Trade, Firm-Delocation, and Optimal Climate Policy

Farid Farrokhi
Purdue University

Ahmad Lashkaripour
Indiana University

December 7, 2020

Abstract

To what extent can trade policy help reduce global carbon emissions? We examine this question using a multi-country multi-industry general equilibrium trade model with transboundary carbon externalities. Our framework accommodates firm-delocation in response to policy, multilateral carbon leakage, and returns to scale in production and abatement. Our central result is a set of simple formulas for unilaterally optimal trade and carbon taxes in an open economy. The optimal policy consists of (i) a uniform carbon tax across all industries; (ii) industry-level production subsidies that restore marginal-cost-pricing independent of the industry's carbon intensity; (iii) industry-level import taxes that penalize carbon-intensive imports but less so in high-returns-to-scale industries; and (vi) industry-level export subsidies that, in addition to improving the terms of trade, promote clean exports against carbon-intensive foreign competition. Mapping our formulas to data, we find that trade taxes can replicate only around 3% of the carbon reduction attainable under (first-best) cooperative global carbon taxes. This lack of effectiveness is partly driven by a tension between the *carbon-reducing* and *terms-of-trade* rationales for trade taxation under scale economies. Trade taxes, however, can be remarkably effective at enforcing international climate agreements even in the presence of scale economies and firm-delocation effects.

1 Introduction

Despite growing concerns over greenhouse gas emissions, international climate 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 carbon taxes and trade penalties (Nordhaus (2015)). Relatedly, some experts have advocated for sub-global agreements wherein a bloc of nations use trade policy to achieve extraterritorial climate objectives.[1]

The basic idea behind these proposals is that governments can tackle cross-border emission 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 carbon tax 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 (e.g. Markusen 1975).[2] On the empirical 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 reducing global emissions unilaterally, or as a device to enforce global climate cooperation.

In this paper, we characterize the unilaterally optimal trade and carbon taxes 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. Using these results, we also characterize the Nash equilibrium in which countries set their taxes non-cooperatively, which we benchmark against the outcome under global cooperation.

To these goals, we produce an intermediate *envelop result* that helps us overcome the main challenges in characterizing optimal policy in general equilibrium. Our result demonstrates 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 a few unresolved questions: First, to what extent could trade taxes correct transboundary carbon externalities and cross-border carbon leakages? Second, how much can unilateralism achieve relative to global cooperation insofar as global emission is concerned? Third, can trade taxes serve as an effective tool to overcome the free-riding problem in international climate agreements?

Our framework exhibits two properties that make it suitable for addressing these questions. First, our model accommodates endogenous entry and allows for the relocation of firms across space and

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

[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. For a recent analysis, see Weisbach and Kortum (2020).

industries in response to policy. In addition to allowing for reallocation across countries and industries, these responses generate scale economies in both production and abatement. The prior literature on trade and environment has often overlooked issues involving endogenous entry, economies of scale, or firm-delocation. However, recent evidence suggests that these features are crucial to how firms respond to changes in environmental and trade policies (Shapiro and Walker (2018)).[3]

Second, our theory delivers sufficient statistics formulas for optimal taxes on emissions, production, and trade (both imports and exports), which can be readily mapped to data. These theoretical developments simplify and enrich our quantitative analysis in several ways. Above all, our analytic formulas simplify the task of computing *optimal* trade taxes under climate change concerns —a task that has often eluded the past literature. Our analytical characterization of optimal policy delivers the sufficient set of elasticities and baseline data required for analyzing the effects of trade and climate policy. Invoking this property, we uncover the full potential of trade taxes at reducing global emissions, while tractably accounting for previously-neglected interactions between firm entry, abatement, and scale economies.

At another level, our theory uncovers new environmental trade-offs facing trade taxes. In particular, we find that endogenous entry creates scale economies in abatement, which dampen the effectiveness of optimal trade taxes at correcting cross-border emission leakage. On one hand, optimal trade taxes include a margin to tackle cross-border emissions by raising taxes on foreign emission-intensive industries. On the other hand, they include a margin to subsidize foreign high-returns-to-scale industries. Overall, the effectiveness of trade policy at reducing transboundary carbon emissions hinges on the empirical cross-industry correlation between emission-intensity and the degree of scale economies.

We map our theory to data on trade, production, and emissions across 15 major regions and 19 tradeable and nontradeable industries in 2009. We estimate the degree of scale economies in each industry using De Loecker and Warzynski's (2012) production-side methodology for estimating markups. In addition, we recover the elasticity of abatement with respect to carbon taxes as well as each country's policy attention towards disutility from emissions using data on currently-applied domestic carbon taxes.

Our quantitative analysis sheds light on several imminent climate policy issues. First, our analysis provides a unique glimpse into the "full potential" of unilateral trade policy at reducing carbon

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

emissions. We find that trade taxes can (at most) replicate only around 3% of the carbon reduction attainable under the global cooperative carbon taxes. Moreover, we find that failing to account for scale economies in production/abatement or for firm-delocation in response to policy will overstate these results.

Second, our quantitative analysis suggests that the previously-neglected tension between the *terms-of-trade* and *emission-reducing* benefits from trade taxation is quantitatively important. While the terms-of-trade rationale incentivizes lower tariffs on high-returns-to-scale industries, the emission-reducing rationale requires high tariffs on emission-intensive industries. Empirically, we find that returns-to-scale and emission-intensity are positively correlated across tradeable industries. For example, typical emission-intensive industries such as Mining, Chemicals, and Minerals also feature high degree of returns to scale. The resulting tension hampers the effectiveness of trade policy at reducing transboundary emissions. In addition, the counter-acting effect between these two rationales partly explains the environmental bias of trade policy, which has been recently documented by Shapiro (2020).

Third, we find that even though trade taxes are an ineffective standalone instrument to reduce emissions, they are remarkably effective at enforcing climate cooperation. This finding contributes to the climate club proposal put forth by Nordhaus (2015). In our analysis, we study a climate club initiated by the US and EU countries, and take a conservative, pessimistic stance that other countries only care about local (as opposed to global) carbon emissions. We find that under optimal trade penalties (which are straightforward to calculate with our theory) nearly all countries are better off joining the US-EU climate club. That is, countries are better off joining the climate club and adopting higher-than-unilaterally-optimal carbon taxes, than to expose themselves to EU-US trade penalties. This outcome is true even if we account for scale economies in abatement/production and firm-delocation in response to policy.

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 differentiation, endogenous abatement, or firm-delocation. As a result, the results from this literature have been rarely used to guide the general equilibrium 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.

More closely to our paper is the work of Weisbach and Kortum (2020) who characterize the optimal trade, production, and carbon taxes in a setting that combines elements from Markusen (1975) and Dornbusch et al. (1977). Our paper complements Weisbach and Kortum (2020) in three ways. First, in terms of modeling emissions, they explicitly specify markets for energy whereas our model is built on Copeland and Taylor (2004) in which energy markets are implicitly defined. Second, in terms of methodology, rather than the primal approach, we adopt the dual approach, which we refine and customize for a class of GE trade models. Third, our theory is designed to be taken to data on multiple countries and industries for quantitative exercises.

Second, our analysis is related to an emerging body of quantitative work that analyzes the efficacy of trade policy at tackling environmental emission (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 emission.

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 and Lugovskyy (2016), Bartelme et al. (2019), Beshkar and Lashkaripour (2020)). These studies have bridged a longstanding 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 and Rodríguez-Clare (2014)). But less attention has been paid to embedding environmental 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 international cooperative and non-cooperative Nash outcomes. In Section 5 we map our theory with our optimal policy formulas to data, which we use in Section 6 to quantify the efficacy of trade policy at reducing global carbon emissions.

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 origin country j, destination country i, and 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 denotes national income, $\bar{D}_i$ denotes the exogenous national debt (i.e., trade deficit), 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 $\tilde{\mathbf{P}}_i = \{\tilde{P}_{ji,k}\}$ denote the entire vector of after-tax consumer prices in country i. The consumer's problem implies an indirect utility function, $V_i(Y_i, \tilde{\mathbf{P}}_i)$, and a Marshallian demand function, $Q_{ji,k} = \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$, for each 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) with respect to the price of variety (ni, g) is,

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

Second, the elasticity of demand for ji,k with respect to income is:

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

While we impose no parametric restrictions on the demand function, we assume that preferences are well-behaved so that the own price elasticity of demand satisfies $\varepsilon_{ji,k} \equiv \varepsilon_{ji,k}^{(ji,k)} \leq 1$.

2.2 Supply

Firms and Market Structure. Production in each *origin j–industry 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 an entry cost $w_j \bar{f}_{j,k}$ to supply their differentiated variety to various destinations. w_j denotes the labor wage rate in origin j 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 an origin–industry specific carbon tax, $\tau_{j,k}$. Firms (from the same origin) can be treated as symmetric: They all choose a common abatement level $a_{j,k} \equiv a_{j,k}(\omega)$, deliver quantity $q_{jik} \equiv q_{ji,k}(\omega)$ to destination i subject to an iceberg trade cost $\bar{d}_{ji,k} \geq 1$ with $\bar{d}_{jj,k} = 1$,[4] and generate a carbon emission $z_{ji,k} \equiv z_{ji,k}(\omega)$ in this process.

Following Copeland and Taylor (2004), a firm's carbon emission per unit of output, $z_{ji,k}/(d_{ji,k} q_{ji,k})$, equals $(1 - a_{j,k})^{1/\alpha_k - 1}$, where $\alpha_k > 0$ is the "emission elasticity" which varies across industries $k \in \mathbb{K}$. Given the choice of abatement $a_{j,k}$, the 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 implies that firms produce less carbon emission and pay less in carbon taxes, but face a higher marginal cost of production.

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

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

with $\gamma_k > 1$ denoting the elasticity of substitution across firm-level varieties from the same origin.

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

Facing with substitution elasticity γ_k, firms charge a constant markup over their marginal cost, which implies the following producer price index for composite variety ji, k:

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

In the above expression, $M_{j,k} \equiv \left|\Omega_{j,k}\right|$ denotes the mass of firms. It is pinned down by the free entry condition, that requires entry costs, $M_{j,k} w_j \bar{f}_{j,k}$, be equal to gross profits across all destinations, $\sum_i \frac{1}{\gamma_k} P_{ji,k} Q_{ji,k}$. Putting these together with $P_{ji,k} = \bar{d}_{ji,k} P_{jj,k}$ and $Q_{j,k} = \sum_i \bar{d}_{ji,k} Q_{ji,k}$, yields the following expression for the mass of firms:

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

Using the Equations (Price) and (Entry), we can express industry-level variables as functions of abatement and output in each origin-industry:

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

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

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

In the above expressions, $\bar{p}_{jj,k}$, $\bar{z}_{j,k}$, $\bar{m}_{j,k}$ are exogenous shifters of price, carbon emission, and mass of firms.[5] Internal economies of scale operate through the endogenous mass of firms, which is given by Equation (4). The resulting scale effects impact both industry-level price and emission, and are reflected in the term $\left(Q_{j,k}/(1-a_{j,k})\right)^{-1/\gamma_k}$ in Equations (2) and (3). This formulation indicates that scale economies are operative through both production and abatement to a common extent, and are governed by γ_k. Finally, choices of abatement are regulated by carbon taxes, which we discuss next.

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

2.3 Policy Instruments

The government in country i has access to a full set of tax instruments necessary to replicate the (unilaterally) first-best outcome. These tax instruments include:[6]

1. An import tax, $t_{ji,k}$, applied to each imported variety ji, k ($t_{ii,k} = 0$ by design)
2. An export subsidy, $x_{ij,k}$, applied to each exported variety ij, k ($x_{ii,k} = 0$ by design)
3. A production subsidy, $s_{i,k}$, applied to all outputs in *origin i-industry k* irrespective of the location of final sales.
4. An carbon tax, $\tau_{i,k}$, applied to all outputs in *origin i-industry 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. Specifically, after-tax consumer prices are related to before-tax producer prices according to[7]:

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

In the case that only home country i sets taxes, 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 $\{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}, \tilde{P}_{ii,k}\}_{j \neq i,k}$,

$$(1 + t_{ji,k}) = \frac{\tilde{P}_{ji,k}}{P_{ji,k}}, \quad (1 + x_{ij,k}) = \frac{P_{ij,k}}{\tilde{P}_{ij,k}} \frac{P_{ii,k}}{\tilde{P}_{ii,k}}, \quad (1 + s_{i,k}) = \frac{P_{ii,k}}{\tilde{P}_{ii,k}} \tag{5}$$

Hence, the government can replicate any choice of trade and production tax-cum-subsidies with the right choice of consumer/producer price wedges. Likewise, the government can choose abatement levels $\{a_{i,k}\}_k$ to replicate their desired carbon tax schedule $\{\tau_{i,k}\}$. Specifically, the optimal choice of

[6] 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.

[7] An alternative way of representing this relationship is $\tilde{P}_{ji,k} = (1 + t_{ji,k})(1 + s^{alt}_{i,k})(1 + x^{alt}_{ij,k}) P_{ji,k}$. Since the policy tools related to production and exports are typically understood as *subsidies*, we have replaced $(1 + s^{alt}_{i,k}) = 1/(1 + s_{i,k})$ and $(1 + x^{alt}_{ij,k}) = 1/(1 + x_{ij,k})$.

abatement by cost-minimizing firms in origin j is given by:[8]

$$(1 - a_{j,k}) = \left(\frac{\alpha_k}{1 - \alpha_k}\right)^{\alpha_k} \left(\frac{w_j / \bar{\varphi}_{j,k}}{\tau_{j,k}}\right)^{\alpha_k}. \tag{6}$$

The above equation indicates that optimal abatement is a function of the wage to carbon tax ratio, with the extent of the relationship controlled by the emission elasticity α_k.[9]

2.4 General Equilibrium

Revenues. Total income in country i, which pins down total expenditure per Equation (1), is the sum of wage payments, $w_i \bar{L}_i$, lump-sum carbon and non-carbon tax revenues, T_i, and the trade deficit, $\bar{D}_i$:

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

The tax revenue T_i is the sum of revenues collected from carbon and import taxes and the revenues exhausted by production and export subsidies:[10]

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

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

[8] Our specification can be alternatively interpreted as one in which production employs a pollutant input (whose unit cost is the carbon 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, emission, and labor for production of ji,k. In this alternative formulation, the unit cost of ji,k is given by $c_{ji,k} = \bar{d}_{ji,k} \alpha_k^{-\alpha_k} (1 - \alpha_k)^{-(1-\alpha_k)} \tau_{j,k}^{\alpha_k} (w_j / \bar{\varphi}_{j,k})^{1-\alpha_k}$, emission 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 emission, $(1 - a_{j,k}) = \left(z_{ji,k} \middle/ \bar{\varphi}_{j,k} l_{ji,k}\right)^{\alpha_k}$, delivers equation (6). 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 $a_{j,k} = 0$.

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

[10] The payment for carbon taxes is equivalent to the compensation of carbon emission as a factor of production.

Definition. For a given vector of taxes $\{t_{ji,k}, x_{ij,k}, s_{i,k}, \tau_{i,k}\}_{j,k}$, *equilibrium* is a vector of wages $\{w_j\}$ such that before-tax prices $\{P_{ji,k}\}$ are given by (2), carbon emission $\{Z_{j,k}\}$ by (3), mass of firms $\{M_{j,k}\}$ by (4), abatement $\{a_{j,k}\}$ by (6), demand quantities by $Q_{ji,k} = \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i)$ in which after-tax prices $\tilde{\mathbf{P}}_i \equiv \{\tilde{P}_{ji,k}\}$ are given by (5) and national expenditure Y_i in equation (1) equals national income according to (7) with tax revenues given by (8), and labor markets clear:[11]

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

Expenditure/Revenue Shares and Carbon 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 $r_{ji,k}$,

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

In addition, we use $v_{j,k}$ to denote the carbon intensity per unit value of output in *origin j–industry 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 present the objective function that governments aim 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

[11] 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).

function of the government in country i is given by:

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

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 carbon emission externalities from global production. δ_{ni} is the disutility to residents of country i from every unit of carbon emission generated in origin n. In this specification, a unit of emission generated in origin n may have a greater negative effect on that country or its neighbors than faraway countries. For instance, we may specify $\delta_{ni} = \delta_A + \delta_B$ if $n = i$, and $\delta_{ni} = \delta_B$ if $n \neq i$, which amounts to the government assigning an additional weight to local emission beyond its care for global emission. These disutility levels can be interpreted as the present discounted externality from global climate change—see Shapiro (2016).

Digging deeper, 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 emission, where $\tilde{P}_i \equiv (\partial V_i(.)/\partial Y_i)^{-1}$. Noting the government's objective function, we can now define the optimal unilateral policy.

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

The elucidate the above definition, Appendix A.1 presents a minimal set of equations that describe the unilateral policy problem.

3 Optimal Unilateral Policy

In this section, we characterize a country's optimal unilateral tax schedule. The unilaterally optimal policy corrects three types of inefficiency from the standpoint of a non-cooperative government who acts in their self interest:

1. [*Carbon Externalities*] The externality imposed on domestic consumers from local and trans-boundary carbon emissions.

2. [*Markup Distortions*] The misallocation caused by cross-industry markup heterogeneity.

3. [*Terms-of-Trade*] The unexploited unilateral gains from exercising national-level export and import market power.

We currently have a limited understanding of how these distinct policy channels interact. To shed light on their interaction, we analytically characterize the optimal unilateral tax schedule for each country. This is a challenging task, which explains why previous characterizations of optimal trade and carbon taxes 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 presenting the optimal tax results, we discuss our methodological approach. Our goal here is to demonstrate that we have a systematic way of characterizing optimal policy with applications beyond this particular 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 of the previous literature. To fully address these issues, we establish an intermediate envelope result which is presented next. Throughout the paper, if not reported in the main text, we report our derivations and proofs in the appendix.

We proceed with presenting our intermediate envelop result in Section 3.1 and derive optimal unilateral tax formulas in Section 3.2. In Subsections 3.3 and 3.4 we discusses special cases of and the key trade-offs in our optimal policy formulas. In Subsection 3.5, we examine second-best scenarios in which a governments is afforded fewer tax instruments than is necessary to attain the first-best outcome.

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 $\{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}, \tilde{P}_{ii,k}\}_{j\neq i,k}$ to replicate any set of trade and production tax-cum-subsidies $\{t_{ji,k}, x_{ij,k}, s_{i,k}\}_{j,k}$ according to Equation (5), and can chose abatement levels $\{a_{i,k}\}_k$ to replicate any set of carbon taxes $\{\tau_{i,k}\}$ according to Equation (6). 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 $\mathbb{P}_i \equiv \{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}, \tilde{P}_{ii,k}, a_{i,k}\}_{j\neq i,k}$ is useful, as it highlight the economic variable each tax instrument directly targets. 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 $\tilde{\mathbf{P}}_{ji} = \left\{\tilde{P}_{ji,k}\right\}_{j\neq i,k}$, $\tilde{\mathbf{P}}_{ij} = \left\{\tilde{P}_{ij,k}\right\}_{j\neq i,k}$, $\tilde{\mathbf{P}}_{ii} = \left\{\tilde{P}_{ii,k}\right\}_k$, and $\mathbf{a}_i = \left\{a_{i,k}\right\}_k$.

We can simplify our optimal policy problem by re-casting it as a problem of choosing consumers prices and abatement levels instead of tax instruments. The following lemma establishes that after solving such a problem, the optimal taxes can be recovered using Equations (5) and (12).

Lemma 1. *Given optimal prices and abatement levels,* $\mathbb{P}_i^\star = \{\tilde{P}^\star_{ji,k}, \tilde{P}^\star_{ij,k}, \tilde{P}^\star_{ii,k}, a^\star_{i,k}\}_{j\neq i,k}$, *optimal taxes* $\mathbb{I}_i^\star = \{t^\star_{ji,k}, x^\star_{ij,k}, s^\star_{i,k}, \tau^\star_{i,k}\}_{j,k}$ *can be recovered according to the following one-to-one mapping:*

$$1 + t^\star_{ji,k} = \frac{\tilde{P}^\star_{ji,k}}{P_{ji,k}}, \quad 1 + x^\star_{ij,k} = \frac{P^\star_{ij,k}}{\tilde{P}_{ij,k}}\frac{P^\star_{ii,k}}{\tilde{P}_{ii,k}}, \quad 1 + s^\star_{i,k} = \frac{P^\star_{ii,k}}{\tilde{P}_{ii,k}}, \quad \tau^\star_{i,k} = \frac{\gamma_k - 1}{\gamma_k}\frac{\alpha_k}{v_{i,k}(a^\star_{i,k})}$$

Following the above lemma, we hereafter formulate the optimal policy problem as a choice of $\mathbb{P}_i$ which maximizes the objective function W_i subject to equilibrium conditions. The next two steps establish neutrality results that greatly simplify this reformulated problem.

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, general equilibrium wage effects are welfare-neutral. To make this point, we formulate all variable outcomes as a function of $\mathbb{P}_i$ and the wage vector $\mathbf{w}$. As detailed in Appendix A.2, this formulation results from a system that solves all equilibrium relationships with the exception of the labor market clearing condition. It 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, though, that all $(\mathbb{P}_i; \mathbf{w})$ pairs are not feasible. Given $\mathbb{P}_i$, a feasible vector of wages must satisfy the labor market clearing condition in each country.

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

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

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

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

where $W_i(\mathbb{P}_i;\mathbf{w}) = V_i(Y_i(\mathbb{P}_i;\mathbf{w}), \tilde{\boldsymbol{P}}_i) - \boldsymbol{\delta}_i \cdot \mathbf{Z}(\mathbb{P}_i;\mathbf{w})$. 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 emission to country i. The necessary condition for the optimality of each 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, i.e., $\mathcal{P} \in \tilde{\mathbf{P}}_i \equiv \{\tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ii}\}$. 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.[12] The last term accounts for the general equilibrium wage effects. By choice of numeraire, we normalize wage in one of the 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 corresponding to own's wage, w_i. Accounting for the change in w_i (relative to $\mathbf{w}_{-i}$) has proven a major obstacle when solving problems like (P1).

[12] 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 $\mathbf{x}$.

The next lemma allows us to overcome this obstacle. 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 opposing but equal-sized 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 origin i's producer prices, which amounts to lower tax revenues. This latter effect arises because the after-tax price of home-made varieties, $\{\tilde{\mathbf{P}}_{ij}, \tilde{\mathbf{P}}_{ii}\}$, are held fixed as part of $\mathbb{P}_i$. Importantly, these two opposing effects 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.

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 the policy vector $\mathbb{P}_i$ and income Y_i. As detailed in Appendix A.3, this formulation derives from solving a system that imposes all equilibrium relationships except the budget constraint, $Y_i = \bar{w}_i \bar{L}_i + \bar{D}_i + T_i(\mathbb{P}_i; Y_i)$. Notice, tax revenues T_i depend on income Y_i because home's demand schedule, which dictates these revenues, depends on income. This brings us to define feasible pairs of policy-income as follows.

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

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

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 the reformulation (P2), we can now explain the welfare neutrality of income effects. The first order condition w.r.t. to policy instrument $\mathcal{P} \in \mathbb{P}_i$ 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 on the above expressions, tax revenues $T_i(.)$ and emission disutility $\delta_i \cdot \mathbf{Z}(.)$

react to policy $\mathcal{P}$ directly holding quantities, and indirectly through changes in demand quantities. Note that, once we hold quantities fixed, we are also holding income fixed. Putting these points 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 instrument $\mathcal{P}$ 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 changes in 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 $\mathcal{P}$ is one of the consumer prices in home $\tilde{P}_{ji,k} \in \tilde{\mathbf{P}}_i$, be it either a domestic ($j = i$) or an imported ($j \neq i$) variety. 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 holding the demand schedule fixed; the second equality follows from Roy's identity; and the third equality holds because emission is fully determined by abatement levels 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), it follows that

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

Noting that $\frac{\partial \mathbf{Q}_{in}}{\partial \tilde{P}_{ji,k}} = \frac{\partial \mathbf{Q}_{in}}{\partial \tilde{P}_{ii,k}} = \frac{\partial \mathbf{Q}_{in}}{\partial Y_i} = 0$ if $n \neq i$, 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

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

which means that the income effect is neutral at the optimum. We show in the next section that this trivial solution is also the unique solution to the system of FOCs. Our above discussion shows that the optimal choice with respect to $\tilde{P}_{ii,k} \in \tilde{\mathbf{P}}_i$ and $\tilde{P}_{ji,k} \in \tilde{\mathbf{P}}_i$ entails that income effects are welfare neutral: $\frac{\partial W_i}{\partial Y_i} = 0$. We summarize this conclusion in the following lemma.

Lemma 3. *Within the feasible policy-income set, $(\mathbb{P}_i; Y_i) \in \mathbb{F}_i^Y$, if $\tilde{\boldsymbol{P}}_i \subset \mathbb{P}_i$ is chosen optimally, then income effects are 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 $\mathcal{P}_i \in \mathbb{P}_i$, 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 (which are in $\mathbb{P}_i$) and income fixed, abatement has not effect on the demand schedule. Accounting for these equal-to-zero terms, $\mathbb{P}_i^\star$ solves the following system according to Proposition 1:

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

Discussion. 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 prohibited from setting any of the instruments that correspond to the after-tax prices of varieties originating from home (namely, $\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ij}\}$), then Lemma 2 fails. The intuition is the following: The government in country i can improve its terms-of-trade by inflating its wage, w_i, relative to $\mathbf{w}_{-i}$. The gains from inflating w_i can be perfectly mimicked with an appropriate adjustment in production and export subsides, $\{s_{i,k}, x_{ij,k}\}_{j\neq i,k}$. This adjustment in our reformulated problem corresponds to an appropriate choice vis-à-vis price vectors, $\tilde{\mathbf{P}}_{ii}$ and $\tilde{\mathbf{P}}_{ij}$. The argument is that a proper adjustment in production and export subsidies can achieve any level of national sales; and, provided that labor markets clear, national sales pin down home's wage, w_i. This argument holds even if the choice with respect to $\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ij}\}$ is not optimal, but it fails if the government is prohibited from manipulating any element of these price vectors. In that case, wage effects become non-neutral and should be properly tracked when solving the optimal policy problem

A similar argument applies to Lemma 3 which states that 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 prohibited from manipulating any element in $\tilde{\mathbf{P}}_i \equiv \{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}\}$, 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 that follows.

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 abatement $\mathbf{a}_i$ is set sub-optimally, Lemmas 2 and 3 continue to hold. Hence, Proposition 1 applies to second-best scenarios where governments cannot tax emission but can manipulate the entire vector of after-tax prices, $\{\tilde{\mathbf{P}}_{ii}, \tilde{\mathbf{P}}_{ji}, \tilde{\mathbf{P}}_{ij}\}$, associated with their economy.

3.2 Characterizing the Optimal Tax Schedule

Proposition 1 describes 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, with detailed derivations provided in Appendixes B.2 and B.4.

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

$$[a_{i,k}] \qquad \tilde{\delta}_{ii} v_{i,k}(a_{i,k}^\star) - \alpha_k \frac{\gamma_k - 1}{\gamma_k} = 0;$$

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

$$[\tilde{P}_{ij,k}] \qquad 1 - \sum_{\ell \neq i} \sum_g \left[\left(\omega_{\ell i,g} + \tilde{\delta}_{\ell i} v_{\ell,g} \frac{\gamma_g - 1}{\gamma_g} \right) \frac{e_{\ell j,g}}{e_{ij,k}} \varepsilon_{\ell j,g}^{(ij,k)} \right]$$
$$+ \sum_g \left[\left(1 - \left(1 - \alpha_g \frac{\gamma_g - 1}{\gamma_g} + \tilde{\delta}_{ii} v_{i,g}\right) \frac{\gamma_g - 1}{\gamma_g} \frac{P_{ij,g}^\star}{\tilde{P}_{ij,g}} \right) \frac{e_{ij,g}}{e_{ij,k}} \varepsilon_{ij,g}^{(ij,k)} \right] = 0$$

where $\omega_{ji,k} \equiv \left(\frac{\partial \ln P_{jj,k}}{\partial \ln Q_{ji,k}}\right)_{\mathbb{P}_i, \boldsymbol{w}, Y_i} = \frac{-r_{ji,k}}{\gamma_k + \sum_{n \neq i} r_{jn,k} \varepsilon_{jn,k}}$ 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 the marginal cost of production and the marginal utility gains associated with lower carbon emissions. Combining the F.O.C w.r.t $a_{i,k}$ with Equation (12) that relates carbon intensity to carbon tax, yields the following formula for the optimal carbon tax:

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

where recall that $\tilde{\delta}_{ii} \equiv \tilde{P}_i \delta_{ii}$. 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 and NK unknowns, which are summarized by the following matrix equation:

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

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 subsidy, $1 + s_{i,k}^\star = P_{ii,k} / \tilde{P}_{ii,k}^\star$ independently from the choice of export subsidies, $1 + x^{\star}_{ij,k} = P_{ij,k}/\tilde{P}^{\star}_{ij,k}$. To solve the above matrix equation, we invoke on another intermediate result that ensures the invertibility of the system.

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

Given Lemma 5, the unique solution to Equation 20 is the trivial solution, which indicates that:

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

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

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

$$\begin{aligned}
&[\textit{import tax}] && 1 + t^{\star}_{ji,k} = 1 + \omega_{ji,k} + \tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k} && \forall j,k \\
&[\textit{export subsidy}] && 1 + x^{\star}_{ij,k} = \left(1 + \frac{1}{\varepsilon_{ij,k}}\right) \chi_{ij,k} && \forall j,k \\
&[\textit{domestic subsidy}] && 1 + s^{\star}_{i,k} = \frac{\gamma_k}{\gamma_k - 1} && \forall k \\
&[\textit{carbon tax}] && \tau^{\star}_{i,k} = \tau^{\star}_{i} = \tilde{\delta}_{ii} && \forall k
\end{aligned} \tag{21}$$

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

To put in words, the optimal unilateral policy for country i includes (i) a uniform Pigouvian tax on emission, $\tau^{\star}_i$, (ii) an industry-specific Pigouvian production subsidy, $s^{\star}_{i,k}$, that eliminates the cross-industry markup heterogeneity, (iii) import taxes, $t^{\star}_{ji,k}$ that penalize high-emission imports and also take advantage of unexploited import market power, and (iv) export subsidies, $x^{\star}_{ij,k}$, that promote low-emission exports and also take advantage of unexploited export market power. The fact that production subsides are *carbon-blind* is a manifestation of the *Targeting Principle*. By this principle, carbon taxes are the optimal instrument to tackle the externality of local emission on domestic consumers, because they correct the externality at its source.

Optimal trade taxes are designed to both improve the terms-of-trade (ToT) and correct transboundary 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:

$$1+t^{\star}_{ji,k} = \underbrace{1+\omega_{ji,k}}_{\text{ToT driven}} \quad + \underbrace{\tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k}}_{\text{crabon reducing}} . \tag{22}$$

The ToT-driven component is motivated by country i's collective import market power vis-á-vis partner j. 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 carbon-reducing component is intended to tackle the transboundary carbon emission externality of goods imported from origin j. Recall that country j fails to internalize this transboundary externality when acting non-cooperatively. Intuitively, the emission-correcting component is higher on carbon-intensive (high-v) import goods, and can be viewed as a *border carbon adjustment tax.*

Likewise, export subsides are designed to both improve the terms-of-trade (ToT) and correct transboundary carbon emissions. The export subsidy on good ij,k, therefore, exhibit two distinct components:

$$1+x^{*}_{ij,k} = \underbrace{\left(1+\frac{1}{\varepsilon_{ij,k}}\right)}_{\text{ToT driven}} \times \underbrace{\chi_{ij,k}}_{\text{carbon reducing}} . \tag{23}$$

The ToT-driven term equals an optimal markup, if country i were pricing its composite export good as a single representative monopolist. The carbon-reducing term subsidizes exports of varieties that compete with carbon-intensive (high-v) foreign varieties in market j. Note that this term internalizes the import tax charged by country i on various trading partners. The intuition can be put as follows. If export subsides prompt more exports from home country i to destination j, then a third country n would reallocate its exports away from j and possibly back to home country i. This gives home country i some leverage to curb emission in every third-country n using its import tax choice. We provide a more detailed discussion when we consider CES-Cobb-Douglas preferences.

3.3 Optimal Policy Formulas in Special Cases

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 a Ricardian trade model. The Ricardian special case of our framework is isomorphic to the multi-industry Eaton and 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, Theorem 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 studied by Costinot et al. (2015).

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,

$$U_i(\boldsymbol{Q}_i) = \prod_k \left(\sum_j b_{ji,k}^{1/\sigma_k} Q_{ji,k}^{\frac{\sigma_k - 1}{\sigma_k}} \right)^{e_{i,k} \frac{\sigma_k}{\sigma_k - 1}}, \tag{24}$$

where $e_{i,k}$ is the expenditure share of country i on industry k, and σ_k is the (Armington) elasticity of substitution between origin countries. The Marshallian demand elasticities in this special case are given by the following formulas:

$$\varepsilon_{ji,k} \equiv \varepsilon_{ji,k}^{(ji,k)} = -1 - (\sigma_k - 1)(1 - \lambda_{ji,k}), \qquad \varepsilon_{ni,k}^{(ji,k)} = (\sigma_k - 1)\lambda_{ji,k} \quad (n \neq j); \qquad \varepsilon_{ni,g}^{(ji,k)} = 0 \quad (g \neq k).$$

Plugging the above values into Theorem 1, yields the following formula for optimal trade taxes,

$$\begin{aligned} 1 + t_{ji,k}^{\star} &= 1 + \tilde{\delta}_{ji}^{\star} v_{j,k}^{\star} \frac{\gamma_k - 1}{\gamma_k} - \frac{r_{ji,k}^{\star}}{\gamma_k - \sum_{n \neq i} r_{jn,k}^{\star} \left[1 + \epsilon_k (1 - \lambda_{jn,k}^{\star}) \right]} \\ 1 + x_{ij,k}^{\star} &= \left[1 + \frac{\gamma_k - 1}{\gamma_k} \sum_{n \neq i} \tilde{\delta}_{ni}^{\star} v_{n,k}^{\star} \frac{\lambda_{nj,k}^{\star}}{1 - \lambda_{ij,k}^{\star}} \right] \left(1 + \frac{1}{(\sigma_k - 1)\left(1 - \lambda_{ij,k}^{\star}\right)} \right)^{-1}. \end{aligned} \tag{25}$$

where the term inside the brackets in the second line corresponds to $\chi_{ij,k}$ in Theorem 1. Absent carbon emission externalities (i.e., set $\nu_{n,k} = 0$ for all n, k), the above formulas collapse to the familiar optimal trade tax formulas in multi-industry quantitative trade models (see Lashkaripour and Lugovskyy (2016)).

3.4 The Trade-Off Between Emission Correction, Scale Economies, and the ToT

The formulas under Equation (25) uncover some primitive trade-offs facing border adjustment carbon taxes. The first trade off is reflected in the emission-correcting term in the optimal import tax formula. This term is the product of (1) $\tilde{\delta}_{ji} v_{j,k}$, which taxes carbon-intensive imports, and (2) $\frac{\gamma_k - 1}{\gamma_k}$, which operates as a tax deflator for industries with high returns-to-scale in emission (low-γ_k). As such, the effectiveness of import taxes at reducing transboundary carbon emission is dictated by $\text{Cov}_k(v_{j,k}, \gamma_k)$. In the case where $\text{Cov}_k(v_{j,k}, \gamma_k) < 0$, 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 $\text{Cov}_k(v_{j,k}, \gamma_k) > 0$ import taxes become quite effective as they hit two birds with one stone.

A similar trade-off faces export subsidies: the optimal export subsidy includes an emission-correcting term (in brackets) that promotes country i's clean exports against its high-emission competition in market j.[13] This term, though, is smaller the higher the degree of scale economies in an industry, i.e., the lower $(\gamma_k - 1)/\gamma_k$ in the second line of Equation (25). The intuition is the same: promoting one's exports against carbon-intensive competition leads to an additional increase in the competition's carbon-intensity through scale effects. So, the optimal level of promotion is accordingly weaker.[14]

The second trade-off occurs between the *ToT-driven* and *emission-correcting* terms. With regards to import taxes, if $\text{Cov}_k(v_{j,k}, \gamma_k) < 0$ the ToT-driven component asks for a lower import tariff on carbon-intensive industries. With regards to export subsidies, if $\text{Cov}_k(v_{j,k}, \sigma_k) > 0$ the ToT-driven component asks for a higher export subsidy (or a lower export tax) on carbon-intensive industries. Hence, in both case, the optimal policy net of climate objectives may exhibit a climate or environmental bias. The direction or magnitude of these trade-offs is ultimately an empirical issue, which we will come back to in Section 5 when our model is mapped to data.[15]

[13]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 high export subsidy (or a lower export tax) to good ij, k to increase its sales in market j against high-emission rivals there. 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.

[14]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 subsidies 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.

[15]We can connect this result to a recent literature that documents an environmental bias in *applied* tariffs. Shapiro (2020) attributes this bias to upstream, carbon-intensive industries being more organized than downstream industries. As such, they can demand a greater degree of tariff protection from governments. Our theory instead indicates that such biases can arise purely from terms-of-trade considerations.

3.5 Discussion: Optimal Policy in Second-Best Scenarios

Theorem 1 concerns a unilaterally first-best scenario in which the government has access to a complete set of policy instruments. In many scenarios, governments may face limitations in using the policy instruments necessary to achieve the first best. In addition to prevalent political economy issues, second-best scenarios may arise from agreements on environment or trade that tie the policy-maker's hands with regards to certain policy tools. In the section, we derive three sets of results that shed light on these second-best scenarios.

First, we examine optimal taxes in a second-best scenario where a country takes carbon taxes as given. This case is relevant for a country under a commitment to international agreements on climate, which prevents it from setting carbon taxes according to unilateral objectives. In that case, the government alters its production subsidies to:

$$1 + s_{i,k}^{\star\star} = \frac{\gamma_k}{\gamma_k - 1} \left[1 + \tilde{\delta}_{ii}(v_{i,k} - v_{i,k}^{\star})\right]^{-1}, \tag{26}$$

where $v_k^{\star}$ is the emission intensity attainable under the first-best unilateral policy. Consider the case where emission intensity, $v_{i,k}$, is smaller than $v_{i,k}^{\star}$ because country i is abiding with an international climate agreement. In that case, $(1 + s_{i,k}^{\star\star})$ includes an extra subsidy that promotes domestic production. This extra subsidy acts against the climate goals in international agreements. Hence, it is important for international agreements on climate to couple carbon taxes with restrictions on production subsidies.[16]

Second, we consider the second-best case in which countries do not have access to export policies, resembling the ban on export subsidies under the WTO. As detailed in Appendix C, optimal production subsidies and carbon taxes remain *ToT-blind* in this case. More specifically, the formula for production subsidies and carbon taxes as well as the emission-correcting term in import taxes remain unchanged. The only alteration is a uniform multiplier applied to the ToT-improving component in the optimal import tax formula under Equation (22). Intuitively, by the Lerner symmetry, import taxes are strictly more effective than carbon taxes or production subsidies at mimicking export subsidies. As such, when import taxes are applicable, there is no rationale for altering emission

[16] On a related note, our model collapses to one with exogenous emission intensity à la Markusen (1975) as the emission elasticity approaches zero, i.e., $\alpha_k \to 0$ (See footnote (8). Here, carbon taxes can be dropped from the model as firms do not undertake abatement. In this case, the optimal production subsidy includes the markup-correcting term $\frac{\gamma_k}{\gamma_k - 1}$ plus an extra term that taxes high-emission (high-v) industries. Namely, $1 + s_{i,k}^{\star\star} = \frac{\gamma_k}{\gamma_k - 1}\left(1 + \tilde{\delta}_{ii} v_{i,k}\right)^{-1}$. This formula can be also derived by setting $v_{i,k}^{\star} = 0$ in equation (26).

or production taxes to compensate for the absence of export subsidies.

Lastly, there is widespread skepticism that environmental policies are occasionally used as protection in disguise. The argument is that when governments are banned from exercising trade or industrial policies, they may turn to carbon taxes as a second best trade-restricting instrument. To make the case, we suppose that all tax instruments aside from carbon taxes are banned. We show in Appendix C, for a simplified version of our model, that in this case, optimal carbon taxes will be no longer uniform. Instead, it is optimal for a country to apply carbon taxes above the first-best rate in industries where the trade elasticity, $(\sigma_k - 1)$, is low. Doing so, enables a country to contract exports in high-market-power industries as an indirect means to extract markups from the rest of the world.

4 Non-cooperative Nash Equilibrium *vs.* Global Climate Cooperation

In this section, we discuss policy outcomes when many countries simultaneously set their policies. We first discuss the case of global climate cooperation in Subsection 4.1. Such a case yields the globally optimum outcome via deep international cooperation. Then, in Subsection 4.2, we characterize the non-cooperative Nash equilibrium where non-cooperative countries who act in their self interest simultaneously apply their optimal unilateral policy. Equipped with these theoretical results, Section 6 quantifies the consequences of global cooperation versus non-cooperation on climate issues.

4.1 Global Climate Cooperation

The globally optimum outcome is attainable when all countries coordinate their carbon taxes, while internalizing their carbon externality on the rest of the world. Such a scenario is akin to a deep multilateral agreement on trade and climate. Below, we formally define this scenario, which we label global climate cooperation.

Definition. *Global Climate Cooperation* corresponds to an equilibrium wherein all governments set their policy instruments cooperatively in order to maximize global welfare, $\sum_i W_i$, subject to equilibrium conditions (1)-(9).

Under global climate cooperation all countries apply zero trade taxes, as these taxes create inefficient distortions from a global perspective. Globally optimal production subsidies solely correct markup distortions, by restoring marginal cost pricing in each industry. Globally optimal carbon taxes are of Pigouvian nature, correcting each origin's local and transboundary carbon externality. In formal terms, the aforementioned policy schedule can be expressed as follows for each country i:

$$\boldsymbol{x}_i^* = \boldsymbol{t}_i^* = \boldsymbol{0}; \qquad 1 + s_{i,k}^* = \frac{\gamma_k}{\gamma_k - 1}; \qquad \tau_{i,k}^* = \tau_i^* = \sum_{j \in \mathbb{C}} \tilde{P}_j \delta_{ij} \tag{27}$$

Globally optimal carbon taxes discriminate by country of origin because (i) the disutility δ_{ij} from origin i's carbon emissions is non-uniform across locations j, and (ii) converting a dollar loss in location j to a welfare loss requires an adjustment by the location-specific consumer price index. To gain 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 carbon tax relative to CPI is uniform across countries, $\tau_i^* / \tilde{P}_j = \bar{\delta} \bar{L}^{world}$.

4.2 Non-Cooperative Nash Equilibrium

As in the previous case, 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 policy choices in the rest of the world as given.[17]

In the Nash equilibrium, the unilaterally optimal carbon tax and production subsidy formulas are still characterized by Theorem 1. However, the trade share, $\lambda_{nj,k}$, and carbon intensities, $v_{j,k}$, in these formulas now depend on policy choices in the rest of the world. Specifically, Consider country i's optimal export subsidies and import taxes. They depend on transboundary carbon intensities, $\{v_{j,k}\}_{j \neq i}$, which are regulated by optimal carbon 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 (25) to arrive at the following optimal trade tax schedule.

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

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

tax schedule:

$$
\begin{aligned}
&[\textit{import tax}] && 1+t^*_{ji,k} = 1+\alpha_k \left(\frac{\gamma_k-1}{\gamma_k}\right)^2 \frac{\tilde{\delta}_{ji}}{\tilde{\delta}_{jj}} \\
&[\textit{export subsidy}] && 1+x^*_{ij,k} = \left[1+\alpha_k \left(\frac{\gamma_k-1}{\gamma_k}\right)^2 \sum_{n\neq i} \frac{\tilde{\delta}_{ni}}{\tilde{\delta}_{nn}} \lambda_{nj,k}\right] \left(\frac{\sigma_k}{\sigma_k-1}\right)^{-1} \\
&[\textit{domestic subsidy}] && 1+s^*_{i,k} = \frac{\gamma_k}{\gamma_k-1} \\
&[\textit{carbon tax}] && \tau^*_{i,k} = \tau^*_i = \tilde{\delta}_{ii}
\end{aligned}
$$

The optimal carbon taxes and production subsidies remain the same (as in Equation 21), even though all countries simultaneously apply taxes and subsidies. That is because, when governments have access to a complete set of trade tax instruments, their unilaterally optimal choice with respect to carbon taxes and production subsidies 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 carbon externalities. But the extent of these externalities, here, depends on cross-national differences in the perceived cost of carbon (or climate change). For instance, suppose country i's government cares significantly more about carbon emissions 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 origin 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 carbon tax on good ji,k is sub-optimal as it does not internalize the transboundary cost of carbon emissions. 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 carbon externality associated with these imports.[18]

[18] 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), carbon 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, carbon intensity is affected by scale economies in both production and abatement, gov-

5 Mapping Theory to Data

In this section, we describe how our equilibrium relationships, including our optimal policy formulas, map to data. Our objective is to use this mapping to quantify the full effectiveness of trade policy at combating carbon emissions. For quantification purposes, we consider the Cobb-Douglas-CES case of our model. To simplify the presentation of our quantitative approach, we define $\rho_{n,k} \equiv L_{n,k}/\bar{L}_n$ as the share of national labor employed in industry k. In addition, we define $\tau^0_{n,k} = \tau^{\text{factual}}_{n,k}/\delta_{nn}\tilde{P}_n$, as the wedge between the observed carbon tax rate and the optimal carbon tax rate in the factual equilibrium.[19]

5.1 Non-cooperative Nash Equilibrium

Using our optimal tax formulas, we fully characterize the change in equilibrium values when moving from the factual (baseline) equilibrium to the counterfactual *non-cooperative* outcome. Our information set consists of baseline values for key economic variables $\mathcal{B}_v \equiv \{\lambda_{ni,k}, r_{ni,k}, \rho_{i,k}, \tilde{\delta}_{ni}, e_{n,k}, w_n\bar{L}_n, Y_n\}_{ni,k}$, applied tax/subsidy rates $\mathcal{B}_t \equiv (s_{n,k}, \tau^0_{n,k}, x_{in,k}, t_{ni,k}\}_{ni,k}$ and structural elasticities $\mathcal{B}_e = \{\sigma_k, \gamma_k, \alpha_k\}_k$. Using the exact hat-algebra notation, for any generic variable z, we denote its counterfactual value in the non-cooperative equilibrium as $z^\star$, and its change as $\hat{z} \equiv z^\star/z$. Invoking this notation and given full information on $\mathcal{B} \equiv \{\mathcal{B}_v, \mathcal{B}_t, \mathcal{B}_e\}$, we can determine the entire vector of optimal tax/subsidy rates, $\mathcal{R}_t \equiv \{s^\star_{n,k}, \tau^\star_{n,k}, x^\star_{in,k}, t^\star_{ni,k}\}_{ni,k}$, using our optimal tax formulas and the change in key economic variables, $\mathcal{R}_v \equiv \{\hat{w}_n, \hat{Y}_n, \hat{\tilde{P}}_n, \hat{\rho}_{n,k}\}_{ni,k}$, using the equilibrium conditions. Then, given $\mathcal{B} \equiv \{\mathcal{B}_v, \mathcal{B}_t, \mathcal{B}_e\}$ and $\mathcal{R} \equiv \{\mathcal{R}_v, \mathcal{R}_t\}$, we can characterize the counterfactual level of welfare, carbon emission, and other key economic outcomes for each country.

For a clearer exposition, we write baseline variables and structural elasticities $\mathcal{B} \equiv \{\mathcal{B}_v, \mathcal{B}_t, \mathcal{B}_e\}$ in blue, the independent unknown variables $\mathcal{R} \equiv \{\mathcal{R}_v, \mathcal{R}_t\}$ in red, and the intermediate unknown variables in black. Following Section 3.2, the optimal tax/subsidy formulas in the Cobb-Douglas-CES

erned by a common parameter γ_k. In 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 carbon-reducing tariff given scale economies in "production". The second $(\gamma_k - 1)/\gamma_k$ is due to the origin country j's carbon taxes interacting with scale economies in "emission".

[19]Recall that the optimal carbon tax is uniform (as shown by the denominator) but that might be not the case in the data (as shown by numerator). We can run different exercises depending on which scenario we choose to measure applied carbon taxes, or equivalently, based on the value assigned to $\tau^0_{n,k}$. For example, one instructive exercise is to adopt a conservative assumption that, under the status quo, home country n applies its optimal carbon taxes and every foreign country $i \neq n$ has zero carbon taxes, meaning that $\tau^0_{n,k} = 1$ and $\tau^0_{i,k} = 0$ for $i \neq n$.

case are given by:

$$\begin{cases} 1+t^{\star}_{ni,k} = 1+\tilde{\delta}_{ni} v_{n,k}\hat{v}_{n,k}\hat{\tilde{P}}_n \frac{\gamma_k-1}{\gamma_k} - \frac{r_{ni,k}\hat{r}_{ni,k}}{\gamma_k-\sum_{\ell\neq i} r_{n\ell,k}\hat{r}_{n\ell,k}\left(1+(\sigma_k-1)(1-\lambda_{n\ell,k}\hat{\lambda}_{n\ell,k})\right)} & \text{a) optimal imp tax } (ni,k) \\ 1+x^{\star}_{in,k} = \left[1+\frac{\gamma_k-1}{\gamma_k}\sum_{\ell\neq i}\tilde{\delta}_{\ell i}\hat{\tilde{P}}_i v_{\ell,k}\hat{v}_{\ell,k}\frac{\lambda_{\ell n,k}\hat{\lambda}_{\ell n,k}}{1-\lambda_{in,k}\hat{\lambda}_{in,k}}\right]\left(1+\frac{1}{(\sigma_k-1)(1-\lambda_{in,k}\hat{\lambda}_{in,k})}\right)^{-1} & \text{b) optimal exp tax } (in,k) \\ \widehat{1+t_{ni,k}} = \frac{1+t^{\star}_{ni,k}}{1+t_{ni,k}};\quad \widehat{1+x_{in,k}} = \frac{1+x^{\star}_{in,k}}{1+x_{in,k}};\quad \widehat{1+s_{n,k}} = \frac{\gamma_k/(\gamma_k-1)}{1+s_{n,k}};\quad \hat{\tau}_{n,k} = \hat{\tilde{P}}_n/\tau^0_{n,k} & \text{c) in changes} \end{cases} \tag{28}$$

The change in the variety-level producer prices and the corresponding change in the CES and Cobb-Douglas consumer price indexes can be described as:

$$\begin{cases} \hat{P}_{ni,k} = \hat{w}_n(\hat{\rho}_{n,k})^{\frac{1}{1-\gamma_k}}(\widehat{1-a_{n,k}})^{-1} & \text{a) producer price } (ni,k) \\ \hat{\tilde{P}}_{ni,k} = \frac{(\widehat{1+t_{ni,k}})}{(\widehat{1+x_{ni,k}})(\widehat{1+s_{n,k}})}\hat{P}_{ni,k} & \text{b) consumer price } (ni,k) \\ \hat{\tilde{P}}_{i,k} = \left[\sum_{n=1}^{N}\lambda_{ni,k}(\hat{\tilde{P}}_{ni,k})^{1-\sigma_k}\right]^{\frac{1}{1-\sigma_k}} & \text{c) consumer price } (i,k) \\ \hat{\tilde{P}}_i = \prod_k(\hat{\tilde{P}}_{i,k})^{e_{i,k}} & \text{d) consumer price } (i) \end{cases} \tag{29}$$

Note that $\hat{P}_{ni,k}$, in the above equation, encompasses changes in wage, $\hat{w}_n$, scale, $\hat{\rho}_{n,k}$, and abatement, $(\widehat{1-a_{n,k}})$, which are each governed by equilibrium conditions. Given the change in consumer prices, the change in within-industry expenditure and revenue shares can be expresses as

$$\begin{cases} \hat{\lambda}_{ni,k} = \left(\hat{\tilde{P}}_{ni,k}/\hat{\tilde{P}}_{i,k}\right)^{1-\sigma_k} & \text{a) within-ind exp share } (ni,k) \\ \hat{r}_{ni,k} = \frac{(\widehat{1+t_{ni,k}})^{-1}(\widehat{1+x_{ni,k}})\hat{\lambda}_{ni,k}\hat{Y}_i}{\sum_\ell r_{j\ell,k}(\widehat{1+t_{n\ell,k}})^{-1}(\widehat{1+x_{n\ell,k}})\hat{\lambda}_{n\ell,k}\hat{Y}_\ell} & \text{b) within-ind rev share } (ni,k) \end{cases} \tag{30}$$

The change in industry-level output, carbon emission, carbon intensity, and abatement are given by:

$$\begin{cases} \hat{Q}_{n,k} = (\hat{\rho}_{n,k})^{1+\frac{1}{\gamma_k-1}}(\widehat{1-a_{n,k}}) & \text{a) output quantity of country-industry } (n,k) \\ \hat{Z}_{n,k} = (\widehat{1-a_{n,k}})^{\frac{1}{\alpha_k}+\frac{1}{\gamma_k}-1}\hat{Q}_{n,k}^{1-\frac{1}{\gamma_k}} & \text{b) emission from country-industry } (n,k) \\ \widehat{1-a_{n,k}} = (\hat{w}_n/\hat{\tau}_{n,k})^{\alpha_k} & \text{d) abatement in country-industry } (n,k) \\ \hat{v}_{n,k} = 1/\hat{\tau}_{n,k} & \text{c) emission intensity of country-industry } (n,k) \end{cases} \tag{31}$$

The expression for $\hat{Q}_{n,k}$ in the above equation derives from the equilibrium price equation, the free entry condition, and the definition of $\rho_{n,k} \equiv L_{n,k}/\bar{L}_n$.[20] The expressions for $\hat{Z}_{n,k}$, $\widehat{1-a_{n,k}}$, and $\hat{v}_{n,k}$

[20]Labor market clearing for country-industry (n,k) requires that: $w_n\rho_{n,k}\bar{L}_n = (1-\alpha_k\frac{\gamma_k-1}{\gamma_k})P_{nn,k}Q_{n,k}$. Replacing for $P_{nn,k}$ from Equation 2, we have: $w_n\rho_{n,k}\bar{L}_n = \left(1-\alpha_k\frac{\gamma_k-1}{\gamma_k}\right)\bar{p}_{nn,k}w_n\left(\frac{Q_{n,k}}{1-a_{n,k}}\right)^{1-1/\gamma_k}$. Applying hat algebra delivers Equation

respectively derive from applying the hat-algebra notation to Equations 3, 6, and 12. The change in wages and industry-level labor shares are governed by the labor market clearing condition expressed in changes:

$$\begin{cases} \hat{w}_n \hat{\rho}_{n,k} \rho_{n,k} w_n \bar{L}_n = \sum_j \left[\frac{(1-\alpha_k \frac{\gamma_k - 1}{\gamma_k})(1+s^\star_{n,k})(1+x^\star_{nj,k})}{(1+t^\star_{nj,k})} \hat{\lambda}_{nj,k} \lambda_{nj,k} e_{j,k} \hat{Y}_j Y_j \right] & \text{a) LMC } (n,k) \\ \sum_k \hat{\rho}_{n,k} \rho_{n,k} = 1 & \text{b) sum of shares=1 } (n) \end{cases} \quad (32)$$

The first line in the above equation ensures that the industry-level wage bill equals total sales net of taxes/subsidies. The second line ensures that the industry-level labor shares add up to one in the counterfactual equilibrium (i.e., $\sum_k \rho^\star_{n,k} = 1$). Finally, the change in national income $\hat{Y}_n$ is governed by the representative consumer's budget constraint (BC) in country i, expressed in changes:

$$\hat{Y}_n Y_n = \hat{w}_n w_n \bar{L}_n + \sum_k \sum_j \left[\frac{(1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k})(1 + s^\star_{n,k})(1 + x^\star_{nj,k})}{(1 + t^\star_{nj,k})} \hat{\lambda}_{nj,k} \lambda_{nj,k} e_{j,k} \hat{Y}_j Y_j \right]$$
$$+ \sum_k \sum_j \left[\frac{\left[1 - (1 + s^\star_{n,k})(1 + x^\star_{nj,k})\right]}{(1 + t^\star_{nj,k})} \hat{\lambda}_{nj,k} \lambda_{nj,k} e_{j,k} \hat{Y}_j Y_j + \frac{t^\star_{jn,k}}{1 + t^\star_{jn,k}} \hat{\lambda}_{jn,k} \lambda_{jn,k} e_{n,k} \hat{Y}_n Y_n \right]. \quad \text{BC } (n) \quad (33)$$

The above equation ensures that total income equals the wage bill plus tax revenues. The first and second sums respectively denote the carbon and non-carbon tax revenues expressed in changes. Recall that $\mathcal{B}_v \equiv \{\lambda_{ni,k}, r_{ni,k}, \rho_{i,k}, \tilde{\delta}_{ni}, e_{n,k}, w_n \bar{L}_n, Y_n\}_{ni,k}$, $\mathcal{B}_t \equiv \{s_{n,k}, \tau^0_{n,k}, x_{in,k}, t_{ni,k}\}_{ni,k}$, and $\mathcal{B}_e = \{\sigma_k, \gamma_k, \alpha_k\}_k$, consist or either observable values or estimable parameters. Also, note that $\hat{\lambda}_{ni,k}$, $\hat{r}_{ni,k}$, $\hat{Z}_{n,k}$, $\hat{P}_{ni,k}$, $\hat{\tilde{P}}_{i,k}$, $\widehat{1 - a_{n,k}}$, and $\hat{v}_{i,k}$ are automatically determined given information on $\mathcal{B} \equiv \{\mathcal{B}_v, \mathcal{B}_t, \mathcal{B}_e\}$, $\mathcal{R}_v \equiv \{\hat{w}_n, \hat{Y}_n, \hat{\tilde{P}}_n, \hat{\rho}_{n,k}\}_{n,k}$ and $\mathcal{R}_t \equiv \{x^\star_{in,k}, t^\star_{ni,k}\}_{ni,k}$.[21] As such, Equations 28-33 constitute a system of $2N(N-1)K + NK + 3N$ *independent* equations and unknowns. The independent unknowns are the elements of $\mathcal{R} \equiv \{\mathcal{R}_v, \mathcal{R}_t\}$, which consist of $N(N-1)K$ optimal import tax rates $t^\star_{ji,k}$, $N(N-1)K$ optimal export subsidy rates $x^\star_{ji,k}$, NK changes in industry-level labor shares $\hat{\rho}_{i,k}$, and $3N$ changes in national wage rates, $\hat{w}_i$, income levels, $\hat{Y}_i$ and consumer price indexes, $\hat{\tilde{P}}_i$. Solving the system characterized by Equations 28-33 fully characterizes the change in all equilibrium values, when moving from the observed baseline to the counterfactual non-cooperative equilibrium. The following proposition summarizes this point.

(31-a).

[21] Since $s^\star_{i,k} = (\gamma_k - 1)/\gamma_k$ and $\tau^\star_i = \tilde{\delta}_{ni} \hat{\tilde{P}}_i$ we can exclude these tax rates from $\mathcal{B}_t$. That is because they are implicitly-determined with information on the remaining elements of $\mathcal{B}$ and $\mathcal{R}$.

Proposition 3. *Solving the system of Equations 28-33, determines* $\mathcal{R}_v \equiv \{\hat{w}_n, \hat{Y}_n, \hat{\tilde{P}}_n, \hat{\rho}_{n,k}\}_{ni,k}$ *and* $\mathcal{R}_t \equiv \{x^\star_{in,k}, t^\star_{ni,k}\}_{ni,k}$ *as a function of observables and structural elasticities,* $\mathcal{B}_v \equiv \{\lambda_{ni,k}, r_{ni,k}, \rho_{i,k}, \tilde{\delta}_{ni}, e_{n,k}, w_n \bar{L}_n, Y_n\}_{ni,k}$, $\mathcal{B}_t \equiv (s_{n,k}, \tau^0_{n,k}, x_{in,k}, t_{ni,k}\}_{ni,k}$, *and* $\mathcal{B}_e = \{\sigma_k, \gamma_k, \alpha_k\}_k$. *Given information on* $\mathcal{B} \equiv \{\mathcal{B}_v, \mathcal{B}_t, \mathcal{B}_e\}$ *and* $\mathcal{R} \equiv \{\mathcal{R}_v, \mathcal{R}_t\}$, *the effect of non-cooperative taxes/subsidies on welfare and carbon emissions can be calculated as*

$$\hat{W}_i = \frac{Y_i}{Y_i - \sum_{n,k} \tilde{\delta}_{ni,k} Z_{n,k}} \left(\frac{\hat{Y}_i}{\hat{\tilde{P}}_i} \right) - \sum_{n,k} \frac{\tilde{\delta}_{ni,k} Z_{n,k}}{Y_i - \sum_{n,k} \tilde{\delta}_{ni,k} Z_{n,k}} \hat{Z}_{n,k}. \tag{34}$$

To put Proposition 3 in perspective, the welfare effect specified by $\hat{W}_i$ can be used to determine the effectiveness of trade policy as an enforcement tool in climate negotiations. $\hat{W}_i$ determines what nations lose or gain from *not* joining a global climate agreement and opting for non-cooperative policy choices. To make sense of this number, we have to contrast it with how much countries gain or lose from joining a global climate agreement.

5.2 Global Climate Cooperation

Applying a similar logic, we can quantify the gains from global climate cooperation using the cooperative tax schedule presented under Equation 27. Cooperative carbon taxes are given by $\tau^*_{i,k} = \tau^*_i = \sum_{j \in \mathbb{C}} \tilde{\delta}_{ij}$. Maintaining our conservative assumption that applied carbon taxes are consistent with the unilaterally optimal rate, we can recover the disutility parameter δ_{ii} from the applied carbon tax $\tau_{i,k} = \delta_{ii}$ in each country. The change in carbon taxes when transitioning from the factual to the counterfactual cooperative equilibrium is, thus, given by

$$\hat{\tau}_{i,k} = \frac{\tau^*_{i,k}}{\tau_{i,k}} = \sum_{n=1}^{N} \left(\frac{\tilde{\delta}_{in} \hat{\tilde{P}}_n}{\tilde{\delta}_{ii}} \right).$$

Given that $1 + s^*_{j,k} = \gamma_k / (\gamma_k - 1)$ and $t_{ji,k} = x_{ji,k} = 0$, the change in non-carbon tax instruments can be expressed as

$$\widehat{1 + s_{j,k}} = \frac{\gamma_k / (\gamma_k - 1)}{1 + s_{i,k}}; \qquad \widehat{1 + x_{ji,k}} = \frac{1}{1 + x_{ji,k}}; \qquad \widehat{1 + t_{ji,k}} = \frac{1}{1 + t_{ji,k}}.$$

Solving the above equations along-side Equations 29-33 determines $\hat{\tau}_{i,k}, \hat{v}_{i,k}, \widehat{1 - a_{i,k}}, \hat{\lambda}_{ji,k}, \hat{\rho}_{i,k}, \hat{r}_{ji,k}, \hat{w}_i, \hat{Y}_i$ and $\hat{\tilde{P}}_i$ as a function of observables $\mathcal{B}_v \equiv \{\lambda_{ni,k}, r_{ni,k}, \rho_{i,k}, \tilde{\delta}_{ni}, e_{n,k}, w_n \bar{L}_n, Y_n\}_{ni,k}$, $\mathcal{B}_t \equiv (s_{n,k}, x_{in,k}, t_{ni,k}\}_{ni,k}$, and estimable parameters, $\mathcal{B}_e = \{\sigma_k, \gamma_k, \alpha_k\}_k$. With knowledge of these variables, we can immediately calculate the change in real income $\hat{V}_i = \hat{Y}_i / \tilde{P}_i$ and carbon emissions $\hat{Z}_{n,k} = \hat{v}_{n,k} \hat{\rho}_{n,k} \hat{w}_n$. Plugging these values into Equation 34 determines the change in (carbon adjusted) welfare when transitioning from the status quo to the cooperative climate policy equilibrium.

5.3 Data Sources

Trade, Production, and Emissions. Data on international emission and expenditure levels are taken from the 2009 WIOD database on Input-Output Tables and Environmental Accounts (Timmer et al. (2012)).[22] The WIOD reports the full matrix of international expenditures across 41 major countries and 35 ISIC-level industries. Since the European Union (EU) acts as one tax-imposing authority, we aggregate all EU members into one tax-imposing region. To merge the WIOD data with our other datasets, we aggregate our sample into 19 industries, the details of which are listed in Table 1. After applying these aggregations, we are left with 15 economic regions ($N = 15$) and 19 industries ($K = 19$), for which we have a full $15 \times 15 \times 19$ matrix describing expenditure levels, $\tilde{P}_{ji,k}Q_{ji,k}$, per origin j–destination i–industry k.

The WIOD environmental accounts report emissions in units of tonnes for CO2 and other gases by origin country and industry. We calculate CO2 equivalent (CO2e) emissions based on global warming potential (GWP-100) from IPCC 2014 report. The WGP-100 measures how much emissions of one tonne of a gas will be absorbed in the atmosphere in a period of 100 years relative to the emissions of one tonne of CO2. Using our data on carbon dioxide, methane, and nitrous oxide, we calculate CO2e as $Z = Z_{\text{CO2}} + 28 \times Z_{\text{CH4}} + 265 \times Z_{\text{N2O}}$. According to the Environmental Protection Agency, CO2, CH4, and N2O gases account for 97% of greenhouse gas emissions worldwide. Accordingly, we construct the carbon emission intensity for *origin i–industry k* as:

$$v_{i,k} = \frac{Z_{i,k}}{P_{ii,k}Q_{i,k}} = \frac{\text{Emission}_{i,k}}{\text{GrossOutput}_{i,k}},$$

where the numerator is measured in tonnes of CO2e, while gross output is measured in US dollars.

Applied Taxes on Trade and Emissions. We compile data on applied tariffs in year 2009 from the United Nations Statistical Division, Trade Analysis and Information System (UNCTAD-TRAINS). The 2009 version of UNCTAD-TRAINS covers 31 two-digit (in ISIC rev.3) sectors, which are aggregated up into our 19 aggregate ISIC industries for which we have compiled international expenditure and

[22] Our baseline year is 2009 as the most recent year with available information on trade & production, emission, and environmentally related taxes. Specifically, 2009 is the last year reported in WIOD Environmental Account and the first year with a large coverage in environmentally-related tax data.

emissions data. For each industry, the UNCTAD-TRAINS reports multiple industry-level measures for applied tariffs. As is standard in the quantitative trade literature, we use the "simple tariff line average" of the "effectively applied tariff" (AHS). The UNCTAD-TRAINS reports *origin–destination–industry*-specific tariffs rates for all 41 countries in the WIOD sample, except for instances where the destination country is a member of the European Union (EU). In such cases we assign applied tariff measures based on the fact that EU tariffs are reported in the UNCTAD-TRAINS, intra-EU trade is subject to zero tariffs, and all EU members impose a common external tariff on non-members. Finally, we assume that applied export and domestic subsidies are negligible in each country, i.e., $s_{i,k} \approx x_{ij,k} \approx 0$.

We obtain data on environmentally-related taxes from Eurostat and OECD. The Eurostat dataset reports environmentally-related taxes at the level of *origin country-industry.* It covers all European countries and is based on NACE rev. 2 industries, which we map to our 19 ISIC industries. The dataset also reports environmentally-related taxes paid by households. In addition to this information, we obtain national-level environmentally-related taxes from the OECD Policy Instruments database (OECD-PINE), which covers both European and non-European countries. In the OECD-PINE dataset, we observe environmentally-related taxes paid nationally without knowing whether they are paid by industries or households. The environmentally-related tax data from both Eurostat and OECD-PINE are reported for four mutually exclusive categories of energy, pollution, resources, and transport, among which we take categories of energy and transport as attributable to carbon emissions.[23]

5.4 Estimation of γ_k, ϵ_k, and α_k

To evaluate policy outcomes, we need the following elasticity parameters per industry, trade elasticity, $\epsilon_k \equiv (\sigma_k - 1)$; emission elasticity, α_k; and degree of firm-level market power, γ_k, which is tied

[23] According to Eurostat data, out of total environmentally-related taxes paid by all industries in the EU in 2009, 81.6% are for energy, 3.1% pollution, 0.7% resources, and 14.5% transport taxes. These numbers, when paid by households are 69.3% for energy, 2.5% pollution, 0.4% resources, and 27.8% transport. According to OECD data that cover non-European countries, out of total environmentally-related taxes paid in the economy (both industries and households), on average 64.1% is paid for energy, 3.1% for pollution, 3.8% for resources, and 30.5% for transport. One reason that in addition to energy taxes, we include environmentally-related transport taxes as attributable to carbon emissions is that in many countries emission taxes are partly reflected in motor vehicle sales, flight tickets, motor vehicle registrations, and so on. Environmentally-related transport taxes capture these emission-related taxes that are not obtained through taxes on fuels. However, environmentally-related taxes on pollution and resources are typically not meant to target CO2 and other greenhouse gas emissions. Specifically, pollution taxes include (1) emissions from non-greenhouse gases such as NOx or SOx that matter for acidification, and other gases that matter for ozone depletion, (2) water pollution, and (3) waste management. Tax on resources include (1) water abstraction, (2) harvesting of biological resources such as fisheries, (3) extraction of raw material such as oil and gas, (4) landscape changes and cutting of trees.

to the markup, $\mu_k \equiv \gamma_k/(\gamma_k - 1)$. We take trade elasticities from Caliendo and Parro (2015), estimate markups based on the production function approach, and calibrate emission elasticities using environmentally-related tax data.

Markups. We use firm-level COMPUSTAT data and closely follow De Loecker et al. (2020) and Baqaee and Farhi (2020). To map COMPUSTAT data to industries as defined in WIOD, we map disaggregated NAICS-level industries from COMPUSTAT to the 19 aggregate 2-digit ISIC-level industries ($k = 1, ..., K = 19$) as well as disaggregated 3-digit ISIC industries. For every industry-year, we first estimate the output elasticity with respect to variable input, based on Olley-Pakes procedure in which, in logs of real value, our dependent variable is sales, the variable input is COGS (Cost of Goods Sold), proxy variable is capital expenditure, state variable is gross capital stock, and following the usual practice in the literature, we control for a firm's share of sales within disaggregated industries (which are 3-digit ISIC in our data). The variable input is what a firm gets to choose in its cost minimization problem. The resulting estimated coefficient of log variable input gives us output elasticities $\theta_{k,t}$ for every industry-year k, t. In addition, for every firm ω in industry k at year t, variable input share is the ratio of variable input (COGS) to sales, $b_{\omega,k,t}$. Using the first order condition of the firm's cost minimization, markups are:

$$\mu_{\omega,k,t} = \frac{\theta_{kt}}{b_{\omega,k,t}}. \tag{35}$$

To obtain markups at the level of industries, we aggregate firm-level markups to the level of 19 industries, with the weight assigned to a firm being equal to within-industry firm's sales share in the three-year period of 2008, 2009, and 2010. We consider a three-year period to make our estimates not particularly sensitive to potential industry-level fluctuations in our baseline year of 2009. Our estimates of the output elasticity $\theta_{k,t}$ are on average 0.82, with 0.80 at 25th and 0.86 at 75th percentile. The variable input shares, $b_{\omega,k,t}$, which we observed in the data, are on average 0.65, with 0.45 at 25th and 0.79 at 75th percentile. The resulting firm-level markup estimates, $\mu_{\omega,k,t}$, are on average 1.58, with 1.07 at 25th and 1.84 at 75th percentile. We report our markup estimates by industries in Table 1.

Emission Elasticity. We recover the emission elasticity from our data on carbon emission taxes, based on Equation 12. Specifically, the emission elasticity is given by the markup $\times$ emission tax per

Table 1: Industry-specific Variables and Elasticities

	Industry	Emission Intensity (v)	Emission Elasticity (α)	Trade Elasticity (σ)	Markup ($\frac{\gamma}{\gamma-1}$)
1	Agriculture	1,589	0.044	8.11	1.464
2	Mining	1,372	0.025	15.72	1.529
3	Food	84	0.011	2.55	1.698
4	Textile	81	0.011	5.56	2.109
5	Wood	109	0.014	10.83	1.278
6	Paper	135	0.008	9.07	1.296
7	Refined Petroleum	376	0.015	51.08	1.178
8	Chemicals	295	0.032	4.75	2.064
9	Plastics	50	0.010	1.66	1.272
10	Nonmetallic Minerals	1,422	0.026	2.76	1.488
11	Metals	372	0.009	6.14	1.239
12	Electronics & Machinery	26	0.007	6.06	1.501
13	Motor Vehicles	30	0.006	0.69	1.211
14	Other Manufacturing	46	0.012	5	1.913
15	Electricity, Gas and Water	3,791	0.021	5	1.119
16	Construction	39	0.012	5	1.098
17	Retail and Wholesale	37	0.018	5	1.137
18	Transportation	503	0.059	5	1.011
19	Other Services	63	0.009	5	1.596

Note: This table shows for every of the 19 industries the global average of Emission Intensity in millions of US dollars per tonne of CO2e, Emission Elasticity, Trade elasticity, and Markup.

unit of emission $\times$ emission intensity,

$$\alpha_{j,k} = \frac{\gamma_k}{\gamma_k - 1} \frac{\tau_{jk} Z_{j,k}}{P_{jj,k} Q_{j,k}} = \mu_{j,k} \tau_{j,k} v_{j,k}. \tag{36}$$

We recover the industry-specific emission elasticity $\{\alpha_k\}_k$ by evaluating the right-hand side of Equation 36. We observe emission tax revenue ($\tau_{jk} Z_{j,k}$) from Eurostat, and gross output ($P_{jj,k} Q_{j,k}$) from WIOD, for $j = \text{EU}$. Together, with our markup estimates, we obtain $\alpha_k = \alpha_{EU,k}$ for every industry k in our sample. The second column in Table 1 reports these estimated values.

The above calculation pins down $\{\alpha_k\}_k$ up to a scale. The reason is that a fraction of the emission tax that affect firms' abatement is paid by households. In the EU, firms pay about half of all emission taxes, that amounts to 1.196% of value-added generated by all EU industries. In comparison, in the EU, emission taxes as percentage of GDP equals 2.401%. As such, we scale up our estimates of $\{\alpha_k\}_k$ by a factor of around two.

Despite relying on different data and identification assumptions, our calibrated emission elasticities are comparable to those estimated by Shapiro and Walker (2018). The vector of $\{\alpha_k\}_k$ across manufacturing industries is on average 0.013 in our case compared to 0.011 in theirs. Moreover, we can also recover emission elasticities for non-manufacturing industries such as agriculture (0.044), mining (0.025), electricity (0.021), and transportation (0.059) that are more emission-intensive than the average manufacturing industry.

Trade Elasticity. [To be completed] We currently employ the trade elasticity estimates produced by Caliendo and Parro (2015). But we are working on the estimation of trade elasticities by applying Caliendo and Parro's (2015) estimation technique to our 2009 data on trade values and applied tariffs.

6 Quantitative Exercises

In this section we use our theory and estimated parameters to shed light on three important trade-related questions facing climate policy. First, what is the full potential of unilateral trade policy at reducing transboundary carbon emissions, when applied by one bloc of nations such as the EU? Second, what would be the effects on global carbon emissions and welfare of nations in the Nash equilibrium where countries set their trade and carbon policies non-cooperatively? How, do these effects compare to those attainable under global cooperation? Third, how effective are trade taxes as a penalty device at forming and expanding the "climate club" system proposed by Nordhaus (2015)?

6.1 Unilateral Optimal Policy by the EU

The failure of international climate agreements has motivated some experts to advocate the unilateral use of trade policy to reduce transboundary carbon emissions. To evaluate this proposal, we study how much carbon-reducing components of unilateral trade policy can reduce global carbon emission, and at what welfare loss (net of the disutility from transboundary emissions).

To this end, we consider a counterfactual scenario wherein the EU adopts a unilaterally optimal policy according to the Cobb-Douglas-CES version of Theorem (1) and other countries remain passive. In this exercise, we set $\delta_{n,EU} = \delta_{EU,EU}$, meaning that the EU cares equally about a unit of emission inside and outside of its borders. We benchmark this exercise against a policy scenario where the only difference is that $\delta_{n,EU} = 0$ for $n \neq EU$. In this case, the carbon-reducing components of trade taxes (both in import taxes and export subsidies) are set to zero. We take the difference between these two policy scenarios as the pure effect of the carbon-reducing content of the optimal trade policy.

We find that with the border-adjustments to the carbon content of goods, the optimal tariffs set by the EU are 0-13% larger across industries with an average of 3.5%, and the optimal export subsidies are 1-11% smaller, with an average of 3.8%. These additional margins in trade taxes can be thought of as the environmental bias of the optimal trade policy. Net of the global scale effect, and beyond the ToT-driven motives, this bias of policy generates a further reallocation of the production from foreign countries into the EU region where the EU can directly tax carbon.

As a result of this bias in the EU trade taxes, the global carbon emission would be lower by 0.4%. Given that the EU is a large region and its care about carbon disutility is among the highest in our sample, the implication is that the unilateral trade policy can reduce global emissions only to a modest degree. However, this global effect comes from a nontrivial reallocation of carbon emissions across countries. The emission rises only slightly in the EU but it decreases virtually elsewhere. That is, the reallocation of production from foreign countries into the EU reduces the overall global emission despite a slight increase in the EU emission. Note that, by the structure of what the EU maximizes in each policy scenario, this reallocation does not improve the EU welfare net of its disutility from transboundary emissions. However, the loss to that portion of EU welfare is only 0.03%.

In this exercise, to single out the role of the carbon-reducing components of trade taxes when exercised unilaterally, we assumed that other countries are passive. We continue to examine the case where all countries set their taxes according to non-cooperative incentives.

6.2 Non-cooperative Nash Equilibrium

In this section, we examine the extent to which non-cooperative trade policies (when adopted simultaneously by all government) can combat global carbon emissions. To put the non-cooperative outcome in perspective, we also compare it to that of global cooperation. For this purpose, we compare the change in welfare and carbon emissions when moving from the baseline equilibrium to (1) the non-cooperative Nash equilibrium in which countries set their unilaterally optimal carbon and trade taxes (as detailed in Section 5.1), and (2) the cooperative equilibrium in which countries set globally optimum carbon taxes (as detailed in Section 5.2). We then contrast outcomes under these two counterfactual scenarios.

To run these exercises, we should take a stand on the disutility parameters $\{\delta_{ni}\}$ that concerns governments' care for local and transboundary emissions. As discussed earlier when calibrating the emission elasticity, α_k, we assume that the weight attached by each country to local emission, $\tilde{\delta}_{ii}$, is consistent with its factual carbon tax rate. Regarding transboundary emission, we assume that all countries attach a similar disutility to transboundary and local carbon emissions, i.e. $\tilde{\delta}_{ni} = \tilde{\delta}_{ii}$. for all $n \neq i$.[24]

Table 2: The consequences of climate policy when all countries care (equally) about global emission

	Increasing Returns to Scale				Constant Returns to Scale			
	Non-Cooperative		Global Cooperation		Non-Cooperative		Global Cooperation	
Country	ΔCO2	ΔW	ΔCO2	ΔW	ΔCO2	ΔW	ΔCO2	ΔW
AUS	-8.7%	-24.4%	-95.2%	67.7%	-7.4%	0.7%	-95.4%	67.1%
EU	0.0%	-32.2%	-91.1%	85.8%	-0.7%	1.9%	-91.2%	83.7%
BRA	-8.8%	-19.2%	-95.7%	73.8%	-8.3%	1.5%	-95.3%	72.2%
CAN	-17.0%	-13.4%	-94.8%	26.3%	-11.1%	-3.6%	-94.8%	24.8%
CHN	1.9%	-5.3%	-97.0%	18.8%	0.8%	0.5%	-97.1%	18.3%
IDN	-5.4%	-3.5%	-96.9%	9.9%	-3.9%	-0.8%	-96.9%	9.6%
IND	-1.2%	-2.5%	-97.8%	17.0%	-3.0%	-1.1%	-97.8%	16.2%
JPN	3.2%	-52.9%	-89.3%	137.0%	0.1%	3.7%	-89.5%	133.2%
KOR	-5.4%	-16.1%	-95.0%	34.0%	-0.5%	-1.3%	-95.2%	32.0%
MEX	-0.6%	-8.0%	-95.6%	23.1%	-3.7%	-2.9%	-95.9%	20.5%
RUS	-11.3%	-6.6%	-97.1%	15.5%	-10.1%	-3.6%	-97.2%	11.8%
TUR	-9.6%	-16.2%	-94.1%	48.9%	-6.4%	-2.5%	-93.8%	44.6%
TWN	22.2%	-20.8%	-96.4%	52.0%	-1.3%	-4.1%	-96.5%	52.4%
USA	0.1%	-13.4%	-95.4%	34.3%	-1.5%	0.3%	-95.6%	32.8%
Global	-3.1%	-16.5%	-95.5%	44.0%	-3.3%	-0.9%	-95.6%	42.4%

Table 2 reports the percentage change in carbon emissions and the corresponding welfare effects. We find that trade taxes, as a standalone device, are not notably effective at combating global carbon emissions. Under the non-cooperative equilibrium, trade policies can lower global carbon emissions by 3.1%. This number corresponds to only 3.2% of the total carbon reduction possible under globally cooperative carbon taxes (i.e., 3.1/95.5%). Furthermore, the welfare consequences of non-cooperation are also relatively bleak. The average country loses more than 16% of it carbon-adjusted real GDP

[24]Our model allows for $\tilde{\delta}_{ni} \neq \tilde{\delta}_{ii}$. However, setting $\tilde{\delta}_{ni} = \tilde{\delta}_{ii}$ enables us to identify an upper bound on the effectiveness of non-cooperative policies. When $\tilde{\delta}_{ni} < \tilde{\delta}_{ii}$, governments care less about transboundary emission by design, and set their non-cooperative policies accordingly.

under the non-cooperative Nash equilibrium. Under global climate cooperation, by comparison, the average country gains 44% in terms of carbon-adjusted real GDP.

Increasing Returns to Scale. Recall from Section 3.4 the possible tension between the terms-of-trade and carbon reduction under increasing returns to scale (IRS). Figure 1 illustrates this tension for the case of EU and the US. In both cases, optimal tariffs are significantly lower (and even negative) in the main model featuring IRS. Consider, for instance, the *Chemical* and *Mining* industries, both of which are carbon-intensive. Absent economies of scale, both the US and EU will apply a relatively high tariff on these industries to tackle transboundary emissions. These same industries, however, exhibit a high degree of scale economies. So, a high tariff in these industries deteriorates the tax-imposing country's ToT. Hence, after accounting for scale economies, the optimal tariff on these industries is slightly negative. This apparent tension perhaps explains why trade policy has a more limited effectiveness at combating carbon emission under IRS—see the first and second panels in 2.

6.3 The Effectiveness of Trade Taxes at Enforcing a Climate Club

Our previous findings indicated that non-cooperative trade and carbon taxes are remarkably less effective at reducing carbon emissions than cooperative carbon taxes. The implementation of cooperative carbon taxes is, however, complicated by the *free-riding* problem or international misalignment in climate concerns. Some governments may find the burden of cooperative carbon taxes too large to join an international climate agreement. Even countries with a high disutility for carbon emissions have an incentive to free-ride on the rest of the world's reduction in carbon emissions without undertaking proportionate abatement measures themselves.

To overcome these problems, Nordhaus (2015) proposes that climate-conscience governments form a *climate club*, and enforce climate cooperation by imposing trade penalties against non-member countries. Quantifying the full effectiveness of the climate club model, though, is a challenging task. It requires knowledge of optimal trade penalties, the computation of which is impractical with standard optimization techniques. Our optimal tax formula, by design, characterizes the extent to which trade taxes can optimally correct for transboundary carbon externalities. They also determine the trade policy schedule that inflicts the greatest terms-of-trade penalty on non-cooperative trading partners. Our quantitative approach that builds on these formulas can, thus, uncover the full effectiveness of the climate club proposal.

Here, we consider a climate club initiated by the US and EU countries. We consider the pessimistic

Figure 1: Optimal EU and US Tariffs

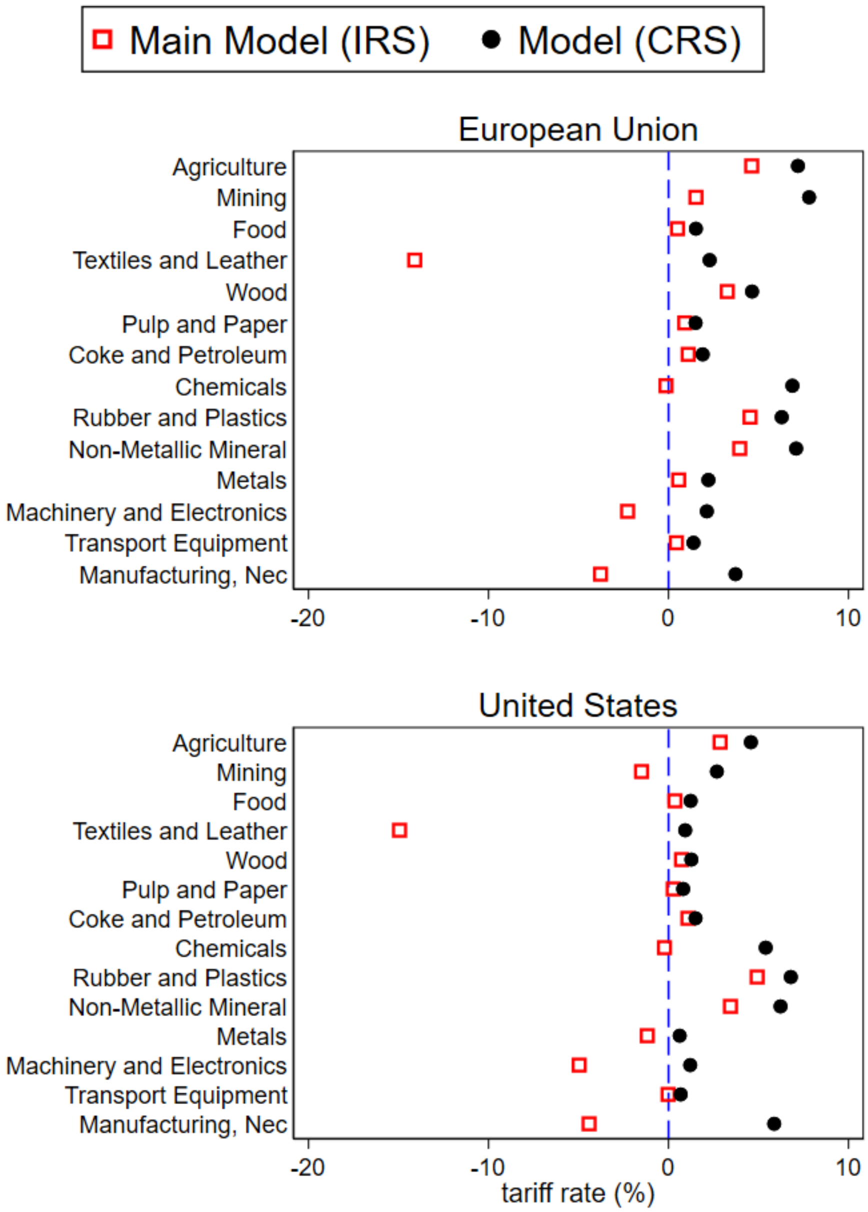
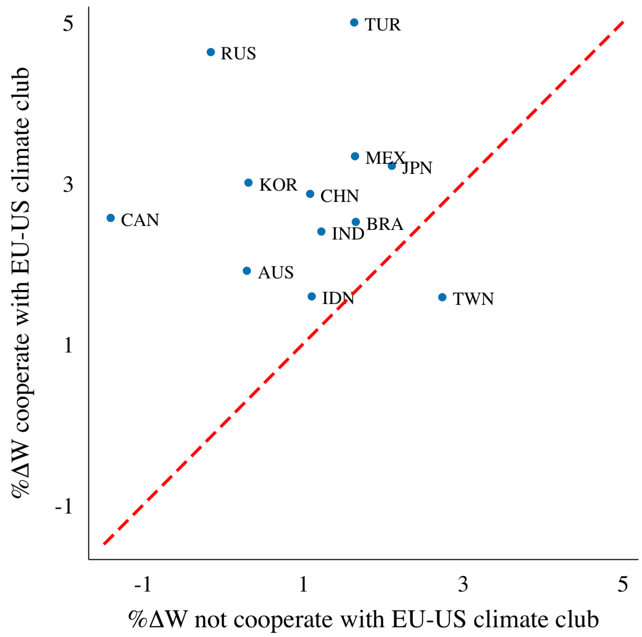


Figure 2: The effectiveness of a US-EU climate club



scenario where countries outside of the US and EU care about local emissions but not transboundary emissions, i.e., for $n \neq \{\text{US}, \text{EU}\}, \delta_{ni} = 0$ if $i \neq n$. In this scenario, the sole rationale for joining the US-EU climate club is to avoid the non-cooperative trade taxes targeted at non-members.

Starting at EU-US club, we compare the welfare gains (or losses) to each non-member country if it joins the club, expecting that other non-member countries stay outside of the club.[25] In each case, countries in the club set their taxes toward each other cooperatively, meaning no trade taxes against each other and emission taxes that correct for their emission externality on each other, whereas otherwise trade and emissions taxes are set non-cooperatively.

The results displayed in Figure 2 suggest that nearly all countries are better off from cooperating with the US-EU climate club. The appeal of joining the climate club is partially reduced when we account for firm-delocation for carbon tax avoidance. Nonetheless, with the exception of Taiwan all countries are better off joining the US-EU climate club.

Altogether, our results indicate that while trade taxes are not effective as stand-alone climate remedy, they are effective at enforcing climate cooperation. Furthermore, the effectiveness of trade taxes as an enforcement tool is not undermined once we account for firm-delocation in repossess to

[25] We have also run a similar exercise in which we compare the gains to non-member countries if all of them join the club. In this case, gains are sufficiently large to form a club of all nations.

strict carbon taxes. Nearly all countries are better off by joining the stricter (globally optimal) carbon taxes than being subjected to non-cooperative trade taxes.

7 Conclusion

[To be Added]

References

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

Baqaee, D. R. and Farhi, E.: 2020, Productivity and misallocation in general equilibrium, *The Quarterly Journal of Economics* **135**(1), 105–163.

Bartelme, D. G., Costinot, A., Donaldson, D. and Rodríguez-Clare, A.: 2019, The textbook case for industrial policy: Theory meets data, *Technical report*, National Bureau of Economic Research.

Beshkar, M. and Lashkaripour, A.: 2020, The cost of dissolving the wto: The role of global value chains.

Böhringer, C., Carbone, J. C. and Rutherford, T. F.: 2016, The strategic value of carbon tariffs, *American Economic Journal: Economic Policy* **8**(1), 28–51.

Caliendo, L. and Parro, F.: 2015, Estimates of the trade and welfare effects of nafta, *The Review of Economic Studies* **82**(1), 1–44.

Cherniwchan, J., Copeland, B. R. and Taylor, M. S.: 2017, Trade and the environment: New methods, measurements, and results, *Annual Review of Economics* **9**, 59–85.

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

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

Costinot, A., Donaldson, D., Vogel, J. and Werning, I.: 2015, Comparative advantage and optimal trade policy, *The Quarterly Journal of Economics* **130**(2), 659–702.

Costinot, A. and Rodríguez-Clare, A.: 2014, Trade theory with numbers: Quantifying the consequences of globalization, *Handbook of international economics*, Vol. 4, Elsevier, pp. 197–261.

Costinot, A., Rodríguez-Clare, A. and Werning, I.: 2016, Micro to macro: Optimal trade policy with firm heterogeneity, *Technical report*, National Bureau of Economic Research.

De Loecker, J., Eeckhout, J. and Unger, G.: 2020, The rise of market power and the macroeconomic implications, *The Quarterly Journal of Economics* **135**(2), 561–644.

De Loecker, J. and Warzynski, F.: 2012, Markups and firm-level export status, *American economic review* **102**(6), 2437–71.

Dornbusch, R., Fischer, S. and Samuelson, P. A.: 1977, Comparative advantage, trade, and payments in a ricardian model with a continuum of goods, *The American Economic Review* **67**(5), 823–839.

Eaton, J. and Kortum, S.: 2002, Technology, geography, and trade, *Econometrica* **70**(5), 1741–1779.

Elliott, J., Foster, I., Kortum, S., Munson, T., Perez Cervantes, F. and Weisbach, D.: 2010, Trade and carbon taxes, *American Economic Review* **100**(2), 465–69.

Horn, R. A. and Johnson, C. R.: 2012, *Matrix analysis*, Cambridge university press.

Lashkaripour, A. and Lugovskyy, V.: 2016, Profits, scale economies, and the gains from trade and industrial policy.

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

Mas-Colell, A., Whinston, M. D., Green, J. R. et al.: 1995, *Microeconomic theory*, Vol. 1, Oxford university press New York.

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

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

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

Shapiro, J. S.: 2020, The environmental bias of trade policy, *Technical report*, National Bureau of Economic Research.

Shapiro, J. S. and Walker, R.: 2018, Why is pollution from us manufacturing declining? the roles of environmental regulation, productivity, and trade, *American Economic Review* **108**(12), 3814–54.

Sturm, D. M.: 2003, Trade and the environment: A survey of the literature, *Environmental policy in an international perspective*, Springer, pp. 119–149.

Timmer, M., Erumban, A. A., Gouma, R., Los, B., Temurshoev, U., de Vries, G. J., Arto, I.-a., Genty, V. A. A., Neuwahl, F., Francois, J. et al.: 2012, The world input-output database (wiod): contents, sources and methods, *Technical report*, Institue for International and Development Economics.

Weisbach, D. and Kortum, S.: 2020, Optimal unilateral carbon policy.

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 $k \in \mathbb{K}$,

$$
\begin{aligned}
&\text{(Optimal Demand)} && Q_{ji,k} = \mathcal{D}_{ji,k}(Y_i, \tilde{\mathbf{P}}_i) \\
&\text{(Producer Price)} && P_{ji,k}(w_j, a_{j,k}; Q_{j,k}) = \bar{d}_{ji,k}\bar{p}_{jj,k}w_j(1-a_{j,k})^{\frac{1}{\gamma_k}-1}Q_{j,k}^{-\frac{1}{\gamma_k}} \\
&\text{(Pollution)} && Z_{j,k}(a_{j,k}; Q_{j,k}) \equiv \bar{z}_{j,k}(1-a_{j,k})^{\frac{1}{\alpha_k}+\frac{1}{\gamma_k}-1}Q_{j,k}^{1-\frac{1}{\gamma_k}} \\
&\text{(Income = Revneue)} && Y_i = w_i\bar{L}_i + \sum_{k,\, j\neq i}\left[\left(\tilde{P}_{ji,k} - P_{ji,k}\right)Q_{ji,k}\right] + \sum_{k,\, j}\left[\left(\tilde{P}_{ij,k} - (1-\alpha_k\frac{\gamma_k - 1}{\gamma_k})P_{ij,k}\right)Q_{ij,k}\right] + \bar{D}_i \\
&\text{(Trade Deficit)} && B_i \equiv \sum_{j\neq i}\sum_k P_{ji,k}Q_{ji,k} - \sum_{j\neq i}\sum_k \tilde{P}_{ij,k}Q_{ij,k} - \bar{D}_i = 0
\end{aligned}
$$

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 emission 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 11). 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 $k \in \mathbb{K}$,

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

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:

$$\begin{cases} f_1(\mathbb{P}_i; \mathbf{w}) \equiv w_1 L_1 - \sum_{j \in \mathbb{C}} \sum_{k \in \mathbb{K}} (1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}) P_{1j,k}(\mathbb{P}_i; \mathbf{w}) Q_{1j,k}(\mathbb{P}_i; \mathbf{w}) = 0 \\ \vdots \\ f_N(\mathbb{P}_i; \mathbf{w}) \equiv w_N L_N - \sum_{j \in \mathbb{C}} \sum_{k \in \mathbb{K}} (1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k}) P_{Nj,k}(\mathbb{P}_i; \mathbf{w}) Q_{Nj,k}(\mathbb{P}_i; \mathbf{w}) = 0 \end{cases} \tag{37}$$

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:

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

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 37. 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 37 w.r.t. $\mathbf{w}$ holding $\mathbb{P}_i$ fixed, yields

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

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

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

Noting that $r_{ij,k}\epsilon_k(1-\lambda_{jj,k}) \ll 1$ if $j \neq i$, we can produce the following approximation:[26]

$$\left(\frac{\partial \boldsymbol{f}}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i}^{-1} = (\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots \approx$$
$$\mathbf{I} + \sum_{m=1}^{\infty} \operatorname{diag}\left(\left[-\sum_{k\in\mathbb{K}} r_{ii,k}\epsilon_k(1-\lambda_{ii,k})\right]^m\right) = \operatorname{diag}\left(\left[\sum_{k\in\mathbb{K}} 1 + r_{ii,k}\epsilon_k(1-\lambda_{ii,k})\right]^{-1}\right).$$

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:

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

---

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

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

Given that (i) $\lambda_{ji,k} r_{ni} \approx 0$ if n and $j \neq i$, and (ii) $\left(\frac{\partial f}{\partial \ln \mathbf{w}}\right)_{\mathbb{P}_i}$ 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:

$$\frac{\partial \ln Z_{j,k}(a_{j,k}, Q_{j,k})}{\partial \ln Q_{j,k}} = 1 - \frac{1}{\gamma_k}, \tag{38}$$

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

$$\frac{\partial \ln Z_{j,k}(a_{j,k}, Q_{j,k})}{\partial \ln(1 - a_{j,k})} = \frac{1}{\alpha_k} + \frac{1}{\gamma_k} - 1. \tag{39}$$

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 ($j \neq i$):

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

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,

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

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

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

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 (9) 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

$$\left(\frac{\partial W_i(\mathbb{P}_i; Y_i)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(.)}{\partial \ln Y_i}\right)_{\mathbb{P}_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln Y_i}\right)_{\mathbb{P}_i}. \tag{40}$$

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

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

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

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

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:

$$\left(\frac{\partial\ln Q_{ji,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=\frac{\partial\ln\mathcal{D}_{ji,k}(.)}{\partial\ln Y_i}=\eta_{ji,k};\qquad\left(\frac{\partial\ln Q_{jn,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=0,\quad(j\neq i)$$

$$\left(\frac{\partial\ln Q_{j,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=\left(\frac{\partial\ln Q_{j,k}}{\partial\ln Q_{ji,k}}\right)_{\mathbb{P}_i}\left(\frac{\partial\ln Q_{ji,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=r_{ji,k}\eta_{ji,k}$$

$$\left(\frac{\partial\ln P_{ij,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=\left(\frac{\partial\ln P_{ii,k}(.)}{\partial\ln Q_{i,k}}\right)_{\mathbb{P}_i}\left(\frac{\partial\ln Q_{i,k}}{\partial\ln Q_{ii,k}}\right)_{\mathbb{P}_i}\left(\frac{\partial\ln Q_{ji,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=-\frac{1}{\gamma_k}r_{ii,k}\eta_{ii,k}$$

$$\left(\frac{\partial\ln P_{ji,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=\left(\frac{\partial\ln P_{jj,k}(.)}{\partial\ln Q_{ji,k}}\right)_{\tilde{P}_{ji,k}}\left(\frac{\partial\ln Q_{ji,k}(.)}{\partial\ln Y_i}\right)_{\mathbb{P}_i}=\omega_{ji,k}\eta_{ji,k}\qquad(j\neq i)$$

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

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

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:

$$\sum_{j=1}^{N}\left[\left(\frac{\partial \ln P_{ij,g}}{\partial \ln Q_{ii,g}}\right)_{\mathbb{P}_i} P_{ij,g}Q_{ij,g}\right] = \sum_{j=1}^{N}\left[\left(\frac{\partial \ln P_{ii,g}}{\partial \ln Q_{i,g}}\right)_{\mathbb{P}_i} \frac{\partial \ln Q_{i,g}(Q_{i1,g}...Q_{iN,g})}{\partial \ln Q_{ii,g}} P_{ij,g}Q_{ij,g}\right]$$
$$= -\sum_{j=1}^{N}\left(\frac{1}{\gamma_g} r_{ii,g}P_{ij,g}Q_{ij,g}\right) = -\frac{1}{\gamma_g} r_{ii,g}\sum_{j=1}^{N}\left(P_{ij,g}Q_{ij,g}\right) = -\frac{1}{\gamma_g}P_{ii,g}Q_{ii,g} \tag{41}$$

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 40, 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

$$\left(\frac{\partial W_i}{\partial \ln Y_i}\right)_{\mathbb{P}_i} = \sum_k\left[\left(\tilde{P}_{ii,k} - \frac{\gamma_k - 1}{\gamma_k}\left(1-\alpha_k\frac{\gamma_k-1}{\gamma_k} + \tilde{\delta}_{ii,k}v_{i,k}\right)P_{ii,k}\right)Q_{ii,k}\eta_{ii,k} + \sum_{j\neq i}\left(\left[\tilde{P}_{ji,k} - (1+\omega_{ji,k} - \frac{\gamma_k-1}{\gamma_k}\delta_{ji}v_{j,k})P_{ji,k}\right]Q_{ji,k}\eta_{ji,k}\right)\right]$$
$$= \sum_k\left[\left(1 - \frac{\gamma_k - 1}{\gamma_k}\left(1-\alpha_k\frac{\gamma_k-1}{\gamma_k} + \tilde{\delta}_{ii,k}v_{i,k}\right)\frac{P_{ii,k}}{\tilde{P}_{ii,k}}\right)e_{ii,k}\eta_{ii,k} + \sum_{j\neq i}\left(\left[1 - (1+\omega_{ji,k} - \frac{\gamma_k-1}{\gamma_k}\delta_{ji}v_{j,k})\frac{P_{ji,k}}{\tilde{P}_{ji,k}}\right]e_{ji,k}\eta_{ji,k}\right)\right]Y_i. \tag{42}$$

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 $P_{ii,k}Q_{i,k} = \sum_{j=1}^N P_{ij,k}Q_{ij,k}$)

$$\frac{\partial V_i(.)}{\partial \ln(1+a_{i,k})} + \frac{\partial V_i(.)}{\partial Y_i} \left(\frac{\partial T_i(\mathbb{P}_i;Y_i)}{\partial \ln(1+a_{i,k})}\right)_{\mathbf{w},Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i;Y_i)}{\partial \ln(1+a_{i,k})}\right)_{\mathbf{w},Y_i}$$
$$= \frac{\partial V_i(.)}{\partial Y_i} \left[1-\alpha_k \frac{\gamma_k-1}{\gamma_k}\right] \frac{\gamma_k-1}{\gamma_k} P_{ii,k} Q_{i,k} - \frac{1}{\alpha_k} \delta_{ii} Z_{i,k} \left[1-\alpha_k \frac{\gamma_k-1}{\gamma_k}\right] = 0. \tag{43}$$

[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:

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

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 41. Accordingly, we can express the different elements in Equation

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

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

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

$$\sum_g \left[\sum_{j \neq i} \left(1 - \left(1 + \omega_{ji,g} + \tilde{\delta}_{ji} v_{j,g} \frac{\gamma_g - 1}{\gamma_g} \right) \frac{P_{ji,g}}{\tilde{P}_{ji,g}} \right) e_{ji,g} \varepsilon_{ji,g}^{(ni,k)} \right] Y_i + \sum_g \left[\left(1 - \left(1 - \alpha_g \frac{\gamma_g - 1}{\gamma_g} + \tilde{\delta}_{ii} v_{i,g} \right) \frac{\gamma_g - 1}{\gamma_g} \frac{P_{ii,g}}{\tilde{P}_{ii,g}} \right) e_{ii,g} \varepsilon_{ii,g}^{(ni,k)} \right] Y_i = \tag{44}$$

For Equation 43 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 44 yields

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

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.

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

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

$$\frac{\tilde{P}^{\star}_{ji,k}}{P_{ji,1}} = 1 + \omega_{ji,k} + \tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k}; \qquad \frac{\tilde{P}^{\star}_{ii,k}}{P_{ii,k}} = \frac{\gamma_k - 1}{\gamma_k}, \tag{45}$$

which when plugged into Equation 42, 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 and 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 45. Combining these expressions, we have:

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

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

$$\frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ij,k}} + \frac{\partial V_i(.)}{\partial Y_i} \left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}} \right)_{\mathbf{w}, Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln \tilde{P}_{ij,k}} \right)_{\mathbf{w}, Y_i} = 0. \tag{46}$$

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 41 (Appendix B.2) and the following relationship:

$$\left(\frac{\partial \ln P_{ni,g}}{\partial \ln Q_{nj,g}}\right)_{\mathbf{w},Y_i} P_{ni,g}Q_{ni,g} = \left(\frac{\partial \ln P_{nn,g}}{\partial \ln Q_{nj,g}}\right)_{\mathbf{w},Y_i} P_{ni,g}Q_{ni,g},$$

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

Doing so yields the following equation:

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

Likewise the last term in Equation 46 (that accounts for emission effects) can be specified as

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

Plugging the above expressions back into Equation 46 (and dividing everything by $\frac{\partial V_i(.)}{\partial Y_i}\tilde{P}_{ij,k}Q_{ij,k}$) yields the following optimality condition:

$$\left\{\frac{\partial V_i(.)}{\partial \ln \tilde{P}_{ij,k}} + \frac{\partial V_i(.)}{\partial Y_i}\left(\frac{\partial T_i(\mathbb{P}_i; Y_i, \mathbf{w})}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i} - \left(\frac{\partial \delta_i \cdot \mathbf{Z}(\mathbb{P}_i; Y_i)}{\partial \ln \tilde{P}_{ij,k}}\right)_{\mathbf{w},Y_i}\right\} \left[\frac{\partial V_i(.)}{\partial Y_i} P_{ij,k}Q_{ij,k}\right]^{-1} =$$

$$1 + \sum_g \left[\left(1 - \left(1 - \alpha_g \frac{\gamma_g - 1}{\gamma_g} + \tilde{\delta}_{ii} v_{i,g}\right) \frac{\gamma_g - 1}{\gamma_g} \frac{P_{ij,g}}{\tilde{P}_{ij,g}}\right) \frac{e_{ij,g}}{e_{ij,k}} \varepsilon^{(ij,k)}_{ij,g}\right] - \sum_{n \neq i} \sum_g \left[\left(\omega_{ni,g} + \tilde{\delta}_{ni} z_{n,k} \frac{\gamma_k - 1}{\gamma_k}\right) \frac{e_{nj,g}}{e_{ij,k}} \varepsilon^{(ij,k)}_{nj,g}\right] = 0$$

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

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

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

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

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 and Johnson (2012)) and we can compute the vector χ_{ij} as

$$\chi_{ij} = \left[\frac{e_{ij,g} \varepsilon_{ij,g}^{(ij,k)}}{e_{ij,k} \varepsilon_{ij,g}}\right]_{k,g}^{-1} \left(\mathbf{1}_K + \left[\frac{\sum_{n \neq i} t_{ni,g}^{*} e_{nj,g} \varepsilon_{nj,g}^{(ij,k)}}{\sum_{n \neq i} \sum_g e_{nj,g} \varepsilon_{nj,g}^{(ij,k)}}\right]_k\right). \tag{47}$$

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

$$\begin{cases} 1 + s_{i,k}^{\star} = \frac{\gamma_k}{\gamma_k - 1}; \qquad \tau_{i,k}^{\star} = \tilde{\delta}_{ii} \\ 1 + t_{ji,k}^{\star} = (1 + \omega_{ni,k}) + \tilde{\delta}_{ni} \left(\frac{\gamma_k - 1}{\gamma_k}\right) v_{n,k} \\ 1 + x_{ij,k}^{\star} = \left(1 + \frac{1}{\varepsilon_{ij,k}}\right) \chi_{ij,k} \end{cases}, \tag{48}$$

where $\chi_{ij,k}$ is given by Equation 47.

C Optimal Emission Policy when Other Taxes are Banned

The following definition puts the case for second-best scenarios 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).

We consider three cases: (1) Emission taxes are unavailable, (2) Export subsidies are unavailable, (3) All taxes but emission taxes are unavailable.

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 emission intensity à la Markusen (1975) (See footnote (8) 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,

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

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 in cases where the government is not afforded choices of emission taxes. Suppose $v_{i,k}$ is the emission intensity under some emission tax that is different from the unilaterally optimal. This might be either because emission is unabated, or quite the opposite, because home country has set its emission tax in line with international agreements at a higher level compared to the unilaterally first best. In either case, production taxes must correct emission externalities that are too little or too much from the unilateral point of view:

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

where $v_k^{\star}$ is the emission intensity attainable under the first-best unilateral policy schedule. Consider a country with sub-optimal emission. In that case, production subsidies/taxes reflect a trade off between promoting scale economies and reducing emissions. More interestingly, consider a country whose $v_{i,k}$ is smaller than $v_{i,k}^{\star}$ because the country is abiding with an international climate agreement. In that case, $(1+s_{i,k}^{\star\star})$ includes an extra subsidy that promotes domestic production.

Case #2: Export Taxes are unavailable In this case, the optimal emission tax remains uniform and follows the same rule as the first-best. Derivations for this case are similar to the ones in the build up to equation (21) in Theorem 1. Specifically, in lemma 4, we consider only the first two sets of FOCs related to abatement and prices faced by home consumers (both domestic purchases and imports). Th resulting optimal tax schedule is given by:

$$\begin{cases} 1+t^{\star}_{ji,k} = (1+\bar{t}_i)(1+\omega_{ji,k}) + \tilde{\delta}_{ji} v_{j,k} \frac{\gamma_k - 1}{\gamma_k} & \forall j,k \\ 1+s^{\star}_{i,k} = \frac{\gamma_k}{\gamma_k - 1} & \forall k \\ \tau^{\star}_{i,k} = \tau^{\star}_i = \tilde{\delta}_{ii} & \forall k \end{cases} \tag{49}$$

Here, $\bar{t}_i \equiv \left(\frac{\partial \ln Y_i}{\partial \ln w_i}\right) \Big/ \left(\frac{\partial B_i}{\partial \ln w_i}\right)$ where B_i is the trade balance condition. $\bar{t}_i$ is zero only when export taxes are available.

Case #2: Emission Taxes are used as protection in disguise 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 paramet1erization given by equation (24). Then, as shown below, the optimal emission tax is given by

$$\tau^*_{i,k} = \left(\frac{\alpha(1-\sigma_k)\,(1-\lambda_{ii,k} r_{ii,k}) + 1}{\widetilde{\alpha}_i (1-\sigma_k)\,(1-\lambda_{ii,k} r_{ii,k}) + r_{ii,k}} \right) \tilde{\delta}_{ii} \qquad \text{(only } \tau \text{ available)} \tag{50}$$

where $\tilde{\alpha}_i > \alpha$ is a country-wide term that depends on the industry-composition of country i's production. The above formula suggests that it is optimal to tax emission above the first-best level in low-σ industries. We continue to show the derivation of equation (50).

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 $Y_i = w_i L_i + \sum_k \alpha_k P_{ii,k} Q_{i,k}$, yields the following formulation of the F.O.C.

$$P_{ii,k} Q_{ii,k} - \alpha_k \sum_j \left[P_{ij,k} Q_{ij,k} \left(1 + \varepsilon_{ij,k} \right) \right] + \tilde{\delta}_{ii} v_{i,k} \sum_j \left[P_{ij,k} Q_{ij,k} \varepsilon_{ij,k} \right] - \left(\frac{1}{\alpha_k} - 1 \right) \tilde{\delta}_{ii} v_{i,k} P_{ii,k} Q_{i,k} - \sum_g \left(\left[\alpha_g - \tilde{\delta}_{ii} v_{i,g} \right] P_{ii,g} Q_{ii,g} \right) \frac{\partial \ln Y_i}{\partial \ln(1 - a_{i,k})} - \bar{\Delta}_i \left[\sum_{j \neq i} \left[P_{ji,k} Q_{ji,k} \varepsilon^{ii}_{ji,k} - P_{ij,k} Q_{ij,k} \left(1 + \varepsilon_{ij,k} \right) \right] + \sum_{j \neq i} \sum_g \left(P_{ji,g} Q_{ji,g} \right) \frac{\partial \ln Y_i}{\partial \ln(1 - a_{i,k})} \right] = 0. \tag{51}$$

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 51 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.

$$r_{ii,k} - \alpha_k \mathcal{E}_{i,k} + \alpha_k \frac{\tilde{\delta}_{ii}}{\tau_{i,k}} \left(\mathcal{E}_{i,k} - 1 \right) - \left(1 - \alpha_k \right) \frac{\tilde{\delta}_{ii}}{\tau_{i,k}} + \sum_g \left(\alpha_g \left[1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}} \right] r_{ii,g} r_{i,g} \right) \frac{\partial \ln Y_i}{\partial \ln(1 - a_{i,k})} r^{-1}_{i,k} - \bar{\Delta}_i \left[\mathcal{E}_{i,k} + \left(1 - \lambda_{ii} \right) \frac{\partial \ln Y_i}{\partial \ln(1 - a_{i,k})} r^{-1}_{i,k} \right] = 0. \tag{52}$$

$\partial Y_i / \partial \ln(1 - a_{i,k})$, 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

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

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 52), which implies

$$\begin{aligned}
\bar{\Delta}_i &= \frac{(1-\bar{\alpha}_i) - \lambda_{ii} + \sum_g \left(\left[r_{ii,g} - (1-\alpha_g) \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{i,g}\right) + \sum_g \left(\alpha_g \left[1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{ii,g} r_{i,g}\right) \left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_g \frac{\partial \ln Y_i}{\partial \ln(1-a_{i,g})}\right]}{(1-\lambda_{ii}) \left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_k \frac{\partial Y_i}{\partial \ln(1-a_{i,k})}\right]} \\
&= \frac{\sum_g \left[(1-\alpha_g)\left(1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right) r_{i,g}\right] + \sum_g \left(\alpha_g \left[1 - \frac{\tilde{\delta}_{ii}}{\tau_{i,g}}\right] r_{ii,g} r_{i,g}\right) \left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_g \frac{\partial \ln Y_i}{\partial \ln(1-a_{i,g})}\right]}{(1-\lambda_{ii}) \left[\frac{\partial Y_i}{\partial \ln w_i} + \sum_k \frac{\partial Y_i}{\partial \ln(1-a_{i,k})}\right]} \qquad (53)
\end{aligned}$$

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

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

Combining the above expression with Equation 53 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

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

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

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

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

D Empirical Exercises

D.1 Constructing $\mathcal{B}_v$ Based on Raw Data and Estimated Parameters

Given data on expenditure, emission, and applied tariff levels $\left\{\tilde{P}_{ji,k}Q_{ji,k}, Z_{i,k}, t_{ji,k}\right\}_{ji,k}$, and estimated parameters, $\{\gamma_k, \sigma_k, \alpha_k\}$, we can construct the vector of observables, $\mathcal{B}_v \equiv \{\lambda_{ni,k}, r_{ni,k}, \rho_{i,k}, \tilde{\tilde{\delta}}_{ni}, e_{n,k}, w_n\bar{L}_n, Y_n\}_{ni,k}$, needed to implement Proposition 3. Values for emission intensities, $v_{i,k}$; and the unit cost of emission on welfare, $\tilde{\tilde{\delta}}_{ni}$, can be calculated as follows

$$v_{i,k} = \frac{Z_{i,k}}{\sum \frac{1}{1+t_{in,k}} \tilde{P}_{in,k} Q_{in,k}}, \qquad \tilde{\tilde{\delta}}_{ni,k} = \tau_{ni,k} = \frac{\alpha_k}{v_{i,k}}$$

Total national expenditure Y_i and expenditure share variables, $\lambda_{ji,k}$, and $e_{i,k}$ can be recovered from variety-level expenditure and tariff data as follows:

$$Y_i = \sum_{j=1}^{15} \sum_{k=1}^{19} \tilde{P}_{ji,k} Q_{ji,k}, \qquad \lambda_{ji,k} = \frac{\tilde{P}_{ji,k} Q_{ji,k}}{\sum_{n=1}^{15} \tilde{P}_{ni,k} Q_{ni,k}}, \qquad e_{i,k} = \frac{\sum_{n=1}^{15} \tilde{P}_{ni,k} Q_{ni,k}}{Y_i}.$$

Finally, the national wage bill, $w_i\bar{L}_i$, industry-level labor shares, $\rho_{i,k}$, and revenue shares, $r_{ji,k}$, can be constructed as follows, given variety-level expenditure and tariff data and the estimated structural parameters:

$$w_i \bar{L}_i = \sum_{j=1}^{15} \sum_{k=1}^{19} \left[\left(1 - \alpha_k \frac{\gamma_k - 1}{\gamma_k} \right) \frac{1}{1+t_{ij,k}} \tilde{P}_{ij,k} Q_{ij,k} \right];$$

$$\rho_{i,k} = \frac{\sum_{j=1}^{15}\left(1-\alpha_k\frac{\gamma_k-1}{\gamma_k}\right)\frac{\tilde{P}_{ij,k}Q_{ij,k}}{1+t_{ij,k}}}{w_i\bar{L}_i}; \qquad r_{ji,k} = \frac{\frac{1}{1+t_{in,k}}\tilde{P}_{in,k}Q_{in,k}}{\sum\frac{1}{1+t_{in,k}}\tilde{P}_{in,k}Q_{in,k}}.$$