g r_{2i,g}\varepsilon_{2i,g}^{(ji,k)} \\ \vdots \\ \sum_g r_{Ni,g}\varepsilon_{Ni,g}^{(ji,k)} \end{bmatrix} \xrightarrow{\text{Cobb-Douglas-CES}} = \begin{bmatrix} r_{1i} \\ \vdots \\ r_{j-1i} \\ r_{j+1i} \\ \vdots \\ r_{Ni} \end{bmatrix} \lambda_{ji,k}\epsilon_k$$

Given that (i) $\lambda_{ji,k} r_{ni} \approx 0$ if n and $j \neq i$, and (ii) $\left(\frac{\partial \boldsymbol{f}}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i}$ is *nearly* diagonal with smaller-than-unity diagonal elements, it immediately follows that

$$\frac{\mathrm{d} \ln \mathbf{w}_{-\{i,j\}}}{\mathrm{d} \ln \tilde{P}_{ji,k}} = \left(\frac{\partial \boldsymbol{f}_{-j}}{\partial \ln \mathbf{w}_{-\{i,j\}}}\right)^{-1}_{\mathbb{P}_i} \frac{\partial \boldsymbol{f}_{-j}}{\partial \ln \tilde{P}_{ij}} \approx 0,$$

where $\mathbf{w}_{-\{i,j\}}$ denotes the wage vector $\mathbf{w}$ excluding elements i and j. The same steps can be taken with regards to nay other price instrument in $\mathbb{P}_i$. Furthermore, the above argument goes through if we allow for a finite γ_k and a non-zero α_k.

A.5 Some Useful Relationships

Before turning to our derivations of optimal policy, we show two sets of useful relationships. The first one is the effects of policy instruments on emission levels. The second one is the effects of policy instruments on producer prices through industry-level scale economies.

Scale Effects in Emission. Recall that total emission, as a function of abatement and output, is given by

$$Z_{j,k}(a_{j,k}; Q_{j,k}) \equiv \bar{z}_{j,k}(1 - a_{j,k})^{\frac{1}{\alpha_k} + \frac{1}{\gamma_k} - 1} Q_{j,k}^{1 - \frac{1}{\gamma_k}}.$$

To track the policy response of emission we use two following partial derivatives. The first one, accounts for scale effects in emission:

$$\frac{\partial \ln Z_{j,k}(a_{j,k}, Q_{j,k})}{\partial \ln Q_{j,k}} = 1 - \frac{1}{\gamma_k}, \tag{25}$$

and, the second one accounts for abatement effects in emission:

$$\frac{\partial \ln Z_{j,k}(a_{j,k}, Q_{j,k})}{\partial \ln(1-a_{j,k})} = \frac{1}{\alpha_k} + \frac{1}{\gamma_k} - 1. \tag{26}$$

Note that $a_{i,k}$ is directly chosen by the policy-maker in our reformulated optimal policy problem. $Q_{j,k}(\mathbb{P}_i; \mathbf{w}, Y_i)$ is implicitly determined by the policy-maker with respect to abatement and the remaining price instruments.

Scale Economies in Production and the Export Supply Elasticity. Below, we define and characterize the export supply elasticity. To that end, we first introduce some intermediate partial derivatives that enter the export supply elasticity formula. These partial derivatives are also independently useful to our subsequent optimal analysis.

Note that total output in origin j–industry k is given by

$$Q_{j,k}(Q_{j1,k}, ...Q_{jN,k}) = \bar{d}_{j1,k}Q_{j1,k} + ... + \bar{d}_{jN,k}Q_{jN,k},$$

The change in total output in response to changes in destination-specific demand can be expressed as

$$\frac{\partial \ln Q_{j,k}(Q_{j1,k}, ...Q_{jN,k})}{\partial Q_{ji,k}} = \frac{\bar{d}_{ji,k}Q_{ji,k}}{Q_{j,k}} \equiv r_{ji,k},$$

where $r_{ji,k}$ is the (within-industry) revenue share that is collected from sales to destination i. Now, consider the producer price index associated with origin j–industry k, which is an explicit function of abatement, wage, and output schedule ($\boldsymbol{Q}_{i,k} \equiv \{Q_{j1,k}, ...Q_{jN,k}\}$):

$$P_{jj,k}(a_{j,k}, \boldsymbol{Q}_{j,k}, w_j) = \bar{p}_{jj,k} w_j (1-a_{j,k})^{\frac{1}{\gamma_k}-1} Q_{j,k}(Q_{j1,k}, ...Q_{jN,k})^{-\frac{1}{\gamma_k}}.$$

Also note that price in various destinations is a constant iceberg cost times the price at origin: $P_{ji,k} = \bar{d}_{ji,k}P_{jj,k}$. Suppose country j is the one setting policy. In that case the elasticity of $P_{ji,k}$ w.r.t. to different elements of the origin j's output vector is given by:

$$\left(\frac{\partial \ln P_{ji,k}(.)}{\partial \ln Q_{jn,g}}\right)_{\mathbb{P}_j, \mathbf{w}, Y_j} = \begin{cases} 0 & g \neq k \\ -\frac{1}{\gamma_k} r_{in,k} & g = k \end{cases}$$

To given intuition, $\mathbb{P}_i$ fixes all price associated with origin j. Hence, a change in output has a direct effect on the price index but no ripple effects. By ripple effects we mean that an increase in $Q_{jn,k}$ lowers $P_{ji,k}$, but this reduction has no further effect on consumer prices which are fixed by $\mathbb{P}_i$. Hence, the reduction in $P_{ji,k}$ has not feedback effect on output through demand effects.

This is no longer the case, if country j is not choosing the policy instruments in our optimal policy problem. Suppose instead that country i is the one choosing its optimal policy vector. In that case an increase in $Q_{jn,k}$ will lower origin j's "consumer" prices in all markets but i. This reduction will further raise demand for origin j's output triggering further scale effects and so forth. To keep track of these ripple effects, we can apply the Implicit Function Theorem to the following equation:

$$G_{j,k}(P_{jj,k}, Q_{jn,k}, ...) = \ln P_{jj,k} - \ln\left(\bar{p}_{jj,k} w_j (1-a_{j,k})^{\frac{1}{\gamma_k}-1}\left[\sum_{j=1}^{N} d_{ji,k} Q_{ji,k}(P_{jj,k})\right]^{-\frac{1}{\gamma_k}}\right) = 0,$$

which yields the following formula for the ***export supply elasticity*** facing variety jn, k ($j \neq i$):

$$\begin{aligned}\left(\frac{\partial \ln P_{ji,k}}{\partial \ln Q_{jn,k}}\right)_{\mathbb{P}_i,\mathbf{w},Y_i} &= \left(\frac{\partial \ln P_{jj,k}}{\partial \ln Q_{jn,k}}\right)_{\mathbb{P}_i,\mathbf{w},Y_i} = -\frac{\partial G_{j,k}(.)/\partial \ln Q_{jn,k}}{\partial G_{j,k}(.)/\partial \ln P_{jj,k}} \\ &= \frac{-\frac{1}{\gamma_k}\frac{\partial \ln Q_{j,k}}{\partial \ln Q_{jn,k}}}{1+\sum_{n\neq i}\frac{1}{\gamma_k}\frac{\partial \ln Q_{j,k}}{\partial \ln Q_{jn,k}}\frac{\partial \ln Q_{jn,,k}}{\partial \ln \tilde{P}_{jn,,k}}\frac{\partial \ln \tilde{P}_{jn,,k}}{\partial \ln P_{jj,,k}}} = \frac{-\frac{1}{\gamma_k} r_{jn,k}}{1+\sum_{n\neq i}\frac{1}{\gamma_k} r_{jn,k}\varepsilon_{jn,k}} \equiv \omega_{jn,k}.\end{aligned}$$

To economize on the notation, we use $\omega_{jn,k}$ to denote the export supply elasticity, noting that this is a variable but estimable reduced-form elasticity.

A.6 Multiplicity of Policy Schedules

[To be added]

B Proofs and Derivations

B.1 Proof of Lemma 2

Step 1. We first show that $\frac{\partial \ln Y_i}{\partial \ln w_i} = 0$ if the policy vector $\mathbb{P}_i$ is fixed and policy-wage is feasible $(\mathbb{P}_i; w_i) \in \mathbb{F}_i^w$. Using the income equation, and holding fixed $\{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}, \tilde{P}_{ii,k}\}_{j\neq i,k} \in \mathbb{P}_i$,

$$\frac{\partial Y_i}{\partial \ln w_i} = w_i \bar{L}_i - \sum_{k,\, j\neq i}\left[\frac{\partial \ln P_{ji,k}}{\partial \ln w_i} + \frac{\partial \ln Q_{ji,k}}{\partial \ln w_i}\right] P_{ji,k} Q_{ji,k} - \sum_{k,\, j}(1-\alpha_k\frac{\gamma_k - 1}{\gamma_k})\left[\frac{\partial \ln P_{ij,k}}{\partial \ln w_i} + \underbrace{\frac{\partial \ln Q_{ij,k}}{\partial \ln w_i}}_{=0 \text{ if } j\neq i}\right] P_{ij,k} Q_{ij,k}$$

Notice that home's wage, w_i, affects price of a variety directly if that variety is produced at home, and also indirectly through scale economies,

$$\begin{aligned}\frac{\partial \ln Y_i}{\partial \ln w_i} = & \frac{w_i \bar{L}_i}{Y_i} - \sum_{k,\, j\neq i} \left[\frac{\partial \ln P_{ji,k}}{\partial \ln Q_{j,k}} \frac{\partial \ln Q_{j,k}}{\partial \ln Q_{ji,k}} \frac{\partial \ln Q_{ji,k}}{\partial \ln w_i} + \frac{\partial \ln Q_{ji,k}}{\partial \ln w_i}\right] \frac{P_{ji,k}Q_{ji,k}}{Y_i} \\ & - \sum_{k,\, j} (1-\alpha_k \tfrac{\gamma_k - 1}{\gamma_k})(1 + \frac{\partial \ln P_{ij,k}}{\partial \ln Q_{i,k}} \frac{\partial \ln Q_{i,k}}{\partial \ln Q_{ii,k}} \frac{\partial \ln Q_{ii,k}}{\partial \ln w_i}) \frac{P_{ij,k}Q_{ij,k}}{Y_i} - \sum_k (1-\alpha_k \tfrac{\gamma_k - 1}{\gamma_k}) \frac{\partial \ln Q_{ii,k}}{\partial \ln w_i} \frac{P_{ii,k}Q_{ii,k}}{Y_i} \\ = & \frac{w_i \bar{L}_i}{Y_i} - \sum_{k,\, j\neq i} (1 - \tfrac{1}{\gamma_k} r_{ji,k}) \eta_{ji,k} \frac{\partial \ln Y_i}{\partial \ln w_i} \frac{P_{ji,k}Q_{ji,k}}{Y_i} \\ & - \sum_{k,\, j} (1-\alpha_k \tfrac{\gamma_k - 1}{\gamma_k})(1 - \tfrac{1}{\gamma_k} r_{ii,k} \eta_{ii,k} \frac{\partial \ln Y_i}{\partial \ln w_i}) \frac{P_{ij,k}Q_{ij,k}}{Y_i} - \sum_k (1-\alpha_k \tfrac{\gamma_k - 1}{\gamma_k}) \eta_{ii,k} \frac{\partial \ln Y_i}{\partial \ln w_i} \frac{P_{ii,k}Q_{ii,k}}{Y_i}\end{aligned}$$

where $\frac{\partial \ln Q_{j,k}}{\partial \ln Q_{ji,k}} = r_{ji,k}$. Reorganizing terms,

$$\Lambda_i^Y \left(\frac{\partial \ln Y_i}{\partial \ln w_i}\right) - \underbrace{\frac{1}{Y_i}\left(w_i \bar{L}_i - \sum_{k,\, j} (1-\alpha_k \frac{\gamma_k - 1}{\gamma_k}) P_{ij,k} Q_{ij,k}\right)}_{=0} = 0$$

where the second term equals zero since the policy-wage pair is feasible, $(\mathbb{P}_i; w_i) \in \mathbb{F}_i^w$, meaning that the labor market clearing condition (XXX) has to hold; and, Λ_i^Y summarizes the coefficient of the wage effect on income,

$$\Lambda_i^Y \equiv 1 + \sum_{k,\, j\neq i} (1 - \frac{r_{ji,k}}{\gamma_k}) \eta_{ji,k} \frac{P_{ji,k}Q_{ji,k}}{Y_i} - \sum_k \frac{r_{ii,k}}{\gamma_k} \eta_{ii,k} \frac{w_i \bar{L}_i}{Y_i} + \sum_k (1-\alpha_k \frac{\gamma_k - 1}{\gamma_k}) \eta_{ii,k} \frac{\partial \ln Y_i}{\partial \ln w_i} \frac{P_{ii,k}Q_{ii,k}}{Y_i}$$

From the fact that Λ_i^Y is generically non-zero, it follows that:

$$\frac{\partial \ln Y_i}{\partial \ln w_i} = 0.$$

Step 2. Within the feasible set of policy-wage, $(\mathbb{P}_i; w_i) \in \mathbb{F}_i^w$, and holding fixed the policy vector $\mathbb{P}_i$, we can express the derivative of $W_i(\mathbb{P}_i; \mathbf{w})$ w.r.t. w_i as follows:

$$\frac{\partial W_i(.)}{\partial w_i} = \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial Y_i}{\partial w_i}\right) - \frac{1}{Y_i}\sum_j \sum_k \left(\delta_{ji} Z_{j,k} \frac{\partial \ln Z_{j,k}(.)}{\partial \ln Q_{j,k}} \frac{\partial \ln Q_{j,k}(.)}{\partial \ln Q_{ji,k}} \frac{\partial \ln D_{ji,k}(.)}{\partial \ln Y_i}\right)\left(\frac{\partial Y_i}{\partial w_i}\right) = 0$$

where we use $\frac{\partial \ln Y_i}{\partial \ln w_i} = 0$ from Step 1.

B.2 Proof of Lemma 3

Notice that we have already sketched a proof for Lemma 3 in the buildup to the formal statement of the lemma. However, here we prove this lemma using a somewhat different approach that allows us to provide more details.

Recall that Applying the chain rule to $W_i(\mathbb{P}_i; Y_i) = V_i(\bar{w}_i \bar{L}_i + T_i(\mathbb{P}_i; Y_i), \tilde{\mathbf{P}}_i) - \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)$, yields the following expression:

$$\frac{\mathrm{d} W_i(\mathbb{P}_i; Y_i)}{\mathrm{d} \ln \mathcal{P}} = \frac{\partial V_i(.)}{\partial \ln \mathcal{P}} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i)}{\partial \ln \mathcal{P}}\right)_{Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln \mathcal{P}}\right)_{Y_i} + \left(\frac{\partial W_i(\mathbb{P}_i; Y_i)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} \frac{\mathrm{d} \ln Y_i}{\mathrm{d} \ln \mathcal{P}}.$$

Before moving forward, let us emphasize two important details:

1. Following Lemma 2, we are treating the vector of wages, $\mathbf{w} = \bar{\mathbf{w}}$, as constant throughout our proof. So, the partial derivatives subject to Y_i can be more-broadly interpreted as partial derivatives subject to holding both Y_i and $\mathbf{w}$ fixed, i.e., $\left(\frac{\partial T_i(\mathbb{P}_i;Y_i)}{\partial \ln \mathcal{P}}\right)_{Y_i} \sim \left(\frac{\partial T_i(\mathbb{P}_i;Y_i)}{\partial \ln \mathcal{P}}\right)_{Y_i,\mathbf{w}}$.
2. Every time we differentiated w.r.t. a $\mathcal{P} \in \mathbb{P}_i$, we are also fixing the remaining elements of $\mathbb{P}_i$. That is because the government is directly choosing every single element of $\mathbb{P}_i$. So, to be even more precise, we may interpret the partial derivative subject to Y_i as derivative subject to fixing Y_i, $\mathbf{w}$, and $\mathbb{P}_i - \{\mathcal{P}\}$, i.e., $\left(\frac{\partial T_i(\mathbb{P}_i;Y_i)}{\partial \ln \mathcal{P}}\right)_{Y_i} \sim \left(\frac{\partial T_i(\mathbb{P}_i;Y_i)}{\partial \ln \mathcal{P}}\right)_{Y_i,\mathbf{w},\mathbb{P}_i-\{\mathcal{P}\}}$.

Noting the above explanation, we now proceed with the proof in two steps.

Step #1: Characterizing $\left(\frac{\partial W_i}{\partial Y_i}\right)_{\mathbb{P}_i}$.

To characterize $\left(\frac{\partial W_i}{\partial Y_i}\right)_{\mathbb{P}_i}$, we can apply the chain rule, which implies

$$\left(\frac{\partial W_i(\mathbb{P}_i;Y_i)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i;Y_i)}{\partial \ln Y_i}\right)_{\mathbb{P}_i}. \tag{27}$$

As outlined in Appendix A.3 , $T_i(.)$ and $Z_{j,k}(.)$ are formulated as

$$\begin{cases} T_i(\mathbb{P}_i;Y_i) = \sum_k \left(\alpha_k \frac{\gamma_k - 1}{\gamma_k} P_{ii,k}(\mathbb{P}_i;Y_i) Q_{i,k}(\mathbb{P}_i;Y_i)\right) \\ \qquad + \sum_{k,j}\left[(\tilde{P}_{ij,k} - P_{ij,k}(\mathbb{P}_i;Y_i))Q_{ij,k}(\mathbb{P}_i)\right] + \sum_{k,j\neq i}\left[(\tilde{P}_{ji,k} - P_{ji,k}(\mathbb{P}_i;Y_i))Q_{ji,k}(\mathbb{P}_i;Y_i)\right]; \\ Z_{j,k}(\mathbb{P}_i;Y_i) = \bar{z}_{j,k}(1-a_{j,k})^{\frac{1}{\alpha_k}+\frac{1}{\gamma_k}-1} Q_{j,k}(\mathbb{P}_i;Y_i)^{1-\frac{1}{\gamma_k}}; \end{cases}$$

with the equilibrium quantity and producer prices given by

$$Q_{jn,k}(\mathbb{P}_i;Y_i) = \begin{cases} \mathcal{D}_{jn,k}(\tilde{\boldsymbol{P}}_{in}, \boldsymbol{P}_{-in}, \bar{Y}_n) & \text{if } n \neq i \\ \mathcal{D}_{ji,k}(\tilde{\boldsymbol{P}}_i, Y_i) & \text{if } n = i \end{cases};$$

$$Q_{i,k}(\mathbb{P}_i;Y_i) = \sum_j d_{ij,k} Q_{ij,k}(\mathbb{P}_i;Y_i);$$

$$P_{jn,k}(\mathbb{P}_i;Y_i) = \bar{\rho}_{jn,k}(1-a_{j,k})^{\frac{1}{\gamma_k}-1} Q_{j,k}(\mathbb{P}_i;Y_i)^{-\frac{1}{\gamma_k}}.$$

where $\overline{\rho}_{jn,k} \equiv \bar{d}_{jn,k}\bar{p}_{jj,k}\bar{w}_j$. Using our definition for the income elasticity of demand, we can produce the following partial derivatives for quantities and producer prices:

$$\begin{aligned}
&\left(\frac{\partial \ln Q_{ji,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \frac{\partial \ln \mathcal{D}_{ji,k}(.)}{\partial \ln Y_i} = \eta_{ji,k}; \qquad \left(\frac{\partial \ln Q_{jn,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = 0, \quad (j \neq i) \\
&\left(\frac{\partial \ln Q_{j,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \left(\frac{\partial \ln Q_{j,k}}{\partial \ln Q_{ji,k}}\right)_{\mathbb{P}_i} \left(\frac{\partial \ln Q_{ji,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = r_{ji,k}\eta_{ji,k} \\
&\left(\frac{\partial \ln P_{ij,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \left(\frac{\partial \ln P_{ii,k}(.)}{\partial \ln Q_{i,k}}\right)_{\mathbb{P}_i} \left(\frac{\partial \ln Q_{i,k}}{\partial \ln Q_{ii,k}}\right)_{\mathbb{P}_i} \left(\frac{\partial \ln Q_{ji,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = -\frac{1}{\gamma_k} r_{ii,k}\eta_{ii,k} \\
&\left(\frac{\partial \ln P_{ji,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \left(\frac{\partial \ln P_{jj,k}(.)}{\partial \ln Q_{ji,k}}\right)_{\tilde{P}_{ji,k}} \left(\frac{\partial \ln Q_{ji,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \omega_{ji,k}\eta_{ji,k} \qquad (j \neq i)
\end{aligned}$$

where $\omega_{ji,k}$ denotes the *export supply elasticity* as defined in Appendix A.5. Using the above expressions and noting that

$$\begin{aligned}
T_i(\mathbb{P}_i; Y_i) = &\sum_{k, j \neq i} \left[(\tilde{P}_{ji,k} - P_{ji,k}(\mathbb{P}_i; Y_i)) Q_{ji,k}(\mathbb{P}_i; Y_i) \right] \\
&+ \sum_{k,j} \left[(\tilde{P}_{ij,k} - (1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}) P_{ij,k}(\mathbb{P}_i; Y_i)) Q_{ij,k}(\mathbb{P}_i; Y_i) \right],
\end{aligned}$$

produces the following formulation for $\left(\frac{\partial T_i(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i}$ and $\left(\frac{\partial Z_{j,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i}$:

$$\begin{cases}
\left(\frac{\partial T_i(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = -\sum_k \left[\left(1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}\right) \frac{1}{\gamma_k} \eta_{ii,k} P_{ii,k} Q_{ii,k} \right] \\
\qquad + \sum_k \left[\left(\tilde{P}_{ii,k} - (1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}) P_{ii,k}\right) Q_{ii,k}\eta_{ii,k} \right] + \sum_{k, j \neq i} \left[(\tilde{P}_{ji,k} - (1 + \omega_{ji,k}) P_{ji,k}) Q_{ji,k}\eta_{ji,k} \right]; \\
\left(\frac{\partial Z_{j,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \left(1 - \frac{1}{\gamma_k}\right) Z_{j,k} r_{ji,k}\eta_{ji,k} = \frac{\gamma_k - 1}{\gamma_k} v_{j,k} P_{ji,k} Q_{ji,k}\eta_{ji,k}.
\end{cases}$$

To provide more detail: The first line in the expression for $\left(\frac{\partial T_i(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i}$ derives from the following intermediate result:

$$\begin{aligned}
\sum_{j=1}^{N} \left[\left(\frac{\partial \ln P_{ij,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbb{P}_i} P_{ij,g} Q_{ij,g} \right] &= \sum_{j=1}^{N} \left[\left(\frac{\partial \ln P_{ii,g}}{\partial \ln Q_{i,g}}\right)_{\mathbb{P}_i} \frac{\partial \ln Q_{i,g}(Q_{i1,g} ... Q_{iN,g})}{\partial \ln Q_{ii,g}} P_{ij,g} Q_{ij,g} \right] \\
&= -\sum_{j=1}^{N} \left(\frac{1}{\gamma_g} r_{ii,g} P_{ij,g} Q_{ij,g} \right) = -\frac{1}{\gamma_g} r_{ii,g} \sum_{j=1}^{N} \left(P_{ij,g} Q_{ij,g} \right) = -\frac{1}{\gamma_g} P_{ii,g} Q_{ii,g}
\end{aligned} \tag{28}$$

Plugging the expressions for $\left(\frac{\partial T_i(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i}$ and $\left(\frac{\partial Z_{j,k}(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i}$ back into Equation 27, and noting that $\sum_j \left(P_{ij,k}Q_{ij,k}r_{ii,k}\right) = P_{ii,k}Q_{i,k}r_{ii,k} = P_{ii,k}Q_{ii,k}$, yields

$$\begin{aligned}\left(\frac{\partial W_i}{\partial \ln Y_i}\right)_{\mathbb{P}_i} &= \sum_k \left[\left(\tilde{P}_{ii,k} - \frac{\gamma_k - 1}{\gamma_k}\left(1 - \alpha_k\frac{\gamma_k - 1}{\gamma_k} + \tilde{\delta}_{ii,k}v_{i,k}\right)P_{ii,k}\right)Q_{ii,k}\eta_{ii,k} + \sum_{j\neq i}\left(\left[\tilde{P}_{ji,k} - (1+\omega_{ji,k} - \frac{\gamma_k - 1}{\gamma_k}\delta_{ji}v_{j,k})P_{ji,k}\right]Q_{ji,k}\eta_{ji,k}\right)\right] \\ &= \sum_k \left[\left(1 - \frac{\gamma_k - 1}{\gamma_k}\left(1 - \alpha_k\frac{\gamma_k - 1}{\gamma_k} + \tilde{\delta}_{ii,k}v_{i,k}\right)\frac{P_{ii,k}}{\tilde{P}_{ii,k}}\right)e_{ii,k}\eta_{ii,k} + \sum_{j\neq i}\left(\left[1 - (1+\omega_{ji,k} - \frac{\gamma_k - 1}{\gamma_k}\delta_{ji}v_{j,k})\frac{P_{ji,k}}{\tilde{P}_{ji,k}}\right]e_{ji,k}\eta_{ji,k}\right)\right]Y_i.\end{aligned} \tag{29}$$

Step #2: Proving that $\left(\frac{\partial W_i}{\partial Y_i}\right)_{\mathbb{P}_i} = 0$ at the optimum.

This step establishes that if for all $\mathcal{P} \in \left\{\mathbf{a}_i, \tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}\right\}$ if

$$\frac{\partial V_i(.)}{\partial \ln \mathcal{P}} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \mathcal{P}}\right)_{\mathbf{w},Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \mathcal{P}}\right)_{\mathbf{w},Y_i} = 0,$$

then $\left(\frac{\partial W_i}{\partial Y_i}\right)_{\mathbb{P}_i} = 0$. For the sake of clarity, our notation indicates explicitly that the partial derivative w.r.t. $\mathcal{P}$ are taken while holding both $\mathbf{w}$ and Y_i (in the demand function) constant.

[Abatement Level: $\mathbf{a}_i$] First, consider the case where $\mathcal{P} = 1 - a_{i,k}$. Keep in mind that the instrument set $\mathbb{P}_i$ includes all consumers prices in the local economy. So, holding all instruments except $a_{i,k}$ (i.e., $\mathbb{P}_i - \{a_{i,k}\}$) fixed, then $a_{i,k}$ has no direct effect on $V_i(Y_i = \bar{w}_i\bar{L}_i + T_i, \tilde{\mathbf{P}}_i)$. However, $a_{i,k}$ does affect tax revenues and local emission levels as indicated below:

$$\begin{cases}\frac{\partial V_i(.)}{\partial \ln(1+a_{i,k})} = 0 \\ \left(\frac{\partial T_i(\mathbb{P}_i;Y_i,\mathbf{w})}{\partial \ln(1+a_{i,k})}\right)_{\mathbf{w},Y_i} = -\sum_{j=1}^{N}\left((1-\alpha_k\frac{\gamma_k-1}{\gamma_k})P_{ij,k}Q_{ij,k}\left(\frac{\partial \ln P_{ii,k}}{\partial \ln(1-a_{i,k})}\right)_{\mathbf{w},Y_i}\right) = (1-\alpha_k\frac{\gamma_k-1}{\gamma_k})\frac{\gamma_k-1}{\gamma_k}\sum_{j=1}^{N}\left(P_{ij,k}Q_{ij,k}\right) \\ \left(\frac{\partial \delta_i\cdot\mathbf{Z}(\mathbb{P}_i;Y_i,\mathbf{w})}{\partial \ln(1+a_{i,k})}\right)_{\mathbf{w},Y_i} = \delta_{ii}\left(\frac{\partial Z_{i,k}(...,1-a_{i,k})}{\partial \ln(1-a_{i,k})}\right)_{\mathbf{w},Y_i} = \left(\frac{1}{\alpha_k} - \frac{\gamma_k-1}{\gamma_k}\right)\delta_{ii}Z_{i,k}\end{cases}.$$

Combining the above equation yields (note that $P_{ii,k}Q_{i,k} = \sum_{j=1}^{N} P_{ij,k}Q_{ij,k}$)

$$\begin{aligned}&\frac{\partial V_i(.)}{\partial \ln(1+a_{i,k})} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(\mathbb{P}_i;Y_i)}{\partial \ln(1+a_{i,k})}\right)_{\mathbf{w},Y_i} - \left(\frac{\partial \delta_i\cdot\mathbf{Z}(\mathbb{P}_i;Y_i)}{\partial \ln(1+a_{i,k})}\right)_{\mathbf{w},Y_i} \\ &= \frac{\partial V_i(.)}{\partial Y_i}\left[1-\alpha_k\frac{\gamma_k-1}{\gamma_k}\right]\frac{\gamma_k-1}{\gamma_k}P_{ii,k}Q_{i,k} - \frac{1}{\alpha_k}\delta_{ii}Z_{i,k}\left[1-\alpha_k\frac{\gamma_k-1}{\gamma_k}\right] = 0.\end{aligned} \tag{30}$$

[Domestic and Import Prices: $\tilde{\mathbf{P}}_{ii}$, and $\tilde{\mathbf{P}}_{ji}$,] Next, consider the case where $\mathcal{P} = \tilde{P}_{ii,k}$ or $\tilde{P}_{ji,k}$ (where $j \neq i$). We are combining both instruments, as the partial derivative w.r.t. to both $\tilde{P}_{ii,k}$ and $\tilde{P}_{ji,k}$ produce similar-looking equation. So, we henceforth use n to denote the origin country with the understanding that either $n = i$ or $n = j$. For this case, we first detail the partial derivative of tax revenues, $T_i(.)$, which is more involved:

$$\begin{aligned}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \tilde{P}_{ni,k}}\right)_{\mathbf{w},Y_i} &= \tilde{P}_{ii,k}Q_{ii,k} + \sum_g \left[\left(\tilde{P}_{ii,g} - [1 - \alpha_g \frac{\gamma_g - 1}{\gamma_g}]P_{ii,g}\right) Q_{ii,g} \left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ni,k}}\right)_{\mathbf{w},Y_i}\right] \\ &- \sum_g \sum_j \left[[1 - \alpha_g \frac{\gamma_g - 1}{\gamma_g}]P_{ij,g}Q_{ij,g} \left(\frac{\partial \ln P_{ij,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbf{w},Y_i} \left(\frac{\partial \ln Q_{ii,g}}{\partial \ln \tilde{P}_{ni,k}}\right)_{\mathbf{w},Y_i}\right] \\ &+ \sum_{j \neq i} \sum_g \left[\left(\tilde{P}_{ji,g} - [1 + \left(\frac{\partial \ln P_{ji,g}}{\partial \ln Q_{ji,g}}\right)_{\mathbf{w},Y_i}]P_{ji,g}\right) Q_{ji,g} \left(\frac{\partial \ln Q_{ji,g}}{\partial \ln \tilde{P}_{ni,k}}\right)_{\mathbf{w},Y_i}\right].\end{aligned}$$

As before, $\left(\frac{\partial \ln Q_{ni,g}}{\partial \ln \tilde{P}_{ni,k}}\right)_{\mathbf{w},Y_i} = \frac{\partial \ln \mathcal{D}_{ni,g}(Y_i, \tilde{\mathbf{P}}_i)}{\partial \ln \tilde{P}_{ni,k}} = \varepsilon_{ni,g}^{(ii,k)}$. The second line can also be simplified the steps outlined under Equation 28. Accordingly, we can express the different elements in Equation

$$\begin{cases} \frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ii,k}} = -P_{ii,k}Q_{ii,k}\frac{\partial V_i(.)}{\partial Y_i} & \text{[Roy's identity]} \\ \left(\frac{\partial T_i(\mathbb{P}_i;Y_i,\mathbf{w})}{\partial \ln \tilde{P}_{ii,k}}\right)_{\mathbf{w},Y_i} = \sum_{n\neq i}\sum_g \left[\left(1 - (1 + \omega_{ni,g})\frac{P_{ni,g}}{\tilde{P}_{ni,g}}\right)\tilde{P}_{ni,g}Q_{ni,g}\varepsilon_{ni,g}^{(ii,k)}\right] + \sum_g \left[\left(1 - \left(1 - \alpha_g\frac{\gamma_g-1}{\gamma_g}\right)\frac{\gamma_g-1}{\gamma_g}\frac{P_{ii,g}}{\tilde{P}_{ii,g}}\right)\tilde{P}_{ii,g}Q_{ii,g}\varepsilon_{ii,g}^{(ii,k} \\ \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i;Y_i,\mathbf{w})}{\partial \ln \tilde{P}_{ii,k}}\right)_{\mathbf{w},Y_i} = \sum_g \sum_j \delta_{ji}\left(\frac{\partial Z_{j,g}(...;Q_{j,g})}{\partial \ln Q_{j,g}}\frac{\partial \ln Q_{j,g}(Q_{j1,k},...,Q_{jN,k})}{\partial \ln Q_{ji,g}}\frac{\partial \ln Q_{ji,g}}{\partial \ln \tilde{P}_{ii,k}}\right)_{\mathbf{w},Y_i} = \sum_g \sum_j \left[\delta_{ji}\frac{\gamma_k - 1}{\gamma_k}v_{j,k}\varepsilon_{ji,g}^{(ni,k)}P_{ji,g}Q_{ji,g}\right] \end{cases}$$

where the last line follows from the fact that (1) $\frac{\partial Z_{j,g}(...;Q_{j,g})}{\partial \ln Q_{j,g}} = \frac{\gamma_k - 1}{\gamma_k}Z_{j,g}$, (2) $\frac{\partial \ln Q_{j,g}(Q_{j1,k},...,Q_{jN,k})}{\partial \ln Q_{ji,g}} = r_{ji,g}$, and (3) $v_{j,k} \equiv Z_{j,k}/P_{jj,k}Q_{j,k}$. Combining the above equations yields

$$\begin{aligned}&\frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ni,k}} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i)}{\partial \ln \tilde{P}_{ni,k}}\right)_{\mathbf{w},Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln \tilde{P}_{ni,k}}\right)_{\mathbf{w},Y_i} = \\ &\sum_g \left[\sum_{j\neq i}\left(1 - \left(1 + \omega_{ji,g} + \tilde{\delta}_{ji}\nu_{j,g}\frac{\gamma_g - 1}{\gamma_g}\right)\frac{P_{ji,g}}{\tilde{P}_{ji,g}}\right)e_{ji,g}\varepsilon_{ji,g}^{(ni,k)}\right]Y_i + \sum_g \left[\left(1 - \left(1 - \alpha_g\frac{\gamma_g - 1}{\gamma_g} + \tilde{\delta}_{ii}\nu_{i,g}\right)\frac{\gamma_g - 1}{\gamma_g}\frac{P_{ii,g}}{\tilde{P}_{ii,g}}\right)e_{ii,g}\varepsilon_{ii,g}^{(ni,k)}\right]Y_i =\end{aligned} \tag{31}$$

For Equation 30 to hold it should be that $\alpha_k\frac{\gamma_k - 1}{\gamma_k}P_{ii,k}Q_{i,k} - \tilde{\delta}_{ii}Z_{i,k} = 0$. Plugging this expression into Equation 31 yields

$$\sum_g \left[\sum_{j\neq i}\left(1 - \left(1 + \omega_{ji,g} + \tilde{\delta}_{ji}\nu_{j,g}\frac{\gamma_g - 1}{\gamma_g}\right)\frac{P_{ji,g}}{\tilde{P}_{ji,g}}\right)e_{ji,g}\varepsilon_{ji,g}^{(ni,k)}\right] + \sum_g \left[\left(1 - \frac{\gamma_g - 1}{\gamma_g}\frac{P_{ii,g}}{\tilde{P}_{ii,g}}\right)e_{ii,g}\varepsilon_{ii,g}^{(ni,k)}\right] = 0.$$

The above equation specifies the optimality condition for $N \times K$ different price instruments, $\tilde{P}_{ni,k}$. Simultaneously solving the above equation for all $\tilde{P}_{ni,k}$ amounts to solving the following matrix equation.

$$\begin{bmatrix} e_{1i,1}\varepsilon_{1i,1}^{(1i,1)} & \cdots & e_{Ni,}\varepsilon_{Ni,1}^{(1i,1)} & \cdots & e_{1i,K}\varepsilon_{1i,K}^{(1i,1)} & \cdots & e_{Ni,K}\varepsilon_{Ni,K}^{(1i,1)} \\ \vdots & & \ddots & \ddots & & & \vdots \\ e_{1i,1}\varepsilon_{1i,1}^{(Ni,K)} & \cdots & e_{Ni,}\varepsilon_{Ni,1}^{(Ni,K)} & \cdots & e_{1i,K}\varepsilon_{1i,K}^{(Ni,K)} & \cdots & e_{Ni,K}\varepsilon_{Ni,K}^{(Ni,K)} \end{bmatrix} \begin{bmatrix} \frac{\tilde{P}^{\star}_{1i,k}}{P_{1i,1}} - \left(1 + \omega_{1i,k} + \tilde{\delta}_{1i}\nu_{1,k}\frac{\gamma_k - 1}{\gamma_k}\right) \\ \vdots \\ \frac{\tilde{P}^{\star}_{ii,k}}{P_{ii,k}} - \frac{\gamma_k - 1}{\gamma_k} \\ \vdots \\ \frac{\tilde{P}^{\star}_{Ni,k}}{P_{Ni,k}} - \left(1 + \omega_{Ni,k} + \tilde{\delta}_{Ni}\nu_{N,k}\frac{\gamma_k - 1}{\gamma_k}\right) \end{bmatrix}_k = \mathbf{0}.$$

As discussed in Section 3.1 and proven in the following appendix, is non-singular. So, the unique solution to the above equation is

$$\frac{\tilde{P}^{\star}_{ji,k}}{P_{ji,1}} = 1 + \omega_{ji,k} + \tilde{\delta}_{ji}v_{j,k}\frac{\gamma_k - 1}{\gamma_k}; \qquad \frac{\tilde{P}^{\star}_{ii,k}}{P_{ii,k}} - \frac{\gamma_k - 1}{\gamma_k}, \tag{32}$$

which when plugged into Equation 29, trivially implies $\left(\frac{\partial W_i}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = 0$.

B.3 Proof of Lemma 5

Following Proposition 2.E.2 in Mas-Colell *et al.* (1995) the Walrasian demand function satisfies $e_{ji,k} = \mid e_{ji,k}\varepsilon_{ji,k}^{(ji,k)} \mid - \sum_{n,g \neq j,k} \mid e_{ni,g}\varepsilon_{ni,g}^{(ji,k)} \mid$. Hence, since there exists a ji, k such that $e_{ji,k} > 0$, the matrix Ξ is strictly diagonally dominant. The Lèvy-Desplanques Theorem (Horn et Johnson (2012)), therefore, ensures that Ξ is non-singular. The lower bound on $\det(\Xi)$ follows trivially from Gerschgorin's circle theorem. Specifically, following Ostrowski (1952),

$$\mid \det(\Xi) \mid \geq \prod_j \prod_k \left(\mid e_{ji,k}\varepsilon_{ji,k}^{(ji,k)} \mid - \sum_{n,g \neq j,k} \mid e_{ni,g}\varepsilon_{ni,g}^{(ji,k)} \mid \right) = \prod_j \prod_k e_{ji,k} > 0.$$

B.4 Proof of Theorem 1

As discussed in Section 3.2, the expression for emission taxes follows from combining cost minimization with the optimal tax condition (refer to Equation 19). Domestic and import taxes were also implicitly derived in Appendix B.2 under Equation 32. Combining these expressions, we have:

$$\tau^{\star}_{i,k} = \tilde{\delta}_{ii}, \qquad 1 + s^{\star}_{i,k} = \frac{\tilde{P}^{\star}_{ii,k}}{P_{ii,k}} = \frac{\gamma_k - 1}{\gamma_k}; \qquad 1 + t^{\star}_{ji,k} = \frac{\tilde{P}^{\star}_{ji,k}}{P_{ji,k}} = 1 + \omega_{ji,k} + \tilde{\delta}_{ji}v_{j,k}\frac{\gamma_k - 1}{\gamma_k}.$$

To determine the export tax we can appeal to Proposition 1, whereby the necessary condition for optimality w.r.t. $\tilde{P}_{ij,k}$ ($j \neq i$) is

$$\frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ij,k}} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i} = 0. \tag{33}$$

First not that $\tilde{P}_{ij,k}$ does not directly enter the indirect utility function, so $\frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ij,k}} = 0$. Recalling the expression for $T_i(\mathbb{P}_i; Y_i, \mathbf{w})$ we can express the term corresponding to tax revenue effects as

$$\begin{aligned}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i} &= \tilde{P}_{ij,k}Q_{ij,k} + \sum_g \left[\left(\tilde{P}_{ij,g} - [1-\alpha_g\frac{\gamma_g - 1}{\gamma_g}]P_{ij,g}\right) Q_{ij,g}\left(\frac{\partial \ln Q_{ij,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i}\right] \\ &- \sum_g \sum_j \left[[1-\alpha_g\frac{\gamma_g-1}{\gamma_g}]P_{ij,g}Q_{ij,g}\left(\frac{\partial \ln P_{ij,g}}{\partial \ln Q_{ij,g}}\right)_{\mathbf{w},Y_i}\left(\frac{\partial \ln Q_{ij,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i}\right] \\ &- \sum_{n\neq i}\sum_g \left[P_{ni,g}Q_{ji,g}\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{nj,g}}\right)_{\mathbf{w},Y_i}\left(\frac{\partial \ln Q_{nj,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i}\right] = 0.\end{aligned}$$

To simplify the above equation, we can appeal to Equation 28 (Appendix B.2) and the following relationship:

$$\begin{aligned}\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{nj,g}}\right)_{\mathbf{w},Y_i} P_{ni,g}Q_{ni,g} &= \left(\frac{\partial \ln P_{nn,g}}{\partial \ln Q_{nj,g}}\right)_{\mathbf{w},Y_i} P_{ni,g}Q_{ni,g}, \\ &= \left(\frac{\partial \ln P_{nn,g}}{\partial \ln Q_{ni,g}}\right)_{\mathbf{w},Y_i} P_{nj,g}Q_{nj,g} \equiv \omega_{ni,g}P_{nj,g}Q_{nj,g}.\end{aligned}$$

Doing so yields the following equation:

$$\left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i} = \tilde{P}_{ij,k}Q_{ij,k} + \sum_g \left[\left(\tilde{P}_{ij,g} - [1-\alpha_g\frac{\gamma_g-1}{\gamma_g}]\frac{\gamma_k-1}{\gamma_k}P_{ij,g}\right)Q_{ij,g}\varepsilon_{ij,g}^{(ij,k)}\right] - \sum_g\sum_j\left[\omega_{ni,g}P_{nj,g}Q_{nj,g}\varepsilon_{nj,g}^{(ij,k)}\right].$$

Likewise the last term in Equation 33 (that accounts for emission effects) can be specified as

$$\left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i} = \sum_{n,g}\left[\delta_{ni}\left(\frac{\partial Z_{n,g}(..., Q_{n,g})}{\partial \ln Q_{nj,g}}\frac{\partial \ln Q_{n,g}(Q_{n1,g}, ..., Q_{nN,g})}{\partial \ln Q_{nj,g}}\frac{\partial \ln Q_{nj,g}}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i}\right] = \frac{\partial V_i(.)}{\partial Y_i}\sum_{n,g}\left[\tilde{\delta}_{ni}\nu_{n,g}\frac{\gamma_g-1}{\gamma_g}P_{nj,g}Q_{nj,g}\right].$$

Plugging the above expressions back into Equation 33 (and dividing everything by $\frac{\partial V_i(.)}{\partial Y_i}\tilde{P}_{ij,k}Q_{ij,k}$) yields the following optimality condition:

$$\begin{aligned}&\left\{\frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ij,k}} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i}\right\}\left[\frac{\partial V_i(.)}{\partial Y_i}P_{ij,k}Q_{ij,k}\right]^{-1} = \\ &1 + \sum_g\left[\left(1 - \left(1-\alpha_g\frac{\gamma_g-1}{\gamma_g} + \tilde{\delta}_{ii}\nu_{i,g}\right)\frac{\gamma_g-1}{\gamma_g}\frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right)\frac{e_{ij,g}}{e_{ij,k}}\varepsilon_{ij,g}^{(ij,k)}\right] - \sum_{n\neq i}\sum_g\left[\left(\omega_{ni,g} + \tilde{\delta}_{ni}z_{n,k}\frac{\gamma_k-1}{\gamma_k}\right)\frac{e_{nj,g}}{e_{ij,k}}\varepsilon_{nj,g}^{(ij,k)}\right] = 0\end{aligned}$$

To detect the optimal export taxes, we guess the following formulation:

$$1 + x_{ij,k} \equiv \frac{\tilde{P}_{ij,k}/P_{ij,g}}{\tilde{P}_{ii,g}/P_{ii,g}} = \frac{\gamma_k - 1}{\gamma_k}\frac{\tilde{P}_{ij,k}}{\tilde{P}_{ij,k}} = \frac{\varepsilon_{ij,k}}{1+\varepsilon_{ij,k}}\chi_{ij,k}^{-1}$$

Plugging the above guess back into the F.O.C. yields the following:

$$1+\sum_g\left[\left(1-\chi_{ij,k}\frac{1+\varepsilon_{ij,k}}{\varepsilon_{ij,k}}\right)\frac{e_{ij,g}\varepsilon_{ij,g}^{(ij,k)}}{e_{ij,k}}\right]-\sum_{n\neq i}\sum_g\left[t^*_{ni,g}\frac{e_{nj,g}\varepsilon_{nj,g}^{(ij,k)}}{e_{ij,k}}\right]=0;\qquad [\tilde{P}_{ij,k}]$$

Noting that $1+\sum_g\left[\frac{e_{ij,g}}{e_{ij,k}}\varepsilon_{ij,g}^{(ij,k)}\right]=-\sum_{n\neq i}\sum_g\left[\frac{e_{nj,g}}{e_{ij,k}}\varepsilon_{nj,g}^{(ij,k)}\right]$ and dividing the above equation by $1+\varepsilon_{ij,k}$,

$$-\sum_g\left[\chi_{ij,k}\frac{e_{ij,g}}{e_{ij,k}}\frac{\varepsilon_{ij,g}^{(ij,k)}}{\varepsilon_{ij,k}}\right]-\sum_{n\neq i}\sum_g\left[\frac{(1+t^*_{ni,g})e_{nj,g}\varepsilon_{nj,g}^{(ij,k)}}{e_{ij,k}\left(1+\varepsilon_{ij,k}\right)}\right]=0$$

Noting that $\left(1+\varepsilon_{ij,k}\right)e_{ij,k}=-\sum_{n\neq i}\sum_g e_{nj,g}\varepsilon_{nj,g}^{(ij,k)}$, we can write the above equation in matrix as

$$\underbrace{\begin{bmatrix}\frac{e_{ij,1}}{e_{ij,1}}\frac{\varepsilon_{ij,1}^{(ij,1)}}{\varepsilon_{ij,1}} & \dots & \frac{e_{ij,K}}{e_{ij,K}}\frac{\varepsilon_{ij,K}^{(ij,1)}}{\varepsilon_{ij,1}}\\ \dots & \dots & \dots\\ \frac{e_{ij,1}}{e_{ij,K}}\frac{\varepsilon_{ij,1}^{(ij,K)}}{\varepsilon_{ij,K}} & \dots & \frac{e_{ij,K}}{e_{ij,K}}\frac{\varepsilon_{ij,K}^{(ij,K)}}{\varepsilon_{ij,K}}\end{bmatrix}}_{\mathbf{E}_{ij}}\begin{bmatrix}\chi_{ij,1}\\ \dots\\ \chi_{ij,K}\end{bmatrix}=\begin{bmatrix}1+\frac{\sum_{n\neq i}\sum_g t^*_{ni,g}e_{nj,g}\varepsilon_{nj,g}^{(ij,1)}}{\sum_{n\neq i}\sum_g e_{nj,g}\varepsilon_{nj,g}^{(ij,1)}}\\ \dots\\ 1+\frac{\sum_{n\neq i}\sum_g t^*_{ni,g}e_{nj,g}\varepsilon_{nj,g}^{(ij,K)}}{\sum_{n\neq i}\sum_g e_{nj,g}\varepsilon_{nj,g}^{(ij,K)}}\end{bmatrix}.$$

Since $\mid e_{ij,k}\varepsilon_{ij,k}^{(ij,k)}\mid-\sum_{k\neq j}e_{ij,g}\varepsilon_{ij,g}^{(ij,k)}=e_{ij,k}+\sum_{n\neq i}\sum_g e_{ij,g}\varepsilon_{nj,g}^{(ij,k)}>0$, then $\mathbf{E}_{ji}\equiv\left[\frac{e_{ij,g}\varepsilon_{ij,g}^{(ij,k)}}{e_{ij,k}\varepsilon_{ij,g}}\right]_{k,g}$ is strict diagonally dominant. Hence, following the Lèvy-Desplanques Theorem, $\mathbf{E}_{ji}$ is invertible (Horn et Johnson (2012)) and we can compute the vector χ_{ij} as

$$\chi_{ij}=\left[\frac{e_{ij,g}\varepsilon_{ij,g}^{(ij,k)}}{e_{ij,k}\varepsilon_{ij,g}}\right]_{k,g}^{-1}\left(\mathbf{1}_K+\left[\frac{\sum_{n\neq i}t^*_{ni,g}e_{nj,g}\varepsilon_{nj,g}^{(ij,k)}}{\sum_{n\neq i}\sum_g e_{nj,g}\varepsilon_{nj,g}^{(ij,k)}}\right]_k\right).\tag{34}$$

Combining the above result with the previously-derived formulas for emission, domestic, and import taxes yields

$$\begin{cases}1+s^\star_{i,k}=\frac{\gamma_k-1}{\gamma_k};\qquad \tau^\star_{i,k}=\tilde{\delta}_{ii}\\ 1+t^\star_{ji,k}=(1+\omega_{ni,k})+\tilde{\delta}_{ni}\left(\frac{\gamma_k-1}{\gamma_k}\right)v_{n,k}\\ 1+x^\star_{ij,k}=\frac{\varepsilon_{ij,k}}{1+\varepsilon_{ij,k}}\chi_{ij,k}^{-1}\end{cases},\tag{35}$$

where $\chi_{ij,k}$ is given by Equation 34.

C Optimal Emission Policy when Other Taxes are Banned

The F.O.C. w.r.t. $1-a_{i,k}$ can be expressed as ($Z_i \equiv \sum_{n,k}\delta_{ni}Z_{n,k}$):

$$\frac{\partial V_i(.)}{\partial Y_i}\frac{\partial Y_i(\boldsymbol{w},\boldsymbol{a})}{\partial \ln(1-a_{i,k})}+\frac{\partial V_i(.)}{\partial \ln \tilde{\boldsymbol{P}}_i}\frac{\partial \ln \tilde{\boldsymbol{P}}_i(\boldsymbol{w},\boldsymbol{a})}{\partial \ln(1-a_{i,k})}+$$
$$\frac{\partial Z_i}{\partial Y_i}\frac{\partial Y_i(\boldsymbol{w},\boldsymbol{a})}{\partial \ln(1-a_{i,k})}+\frac{\partial Z_i}{\partial \tilde{\boldsymbol{P}}_i}\frac{\partial \tilde{\boldsymbol{P}}_i(\boldsymbol{w},\boldsymbol{a})}{\partial(1-a_{i,k})}+\frac{\partial \mathcal{V}_i(.)}{\partial \ln \boldsymbol{w}}\frac{\mathrm{d}\ln \boldsymbol{w}}{\mathrm{d}\ln(1-a_{i,k})}=0$$

To simplify the above problem, we impose the following additional assumptions:

1. Preferences are given by the Cobb-Douglas-CES specification;
2. Country i is a small open economy with $\delta_{-ii}=0$; and
3. All industries are perfectly competitive, i.e., $\gamma_k \to \infty$.

Noting that $\partial \ln P_{in,k}/\partial \ln(1-a_{i,k})=-1$ and noting that $Z_{i,k}=v_{i,k}P_{ii,k}Q_{i,k}$, it follows that:

$$\frac{\partial Z_i}{\partial \ln(1-a_{i,k})}=\frac{\partial \delta_{ii}Z_{i,k}}{\partial \ln(1-a_{i,k})}=-\delta_{ii}v_{i,k}\sum_j\left[P_{ij,k}Q_{ij,k}\varepsilon_{ij,k}\right]$$
$$+\left(\frac{1}{\alpha_k}-1\right)\delta_{ii}v_{i,k}P_{ii,k}Q_{i,k}+\delta_{ii}v_{i,k}P_{ii,k}Q_{ii,k}\frac{\partial Y_i}{\partial \ln(1-a_{i,k})}$$

Wage effects can be characterized by applying the $D_i(\boldsymbol{a}_i,w_i)=\sum_{j\neq i}\sum_g\left(P_{ji,g}Q_{ji,g}-P_{ij,g}Q_{ij,g}\right)$

$$\frac{\mathrm{d}\ln w_i}{\mathrm{d}\ln(1-a_{i,k})}=-\left(\sum_{j\neq i}\left[P_{ji,k}Q_{ji,k}\varepsilon^{ii}_{ji,k}-P_{ij,k}Q_{ij,k}\left(1+\varepsilon_{ij,k}\right)\right]+\sum_{j\neq i}\sum_g\left(P_{ji,g}Q_{ji,g}\right)\frac{\partial Y_i}{\partial \ln(1-a_{i,k})}\right)\left(\frac{\partial D_i}{\partial \ln w_i}\right)^{-1}$$

Using the above expression, invoking Roy's identity, and noting that $Y_i=w_iL_i+\sum_k\alpha_kP_{ii,k}Q_{i,k}$, yields the following formulation of the F.O.C.

$$P_{ii,k}Q_{ii,k}-\alpha_k\sum_j\left[P_{ij,k}Q_{ij,k}\left(1+\varepsilon_{ij,k}\right)\right]+\tilde{\delta}_{ii}v_{i,k}\sum_j\left[P_{ij,k}Q_{ij,k}\varepsilon_{ij,k}\right]$$
$$-\left(\frac{1}{\alpha_k}-1\right)\tilde{\delta}_{ii}v_{i,k}P_{ii,k}Q_{i,k}-\sum_g\left(\left[\alpha_g-\tilde{\delta}_{ii}\nu_{i,g}\right]P_{ii,g}Q_{ii,g}\right)\frac{\partial \ln Y_i}{\partial \ln(1-a_{i,k})}$$
$$-\bar{\Delta}_i\left[\sum_{j\neq i}\left[P_{ji,k}Q_{ji,k}\varepsilon^{ii}_{ji,k}-P_{ij,k}Q_{ij,k}\left(1+\varepsilon_{ij,k}\right)\right]+\sum_{j\neq i}\sum_g\left(P_{ji,g}Q_{ji,g}\right)\frac{\partial \ln Y_i}{\partial \ln(1-a_{i,k})}\right]=0. \qquad (36)$$

where $\bar{\Delta}_i \equiv \frac{\partial \mathcal{V}_i/\partial \ln w_i}{\partial D_i/\partial \ln w_i}$ is a uniform term without industry subscript. Dividing Equation 36 by $R_{i,k} = \sum_n P_{in,k}Q_{in,k}$ and defining $\mathcal{E}_{i,k} = \sum_j \left[r_{ij,k}\left(1+\varepsilon_{ij,k}\right)\right] = -\epsilon_k\left(1 - r_{ii,k}\lambda_{ii,k}\right)$, we can simplify the F.O.C.

$$\begin{aligned} & r_{ii,k} - \alpha_k \mathcal{E}_{i,k} + \alpha_k \frac{\tilde{\delta}_{ii}}{\tau_{i,k}}\left(\mathcal{E}_{i,k} - 1\right) - \left(1-\alpha_k\right)\frac{\tilde{\delta}_{ii}}{\tau_{i,k}} \\ & \quad + \sum_g \left(\alpha_g \left[1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{ii,g} r_{i,g}\right) \frac{\partial \ln Y_i}{\partial \ln(1-a_{i,k})} r_{i,k}^{-1} - \bar{\Delta}_i \left[\mathcal{E}_{i,k} + \left(1-\lambda_{ii}\right)\frac{\partial \ln Y_i}{\partial \ln(1-a_{i,k})} r_{i,k}^{-1}\right] = 0. \end{aligned} \tag{37}$$

$\partial Y_i/\partial \ln(1-a_{i,k})$, in the above expression, can be obtained by applying the Implicit Function Theorem to $Y_i = w_i L_i + \sum_k \alpha_k P_{ii,k} Q_{i,k}$, while noting that $\eta_{in,k} = 1$ given our parametric assumption with regards to preferences. Namely,

$$\frac{\partial Y_i}{\partial \ln(1-a_{i,k})} = \frac{-\alpha_k \sum_j \left[P_{ij,k}Q_{ij,k}\left(1+\varepsilon_{ij,k}\right)\right]}{Y_i - \sum_g \alpha_g \eta_{ii,g} P_{ii,g} Q_{ii,g}} = \frac{-\alpha_k \mathcal{E}_{i,k}}{1 - \bar{\alpha}_i \lambda_{ii}} r_{i,k}$$

Plugging the above equation back into the F.O.C. implies

$$\frac{\tilde{\delta}_{ii}}{\tau_{i,k}} - 1 = \frac{(\widetilde{\alpha}_{i,k} - \alpha_k)\mathcal{E}_{i,k} + 1 - r_{ii,k}}{\alpha_k \mathcal{E}_{i,k} - 1} \implies \tau_{i,k} = \left(\frac{\alpha_k \mathcal{E}_{i,k} - 1}{\widetilde{\alpha}_{i,k}\mathcal{E}_{i,k} - r_{ii,k}}\right)\tilde{\delta}_{ii}$$

where

$$\widetilde{\alpha}_{i,k} - \alpha_k \equiv \bar{\Delta}_i \left[\frac{1-\alpha_k}{1-\bar{\alpha}_i\lambda_{ii}}\right] - \frac{\alpha_k}{1-\bar{\alpha}_i\lambda_{ii}} \sum_g \left(\alpha_g\left[1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{ii,g} r_{i,g}\right).$$

To finalize the proof, we need to characterize $\bar{\Delta}_i$, which will in turn pin down $\widetilde{\alpha}_{i,k}$. To this end, we can appeal to the definition $\bar{\Delta}_i \equiv \frac{\partial \mathcal{V}_i/\partial \ln w_i}{\partial D_i/\partial \ln w_i}$, which implies that

$$\bar{\Delta}_i = \frac{(1-\bar{\alpha}_i) - \lambda_{ii} + \sum_k \left(\left[\alpha_k \mathcal{E}_{i,k} - \alpha_k \frac{\tilde{\delta}_{ii}}{\tau_{i,k}}\left(\mathcal{E}_{i,k}-1\right)\right] r_{i,k}\right) + \sum_k \left(\left[\alpha_k - \delta_{ii} v_{i,k}\right] r_{ii,k} r_{i,k}\right) \frac{\partial Y_i}{\partial \ln w_i}}{\left(1-\lambda_{ii}\right)\frac{\partial Y_i}{\partial \ln w_i} - \mathcal{E}_i}$$

We can replace for $\alpha_k \mathcal{E}_{i,k} - \alpha_k \frac{\tilde{\delta}_{ii}}{\tau_{i,k}}\left(\mathcal{E}_{i,k} - 1\right)$ from the F.O.C. (Equation 37), which implies

$$\begin{aligned} \bar{\Delta}_i &= \frac{(1-\bar{\alpha}_i) - \lambda_{ii} + \sum_g \left(\left[r_{ii,g} - \left(1-\alpha_g\right)\frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{i,g}\right) + \sum_g \left(\alpha_g\left[1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{ii,g} r_{i,g}\right)\left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_g \frac{\partial \ln Y_i}{\partial \ln(1-a_{i,g})}\right]}{\left(1-\lambda_{ii}\right)\left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_k \frac{\partial Y_i}{\partial \ln(1-a_{i,k})}\right]} \\ &= \frac{\sum_g \left[\left(1-\alpha_g\right)\left(1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right) r_{i,g}\right] + \sum_g \left(\alpha_g\left[1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{ii,g} r_{i,g}\right)\left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_g \frac{\partial \ln Y_i}{\partial \ln(1-a_{i,g})}\right]}{\left(1-\lambda_{ii}\right)\left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_k \frac{\partial Y_i}{\partial \ln(1-a_{i,k})}\right]} \end{aligned} \tag{38}$$

Reapplying the Implicit Function Theorem to $Y_i = w_i L_i + \sum_k \alpha_k P_{ii,k} Q_{i,k}$ implies that

$$\frac{\partial \ln Y_i}{\partial \ln w_i} + \sum_k \frac{\partial \ln Y_i}{\partial \ln(1-a_{i,k})} = \frac{1 - \sum_k \left(\alpha_k r_{i,k}\right) + \sum_k \left(\alpha_k \mathcal{E}_i r_{i,k}\right)}{1 - \bar{\alpha}_i \lambda_{ii}} - \sum_k \frac{\alpha_k \mathcal{E}_i r_{i,k}}{1 - \bar{\alpha}_{ii}\lambda_{ii}} = \frac{1-\bar{\alpha}_i}{1-\bar{\alpha}_{ii}\lambda_{ii}}.$$

Combining the above expression with Equation 38 and assuming that $\alpha_k = \alpha$ for all k, yields the following:

$$(1-\lambda_{ii})\,\frac{1-\alpha}{1-\alpha\lambda_{ii}}\bar{\Delta}_i = (1-\alpha)\sum_g\left[\left(1-\frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right)r_{i,g}\right] + \frac{1-\alpha}{1-\alpha\lambda_{ii}}\sum_g\left(\alpha\left[1-\frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right]r_{ii,g}r_{i,g}\right),$$

Finally, noting the definition for $\widetilde{\alpha}_{i,k}-\alpha$, delivers the following expression

$$\begin{aligned}\widetilde{\alpha}_{i,k}-\alpha &= \left[(1-\alpha)\sum_g\left[\left(1-\frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right)r_{i,g}\right] + \alpha\sum_g\left(\left[1-\frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right]r_{ii,g}r_{i,g}\right)\right](1-\lambda_{ii})^{-1} \\ &= \sum\left[\left(1-\frac{\tilde{\delta}_{ii}}{\tau_{i,k}}\right)\frac{1-\alpha(1-r_{ii,g})}{1-\lambda_{ii}}r_{i,g}\right] = -\sum_g\left[\left(\frac{(\widetilde{\alpha}_{i,g}-\alpha)\mathcal{E}_{i,g}+1-r_{ii,g}}{\alpha\mathcal{E}_{i,g}-1}\right)\frac{1-\alpha(1-r_{ii,g})}{(1-\lambda_{ii})}r_{i,g}\right].\end{aligned}$$

The above system implies that $\tilde{\alpha}_{i,k} = \tilde{\alpha}_i$ is uniform. So, given that $\mathcal{E}_{i,g} = -\epsilon_g\left(1-r_{ii,g}\lambda_{ii,g}\right)$, we can solve for $\tilde{\alpha}_i$ as

$$\tilde{\alpha}_i - \alpha = \frac{\sum_g\left[\frac{1-r_{ii,g}}{\epsilon_k\left(1-r_{ii,g}\lambda_{ii,g}\right)+1}\frac{1-\alpha(1-r_{ii,g})}{(1-\lambda_{ii})}r_{i,g}\right]}{\sum_g\left[\left(1+\frac{\epsilon_g\left(1-r_{ii,g}\lambda_{ii,g}\right)}{\epsilon_g\left(1-r_{ii,g}\lambda_{ii,g}\right)+1}\frac{1-\alpha(1-r_{ii,g})}{(1-\lambda_{ii})}\right)r_{i,g}\right]} > 0$$