

The Cost of Dissolving the WTO: *The Role of Global Production Networks*

Mostafa Beshkar Ahmad Lashkaripour
Indiana University Indiana University

April 2020

Abstract

As trade agreements face renewed pressure, we show that the rise of global production networks has multiplied the value of trade agreements to unprecedented levels. We cast our argument using a non-parametric neoclassical trade model that accommodates global production networks and nests a wide class of quantitative trade models as a special case. To guide our analysis, we derive analytic formulas for optimal non-cooperative trade taxes in this general framework. These formulas govern the extent of trade restriction if global trade agreements were to dissolve. Mapping these formulas to data, we quantify the value of trade agreements for various countries. We find that the disintegration of existing trade agreements will erase 30% of the overall gains from trade, which amounts to a \$2.7 trillion loss in global GDP. Around 41% of this value is driven by the agreements' promotion of global production networks.

1 Introduction

Over the past few years, the tides have slowly turned against the WTO and other free trade agreements (Irwin (2020)). The rebirth of American unilateralism and concerns over the erosion of national sovereignty have put the WTO and other trade agreements at a critical cross road.¹

¹See <https://www.cfr.org/backgrounder/whats-next-wto> for a synthesis of common criticisms facing the WTO and other free trade agreements.

Experts have reacted to these developments with two opposing views: (1) One opinion is that trade agreements are more crucial than ever because a return to protectionism will disrupt global production networks. (2) A somewhat contrary opinion is that, given their increased reliance on global production networks, countries have little incentive to restrict foreign trade. So, it is conceivable that sovereign governments would maintain the free flow of goods in and out of their country without a need for existing agreements.

While these opinions gain traction in political circles, we lack the necessary evidence to assess them. On one hand, the trade literature has directed most of its attention to measuring the gains from trade relative to *autarky*. These measures do not speak to the value of free trade agreements, since a collapse of existing agreements will not necessarily lead to autarky (Costinot and Rodríguez-Clare (2014)). On the other hand, existing studies that focus on the pure gains from trade agreements overlook the role of global production networks.²

This paper aims to fill this gap. We present a general theory that describes how the globalization of production networks has altered governments' incentives *vis-à-vis* trade restrictions. We map our theory to data, to determine the counterfactual level of trade restriction and the efficiency loss that will occur without existing agreements. We find that the disintegration of existing trade agreements will erase 30% of the overall gains from trade, which amounts to a \$2.7 trillion loss in global GDP. Around 41% of this value is driven by trade agreements facilitating the globalization of production networks.

Our analysis builds on Bagwell and Staiger's (1999) insight that trade agreements solely "*remedy the inefficient terms-of-trade-driven restrictions in trade that arise when trade policies are set unilaterally.*" Based on this insight, the counterfactual level of trade restriction (without trade agreements) can be determined using the optimal non-cooperative trade tax schedule in each country. Such a tax schedule describes the best-response of each country to the vector of trade taxes chosen by its trading partners. Solving the best-response functions simultaneously for all countries, determines the *terms-of-trade-driven* trade restrictions that will occur in the post-agreement Nash equilibrium.

²Ossa (2014) provides the most sophisticated analysis of the gains from trade agreements to date. His analysis presents a notable advance over prior analyses of the "gains from trade agreements" but it abstracts from global production networks.

Unfortunately, the existing literature does not provide a formal characterization of optimal non-cooperative trade taxes in the presence of global production networks. Existing characterizations of optimal trade taxes are either derived within partial equilibrium frameworks or abstract from global production networks.

To bridge this gap, we first present an analytical characterization of optimal trade policy in a general equilibrium non-parametric neoclassical trade model that accommodates (i) an arbitrary number of industries, (ii) an arbitrary demand structure, (iii) an arbitrary global input-output network, and (vi) diseconomies of scale due to industry-specific factors of production. Our general model nests an important class of quantitative trade models, including the widely-used multi-industry gravity model with input-output linkages in [Caliendo and Parro \(2014\)](#).

Our characterization delivers simple sufficient statistics optimal tax formulas that depend on only (i) industry-level trade elasticities, (ii) share of industry-specific capital in production, and (iii) observable expenditures, revenue, and input-output shares. These formulas highlight three stark patterns:

- (a) Absent diseconomies of scale, the optimal non-cooperative policy consists of (i) uniform or zero import taxes, and (ii) non-zero export taxes that are decreasing in the industry-level trade elasticity and level of upstreamness.³
- (b) But if export taxes cannot be used due to institutional constraints, import taxes can be employed as a *second-best* instrument to manipulate export market power through the global production network.
- (c) Optimal second-best import taxes are targeted primarily at imported intermediates, which are subsequently used in the production of highly-differentiated exports.⁴ They are uniform across all final and intermediate

³The redundancy of import taxes drives from the *targeting principle* and the *Lerner symmetry*. Absent diseconomies of scale, Home can only manipulate its import market power on an economy-wide basis. Doing so asks for a uniform tariff that (by the Lerner symmetry) can be perfectly mimicked with a uniform shift in all export taxes. By the targeting principle, import taxes are not the first-best instrument for manipulating export market power either. So, altogether, they can be discarded. Export taxes, meanwhile, are the first-best instrument for manipulating industry-level export market power. Optimal export taxes are decreasing in the industry-level trade elasticity, which governs the degree of export market power. They are relatively lower on upstream industries that supply intermediate inputs, because an export tax on intermediate inputs is partially passed back to domestic consumers. Finally, the absence of diseconomies of scale is akin to the complete tariff passthrough estimated by [Fajgelbaum et al. \(2020\)](#) and [Amiti et al. \(2019\)](#).

⁴To elaborate, import taxes on intermediates are partially passed on to foreign consumers

goods that are never employed in exporting industries.

The above tax structure optimizes a non-cooperative government's ability to correct two *unilateral* terms-of-trade-related inefficiencies: (i) unexploited import market power, which concerns the unexploited ability of the Home country to charge a *mark-down* on the price of imported goods; and (ii) unexploited export market power, which concerns the unexploited ability of the Home country to charge a *mark-up* on the price of exported goods. Correcting these two inefficiencies is unilaterally optimal, but is detrimental to global efficiency and has beggar-thy-neighbor effects on the rest of the world.

Our theory indicates that the globalization of production networks has tilted the terms-of-trade-driven motives for taxation in two opposite directions: On one hand, global production networks prompt governments to set a lower export tax on upstream industries, because a fraction of this tax is passed back to domestic consumers. On the other hand, if governments cannot use export taxes, global production networks prompt them to set a higher import tariff on upstream industries, because these tariffs are partially passed on to Foreign consumers.⁵ In that case, import tariffs can serve as a *second-best* export policy measure. The empirical significance of these two effects depends on the global input-output structure and the size of the tax-imposing country compared to the rest of the world.

In Section 6 we map our analytic formulas to data to measure both the consequences of non-cooperative trade policies and the value of the agreements that remedy them. To this end, we employ data on applied tariffs, expenditure, production, and input-output shares across 16 industries and 42 major economies. Our data includes all 27 members of the European Union, Australia, Brazil, China, Indonesia, India, Japan, Korea, Mexico, Norway, Russia, Switzerland, Taiwan, Turkey, the United States, and an aggregate of the rest of the world.

With the aid of our sufficient statics tax formulas, the task of computing the value of trade agreements reduces to solving a system of equations that requires data on only (i) industry-level trade elasticities, (ii) observable expendi-

through the global production network. The more-intensively an imported intermediate is used in the production of differentiated export goods, the higher the incidence of its tax on Foreign consumers. The trade-off, though, is that taxing intermediate inputs to mimic export taxes, distorts domestic production in a way that is avoidable with first-best export taxes.

⁵Bown, Conconi, Erbahar, and Trimarchi (2020) confirm these cascading effects empirically, showing that tariffs in upstream sectors have large effects on prices, employment, sales, and investment in downstream sectors.

ture/revenue shares, and (iii) global input-output shares. Our quantitative analysis delivers four basic results:

- i. The globalization of production networks has only modestly reduced the *terms-of-trade-driven* incentives to restrict trade. Correspondingly, due to cascading effects, the gains from non-cooperative trade taxation and the externality inflicted by these taxes on the rest of the world have *doubled* in the presence of global production networks.
- ii. Import taxes are significantly less effective at manipulating the terms-of-trade compared to first-best export taxes. They can only replicate 47% of the unilateral gains attainable under first-best export tax schedule. The ineffectiveness of second-best import taxes is driven by an innate trade-off between preserving allocative efficiency in the local economy and exploiting export market power through the production network. The former discourages taxing differentiated intermediate inputs, while the latter requires it.
- iii. Existing trade agreements contribute \$2.7 trillion to global GDP by remedying the inefficient terms-of-trade-driven restrictions in trade. To put this figure in perspective, it amounts to 30% of the total gains from trade relative to autarky and equals the GDP of France.
- iv. Around 41% of the above value drives from the agreements' promotion of global production networks. So, as the tides are starting to turn against free trade agreements, their value has multiplied from the rise of global production networks.

Related Literature. To the best of our knowledge, this paper is the first to analytically characterize the (unilaterally) optimal trade policy in an important class of general equilibrium quantitative trade models with global production networks. Several studies have adopted similar models to study the effect of global production networks on the gains from an *exogenous* tariff liberalization (e.g., [Costinot and Rodríguez-Clare \(2014\)](#); [Caliendo and Parro \(2014\)](#); [Caliendo, Feenstra, Romalis, and Taylor \(2015\)](#); and [Baqaee and Farhi \(2019\)](#)). But we are not aware of any prior characterization of *optimal* trade policy in these canonical models.⁶

⁶There is an old literature that characterizes the optimal trade tax on intermediate inputs (e.g.,

Our characterization of optimal trade policy is related to [Blanchard, Bown, and Johnson \(2016\)](#) who characterize the optimal *final-good* tariff in a partial equilibrium trade model under political economy considerations.⁷ Relative to the aforementioned study, our conceptual framework purposely precludes political pressures to isolate the terms-of-trade motives behind trade taxation. We, however, accommodate a wide range of general equilibrium linkages, analyzes both import and export taxation, and place no restrictions on how trade taxes discriminate between final or intermediate input varieties of the same good. These extensions allow us to identify a set of new insights that have eluded the prior literature on this topic. Relatedly, our result on the uniformity (and therefore redundancy) of optimal tariffs generalizes similar results in [Costinot, Donaldson, Vogel, and Werning \(2015\)](#) and [Opp \(2010\)](#) to a non-parametric neoclassical trade model that accommodates global input-output networks.

We also contribute to a recent literature that quantifies the gains from multilateral trade agreements. [Ossa \(2014, 2016\)](#) highlight the importance of market imperfections when quantifying the gains from trade agreements. [Bagwell, Staiger, and Yurukoglu \(2018\)](#) highlight the importance of inter-connected bilateral negotiations. [Carballo, Handley, and Limão \(2018\)](#) analyze the role of trade agreements at mitigating policy uncertainty. We contribute to this literature by measuring the contribution of global production networks to the value of trade agreements.

Another strand of literature has used the structural gravity approach to measure the value of free trade agreements (e.g., [Baier and Bergstrand \(2007\)](#); [Egger, Larch, Staub, and Winkelmann \(2011\)](#); [Anderson and Yotov \(2016\)](#)). This literature takes a different approach from ours and the quantitative studies listed above. It (a) models the economy as one industry, and (b) uses the estimated gravity coefficients to infer the counterfactual level of trade restriction in the absence of trade agreements. While this approach has many merits, it is not optimized to account for global production networks—mostly, due to the single-industry nature of the structural gravity model (see [Limão \(2016\)](#) for a review of this approach).

[Suzuki \(1978\)](#); [Das \(1983\)](#)). This literature typically assumes a three-good, partial equilibrium economy with only one intermediate good that is produced in only one country. These frameworks, as a result, do not feature a global production network in its formal sense, and deliver starkly different predictions relative to our framework.

⁷Relatedly, [Grant \(2019\)](#) studies the role of *Special Economic Zones* using an enriched partial equilibrium trade policy model where buyers are possibly heterogenous and the government applies import tariffs on many intermediate and final good varieties.

2 The Economic Environment

We consider a global economy that consists of $i = 1, \dots, N$ countries (with \mathbb{C} denoting the set of countries) and $k = 1, \dots, K$ industries (with \mathbb{K} denoting the set of industries). Each country i is populated with L_i workers that are perfectly mobile across industries but immobile across countries. $L_{i,k}$ denotes the number of workers employed in industry k . In addition to labor, each industry in country i is endowed with $\bar{K}_{i,k}$ units of industry-specific capital that is inelastically supplied.

Industry $k \in \mathbb{K}$ in country $i \in \mathbb{C}$ produces a differentiated variety that is ultimately sold in some market $j \in \mathbb{C}$. We use ij, k (origin i –destination j –industry k) to index these product varieties. We use the superscript \mathcal{C} to designate if a good is used as a final consumption good and superscript \mathcal{I} to designate if it is used as an intermediate input. Since we impose no restrictions on the size or the number of industries, we can interpret index k as denoting narrowly-defined product categories rather than broadly-defined industries.

2.1 Preferences

The representative consumer in country i chooses the vector of consumption quantities, $\mathbf{q}_i^{\mathcal{C}} \equiv \{q_{ji,k}^{\mathcal{C}}\}$, to maximize a non-parametric utility function, $U_i(\mathbf{q}_i^{\mathcal{C}})$, subject to their budget constraint. The superscript \mathcal{C} , as noted earlier, differentiates between final consumption varieties and intermediate input varieties, with the latter denoted by \mathcal{I} . The optimal consumption choice yields the following indirect utility function,

$$\begin{aligned} V_i(Y_i, \tilde{\mathbf{p}}_i^{\mathcal{C}}) &\equiv \max_{\mathbf{q}_i^{\mathcal{C}}} U_i(\mathbf{q}_i^{\mathcal{C}}) \\ \text{s.t.} \quad &\sum_{k \in \mathbb{K}} \sum_{j \in \mathbb{C}} \left(\tilde{p}_{ji,k}^{\mathcal{C}} q_{ji,k}^{\mathcal{C}} \right) = Y_i. \end{aligned} \quad (1)$$

The same problem also yields a non-parametric Marshallian demand function,

$$\mathbf{q}_i^{\mathcal{C}} = \mathcal{D}_i(Y_i, \tilde{\mathbf{p}}_i^{\mathcal{C}}),$$

which summarizes the demand-side of the economy as a function of net consumption income Y_i and the vector of consumer prices $\tilde{\mathbf{p}}_i^{\mathcal{C}} \equiv \{\tilde{p}_{ji,k}^{\mathcal{C}}\}$ in country i . The tilde notation on the price variables is used to differentiate the *after-tax* final price

from *pre-tax* producer price. To keep track of optimal demand choices, we define the price and income elasticities of demand as follows.

D1. [Marshallian Demand Elasticities]

- (i) [own price elasticity] $\varepsilon_{ji,k} \equiv \partial \ln q_{ji,k}^C / \partial \ln \tilde{p}_{ji,k}$;
- (ii) [cross-price elasticity] $\varepsilon_{ji,k}^{j,g} \equiv \partial \ln q_{ji,k}^C / \partial \ln \tilde{p}_{j,g}$ for $j, g \neq ji, k$;
- (iii) [income elasticity] $\eta_{ji,k} \equiv \partial \ln q_{ji,k}^C / \partial \ln Y_i$.

Throughout this paper we restrict our attention to well-behaved demand functions that are continuous and locally non-satiated. We also assume that demand for each traded variety exhibits an elastic region where $|\varepsilon_{ji,k}| > 1$. As in monopoly problems, this condition will be necessary for obtaining a bounded solution for optimal trade taxes.

2.2 Technology

We assume that firms are competitive and operate with a non-parametric production function that employs (i) labor, (ii) intermediate inputs, and (iii) industry-specific capital. To elaborate, let $Q_{i,k} = \sum_j \tau_{ij,k} q_{ij,k}$ denote country i 's total output in industry k , where $\tau_{ij,k} \geq 1$ accounts for the iceberg melt cost associated with transporting variety ij, k —as is standard in literature, we normalize $\tau_{ii,k} = 1$. The industry-level output is produced using a general constant returns to scale production function,

$$Q_{i,k} = F_{i,k}(L_{i,k}, \bar{K}_{i,k}, \mathbf{q}_{i,k}^I),$$

that combines labor, $L_{i,k}$, industry-specific capital, $\bar{K}_{i,k}$, and intermediate inputs, $\mathbf{q}_{i,k}^I$.⁸ We assume that the share of labor plus industry-specific capital is constant in total production and equal to $\bar{\alpha}_{i,k}$. However, we take no parametric stance on how labor and capital are combined or how different intermediate inputs are combined in the production function.

Facing the above production structure, cost-minimizing firms charge a competitive “producer” price that is a function of (i) the wage rate in economy i , w_i , (ii) the vector of intermediate input prices employed by producers in country i ,

⁸The above production structure implicitly assumes a non-finite elasticity of transformation between varieties sold in different markets. Relaxing this assumption will lead to diseconomies of scale at the good rather than industry level (see Powell and Gruen (1968)).

$\tilde{\mathbf{p}}_i^{\mathcal{I}} \equiv \{\tilde{p}_{ji,g}^{\mathcal{I}}\}$, and (iii) total industry-level output, $Q_{i,k}$. Namely,

$$p_{ij,k} = \tau_{ij,k} C_{i,k}(w_i, \tilde{\mathbf{p}}_i^{\mathcal{I}}; Q_{i,k}), \quad (2)$$

where $C_{i,k}(\cdot)$ is homogeneous of degree 1 with respect to w_i and $\tilde{\mathbf{p}}_i^{\mathcal{I}}$. A familiar special case of this general structure is the Ricardian case in which $C_{i,k}(\cdot) = a_{i,k} w_i$, where $a_{i,k}$ is a constant unit labor cost that is invariant to industry-level output.

Cost minimization by suppliers in industry k of country i yields an industry-level demand function for intermediate inputs, $\tilde{\mathcal{D}}_{i,k}(\mathcal{Y}_{i,k}, \tilde{\mathbf{p}}_i^{\mathcal{I}})$, which depends on gross expenditure on intermediate inputs, $\mathcal{Y}_{i,k} \equiv \sum_n (1 - \bar{\alpha}_{i,k}) p_{in,k} q_{in,k}$, and the entire vector of intermediate input prices in country i . Country i 's overall demand for intermediate inputs, $\mathbf{q}_i^{\mathcal{I}} \equiv \{q_{ji,k}^{\mathcal{I}}\}$, is determined as the sum of demands across all industries:

$$\mathbf{q}_i^{\mathcal{I}} = \sum_k \tilde{\mathcal{D}}_{i,k}(\mathcal{Y}_{i,k}, \tilde{\mathbf{p}}_i^{\mathcal{I}}).$$

For notational convenience, we assume that the *after-tax* price of product, ji, k , is the same whether it is used as an intermediate input (indexed \mathcal{I}) or a consumption good (indexed \mathcal{C}):

$$\tilde{p}_{ji,k}^{\mathcal{I}} = \tilde{p}_{ji,k}^{\mathcal{C}} = \tilde{p}_{ji,k}.$$

This assumption is innocuous since we can always extend the set of goods so that ji, k indexes only the final good varieties, while ji, k' indexes only the intermediate input varieties supplied by the same industry.⁹ With the same rationale, we also assume that for all k , $\tilde{\mathcal{D}}_{i,k}(\cdot) / \partial \ln \mathcal{Y}_{i,k} = \partial \ln \mathcal{D}_i(\cdot) / \partial \ln Y_i$ and $\partial \ln \tilde{\mathcal{D}}_{i,k}(\cdot) / \partial \ln \tilde{\mathbf{p}}_i^{\mathcal{I}} = \partial \ln \mathcal{D}_i(\cdot) / \partial \ln \tilde{\mathbf{p}}_i^{\mathcal{C}}$.

We use input-output shares to keep track of global production networks. To define these shares, let $q_{\ell i, g}^{\mathcal{I}}(k)$ denote the optimal amount of input $\ell i, g$ used by industry k in country i —by construction, $q_{\ell i, g}^{\mathcal{I}} = \sum_k q_{\ell i, g}^{\mathcal{I}}(k)$. Considering this choice of notation, we define the global input-output shares as follows.

D2. [Input-Output Shares] The share of intermediate input goods from “country $\ell \times$ industry g ” that are used in the production of output goods in “country $j \times$ industry k ” is defined as $\alpha_{i,k}^{\ell, g} \equiv \frac{\tilde{p}_{\ell i, g}^{\mathcal{I}} q_{\ell i, g}^{\mathcal{I}}(k)}{\sum_j p_{ij,k} q_{ij,k}}$.

⁹Following this argument, our model still accommodates cases where the government imposes differential tax rates on final good varieties (k) versus intermediate input varieties (k').

To be clear, we do not impose that input-output shares be constant. Instead, α 's can be variable, and will presumably change in response to trade taxation, e.g., taxing input $\ell_{i,g}$ will lower $\alpha_{i,k}^{\ell,g}$. Our input-output setup is also flexible enough to allow for the expenditure share on the intermediate input varieties of a given good to diverge from the expenditure share on its final good varieties. A well-known special case of our general input-output structure is [Caliendo and Parro \(2014\)](#), which we formally discuss in Section 4.1.

We use the supply elasticity, as defined below, to keep track of diseconomies of scale at the industry level.

D3. [Industry-Level Supply Elasticity] $\gamma_{i,k} \equiv \partial \ln C_{i,k}(\cdot) / \partial \ln Q_{i,k}$

The above definition is motivated by the observation that $C_{i,k}(\dots, Q_{i,k})$ characterizes the industry-level supply curve for country i . To gain intuition about how this elasticity relates to diseconomies of scale, consider the case where the share of the industry-specific capital in production is constant and equal to $\beta_{i,k}$. In that case, it is straightforward to show that $\gamma_{i,k} = \beta_{i,k} / (1 - \beta_{i,k})$. Beyond this special, $\gamma_{i,k}$ still reflects the importance of the inelastically-supplied capital in production. As we will see shortly, $\gamma_{i,k}$ also governs the degree of national-level import market power in each industry.

2.3 Trade Policy Instruments

The government in country i has access to a full set of industry-level export tax-cum-subsidy instruments (denoted by x) and import tax-cum-subsidy instruments (denoted by t). Together, these policy instruments create a wedge between the after-tax price, $\tilde{p}_{j,i,k}$, and the producer price, $p_{j,i,k}$, of each good j,i,k as follows:

$$\tilde{p}_{j,i,k} = (1 + t_{j,i,k}) (1 + x_{j,i,k}) p_{j,i,k}. \quad (3)$$

In the above equation, $t_{j,i,k}$ denotes the import tax applied by country i on good j,i,k , while $x_{j,i,k}$ denotes the export tax applied by country j on the same good. The combination of these tax instruments raises the following tax revenue for the government in country i :

$$\mathcal{R}_i = \sum_{k \in \mathbb{K}} \sum_{j \in \mathbb{C}} [t_{j,i,k} (1 + x_{j,i,k}) p_{j,i,k} q_{in,k} + x_{j,i,k} p_{ij,k} q_{ij,k}]. \quad (4)$$

Tax revenues are rebated to the consumers in a lump-sum fashion.¹⁰ Throughout this paper, we assume that domestic policies are unavailable, which amounts to $t_{ii,k} = x_{ii,k} = 0$ for all i and k . As noted earlier, our product space is general enough that k can index only the final good version or the intermediate input version of each industry's output. So, in principle, trade taxes can arbitrarily discriminate between final consumption and intermediate input varieties of the same good.

2.4 General Equilibrium

We assume throughout this paper that equilibrium is unique—noting that uniqueness can be formally established using the procedure in [Alvarez and Lucas \(2007\)](#). Below, we formally define the general equilibrium in our setup.

Definition. For any given vector of taxes, \mathbf{t} , and, \mathbf{x} , equilibrium is a vector of wages, w , (pre-tax) producer and (after-tax) final prices, \mathbf{p}_i , and $\tilde{\mathbf{p}}_i$, final consumption and input demand choices, \mathbf{q}_i^C and \mathbf{q}_i^T , total surplus paid to industry-specific capital, Π , net consumer expenditure, \mathbf{Y} , and gross industry-level output, \mathcal{Y} , such that (i) the producer price for each good is characterized by Equation 2; (ii) the consumer price for each good is given by Equation 3; (iii) consumption choices, $\mathbf{q}_i^C = \mathcal{D}_i(Y_i, \tilde{\mathbf{p}}_i)$, are a solution to 1; (iv) demand for intermediate inputs, $\mathbf{q}_i^T = \sum_k \tilde{\mathcal{D}}_i(\mathcal{Y}_{i,k}, \tilde{\mathbf{p}}_i^T)$, is chosen to minimize cost; (v) factor markets clear

$$w_i L_i + \Pi_i = \sum_j \sum_k p_{ij,k} q_{ij,k} - \sum_j \sum_k p_{ji,k}^T q_{ji,k}^T \quad \forall i,$$

and (v) net consumption income equals factor income plus tax revenue

$$Y_i = w_i L_i + \Pi_i + \mathcal{R}_i \quad \forall i,$$

where the tax revenue, \mathcal{R}_i , is given by Equation 4.

To streamline the presentation of our theory, we henceforth express aggregate welfare in country i as a function of taxes \mathbf{t} , and, \mathbf{x} , and wages, w ,

$$W_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{t}_{-i}, \mathbf{x}_{-i}, w) \equiv V_i(Y_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{t}_{-i}, \mathbf{x}_{-i}, w), \tilde{\mathbf{p}}_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{t}_{-i}, \mathbf{x}_{-i}, w)).$$

¹⁰Since labor is inelastically supplied in our framework, lump-sum rebates are observationally equivalent to a wage subsidy or a uniform consumption subsidy.

w is an equilibrium outcome and, thus, an implicit function of trade taxes, So, we use \mathbb{A} to denote the set of all wage \times policy combinations, $A = (t, x; w)$, that are feasible. The reason we express W_i as a function of $(t, x; w)$ rather than just trade taxes is that an across-the-board shift in trade taxes combined with an equal-proportional adjustment to nominal wage rates preserves welfare. Expressing equilibrium outcomes in terms of the triplet $(t, x; w)$ enables us to track this kind of tax neutrality.

In equilibrium, the importance of each good for taxation purposes is determined (among other things) by its share in gross expenditure and output. So before concluding this section, it is useful to define the aforementioned shares for good ji, k (origin j -destination i -industry k).

D4. [Gross Expenditure and Output Shares]

$$[\text{within-industry expenditure share}] \lambda_{ji,k} \equiv \frac{\tilde{p}_{ji,k} q_{ji,k}}{\sum_{j \in \mathbb{C}} \tilde{p}_{ji,k} q_{ji,k}}$$

$$[\text{overall expenditure share}] \hat{\lambda}_{ji,k} \equiv \frac{\tilde{p}_{ji,k} q_{ji,k}}{\sum_{j \in \mathbb{C}} \sum_{g \in \mathbb{K}} \tilde{p}_{ji,g} q_{ji,g}}.$$

$$[\text{within-industry output share}] r_{ji,k} \equiv \frac{p_{ji,k} q_{ji,k}}{\sum_{i \in \mathbb{C}} p_{ji,k} q_{ji,k}}$$

$$[\text{overall output share}] \hat{r}_{ji,k} \equiv \frac{p_{ji,k} q_{ji,k}}{\sum_{i \in \mathbb{C}} \sum_{g \in \mathbb{K}} p_{ji,g} q_{ji,g}}$$

Gross expenditure and output shares, as defined above, are directly observable in standard trade datasets. As we will see shortly, optimal trade taxes can be fully characterized in terms of these observable shares plus input-output shares and the reduced-form demand and supply elasticities defined earlier.

3 Unilaterally Optimal Trade Taxes

We begin our analysis by characterizing the unilaterally optimal tax schedule in each country given applied taxes in the rest of the world. Since our analysis precludes political economy motives, the aforementioned taxes are purely terms-of-trade-driven. So, following [Bagwell and Staiger \(1999\)](#), they govern the efficiency loss that free trade agreements are designed to remedy.¹¹

¹¹Our analysis fits into the terms-of-trade framework, wherein trade agreements “remedy the inefficient terms-of-trade-driven restrictions in trade that arise when trade policies are set unilaterally, ([Bagwell and Staiger \(1999\)](#))” Based on this viewpoint, the value of free trade agreements can be measured as the reduction in trade that would otherwise occur if countries were imposing terms-of-trade-driven restrictions in trade. See also [Maggi and Rodriguez-Clare \(1998\)](#); ? for a domestic-commitment theory of trade agreements.

For expositional purposes, we cast our theory first using a *two-country* setup where h indexes the Home country and f indexes Foreign that represents an aggregate of the rest of the world, i.e., $\mathbb{C} = \{h, f\}$. *Later, in Section 4.2 we show how our baseline results extend to a multi-country setup with arbitrarily many countries.* The Home government's optimal non-cooperative policy, $\mathbf{t}_h^* \equiv \{t_{fh,k}^*\}$, and $\mathbf{x}_h^* \equiv \{x_{hf,k}\}$, solves the following problem taking the vector of Foreign taxes, $\mathbf{t}_f \equiv \{t_{hf,k}\}$, and $\mathbf{x}_f \equiv \{x_{fh,k}\}$, as given:

$$\max_{(\mathbf{t}_h, \mathbf{x}_h; \mathbf{t}_f, \mathbf{x}_f, \mathbf{w}) \in \mathbb{A}} W_h(\mathbf{t}_h, \mathbf{x}_h; \mathbf{t}_f, \mathbf{x}_f, \mathbf{w}). \quad (5)$$

It is needless to say that the above problem is complicated by a myriad of general equilibrium interrelations. We can nonetheless simplify the problem by appealing to several intermediate results, the first of which is the Lerner symmetry.¹²

Lemma 1. [*The Lerner Symmetry*] For any $a \in \mathbb{R}_+$, combinations $A = (\mathbf{1} + \mathbf{t}_h, \mathbf{1} + \mathbf{x}_h; \mathbf{t}_f, \mathbf{x}_f, w_h, w_f)$ and $A' = (a(\mathbf{1} + \mathbf{t}_h), (\mathbf{1} + \mathbf{x}_h)/a; \mathbf{t}_f, \mathbf{x}_f, aw_h, w_f)$ represent identical equilibria, i.e., $W_i(A) = W_i(A')$ for all i .

The above lemma simplifies our analysis as follows: Optimal import and export taxes both feature a uniform term that accounts for the ability of trade taxes to increase Home's wage relative to Foreign (w_h/w_f). Following the above lemma, we need not to formally characterize this term as it is redundant. There is another way to cast this redundancy: When both export and import taxes are applicable, Lemma 1 indicates that we can normalize wages in both economies (i.e., set $w_h = w_f = 1$) and still identify one of the multiple optimal tax combinations. This result, however, follows only if a full set of industry-level export and import tax instruments are applicable. If the policy space is restricted in any way, the uniform tax component that accounts for the general equilibrium changes in w_h/w_f should be formally characterized—see Section 3.1 for further details.

In our general equilibrium setup, trade taxes have a non-trivial passthrough onto the vector of final prices, $\tilde{\mathbf{p}}_i$, in each country. A tax on one good can alter the entire vector of prices through its effect on country-level wages, input prices, and industry-wide output. To handle these complex interrelations, we can

¹²See Costinot and Werning (2019) for a more comprehensive treatment of the Lerner symmetry in the presence of multinational firms, global imbalances, and imperfect competition.

cast our original optimal policy problem as one where the government directly chooses final prices, \tilde{p}_{fh} , \tilde{p}_{hf} , and \tilde{p}_{hh} , to maximize Home's welfare $W_h(\tilde{p}; w) \equiv V_h(Y_h(\tilde{p}; w), \tilde{p})$. Stated formally, the optimal policy problem can be reformulated as

$$\max_{(\tilde{p}; w) \in \tilde{\mathbb{A}}} W_h(\tilde{p}; w),$$

where $\tilde{\mathbb{A}}$ denotes the set of feasible wage-price combinations that is defined analogously to \mathbb{A} . Implicit in the above formulation is the observation that (i) $W_h(\cdot)$ does not *explicitly* dependent on \tilde{p}_{ff} , and (ii) the market equilibrium is efficient, so it is optimal to set $\tilde{p}_{hh} = p_{hh}$. The above problem, as a result, corresponds to one where the government chooses the consumer-to-producer price wedges that pin down the trade taxes: $\mathbf{1} + \mathbf{t}_h = \tilde{p}_{fh}/(\mathbf{1} + \mathbf{x}_f)p_{fh}$ and $\mathbf{1} + \mathbf{x}_h = \tilde{p}_{hf}/(\mathbf{1} + \mathbf{t}_f)p_{hf}$.

Finally, we can appeal to supply- and demand-side envelop conditions to handle general equilibrium behavioral responses. On the demand side, we can appeal to Roy's identity, whereby the direct welfare effect of a change on consumer price, $\tilde{p}_{fh,k}^C$, is reduced to $\partial V_h(\cdot)/\partial \tilde{p}_{fh,k}^C = -q_{fh,k}^C$. Similarly, we can appeal to Shepard's lemma to account for input-output-driven price linkages. Specifically, a change in input price $\tilde{p}_{li,g}^I$, holding the wage and all other input prices fixed, has the following effect on the output price of a cost-minimizing supplier:

$$\frac{\partial \ln p_{ij,k}}{\partial \ln \tilde{p}_{li,g}^I} = \frac{\partial \ln C_{i,k}(\cdot)}{\partial \ln \tilde{p}_{li,g}^I} = \alpha_{i,k}^{\ell,g},$$

where $\alpha_{i,k}^{\ell,g}$ denotes the share of input li,g in output ij,k as defined by D3. The change in $\tilde{p}_{li,g}^I$ has an additional general equilibrium effect, which operates through changes in w , \mathcal{Y} , and \mathcal{Y} . As noted earlier, though, the handling of these general equilibrium effects can be simplified with the application of Lemma 1.

Trade taxes also affect, Π_i , which is the surplus paid to industry-specific capital in country i . To track this effect, we can appeal to Hotelling's lemma, whereby the effect of a change in output prices on surplus (holding all input prices constant) can be stated as: $\partial \Pi_i(\cdot)/\partial p_{ij,k} = q_{ij,k}$. Likewise, the effect of a change in the price of intermediate input ji,k (holding the price of all other inputs fixed) can be expressed as: $\partial \Pi_i(\cdot)/\partial p_{ji,k}^I = q_{ji,k}^I$.

By combining the aforementioned envelope conditions and appealing to Lemma 1, we can produce the following theorem that characterizes the optimal

policy as a function of gross expenditure shares, λ , gross revenue shares, r , global input-output shares, α , reduced-form demand elasticities, ε , and reduced-form supply elasticities, γ . The former three statistics are directly observable, while the latter two can be locally estimated.

Theorem 1. *Home's unilaterally optimal trade taxes are unique up-to a uniform tax shifter, \bar{t} , and given by*

$$1 + t_{fh,k}^* = \left(1 + \frac{\gamma_{f,k} r_{fh,k}}{1 - \gamma_{f,k} r_{ff,k} \varepsilon_{ff,k}} \right) (1 + \bar{t})$$

$$1 + x_{hf,k}^* = \frac{\varepsilon_{hf,k}}{1 + \varepsilon_{hf,k} + \zeta_{hf,k} - \sum_g \alpha_{f,g}^{h,k} \frac{r_{fh,g}}{\lambda_{hf,k}}} (1 + \bar{t})^{-1}$$

where $[\zeta_{hf,k}]_k = [\Xi^{-1} - \mathbf{I}_K] \mathbf{\Omega}$ accounts for input-output adjusted cross-demand effects between industries, with $\Xi \equiv \left[\frac{\lambda_{hf,g} \varepsilon_{hf,g}^{hf,k}}{\lambda_{hf,k} \varepsilon_{hf,k}} \right]_{k,g}$ and $\mathbf{\Omega} \equiv \left[1 - \sum_g \alpha_{f,g}^{h,k} \frac{r_{fh,g}}{\lambda_{hf,k}} \right]_k$.

Foreign's unilaterally optimal taxes are, analogously, described by a similar formula that swaps the country indices. As noted earlier, per the Lerner symmetry, the unilaterally optimal policy is *unique* only up to a uniform tariff, \bar{t} . For a high-enough choice of \bar{t} , the optimal policy will consist of import taxes paired with export subsidies. Aside from the uniform tax shifter, Home's unilaterally optimal import tax, $t_{fh,k}^*$, equals the optimal industry-level *mark-down* on $p_{fh,k}$ (i.e., the inverse of the export supply elasticity). Home's unilaterally optimal export tax, $x_{hf,k}^*$, is equal the optimal monopoly *mark-up* on $\tilde{p}_{hf,k}$ that internalizes cross-demand effects and tax propagation through the production network. In the special case with zero cross-substitutability between industries (i.e., $\Xi = \mathbf{I}_K \iff \zeta_{hf,k} = 0$) and no input-output networks (i.e., $\alpha_{f,g}^{h,k} = 0$), the optimal export tax reduces to the familiar *single-product* optimal monopoly markup, $1 + x_{hf,k}^* = \varepsilon_{hf,k} / (1 + \varepsilon_{hf,k})$.

Importantly, the optimal tax schedule specified by Theorem 1 internalizes Foreign's applied taxes, t_f , and x_f . For instance, an increase in Foreign's import taxes will shrink Home's exports to Foreign. This effect will reflect itself as a reduction in $\varepsilon_{hf,k}$ (and therefore $x_{hf,k}^*$) if the demand for variety hf,k is sub-convex. We can, thus, express $t_{fh,k}^*$ and $x_{hf,k}^*$ as an implicit function of Foreign's applied taxes. We will invoke this property in Section 5, where we solve for Nash taxes in the non-cooperative equilibrium.

Theorem 1 has an attractive feature: It characterizes the optimal policy in terms of estimable or observable sufficient statistics. In the words of [Piketty and Saez \(2013\)](#), such sufficient statistic formulas have two broad merits. First, they allow us “to understand the key economic mechanisms behind the formulas.” Second, they “are also often robust to changing the primitives of the model.” In the present context, the formula characterized by Theorem 1 can be empirically evaluated with readily-available trade statistics. Given these qualities and as shown later in Section 6, Theorem 1 streamlines the quantitative analysis of trade policy to a great degree.

There is a simple intuition for why optimal import taxes (unlike export taxes) do not *explicitly* depend on global input-output shares. From the perspective of the Home country, trade taxes can correct two *terms-of-trade*-related distortions:

- i. Unexploited import market power, which concerns the unexploited ability of the Home economy to charge a *mark-down* on Foreign producer prices, $p_{fh,k}$.
- ii. Unexploited export market power, which concerns the unexploited ability of the Home economy to charge a *mark-up* on Foreign consumer prices, $\tilde{p}_{hf,k}$.

By the targeting principle, industry k 's optimal import tax, $t_{fh,k}^*$, is targeted exclusively at lowering $p_{fh,k}$ (Margin 1). For this reason, $t_{fh,k}^*$ does not *explicitly* depend on input-output shares. To make this point, consider the following thought experiment: Fix the price of all of Home's export goods as these prices can be directly manipulated with export taxes. In that case, $t_{fh,k}$ cannot affect either $\{p_{fh,g}\}$ or $\{\tilde{p}_{hf,g}\}$ through the input-output network. Instead, $t_{fh,k}$ can only lower $p_{fh,k}$ by shrinking Foreign's output in industry k or by lowering w_f . These effects, by construction, operates independent of the input-output network.

Based on the same rationale, optimal export taxes depend explicitly on global input-output shares. An export tax on intermediate input, hf, k , will raise the price of any Foreign-produced good employing that input, including goods that are sold back to the Home country. That is, An export tax on intermediate inputs is partially passed back to Home consumers through re-importation. To mitigate these adverse feedback effects, the optimal export tax is lower on more upstream industries where re-importation of export taxes is more of an issue. This claim is implicit

in the formula under Theorem 1, considering that¹³

$$\sum_g \alpha_{f,g}^{h,k} \frac{\hat{r}_{fh,g}}{\hat{\lambda}_{hf,k}} \in \begin{cases} \{0\} & \text{if } fh,k \text{ is a final good} \\ (0,1) & \text{otherwise} \end{cases}$$

That export taxes are lower on upstream industries, however, should not be confused with smaller gains from taxation. On the contrary, the unilateral gains from taxing upstream exports are actually larger. We formally document and discuss this point in Section 6.

Theorem 1 indicates that import taxes are a necessary instrument only if there are diseconomies of scale arising from industry-specific capital. Otherwise, if $\gamma_{f,k} \approx 0$, Home can discard import taxes by choice of $\bar{t} = 0$, and attain the first-best non-cooperative outcome with only export taxes. The following corollary synthesizes this claim and others discussed earlier.

Corollary 1. *Holding λ and r constant, the globalization of production networks lowers the unilaterally optimal export tax on more-upstream industries. Moreover, absent diseconomies of scale, import taxes remain a redundant terms-of-trade-improving policy instrument even after countries become immersed in global production networks.*

The globalization of production networks, in the above corollary, corresponds to a counterfactual thought experiment where λ and r are fixed, but α is risen from zero to its factual level. This experiment aligns with the quantitative exercise presented later in Section 6. Corollary 1, therefore, indicates that overlooking global production networks will lead to an over-estimation of non-cooperative export taxes in upstream industries.

Finally, the absence of diseconomies of scale is a sufficient but *not* necessary condition for the redundancy of import taxes. Import taxes can be redundant under much weaker conditions. For instance, suppose Foreign employs industry-specific capital but Home accounts for a small share of global demand for the Foreign industry. Then, $r_{fh,k} \approx 0$ and import taxes are once again redundant. The aforementioned situation provides an accurate description of any individual country's position relative to the rest of the world. It also aligns with the *complete* tariff passthrough documented by [Amiti et al. \(2019\)](#) and [Fajgelbaum et al. \(2020\)](#).

¹³More generally, $\sum_g \alpha_{f,g}^{h,k} \hat{r}_{fh,g} / \hat{\lambda}_{hf,k} = 0$ if good fh,k is not used as an input in any of Foreign's export goods. Final goods automatically satisfy this qualification.

3.1 Second-Best Non-Cooperative Import Taxes

We now consider a scenario that has received considerable attention in the prior literature. In this scenario, governments cannot apply export taxes due to institutional constraints like those embedded in the United States' constitution. But they have the discretion to apply import taxes. We show that, in such circumstances, it is unilaterally optimal for governments to use import taxes as a second-best instrument to manipulate export market power. To make this point succinctly, we abstract from diseconomies of scale and cross-industry demand effects; but these channels are formally accounted for in Appendix C.

In a second-best scenario where the government cannot apply export taxes, Home's unilaterally optimal import taxes solve the following problem (taking \mathbf{t}_f and \mathbf{x}_f as given):

$$\max_{(\mathbf{t}_h, \mathbf{0}; \mathbf{t}_f, \mathbf{x}_f, \mathbf{w}) \in \mathbb{A}} W_h(\mathbf{t}_h, \mathbf{0}; \mathbf{t}_f, \mathbf{x}_f, \mathbf{w}).$$

Since export taxes are restricted to zero, i.e., $\mathbf{x}_h = \mathbf{0}$, the Lerner symmetry no longer implies a multiplicity of optimal tax schedules. Instead, the above problem identifies a unique vector of optimal import taxes. Since the Home economy possesses no import market power by assumption (i.e., $\gamma_{f,k} r_{fh,k} \approx 0$), second-best import taxes pursue one objective: to *indirectly* manipulate Home's export market power.

Export market power manipulation can be carried through two distinct channels: First, import taxes can raise Home's wage relative to Foreign, w_h/w_f (i.e., they can charge a markup on the wage rate embedded in Home's exports). Second, import taxes can be used to charge a markup on $\tilde{p}_{hf,k}$ through the input-output network. The former channel was previously-redundant due to the Lerner symmetry-driven multiplicity of optimal taxes. The latter channel was also irrelevant due to the targeting principle.

To account for the now-relevant general equilibrium wage effects, we use $\mathcal{L}_{ij}(\mathbf{t}, \mathbf{x}; \mathbf{w})$ to denote country j 's demand for country i 's labor. The elasticity of \mathcal{L}_{ij} with respect to w_i depends on the value-added content of sales, which determines the overall contribution of country i 's labor to gross revenue in each industry. We can measure the value-added content of country i 's output in industry k , namely, $\delta_{i,k}$, using the input-output matrix. Specifically, let

$$\boldsymbol{\alpha}_{ji} \equiv \left[\alpha_{i,k}^{j,g} \right]_{k,g}$$

denote the $K \times K$ input-output matrix that describes the share of country j 's inputs used in country i 's outputs. We can apply the Implicit Function Theorem to calculate $\delta_i \equiv [\delta_{i,k}]_k$ as follows:

$$\delta_i = (\mathbf{I}_K - \alpha_{ii})^{-1} \bar{\alpha}_i,$$

where $\bar{\alpha}_i \equiv [\bar{\alpha}_{i,k}]_k$ is a $K \times 1$ vector denoting industry-level labor shares in country i . Using the above notation, we can calculate the elasticity of labor demand in terms of gross reduced-form demand elasticities, $\varepsilon_{ij,k}$, and value-added shares, $\delta_{i,k}$.

D4. [Elasticity of Labor Demand] $\bar{\varepsilon}_{ij} \equiv \frac{\partial \ln \mathcal{L}_{ij}}{\partial \ln w_i} = \sum_k \omega_{ij,k} \varepsilon_{ij,k}$, where $\omega_{ij,k} \equiv \frac{\delta_{j,k} \lambda_{ij,k}}{\sum_g \delta_{j,k} \lambda_{ji,k}}$ is a weight that reflects the value-added contribution of good ij, k to country i 's exports.

In principle, $\bar{\varepsilon}_{ij}$ measures country i 's economy-wide export market power net of input-output linkages. As noted by the following Theorem, Home's optimal import tax in each industry is determined by $\bar{\varepsilon}_{hf}$ and an industry-specific term that accounts for the ability of import taxes to manipulate industry-specific export market power through the production network.

Theorem 2. *In a second-best scenario where export taxes are inapplicable, unilaterally optimal import taxes feature of a uniform component and an industry-specific component that captures the extent to which import taxes are passed on to Foreign consumers. Namely,*

$$1 + t_{fh,k}^* = \frac{\bar{\varepsilon}_{hf}}{1 + \bar{\varepsilon}_{hf}} \left(1 + \frac{\tau_k}{\varepsilon_{fh,k}} \right),$$

where $\tau \equiv [\tau_k]$ is given by

$$\tau = \left[\mathbb{1}_{k=g} + \frac{\varepsilon_{fh,g}^{hh,g}}{\varepsilon_{fh,k}} \tilde{\alpha}_{h,g}^{f,k} \right]_{k,g}^{-1} \left[\sum_{g \in \mathbb{K}} \left(\tilde{\alpha}_{f,g}^{f,k} - \left[1 + \frac{\varepsilon_{hf,g}}{\bar{\varepsilon}_{hf}} \right] \frac{\lambda_{hf,g}}{r_{fh,g}} \tilde{\alpha}_{h,g}^{f,k} \right) \right]_k,$$

with $[\tilde{\alpha}_{f,g}^{f,k}]_{k,g} = (\mathbf{I}_K - \alpha_{ff})^{-1} \alpha_{fh} \alpha_{hf}$, and $[\tilde{\alpha}_{h,g}^{f,k}]_{k,g} = (\mathbf{I}_K - \alpha_{fh})^{-1} \alpha_{fh}$. Correspondingly, $\tau_k = 0$ for import goods that are not used as intermediate inputs in export goods and $\tau_k \neq 0$ otherwise.

The uniform term, $\bar{\varepsilon}_{hf}/(1 + \bar{\varepsilon}_{hf})$, in the above formula, corresponds to an optimal export mark-up on the wage rate, w_h . To elaborate, a uniform import tax is

equivalent to a uniform export tax, which is itself analogous to a uniform markup applied to the wage rate, w_h , in Foreign markets. The optimal level of this markup is proportional to the elasticity of Foreign demand for Home's labor, $\bar{\epsilon}_{hf}$.¹⁴

The industry-specific component, $1 + \tau_k / \epsilon_{fh,k}$, accounts for the fact that an import tax on intermediate inputs is partially passed on to Foreign consumers. So, such a tax can be used to imperfectly mimic (the unavailable first-best) export taxes. To see this point from the lens of Theorem 2, note that the optimal import tax on good fh,k depends primarily on $\tilde{\alpha}_{h,g}^{f,k}$, which measures the degree to which the tax is passed on to Foreign consumers. By construction,

$$\left[\tilde{\alpha}_{h,g}^{f,k} \right] = \alpha_{fh} + \alpha_{fh}^2 + \alpha_{fh}^3 + \dots$$

which implies that if (i) industry k is a strictly downstream industry in a vertical economy, or (ii) the inputs produced by industry k are never used in exporting industries, then $\tilde{\alpha}_{h,g}^{f,k} = \tilde{\alpha}_{f,g}^{f,k} = 0$ and, as a result, $\tau_k = 0$. Plugging this value back into Theorem 2, the optimal import tax on such industries is uniform and equal to $t_{fh,k}^* = \bar{\epsilon}_{hf} / (1 + \bar{\epsilon}_{hf})$. The intuition is that an import tax on the aforementioned industries is never passed on to Foreign consumers beyond general equilibrium wage effects. So, there is no rationale for taxing such imports beyond what is justified by the flat wage markup.

If industry k supplies intermediate inputs to differentiated and export-oriented industries, then $\tilde{\alpha}_{h,g}^{f,k} > 0$, which implies that $\tau_k / \epsilon_{fh,k} > 0$.¹⁵ Plugging this value into Theorem 2 indicates that industry k imports should be taxed above the uniform rate. The reason is that an import tax on intermediate inputs is partly passed on to Foreign consumers, so $t_{fh,k}$ can be used to extract an additional markup, $\tau_k / \epsilon_{fh,k}$, from Foreign consumers besides the flat wage markup, $\bar{\epsilon}_{hf} / (1 + \bar{\epsilon}_{hf})$. In other words, a tax on imported intermediates can be used to (imperfectly) mimic

¹⁴To gain further perspective, consider a basic CES gravity model without input-output linkages (as presented in Section 4.1). In this widely-used framework, $\tilde{\alpha}_{i,g}^{f,k} = 0$ and $\delta_{i,k} = 1$, which implies a uniform optimal import tax across all industries:

$$\bar{t}_{fh,k}^* = 1 / \sum_k \left(\frac{\hat{r}_{hf,k}}{\hat{r}_{hf}} \epsilon_k \lambda_{ff,k} \right),$$

with ϵ_k denoting the trade elasticity in industry k . The above formula itself strictly generalizes the optimal single-industry tariff formula in Gros (1987) to many asymmetric industries.

¹⁵This result follows from $\tilde{\alpha}_{h,g}^{f,k} > \tilde{\alpha}_{f,g}^{f,k} \approx 0$.

the first-best export taxes.¹⁶ The trade-off is that taxing intermediate inputs distorts production decisions in a way that is avoidable with first-best export taxes.

In summary, these arguments indicate that the globalization of production networks creates an additional incentive for import taxation in second-best scenarios. More specifically, if production networks become globalized and export taxes are inapplicable, it is unilaterally optimal to increase the import tariff on more-upstream industries. The following corollary synthesizes these arguments.

Corollary 2. *Suppose export taxes are inapplicable: Holding λ and r constant, the globalization of production networks increases the unilaterally optimal tariff on upstream industries that supply intermediates inputs to export-oriented sectors. The globalization of production networks, however, has no effect on the unilaterally optimal tariff in industries that supply final goods or intermediate goods that are never used in export-oriented sectors.*

The above corollary raises a basic question: *How much are governments willing to raise the tariff on intermediate inputs to mimic export taxes?* Answering this question is ultimately a quantitative matter, which is explored in Section 6. But we can gain valuable insights by approaching the question theoretically. If imported goods are re-exported without being processed, then an import tax on these goods is identical to an export tax. So the globalization of production-networks prompts a sharp increase in optimal import taxes.

Beyond this extreme case, import taxes on intermediate inputs are an extremely inefficient substitute for export taxes: First, they directly distort production-related choices in the Home economy. Second, they may be partially passed on to domestic consumers. Accordingly, the effectiveness of second-best trade taxes is hampered by an innate trade-off between preserving allocative efficiency in the local economy and exploiting export market power through the production network. The former discourages taxing differentiated intermediate inputs, while the latter requires it. As a result, for all practical purposes, the globalization of production networks should have a modest effect on non-cooperative import taxes.

¹⁶Aside from $\tilde{\alpha}_{h,g}^{f,k}$, the effectiveness of $t_{fh,k}$ at mimicking first-best export taxes is governed by two additional factors: (a) the degree to which intermediate input variety fh,k is substitutable with other inputs (i.e., $\varepsilon_{fh,k}$ in the formula specified by Theorem 2); and (b) the intensity at which input fh,k is used in highly-differentiated export-oriented industries (i.e., $\tilde{\alpha}_{h,g}^{f,k} \lambda_{hf,g} \varepsilon_{hf,g}$ in the formula specified by Theorem 2). Specifically, $t_{fh,k}^*$ is higher if intermediate input fh,k is low-demand elastic and is used more-intensively in highly-differentiated, export-oriented industries. The following corollary summarizes these arguments.

4 Canonical Special Cases and Extensions

In this section, we first highlight two widely-used quantitative trade models that are nested by our theory. We then demonstrate how our baseline results readily extend to a multi-country setup.

4.1 Canonical Special Cases

To highlight the practicality of our optimal tax formula, we outline two canonical special cases: First, a basic multi-industry gravity model without input-output linkages à la [Costinot et al. 2011](#). Second, a multi-industry gravity model with flexible input-output linkages à la [Caliendo and Parro \(2014\)](#).

(i) Basic Multi-Industry Gravity Model (Costinot et al. 2011). This model corresponds to a special case of our framework where labor is the only factor of production and preferences have a Cobb-Douglas-CES parameterization: $U_i = \prod_k Q_{i,k}^{\epsilon_{i,k}}$, where $Q_{i,k} = \left(\sum_{j=h,f} \chi_{ji,k} q_{ji,k}^{\rho_k} \right)^{1/\rho_k}$. The Cobb-Douglas-CES demand structure implies that $\epsilon_{hf,k} = -1 - \epsilon_k \lambda_{ff,k}$, where $\epsilon_k \equiv \rho_k / (1 - \rho_k)$. The Cobb-Douglas assumption also eliminates cross-price elasticity effects, which amounts to $\zeta_{hf,k} = 0$. The absence of input-output networks implies that $\alpha_{h,g}^{f,k} = 0$ for a k and g . Plugging these values into the optimal tax formula specified by [Theorem 1](#) yields the following:

$$\begin{aligned} 1 + t_{fh,k}^* &= 1 + \bar{t} \\ 1 + x_{hf,k}^* &= \left(1 + \frac{1}{\epsilon_k \lambda_{ff,k}} \right) (1 + t)^{-1}. \end{aligned} \quad (6)$$

Stated verbally, the optimal trade tax consists of a uniform tariff, \bar{t} , and an industry-specific export tax that varies primarily with the industry-level trade elasticity, ϵ_k . If the economy is modeled as a single industry, the above formula reduces to $x_{hf}^* = 1/\epsilon \lambda_{ff}$, which by the Lerner symmetry is equivalent to [Gros' \(1987\)](#) optimal tariff formula, $t_{fh}^* = 1/\epsilon \lambda_{ff}$.

(ii) Multi-Industry Gravity Model with I-O Linkages (Caliendo and Parro (2014))
This model features the same CES-Cobb-Douglas utility function described above.

Production, though, combines intermediate inputs and labor using a Cobb-Douglas-CES aggregator, which implies that

$$p_{ij,k} = w_i^{\bar{\alpha}_{i,k}} \prod_g \tilde{P}_{i,g}^{\alpha_{i,gk}},$$

where $\alpha_{i,kg}$ denotes the constant share of industry k 's inputs in industry g 's output, with $\bar{\alpha}_{i,k} \equiv 1 - \sum_g \alpha_{i,gk}$. $\tilde{P}_{i,g}$ denotes the (after-tax) price index of industry g 's composite intermediate input, which is by assumption identical to the price index of industry g 's composite consumption good: $\tilde{P}_{i,g} = \left(\sum_j \tilde{p}_{ji,g}^{1-\epsilon_g} \right)^{1/(1-\epsilon_g)}$. The Cobb-Douglas-CES demand for the consumption variety of good ji,k is $q_{ji,k}^c = (\tilde{p}_{ji,k}/\tilde{P}_{i,k})^{-\epsilon_k} e_{i,k} Y_i$. The demand for the intermediate input variety of good ji,k is given by $q_{ji,k}^I = (\tilde{p}_{ji,k}/\tilde{P}_{i,k})^{-\epsilon_k} \sum_g \alpha_{i,kg} \mathcal{Y}_{i,g}$.

We can apply Theorem 1 along with the following steps to determine the optimal trade tax in this setup: Since there are no industry-specific factors of production, we can set $\gamma_{f,k} = 0$ for all k . The Cobb-Douglas assumption meanwhile ensures that $\zeta_{hf,k} = 0$. The within-industry CES demand system implies that $\varepsilon_{hf,k} = -1 - \epsilon_k \lambda_{ff,k}$. The assumption that the demand aggregator for final and intermediate goods are identical, entails that $\alpha_{f,g}^{h,k} = \alpha_{f,kg} \lambda_{hf,k}$. Plugging these values into Theorem 1 yields the following formula for optimal trade taxes in the [Caliendo and Parro \(2014\)](#) model:

$$\begin{aligned} 1 + t_{fh,k}^* &= 1 + \bar{t} \\ 1 + x_{hf,k}^* &= \left(\frac{1 + \epsilon_k \lambda_{ff,k}}{\epsilon_k \lambda_{ff,k} - \sum_g \alpha_{f,kg} r_{fh,g}} \right) (1 + t)^{-1}. \end{aligned} \quad (7)$$

In-line with our earlier discussion, the optimal export tax implied by the above formula is strictly lower than that implied by the baseline gravity model (Equation 6). However, the gains from optimal trade taxation are higher relative to the baseline model. This point is quantitatively established in Section 6 where we calibrate the above formulas to actual data.

4.2 Extension to a Multi-Country Setup

We now extend our optimal tax formulas to a multilateral setup with $N > 2$ countries. The optimal policy problem in this setup is a basic extension of 5: Country

i chooses $(N - 1)K$ import tax instruments, $\mathbf{t}_i = \{t_{ji,k}\}$, and $(N - 1)K$ export tax instruments, $\mathbf{x}_i = \{x_{ij,k}\}$, given the vector of applied ‘taxes in the rest of the world, \mathbf{t}_{-i} , and \mathbf{x}_{-i} .¹⁷ We can follow the same steps outlined earlier to derive sufficient statistics formulas for optimal trade taxes in this setup (see Appendix D):

$$1 + t_{ji,k}^* = \left(1 + \frac{\gamma_{j,k} r_{ji,k}}{1 - \sum_{l \neq i} \gamma_{j,k} r_{jl,k} \varepsilon_{jl,k}} \right) (1 + \bar{t})$$

$$1 + x_{ij,k}^* = \frac{\varepsilon_{ij,k}}{1 + \varepsilon_{ij,k} + \zeta_{ij,k} - \sum_g \sum_{j \neq i} \frac{\Lambda_{ji,g}}{\tilde{\Lambda}_{ij,k}} \tilde{\alpha}_{jj,g}^{i,k}} (1 + \bar{t})^{-1}. \quad (8)$$

In the above formula, $\zeta_{ij,k}$ accounts for cross-industry demand effects and is defined analogously to the cross-demand term specified under Theorem 1;¹⁸ $\Lambda_{ji,g} \equiv p_{ji,g} q_{ji,g} / \sum_{j \neq i} \sum_k p_{j,i,k} q_{j,i,k}$ and $\tilde{\Lambda}_{ij,g} \equiv \tilde{p}_{ij,g} q_{ij,g} / \sum_{j \neq i} \sum_k \tilde{p}_{ij,k} q_{ij,k}$ denote import and export shares; and $[\tilde{\alpha}_{jj,g}^{i,k}] = \alpha_j^i \alpha_j^j$ if $j \neq j$, with $[\tilde{\alpha}_{jj,g}^{i,k}] = \alpha_j^i$.

The optimal tax formula specified above differs from the baseline formula (presented under Theorem 1) in one basic aspect. The optimal export tax accounts for the fact that a tax on any individual export variety may be passed back to domestic consumers through multiple partners. The term $\tilde{\alpha}_{jj,g}^{i,k}$ in the denominator, accounts for the extent to which $x_{ij,k}$ is passed back to domestic consumers through country j ’s export of industry g goods. The higher the $\tilde{\alpha}_{jj,g}^{i,k}$ ’s from various j ’s, the lower the optimal export tax. The same formula also indicates that the burden of an export tax on one partner may be borne primarily by a third country that is not directly involved in the export transaction.

This subtle difference notwithstanding, the claims presented under Corollaries 1 and 2 readily extend to the multi-country case. As before, the optimal import taxes do not *explicitly* depend on the global input-output shares, which is a manifestation of the targeting principle. Absent diseconomies of scale, import taxes are also uniform and, therefore, redundant under the first-best scenario.¹⁹ We

¹⁷An added complication here is that a tax on good ji,k can alter the entire vector of national-level wages $\{w_1, \dots, w_N\}$. This complication can be handled by noting that $t_{ji,k}$ ’s effect on W_i has a first-order component that operates through changes in w_j/w_i and a second-order component that operates through changes in w_j/w_i (where $j \neq j$)—see Appendix D.

¹⁸Stated formally, $\zeta_{ij,k}$ is given by $[\zeta_{ij,k}]_k = \left(\left[\frac{\hat{\lambda}_{ij,g} \varepsilon_{ij,g}^{i,k}}{\tilde{\lambda}_{ij,k} \varepsilon_{ij,k}} \right]_{k,g}^{-1} - \mathbf{I}_K \right) \left[1 - \sum_g \sum_j \frac{\Lambda_{ji,g}}{\tilde{\Lambda}_{ij,k}} \tilde{\alpha}_{jj,g}^{i,k} \right]$.

¹⁹The uniformity of optimal tariffs across industries is independent of the first-order approximation discussed in Footnote 10—see Appendix D for details.

also demonstrated how the characterization under Theorem 1 readily extends to a multi-country setup.

5 The Non-Cooperative Trade Equilibrium

As noted earlier, the value of free trade agreements (FTAs) can be measured as the prevented welfare loss that would otherwise occur if countries adopted non-cooperative trade taxes. In this section we discuss how our earlier results can help characterize these losses. We do so by solving the non-cooperative tax equilibrium that will ensue after the dissolution of trade agreements.

Theorem 1 described Home's optimal non-cooperative tax schedule given the vector of taxes in the rest of the world. Section 4.2 extended this characterization to a setup where Country i sets optimal non-cooperative taxes against arbitrarily many trading partners. These taxes are characterized by Equation 8, and are an implicit function of applied taxes in the rest of the world. To formalize this dependence, let $t_{ji,k}^* = \mathcal{T}_{ji,k}(t_{-i}, x)$ and $x_{ij,k}^* = \mathcal{X}_{ij,k}(t, x_{-i})$ denote Country i 's optimal non-cooperative trade taxes as a function of applied taxes in the rest of the world (i.e., t_{-i} and x_{-i}). If trade agreements disintegrate and all countries set their trade taxes non-cooperatively, the counterfactual vector of trade taxes in the non-cooperative Nash equilibrium can be determined using the following system:²⁰

$$\begin{cases} x_{1j,k}^* = \mathcal{X}_{1j,k}(t^*, x_{-1}^*) \\ t_{j1,k}^* = \mathcal{T}_{j1,k}(t_{-1}^*, x^*) \\ \vdots \\ x_{Nj,k}^* = \mathcal{X}_{1N,k}(t^*, x_{-N}^*) \\ t_{jN,k}^* = \mathcal{T}_{jN,k}(t_{-N}^*, x^*) \end{cases} \quad (9)$$

Let t and x denote the *factual* vector of trade taxes, which are the outcome of existing agreements. The value of trade agreements for country i can, thus, be measured as the welfare change when moving from applied trade taxes to Nash trade taxes:

$$\text{Value of FTAs}_i = W_i(t, x, w) / W_i(t^*, x^*, w^*).$$

²⁰This approach is analogous to Grossman and Helpman's (1995) approach to characterizing Nash tariff in a partial equilibrium model with lobbying.

To take stock, we first derived sufficient statistics formulas for unilaterally optimal trade taxes. These formulas determine the inefficient *terms-of-trade-driven* restrictions on trade that will occur if trade agreements were to dissolve. With knowledge of these counterfactual trade restrictions, we can then measure the deadweight loss associated with unilateralism and the value of trade agreements that remedy these losses. Before mapping out theory to data, two additional remarks about this approach are in order.

The Role of Political Economy Considerations. The Nash taxes implied by System 9, do not necessarily coincide with the tax levels that will prevail without agreements. Instead, the counterfactually applied taxes will reflect the sum of politically-driven and terms-of-trade-driven restriction on trade (Grossman and Helpman (1995)). As argued by Bagwell and Staiger (1999), however, the politically optimal trade taxes are Pareto efficient. So, the only inefficiency trade agreements correct are terms-of-trade-driven restrictions. These are the exact restrictions identified by System 9, which also motivates the expression for “Value of FTAs_{*i*}.”

Sequential Trade Policy Game. In System 9, the best response of each country is derived taking taxes in the rest of the world as given. This approach is perhaps appropriate if non-cooperative taxes are adopted simultaneously. Alternatively, consider a Stackelberg game where one country is sufficiently large compared to the rest of the world and has a first-mover advantage. In that case, the first-mover may strategically choose non-uniform tariffs to influence export taxes in the rest of the world. This situation is conceptually interesting, but not representative. As we will see in the next section, most countries are effectively small open economies *vis à vis* the rest of the world.

6 Quantitative Analysis

In this section, we map our theoretical model to global trade and production data. Our analysis pursues two distinct objectives. First, we want to measure the consequences of unilateralism. Second, we wish to quantify the value of existing trade agreements and determine what fraction of it is driven by global production networks. In this process, we also highlight how our analytic tax formulas streamline

the computational analysis of trade policy.

6.1 Mapping Non-Cooperative Tax Formulas to Data

Below, we discuss how the non-cooperative optimal tax formulas specified under Theorem 1 can be mapped to data. With the aid of these formulas, the gains from trade agreements or the consequences of unilateral policies can be computed using a system of equations that depend on observables and trade elasticities. We demonstrate this point first with a baseline model that overlooks global input-output networks. We then move on to our main model that formally accounts for global input-output networks. Comparing the predictions of these two models will help us isolate the role of production networks.

Baseline Model without Input-Output Networks. We first consider the basic multi-industry gravity model without input-output networks as outlined in Section 4.1. To present our methodology succinctly, we focus our presentation here on the two-country model. The multi-country implementation is presented in Appendix E. Recall that the baseline gravity model features a Cobb-Douglas utility aggregator across industries paired with a CES demand aggregator within industries. This preference structure implies that $V(Y_i, \tilde{P}_i) = Y_i / \tilde{P}_i$, where the aggregate price index in economy $i \in \mathbb{C}$ is given by

$$\tilde{P}_i = \prod_{k \in \mathbb{K}} \left(\sum_{j \in \mathbb{C}} \chi_{ji,k} \tilde{p}_{ji,k}^{-\epsilon_k} \right)^{-e_{i,k}/\epsilon_k}. \quad (10)$$

In the above formulation, $\chi_{ji,k}$ accounts for policy invariant taste shifters and $\tilde{p}_{ij,k} = (1 + t_{ij,k})(1 + x_{ij,k})\tau_{ij,k}a_{i,k}w_i$. Consider a counterfactual policy change, whereby the vector of trade taxes changes from its applied rate $\{t_{ij,k}\}$, and $\{x_{ji,k}\}$ to a counterfactual rate, $\{t'_{ij,k}\}$, and $\{x'_{ji,k}\}$. In response to this policy change, let z' denote the counterfactual value of any variable z and let $\hat{z} = z'/z$ denote the corresponding change in that variable. The CES demand structure implies that

$$\hat{P}_{i,k} = \sum_j \left(\left[\hat{w}_j (1 + \hat{t}_{ji,k}) (1 + \hat{x}_{ji,k}) \right]^{-\epsilon_k} \lambda_{ji,k} \right)^{-1/\epsilon_k}$$

where $1 + \widehat{t}_{ji,k} \equiv (1 + t'_{ji,k}) / (1 + t_{ji,k})$ and $1 + \widehat{x}_{ij,k} = (1 + x'_{ij,k}) / (1 + x_{ij,k})$ denote the change in import and export taxes. The change in trade shares is then given by

$$\widehat{\lambda}_{ji,k} = \frac{\left[(1 + \widehat{t}_{ji,k})(1 + \widehat{x}_{ji,k})\widehat{w}_j \right]^{-\epsilon_k}}{\sum_n \lambda_{ni,k} \left[(1 + \widehat{t}_{ni,k})(1 + \widehat{x}_{ni,k})\widehat{w}_n \right]^{-\epsilon_k}}. \quad (11)$$

The new vector of wages and income levels in the counterfactual equilibrium are determined by the labor market clearing condition:

$$w'_i L_i = \sum_k \sum_j \left(\frac{\lambda'_{ij,k}}{(1 + x'_{ij,k})(1 + t'_{ij,k})} e_{j,k} Y'_j \right)$$

and the condition that total income equals wage income plus tax revenue:

$$Y'_i = w'_i L_i + \sum_k \sum_j \left(\frac{t'_{ji,k}}{1 + t'_{ji,k}} \lambda'_{ji,k} e_{i,k} Y'_i + \frac{x'_{ij,k}}{1 + x'_{ij,k}} \lambda'_{ij,k} e_{j,k} Y'_j \right).$$

Both of the above equations can be written in terms of changes, by noting that $Y'_i = \widehat{Y}_i Y_i$, $w'_i = \widehat{w}_i w_i$, and $\lambda'_{ij,k} = \widehat{\lambda}_{ij,k} \lambda_{ij,k}$. Likewise, the post-change tax rates in the above equations can be expressed as $t'_{ji,k} = (1 + \widehat{t}_{ji,k})(1 + t_{ji,k}) - 1$ and $x'_{ij,k} = (1 + \widehat{x}_{ij,k})(1 + x_{ij,k}) - 1$.

With the above background in mind, consider our counterfactual policy experiment of interest: One in which Home and Foreign simultaneously apply their optimal non-cooperative trade taxes. Following Equation 6, Home's optimal non-cooperative optimal taxes are given by: $t'_{fh,k} = t^*_{fh,k} = 0$ and $x'_{hf,k} = x^*_{fh,k} = 1/\epsilon_k \lambda'_{ff,k}$. Likewise, Foreign's non-cooperative optimal taxes are $t'_{hf,k} = t^*_{hf,k} = 0$ and $x'_{fh,k} = x^*_{fh,k} = 1/\epsilon_k \lambda'_{hh,k}$. Inserting these counterfactual tax changes into the formulas outlined earlier, we can compute $\widehat{\lambda}_{ij,k}$, \widehat{w}_i , and \widehat{Y}_i , as well as the welfare consequences that result from this policy change. The following proposition outlines this procedure.

Proposition 1. *Suppose the observed data is generated by the baseline gravity model without input-output networks, as outlined in Section 3. The consequences of moving from status-quo to the non-cooperative trade equilibrium can be fully characterized by solving*

the following system of equations

$$\left\{ \begin{array}{l} 1 + x_{ji,k}^* = 1 + \frac{1}{\epsilon_k \hat{\lambda}_{ii,k} \lambda_{ii,k}}; \quad 1 + t_{ji,k}^* = 1 \\ \hat{\lambda}_{ji,k} = \left[\frac{(1+t_{ji,k}^*)(1+x_{ji,k}^*)}{(1+t_{ji,k})(1+x_{ji,k})} \hat{w}_j \right]^{-\epsilon_k} \hat{P}_{i,k}^{\epsilon_k} \\ \hat{P}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[\frac{(1+t_{ji,k}^*)(1+x_{ji,k}^*)}{(1+t_{ji,k})(1+x_{ji,k})} \hat{w}_j \right]^{-\epsilon_k} \lambda_{ji,k} \right) \\ \hat{w}_i w_i L_i = \sum_k \sum_j \left[\frac{\hat{\lambda}_{ij,k} \lambda_{ij,k}}{(1+x_{ij,k}^*)(1+t_{ij,k}^*)} e_{j,k} \hat{Y}_j Y_j \right] \\ \hat{Y}_i Y_i = \hat{w}_i w_i L_i + \sum_k \sum_j \left(\frac{t_{ji,k}^*}{1+t_{ji,k}^*} \hat{\lambda}_{ji,k} \lambda_{ji,k} e_{i,k} \hat{Y}_i Y_i + \frac{x_{ij,k}^*}{1+x_{ij,k}^*} \hat{\lambda}_{ij,k} \lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j \right) \end{array} \right. ,$$

The above system solves $N(K + 2)$ independent unknowns, namely, $\{x_{ji,k}^*\}$, $\{\hat{w}_i\}$, and $\{\hat{Y}_i\}$, as a function of industry-level trade elasticities, ϵ_k , and the following set of observables: (i) applied trade taxes; (ii) within- and across-industry expenditure shares $\lambda_{ji,k}$, and $e_{i,k}$; (iii) wage income, $w_i L_i$; and (iv) total income, $Y_i = w_i L_i + \mathcal{R}_i$.

Solving the system specified by Proposition 1, determines the welfare consequences of dissolving the existing trade agreements as

$$\text{Value of FTAs}_i = \hat{W}_i^{-1} = \left(\frac{\hat{Y}_i}{\prod_k \hat{P}_{i,k}} \right)^{-1}.$$

The consequences of unilateralism can also be measured using a similar approach. To that end, the system specified by Proposition 1 should be modified so that taxes in one country are assigned their unilaterally optimal rate, whereas taxes in the rest of the world are assigned their applied rate.

Importantly, the aforementioned procedure is implementable without appealing to any global optimization routine or without knowledge of structural parameters like $\tau_{ij,k}$, $a_{i,k}$, or $\chi_{ij,k}$. To put our approach in perspective, compare it to the standard approach that involves a constrained global optimization subject to equilibrium constraints (Costinot and Rodríguez-Clare (2014); Ossa (2014)). The standard approach can be implemented using a nested fixed-point procedure, which is impractical with many industries (i.e., a high K). Or, alternatively, it can be implemented using the MPEC procedure in Su and Judd (2012), which requires specialized commercial solvers like SNOPT or KNITRO to attain credible results. The procedure outlined by Proposition 1 bypasses these challenges by reducing

the global optimization problem to a system of equations that is straightforward to solve.

Main Model with Input-Output Networks. Now, consider the multi-industry gravity model with input-output networks as presented in Section 4.1. In this extended model, preferences are still characterized by Equation 10; but the price of good ij, k depends on the entire vector of industry-level price indexes in country i :

$$\tilde{p}_{ij,k} = (1 + t_{ij,k})(1 + x_{ij,k})\tau_{ij,k}a_{i,k}w_i^{\bar{\alpha}_{i,k}} \prod_g \tilde{P}_{i,g}^{\alpha_{i,gk}}.$$

The change in expenditure shares in response to a change in trade taxes can be, accordingly, specified as

$$\hat{\lambda}_{ij,k} = \frac{\left[(1 + \widehat{t}_{ij,k})(1 + \widehat{x}_{ij,k})\hat{w}_i^{\bar{\alpha}_{i,k}} \prod_g \hat{P}_{i,g}^{\alpha_{i,gk}} \right]^{-\epsilon_k}}{\sum_n \lambda_{nj,k} \left[(1 + \widehat{t}_{nj,k})(1 + \widehat{x}_{nj,k})\hat{w}_n^{\bar{\alpha}_{n,k}} \prod_g \hat{P}_{n,g}^{\alpha_{n,gk}} \right]^{-\epsilon_k'}}$$

where the change in price indexes are given by

$$\hat{P}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[(1 + \widehat{t}_{ni,k})(1 + \widehat{x}_{ni,k})\hat{w}_n^{\bar{\alpha}_{n,k}} \prod_g \hat{P}_{n,g}^{\alpha_{n,gk}} \right]^{-\epsilon_k} \lambda_{ji,k} \right).$$

In the counterfactual equilibrium that arises after the tax change, gross industry-level revenues can be calculated as the sum of sales net of taxes:

$$\mathcal{Y}'_{i,k} = \sum_j \left(\frac{\lambda'_{ij,g}}{(1 + x'_{ij,g})(1 + t'_{ij,g})} E'_{j,k} \right),$$

where $E'_{i,k} \equiv e_{i,k}Y'_i + \sum_g \alpha_{i,gk}\mathcal{Y}'_{i,g}$ denotes gross expenditure on industry k in country i . Net consumption income is equal to wage income plus tax revenues:

$$Y'_i = w'_i L_i + \sum_k \sum_j \left(\frac{t'_{ji,k}}{1 + t'_{ji,k}} \lambda'_{ji,k} E'_{i,k} + \frac{x'_{ij,k}}{1 + x'_{ij,k}} \lambda'_{ij,k} E'_{j,k} \right).$$

Wages income in the above expression is itself equal to the sum of labor compensation across all industries: $w'_i L_i = \sum_k \bar{\alpha}_{i,k} \mathcal{Y}'_{i,k}$. As before we can write the

above equilibrium conditions in changes by noting that $\lambda'_{ji,k} = \hat{\lambda}_{ji,k}\lambda_{ji,k}$, $w'_i = \hat{w}_i w_i$, $\mathcal{Y}'_{i,k} = \hat{\mathcal{Y}}_{i,k}\mathcal{Y}_{i,k}$, and $Y'_i = \hat{Y}_i Y_i$.

To compute the gains from existing trade agreements, we need to apply the above procedure to a counterfactual tax change whereby Home and Foreign simultaneously apply their optimal non-cooperative taxes. Specifically (as specified by Equation 7) Home applies $1 + x'_{hf,k} = 1 + x^*_{hf,k} = \frac{1 + \epsilon_k \lambda'_{ff,k}}{\epsilon_k \lambda'_{ff,k} - \sum_g \alpha_{f,k,g} r'_{fh,g}}$ and $t'_{fh,k} = t^*_{fh,k} = 0$, while Foreign analogously applies $1 + x'_{fh,k} = 1 + x^*_{fh,k} = \frac{1 + \epsilon_k \lambda'_{hh,k}}{\epsilon_k \lambda'_{hh,k} - \sum_g \alpha_{h,k,g} r'_{hf,g}}$ and $t'_{hf,k} = t^*_{hf,k} = 0$. The following proposition summarizes these arguments.

Proposition 2. *Suppose the observed data is generated by a multi-industry gravity model with global input-output networks, as outlined in Section 4.1. The consequences of moving from status-quo to the non-cooperative trade equilibrium can be fully determined by solving the following system of equations*

$$\left\{ \begin{array}{l} 1 + x^*_{ji,k} = \frac{1 + \epsilon_k \hat{\lambda}_{ii,k} \lambda_{ii,k}}{\epsilon_k \hat{\lambda}_{ii,k} \lambda_{ii,k} - \sum_g \alpha_{j,k,g} \hat{r}_{ij,g} r_{ij,g}}; \quad 1 + t^*_{ji,k} = 1 \\ \hat{\lambda}_{ji,k} = \left[\frac{(1 + t^*_{ji,k})(1 + x^*_{ji,k})}{(1 + t_{ji,k})(1 + x_{ji,k})} \hat{w}_j \prod_g \hat{P}_{i,g}^{\alpha_{i,gk}} \right]^{-\epsilon_k} \hat{P}_{i,k}^{\epsilon_k} \\ \hat{P}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[\frac{(1 + t^*_{ji,k})(1 + x^*_{ji,k})}{(1 + t_{ji,k})(1 + x_{ji,k})} \hat{w}_j \prod_g \hat{P}_{i,g}^{\alpha_{i,gk}} \right]^{-\epsilon_k} \lambda_{ji,k} \right) \\ \hat{r}_{ji,k} r_{ji,k} = \frac{\hat{\lambda}_{ji,k} \lambda_{ji,k} \hat{E}_{i,k} E_{i,k}}{(1 + x^*_{ji,k})(1 + t^*_{ji,k})} / \sum_n \left(\frac{\hat{\lambda}_{jn,k} \lambda_{jn,k} \hat{E}_{n,g} E_{n,g}}{(1 + x^*_{jn,k})(1 + t^*_{jn,k})} \right) \\ \hat{E}_{i,k} E_{i,k} = e_{i,k} \hat{Y}_i Y_i + \sum_g \sum_j \alpha_{i,gk} \frac{\hat{\lambda}_{ij,g} \lambda_{ij,g} \hat{\mathcal{Y}}_{j,g} \mathcal{Y}_{j,g}}{(1 + x^*_{ij,g})(1 + t^*_{ij,g})} \\ \hat{w}_i w_i L_i = \sum_k \sum_j \bar{\alpha}_{i,k} \frac{\hat{\lambda}_{ij,k} \lambda_{ij,k} \hat{\mathcal{Y}}_{j,k} \mathcal{Y}_{j,k}}{(1 + x^*_{ij,k})(1 + t^*_{ij,k})} \\ \hat{Y}_i Y_i = \hat{w}_i w_i L_i + \sum_k \sum_j \left(\frac{t^*_{ji,k}}{1 + t^*_{ji,k}} \hat{\lambda}_{ji,k} \lambda_{ji,k} \hat{E}_{i,k} E_{i,k} + \frac{x^*_{ij,k}}{1 + x^*_{ij,k}} \hat{\lambda}_{ij,k} \lambda_{ij,k} \hat{E}_{j,k} E_{j,k} \right) \end{array} \right. ,$$

The above system solves $N(3K + 2)$ independent unknowns, namely, $\{x^*_{ji,k}\}$, $\{\hat{P}_{i,k}\}$, $\{\hat{w}_i\}$, $\{\hat{Y}_i\}$, and $\{\hat{\mathcal{Y}}_{i,k}\}$, as a function of industry-level trade elasticities, ϵ_k , and the following set of observables: (i) gross expenditure and revenue shares, $\lambda_{ji,k}$, and $r_{ij,k}$; (ii) industry-level consumption shares, $e_{i,k}$; (iii) wage income, $w_i L_i$; (iv) final good expenditure, $Y_i = w_i L_i + \mathcal{R}_i$; (v) gross industry-level revenue and expenditure levels, $\mathcal{Y}_{i,k}$ and $E_{i,k}$; and (vi) input-output shares, $\alpha_{i,k,g}$.

As with the baseline model, the system specified by Proposition 2 also

determines the gains from trade agreements, which are given by $\hat{W}_i^{-1} = \left(\hat{Y}_i / \prod_k \hat{P}_{i,k} \right)^{-1}$. Finally, an amended version of Proposition 2 where taxes in one country are assigned their unilaterally optimal rate, whereas taxes in the rest of the world are assigned their applied rate, can determine the consequences of unilateralism.

6.2 Data Description

Our main data source is the 2014 edition of the World Input-Output Database (WIOD, [Timmer et al. 2012](#)). The WIOD database covers 56 industries and 42 countries that account for more than 85% of world GDP, plus an aggregate of the rest of the world. The countries in the sample include all 27 members of the European Union (EU) and 15 other major economies: Australia, Brazil, Canada, China, India, Indonesia, Japan, Mexico, Norway, Russia, South Korea, Switzerland, Taiwan, Turkey, and the United States—see [Table 4](#) in the appendix for a full list of countries. Following [Costinot and Rodríguez-Clare \(2014\)](#), we aggregate the original 56 industries in the WIOD into 16 industrial categories, which are listed in [Table 1](#). Finally, we make the WIOD database consistent with our theoretical model by purging it from trade imbalances. In this process, we closely follow the methodology in [Costinot and Rodríguez-Clare \(2014\)](#) who apply [Dekle et al.’s \(2007\)](#) hat-algebra methodology to purge the 2008 edition of the WIOD.²¹

We take data on applied tariffs, $t_{j,i,k}$, from the United Nations Statistical Division, Trade Analysis and Information System (UNCTAD-TRAINS). The 2014 version of the UNCTAD-TRAINS reports the *simple tariff line average* of the *effectively applied tariff* (AHS) across 31 two-digit (in ISIC rev.3) sectors, 185 importers, and 243 export partners. When tariff data are missing in a given year, we use tariff data for the nearest available year, giving priority to earlier years. We aggregate the UNCTAD-TRAINS data into individual WIOD industries, following the methodology in [Kucheryavyy et al. \(2016\)](#). Since individual EU member countries are not represented in the UNCTAD-TRAINS data during the 2000-2014 period, our analysis treats the 27 EU members as one taxing authority.

To map our model to the WIOD and UNCTAD-TRAINS datasets, we treat good

²¹A similar approach is also applied by [Ossa \(2014\)](#) to eliminate trade imbalances from the GTAP database.

ij, k as a good pertaining to WIOD industry k that is supplied by country i to market j . Under this interpretation, our data reports the following information:

- i. The gross expenditure on good ij, k : $\tilde{p}_{ij,k}q_{ij,k}$
- ii. The applied import tax on good ij, k : $t_{ij,k}$
- iii. The share of industry k inputs used in industry g 's output in country i : $\alpha_{i,k,g}$.

Using the above data points, we can construct the within-industry gross expenditure shares as

$$\lambda_{ij,k} = \frac{\tilde{p}_{ij,k}q_{ij,k}}{\sum_n \tilde{p}_{in,k}q_{in,k}}.$$

Similarly and assuming $x_{ij,k} = 0$, the within-industry gross output shares can be constructed as

$$r_{ij,k} = \frac{\tilde{p}_{ij,k}q_{ij,k}/(1+t_{ij,k})}{\sum_n \tilde{p}_{in,k}q_{in,k}/(1+t_{ij,k})}.$$

The total gross output of industry k in country i can be calculated as the sum of gross sales net of taxes:

$$\mathcal{Y}_{i,k} = \sum_n [\tilde{p}_{in,k}q_{in,k}/(1+t_{ij,k})],$$

With information on gross output, country i 's net spending on final goods can be calculated as the gross minus intermediate input expenditure:

$$Y_i = \sum_j \sum_k (\tilde{p}_{ji,k}q_{ji,k}) - \sum_k \sum_g (\alpha_{i,k,g}\mathcal{Y}_{i,g}).$$

The industry-level consumption shares are, accordingly, given by

$$e_{i,k} = \left[\sum_j (\tilde{p}_{ji,k}q_{ji,k}) - \sum_g (\alpha_{i,k,g}\mathcal{Y}_{i,g}) \right] / Y_i.$$

Finally, wage income can be calculated as the sum of labor compensation across all industries: $w_i L_i = \sum \bar{\alpha}_{i,k}\mathcal{Y}_{i,k}$, where $\bar{\alpha}_{i,k} = 1 - \sum_g \alpha_{i,g,k}$.²²

²²To be clear, when applying the above data points to Propositions 1 and 2, we need to assign one country as Home and aggregate the remaining countries into the rest of world.

Estimating the Industry-Level Trade Elasticities We estimate the industry-level trade elasticities using the triple-difference estimator developed by [Caliendo and Parro \(2014\)](#). To present this procedure, note that the multi-industry gravity model (*with* or *without* input-output networks) predicts the following formulation for trade shares:

$$\lambda_{ij,k} = \Phi_{i,k} \Omega_{j,k} \tau_{ij,k}^{-\epsilon_k} (1 + t_{ij,k})^{-\epsilon_k},$$

where $\Phi_{i,k} \equiv a_{i,k}^{-\epsilon_k} w_{i,k}^{-\bar{\alpha}_{i,k} \epsilon_k} \prod_g \bar{P}_{i,g}^{-\alpha_{i,gk} \epsilon_k}$ and $\Omega_{j,k} \equiv \sum_n [\tau_{nj,k}^{-\epsilon_k} (1 + t_{nj,k})^{-\epsilon_k} \Phi_{n,k}]$ can be viewed as *exporter* \times *industry* and *importer* \times *industry* fixed effects. Suppose the iceberg trade cost, $\ln \tau_{ij,k} = \ln d_{ij,k} + \varepsilon_{ij,k}$, is composed of two components: (i) a systematic and symmetric component, $d_{ij,k} = d_{ji,k}$, that accounts for the effect of distance, common language, and common border, and (ii) a random disturbance term, $\varepsilon_{ij,k}$, that represents deviation from symmetry. Using this decomposition, we can produce the following estimating equation for any triplet (j, i, n) :

$$\ln \frac{\lambda_{ji,k} \lambda_{in,k} \lambda_{nj,k}}{\lambda_{ij,k} \lambda_{ni,k} \lambda_{jn,k}} = -\epsilon_k \ln \frac{(1 + t_{ji,k})(1 + t_{in,k})(1 + t_{nj,k})}{(1 + t_{ij,k})(1 + t_{ni,k})(1 + t_{jn,k})} + \tilde{\varepsilon}_{jin,k},$$

where $\tilde{\varepsilon}_{jin,k} \equiv \epsilon_k (\varepsilon_{ij,k} - \varepsilon_{ji,k} + \varepsilon_{in,k} - \varepsilon_{ni,k} + \varepsilon_{nj,k} - \varepsilon_{jn,k})$. Using the above estimating equation, we can attain unbiased and consistent estimates for ϵ_k under the identifying assumption that $\text{Corr}(t_{ji,k}, \varepsilon_{ji,k}) = 0$. We estimate the above equation separately for each industry, using data on $\{\lambda_{ji,k}\}$ from the 2014 version of the WIOD and data on $\{t_{ji,k}\}$ from the UNCTAD-TRAINS database. The estimation results are reported in Table 1 and broadly align with those produced by [Caliendo and Parro \(2014\)](#), who use data on a smaller sample of countries from 1993.

6.3 The Consequences of Unilateralism

As an intermediate step we discuss the consequences of unilateralism. For each country, the computed gains from a unilateral adoption of optimal trade taxes are reported in Table 2.²³ The same table also reports the externality imposed by these taxes on the rest of the world—all in terms of a percentage change in real GDP. The

²³Each iteration of our analysis treats one of the countries listed in Table 2 as the Home country and aggregates the remaining countries into one Foreign economy. Alternatively, we can use the multilateral formulas presented in Section 4.2 to avoid aggregating other countries into one Foreign economy. Doing so delivers welfare gains that are practically indistinguishable from our benchmark results.

Table 1: List of industries and estimated trade elasticities.

Number	Description	ϵ_k	std. err.	Obsv.
1	Crop and animal production, hunting Forestry and logging Fishing and aquaculture	0.93	0.19	12,341
2	Mining and Quarrying			
3	Food, Beverages and Tobacco	0.53	0.13	12,300
4	Textiles, Wearing Apparel and Leather	2.71	0.51	12,341
5	Wood and Products of Wood and Cork	5.64	0.87	12,183
6	Paper and Paper Products Printing and Reproduction of Recorded Media	4.65	1.49	12,300
7	Coke, Refined Petroleum and Nuclear Fuel	13.38	1.94	9,538
8	Chemicals and Chemical Products Basic Pharmaceutical Products	2.36	0.91	12,300
9	Rubber and Plastics			
10	Other Non-Metallic Mineral	1.51	0.89	12,341
11	Basic Metals Fabricated Metal Products			
12	Computer, Electronic and Optical Products Electrical Equipment	4.07	1.02	12,341
13	Machinery and Equipment n.e.c	5.65	1.34	12,341
14	Motor Vehicles, Trailers and Semi-Trailers Other Transport Equipment	2.70	0.45	12,341
15	Furniture; other Manufacturing	2.04	0.59	12,341
16	All Service-Related Industries (WIOD Industry No. 23-56)	3.80	0.84	12,341

Note: This table estimates the industry-level trade elasticities using the [Caliendo and Parro \(2014\)](#) methodology. The original WIOD industry classification features 56 industries, which we aggregate into 16 industrial categories.

first two columns in [Table 2](#) report the gains implied by the baseline gravity model that overlooks global production networks. The second two columns report the gains implied by our main model that accounts for global production networks. In both cases, we separately report the consequences of *first-best* unilateral export taxes and *second-best* unilateral import taxes.

The numbers reported in Table 2 indicate that for a non-cooperative country first-best export taxes deliver welfare gains that are (on average) 50% larger than second-best import taxes. On the flip side, non-cooperative export taxes also inflict a greater negative externality on the rest of the world. This observation speaks to the weight of *terms-of-trade* versus *political economy* considerations in the governments' objective function. Analyzing import taxes in isolation implies a greater role for terms-of-trade considerations (as estimated by Goldberg and Maggi (1999)).²⁴ But in that case, our findings indicate that governments should be applying export taxes more vigorously as they deliver sizable terms-of-trade gains and are not prohibited by the WTO. That export taxes are underexploited implies that export-oriented industries have a *greater-than-previously-estimated* political weight in the governments objective function.

Second-best non-cooperative import taxes are only modestly more effective at mimicking export taxes in the presence of global production networks. As argued earlier, this outcome is an artifact of an innate tension between two welfare considerations: (a) to maintain allocative efficiency in the domestic economy, non-cooperative governments should not tax highly-differentiated intermediate inputs, but (ii) to exploit export market power with second-best import taxes, governments should levy taxes on highly-differentiated intermediate imports. This tension renders import taxes as an ineffective second-best substitute for export taxes.

Global production networks multiply the gains from non-cooperative taxation and the externality of such policies on the rest of the world by 50%. Intuitively, in the presence of global production networks, the excess burden per dollar of tax revenue falls relatively more on the rest of the world. Consider, for instance, an export tax on rare earth minerals that are used as an input in smartphone production abroad. A fraction of this export tax is borne on transactions involving foreign producers and foreign consumers of smartphones. So, through the global production network, the tax-imposing country can effectively raise revenue on transactions occurring outside its jurisdiction—i.e., global production networks present governments with *extra-territorial taxing power*.

²⁴To elaborate, Goldberg and Maggi (1999) study the cross-industry variation in import taxes and find that “the weight of welfare in the government’s objective function is many times larger than the weight of contributions.”

Table 2: The Welfare Consequences of Unilateralism (% change in real GDP)

Country	Baseline Model				Model with IO Networks			
	Export taxation		Import taxation		Export taxation		Import taxation	
	$\% \Delta W_i$	$\% \Delta W_{-i}$	$\% \Delta W_i$	$\% \Delta W_{-i}$	$\% \Delta W_{-i}$	$\% \Delta W_i$	$\% \Delta W_i$	$\% \Delta W_{-i}$
AUS	1.7%	-0.1%	1.3%	-0.0%	3.3%	-0.1%	2.2%	-0.1%
EU	0.7%	-0.4%	0.5%	-0.3%	1.3%	-0.8%	1.0%	-0.6%
BRA	1.2%	-0.1%	0.9%	-0.0%	2.1%	-0.1%	1.4%	-0.1%
CAN	1.9%	-0.1%	1.2%	-0.1%	3.6%	-0.2%	2.3%	-0.1%
CHE	1.5%	-0.0%	1.2%	-0.0%	3.0%	-0.1%	2.2%	-0.0%
CHN	0.5%	-0.3%	0.4%	-0.2%	1.6%	-0.5%	1.4%	-0.4%
IDN	1.9%	-0.0%	1.3%	-0.0%	3.7%	-0.1%	2.3%	-0.0%
IND	0.8%	-0.0%	0.4%	-0.0%	1.6%	-0.1%	0.9%	-0.1%
JPN	0.7%	-0.1%	0.5%	-0.1%	1.4%	-0.2%	1.1%	-0.1%
KOR	1.2%	-0.1%	0.9%	-0.0%	3.0%	-0.1%	2.4%	-0.1%
MEX	1.7%	-0.0%	1.1%	-0.0%	3.0%	-0.1%	0.9%	-0.0%
NOR	3.0%	-0.0%	2.3%	-0.0%	5.2%	-0.1%	3.3%	-0.0%
RUS	1.6%	-0.1%	1.0%	-0.0%	3.2%	-0.2%	1.9%	-0.1%
TUR	1.6%	-0.0%	1.1%	-0.0%	3.4%	-0.1%	1.8%	-0.0%
TWN	1.7%	-0.0%	1.3%	-0.0%	3.8%	-0.1%	3.1%	-0.0%
USA	0.6%	-0.3%	0.4%	-0.2%	0.9%	-0.5%	0.6%	-0.4%
Average	1.39%	-0.10%	0.99%	-0.07%	2.75%	-0.19%	1.80%	-0.14%

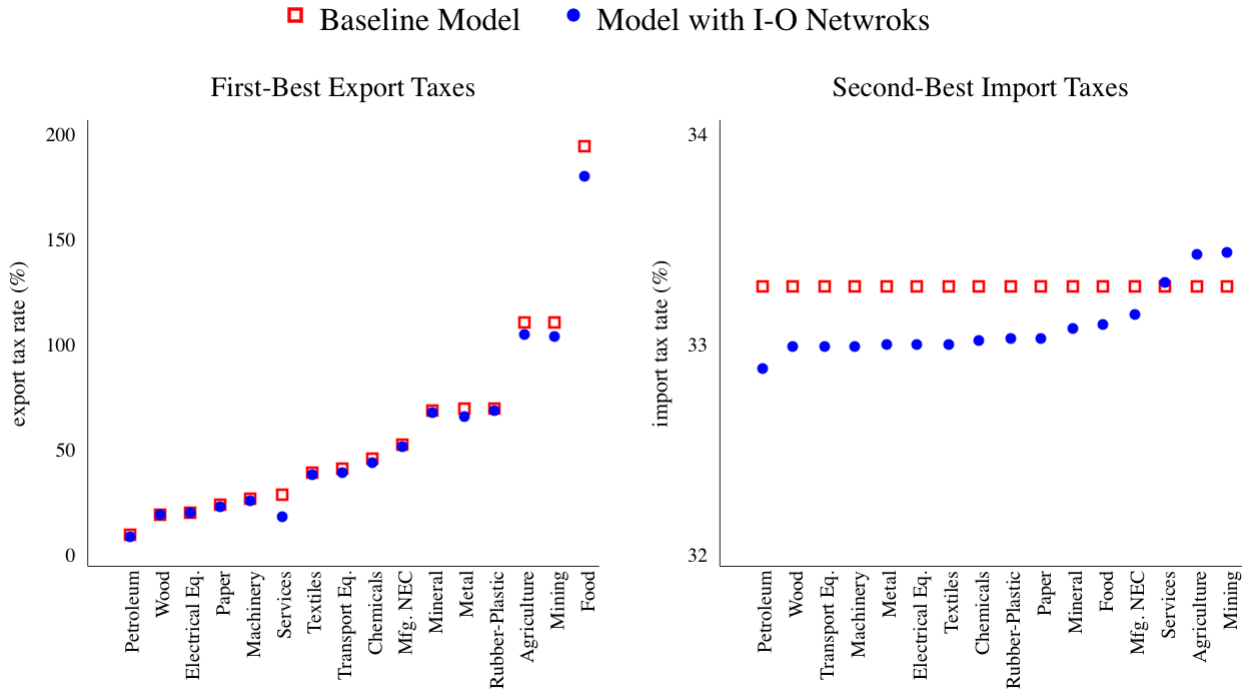
Note: This paper reports the change in a country's welfare from applying its optimal trade taxes. The baseline model is the perfectly competitive multi-industry gravity model 'a la [Costinot et al. \(2011\)](#). The model with IO networks is the perfectly competitive multi-industry gravity model with input-output networks 'a la [Caliendo and Parro \(2014\)](#). The gains from unilateral export taxation are solved using an amended version of Propositions 1 and 2, where one country acts non-cooperatively and the rest of the world retains their applied tax rates. The gains from (second-best) import taxes concern a scenario where export taxes are inapplicable.

6.3.1 Have Global Production Networks Reduced Protectionist Motives?

Global production networks influence the consequences of unilateralism (a) by changing the non-cooperative tax level in each country, and (b) by multiplying welfare consequences conditional on the chosen tax level. The latter channel is relatively well understood and echos the results from the standard gains from trade literature. The former channel is new and merits further elaboration. In fact, as noted in the introduction, some experts argue that existing trade agreements are obsolete because countries that are immersed in global production networks have little (terms-of-trade-driven) incentive to raise trade restrictions.

Figure 1 illustrates the optimal non-cooperative trade tax schedule for the European Union (E.U.). These are the non-cooperative E.U. taxes given the vector of applied taxes by trading partners. As implied by Theorem 1, the optimal export tax is lower once we account for input-output networks. The intuition is that export taxes cascade through the input-output network, and a fraction of the tax is ultimately passed back to domestic consumers. The optimal export tax is set lower to mitigate these adverse cascading effects. This reduction is especially meaningful in more upstream industries like Metal, Mining, and Services.

Figure 1: The European Union’s Optimal Non-Cooperative Trade Taxes



Note: This graph plots the optimal trade tax for the E.U. economy against the rest of the world. Industries are ordered by the trade elasticity estimates reported in Table 1. The left panel reports the first-best export taxes, normalizing the redundant import taxes to zero. The hollow squares correspond to the optimal export taxes predicted by the multi-industry competitive gravity model without input-output networks. Second best import taxes correspond to a scenario where export taxes are prohibited.

Second-best non-cooperative import taxes are uniform if we do not account for input-output networks. Otherwise, they vary across industries in a way that partially resembles the optimal first-best export taxes. In particular, the optimal import tax is higher on low trade elasticity industries. But, unlike first-best export

taxes, it is also relatively higher on upstream industries. The optimal import tax is, for example, higher on Mining and Service industries that are both relatively more upstream and have a relatively low trade elasticity. The Petroleum industry, in comparison, is relatively upstream but exhibits the highest trade elasticity of all industries. As a result, the second-best import tax on the Petroleum industry is lower than others.²⁵

Importantly, these patterns can reconcile the terms-of-trade thesis with evidence that governments assign a low political weight to protecting upstream industries because consumers of intermediate inputs are more organized (see [Shapiro \(2019\)](#)). To see this, note that applied tariffs -in absence of external restrictions- convey information about the sum of *political economy* and *terms-of-trade* considerations ([Grossman and Helpman \(1992\)](#)). Correspondingly, the political weight of industry-level protection can be inferred as the wedge between the socially optimal tariff and the applied tariff rate in each industry. A high political weight, under this approach, indicates more lobbying for protective tariffs a given industry.

Suppose we infer political weights from optimal tariffs that do *not* internalize global production networks. In that case, we will find no systematic relationship between the level of upstreamness and political weights, which defies the conventional wisdom.²⁶ However, if we apply our theory that implies a higher optimal tariff on upstream industries, we will infer a low political weight for protection in these industries, which aligns with the aforementioned evidence in [Shapiro \(2019\)](#).

Above all, [Figure 1](#) indicates that the globalization of production networks

²⁵The fact that second-best import taxes vary modestly across industries reflects a tension between *misallocation-worsening* and *terms-of-trade-improving* effects. To explain this tension again, consider the extreme case where all imported intermediates are re-exported without processing in the domestic economy. In that case, the second-best import tax schedule converge to the first-best export tax schedule. Because, in this extreme case, neither tax instrument distorts input choices in the domestic economy. This is far from what happens in reality, though—as is clearly manifested in [Figure 1](#). In practice, industries like Chemicals and Mining that should be targeted with second-best import taxes are both high-value-added and more upstream. So, disrupting them with taxes will be quite detrimental to allocative efficiency.

²⁶A notable example undertaking this approach is [Ossa \(2014\)](#). He infers political weights using optimal tariffs implied by the multi-industry [Krugman \(1980\)](#) model. Optimal tariffs in the aforementioned model are determined primarily by the industry-level trade elasticities (see [Lashkaripour and Lugovskyy \(2019\)](#)). Accordingly, [Ossa \(2014\)](#) estimates political weights that are strongly increasing in an industry's trade elasticity but are less dependent on an industry's level of upstreamness.

has *not* reduced the governments' incentive to raise trade restrictions. This outcome is driven by all economies, including the E.U., being effectively small open economies in relation to the rest of the world. In that case, E.U. taxes on intermediate inputs have a negligible (albeit non-zero) effect on prices in the rest of the world. Consequently, the optimal non-cooperative taxes implied with and without global production networks are rather similar and resemble the optimal tax of a small open economy supplying differentiated products.

6.4 The Value of Trade Agreements

As a final step, we quantify the value of existing trade agreements for the global economy. Recall that the value of these agreements is driven by their ability to remedy *non-political* trade restrictions that ensue in the non-cooperative Nash equilibrium. The extent of these counterfactual restrictions can be measured using the system of equations specified under Propositions 1 and 2.²⁷

Table 3 reports the value of existing trade agreements for each country, and contrasts it to the value implied by the baseline model that overlooks global production networks. In all cases, the figures are reported under two different scenarios: First, a scenario where countries can freely raise all trade tax instruments, in which case agreements remedy the first-best non-cooperative export restrictions. Second, a scenario where countries cannot raise export taxes due to institutional barriers like those in the U.S. constitution. In this latter case, agreements remedy the second-best non-cooperative import restrictions.

To put the estimated gains from trade agreements in perspective, the last two columns of Table 3 report the overall gains from trade relative to *autarky*. These gains are calculated using the sufficient statistics formulas presented by Costinot and Rodríguez-Clare (2014).

The value of trade agreements, as far as remedying unilateral export restriction is concerned, is around 3.5% of the real GDP for the average country. This amounts to 30% of the overall gains from trade. In other words, if existing trade agreements disintegrate, the average country will lose close to a third of the underlying gains from trade. To put a dollar value on these effects, trade agreements add \$2.7 trillion

²⁷See Appendix E for how the system specified by Propositions 1 and 2 extends to an environment with arbitrarily many countries.

Table 3: The Gains from Trade Agreements (% change in real GDP)

Country	Gains from Remedying Export Restrictions		Gains from Remedying Import Restrictions		Overall Gains From Trade	
	Baseline	IO Networks	Baseline	IO Networks	Baseline	IO Networks
AUS	0.1%	0.7%	0.5%	0.9%	5.7%	0.9%
EU	1.1%	2.1%	1.1%	1.2%	3.0%	1.2%
BRA	0.2%	0.5%	0.4%	0.4%	2.8%	0.4%
CAN	2.7%	4.4%	2.1%	2.7%	10.7%	2.7%
CHE	3.3%	5.1%	2.8%	3.5%	11.8%	3.5%
CHN	0.9%	2.7%	0.9%	1.3%	1.9%	1.3%
IDN	0.5%	1.3%	0.8%	0.9%	4.7%	0.9%
IND	3.1%	4.3%	1.9%	1.9%	2.5%	1.9%
JPN	3.2%	4.7%	1.9%	1.9%	3.5%	1.9%
KOR	5.4%	8.6%	3.0%	3.8%	4.4%	3.8%
MEX	2.0%	3.0%	2.0%	2.2%	19.8%	2.2%
NOR	0.2%	0.3%	0.4%	0.9%	8.8%	0.9%
RUS	1.2%	2.4%	1.4%	2.1%	11.1%	2.1%
TUR	2.1%	3.5%	2.0%	2.2%	7.9%	2.2%
TWN	6.0%	10.2%	4.4%	5.4%	8.1%	5.4%
USA	1.0%	1.8%	1.0%	1.0%	2.9%	1.0%
Average	2.06%	3.48%	1.66%	2.01%	6.84%	12.70%

Note: This paper reports the change in a country's welfare from applying its optimal trade taxes. The baseline model is the perfectly competitive multi-industry gravity model 'a la [Costinot et al. \(2011\)](#). The model with IO networks is the perfectly competitive multi-industry gravity model with input-output networks 'a la [Caliendo and Parro \(2014\)](#). The value of trade agreements at remedying export taxation are solved using the systems specified by Propositions 1 and 2. The gains from remedying import taxation concern a scenario where non-cooperative export taxes are inapplicable.

to global GDP , which is equal to the size of France's economy.

The above numbers can be viewed as an upper bound on the value of trade agreements. As note earlier, governments may be obliged to erect only import restriction if existing agreements dissolve. In that case, by remedying non-cooperative import restrictions, existing trade agreements add 2% to the average country's real GDP (see fourth column in Table 3). Altogether, these effects contribute a total of \$1.6 trillion to global GDP. This second figure can be viewed as a lower bound on the value of existing trade agreements.

Around 41% of the estimated value of trade agreements is driven by their pro-

motion of global production networks. This number can be inferred by comparing the value implied by the baseline model that overlooks global production networks and the main model that accommodates them. In other words, existing estimates that overlook global production networks may understate the value of trade agreements by more than 40%. More importantly, these numbers suggest that the rise of global production networks has rendered trade agreements more valuable than ever. This is a crucial observation in a time when the tides have turned against free trade agreements like the WTO.

7 Concluding Remarks

Our understanding of global trade has improved drastically over the past two decades; thanks in part to the emergence of tractable quantitative trade models that admit a rich geography of trade costs and global production networks. Surprisingly, these canonical models have had little impact on how we think about trade policy and trade agreements. To address this gap, we presented a full analytical characterization of optimal non-cooperative trade policy in the aforementioned class of quantitative trade models. We demonstrated that the globalization of production networks have barely reduced the terms-of-trade-driven incentives to restrict trade; but it has increased the cost of such restrictions quite remarkably. Accordingly, we measured that a disintegration of existing agreements will erase 30% of the total gains from trade.

Our theory presents researchers with a new benchmark to study the political economy of trade taxation in the age of global production networks. The current practice in the literature is to infer political economy motives by contrasting the applied trade taxes with their socially optimal level. Under this approach, our general theory of optimal trade policy will enable researchers to credibly estimate the political weight of protecting upstream versus downstream industries. Likewise, our theory can shed fresh light on the relative importance of *terms-of-trade* versus *political economy* considerations in the conduct of trade policy. Previous estimates that focus only on import taxation indicate that political considerations are dwarfed by terms-of-trade considerations. Our theory suggests that accounting for export taxes may revise this prediction.

On a broader level, our findings highlight an important dilemma facing policy-

makers in the current political climate. On one hand, critics argue that trade agreements infringe on national sovereignty, and limit governments' ability to tackle domestic distortions, inequality, and environmental issues. On the other hand, we find that dissolving existing trade agreements will result in extensive losses from disrupting global production networks. Future research can perhaps resolve this dilemma, at the core of which lies a basic question: Can existing agreements be reformed to maintain international cooperation in trade policies but with less infringement on national sovereignty?

References

- Alvarez, F. and R. E. Lucas (2007). General equilibrium analysis of the eaton–kortum model of international trade. *Journal of monetary Economics* 54(6), 1726–1768.
- Amiti, M., S. J. Redding, and D. E. Weinstein (2019). The impact of the 2018 tariffs on prices and welfare. *Journal of Economic Perspectives* 33(4), 187–210.
- Anderson, J. E. and Y. V. Yotov (2016). Terms of trade and global efficiency effects of free trade agreements, 1990–2002. *Journal of International Economics* 99, 279–298.
- Bagwell, K. and R. W. Staiger (1999). An economic theory of gatt. *American Economic Review* 89(1), 215–248.
- Bagwell, K., R. W. Staiger, and A. Yurukoglu (2018). Quantitative analysis of multi-party tariff negotiations. Technical report, National Bureau of Economic Research.
- Baier, S. L. and J. H. Bergstrand (2007). Do free trade agreements actually increase members' international trade? *Journal of international Economics* 71(1), 72–95.
- Baqae, D. and E. Farhi (2019). Networks, barriers, and trade. Technical report, National Bureau of Economic Research.
- Blanchard, E. J., C. P. Bown, and R. C. Johnson (2016). Global supply chains and trade policy. Technical report, National Bureau of Economic Research.
- Bown, C., P. Conconi, A. Erbahar, and L. Trimarchi (2020). Trade protection along supply chains.
- Caliendo, L., R. C. Feenstra, J. Romalis, and A. M. Taylor (2015). Tariff reductions, entry, and welfare: Theory and evidence for the last two decades. Technical report, National Bureau of Economic Research.

- Caliendo, L. and F. Parro (2014). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, rdu035.
- Carballo, J., K. Handley, and N. Limão (2018). Economic and policy uncertainty: Export dynamics and the value of agreements. Technical report, National Bureau of Economic Research.
- Costinot, A., D. Donaldson, and I. Komunjer (2011). What goods do countries trade? a quantitative exploration of ricardo's ideas. *The Review of economic studies*, rdr033.
- Costinot, A., D. Donaldson, J. Vogel, and I. Werning (2015). Comparative advantage and optimal trade policy. *The Quarterly Journal of Economics* 130(2), 659–702.
- Costinot, A. and A. Rodríguez-Clare (2014). Trade theory with numbers: Quantifying the consequences of globalization. In *Handbook of international economics*, Volume 4, pp. 197–261. Elsevier.
- Costinot, A. and I. Werning (2019). Lerner symmetry: A modern treatment. *American Economic Review: Insights* 1(1), 13–26.
- Das, S. P. (1983). Optimum tariffs on final and intermediate goods. *International Economic Review*, 493–508.
- Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced trade. *American Economic Review* 97(2), 351–355.
- Egger, P., M. Larch, K. E. Staub, and R. Winkelmann (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy* 3(3), 113–43.
- Fajgelbaum, P. D., P. K. Goldberg, P. J. Kennedy, and A. K. Khandelwal (2020). The return to protectionism. *The Quarterly Journal of Economics* 135(1), 1–55.
- Goldberg, P. K. and G. Maggi (1999). Protection for sale: An empirical investigation. *American Economic Review* 89(5), 1135–1155.
- Grant, M. (2019). Why special economic zones? using trade policy to discriminate across importers.
- Gros, D. (1987). A note on the optimal tariff, retaliation and the welfare loss from tariff wars in a framework with intra-industry trade. *Journal of international Economics* 23(3), 357–367.
- Grossman, G. M. and E. Helpman (1992). Protection for sale. Technical report, National Bureau of Economic Research.
- Grossman, G. M. and E. Helpman (1995). Trade wars and trade talks. *Journal of political Economy* 103(4), 675–708.
- Irwin, D. A. (2020). *Free trade under fire*. Princeton University Press.

- Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review* 70(5), 950–959.
- Kucheryavyi, K., G. Lyn, and A. Rodríguez-Clare (2016). Grounded by gravity: A well-behaved trade model with industry-level economies of scale. Technical report, National Bureau of Economic Research.
- Lashkaripour, A. and V. Lugovskyy (2019). Profits, scale economies, and the gains from trade and industrial policy.
- Limão, N. (2016). Preferential trade agreements. In *Handbook of commercial policy*, Volume 1, pp. 279–367. Elsevier.
- Maggi, G. and A. Rodríguez-Clare (1998). The value of trade agreements in the presence of political pressures. *Journal of Political Economy* 106(3), 574–601.
- Opp, M. M. (2010). Tariff wars in the ricardian model with a continuum of goods. *Journal of International Economics* 80(2), 212–225.
- Ossa, R. (2014). Trade wars and trade talks with data. *The American Economic Review* 104(12), 4104–4146.
- Ossa, R. (2016). Quantitative models of commercial policy. *Handbook of Commercial Policy* 1, 207–259.
- Piketty, T. and E. Saez (2013). Optimal labor income taxation. In *Handbook of public economics*, Volume 5, pp. 391–474. Elsevier.
- Powell, A. A. and F. Gruen (1968). The constant elasticity of transformation production frontier and linear supply system. *International Economic Review* 9(3), 315–328.
- Shapiro, J. S. (2019). The environmental bias of trade policy. *UC Berkeley and NBER (Ed.)*, 1–76.
- Su, C.-L. and K. L. Judd (2012). Constrained optimization approaches to estimation of structural models. *Econometrica* 80(5), 2213–2230.
- Suzuki, K. (1978). The welfare effects of an export tax levied on an intermediate good. *The Quarterly Journal of Economics*, 55–69.
- Timmer, M., A. A. Erumban, R. Gouma, B. Los, U. Temurshoev, G. J. de Vries, I.-a. Arto, V. A. A. Genty, F. Neuwahl, J. Francois, et al. (2012). The world input-output database (wiod): contents, sources and methods. Technical report, Institute for International and Development Economics.

Online Appendix

A Proof of Lemma 1

The proof of Lemma 1 follows from two intermediate claims, labeled C1 and C2. To present these claims, note that $A = (\mathbf{1} + \mathbf{t}, \mathbf{1} + \mathbf{x}, w_h, w_f)$ and $A' = (a(\mathbf{1} + \mathbf{t}), (\mathbf{1} + \mathbf{x})/a, aw_h, w_f)$. Accordingly, all equilibrium variables corresponding to the wage-tax combination A' will be denoted by the prime notation. Claim C1 is based on the demand function $\mathcal{D}_{ji,k}(\tilde{\mathbf{p}}_i, Y_i)$ being homogeneous of degree zero. Specifically,

$$\begin{cases} \tilde{\mathbf{p}}'_h = a\tilde{\mathbf{p}}_h, & Y'_h = aY_h \\ \tilde{\mathbf{p}}'_f = \tilde{\mathbf{p}}_f, & Y'_f = aY_f \end{cases} \implies q'_{ji,k} = q_{ji,k} \quad \forall ji,k. \quad (\text{C1})$$

where recall that $\tilde{\mathbf{p}}_i \equiv \{\tilde{p}_{ji,k}\}_{j,k}$ denotes the vector of final prices in country i . Claim C2 can be stated as follows

$$q'_{ji,k} = q_{ji,k} \quad \forall ji,k \implies \begin{cases} \tilde{\mathbf{p}}'_h = a\tilde{\mathbf{p}}_h, & Y'_h = aY_h \\ \tilde{\mathbf{p}}'_f = \tilde{\mathbf{p}}_f, & Y'_f = aY_f \end{cases}. \quad (\text{C2})$$

The claim that $\tilde{\mathbf{p}}'_h = a\tilde{\mathbf{p}}_h$ and $\tilde{\mathbf{p}}'_f = \tilde{\mathbf{p}}_f$ follow trivially from the price equation:

$$\tilde{p}_{ji,k} = (1 + x_{ji,k})(1 + t_{ji,k})C_{j,k}(w_j, \tilde{\mathbf{p}}_j; \sum_n \tau_{jn,k} q_{jn,k}),$$

given that $C_{i,k}$ is homogeneous of degree 1 w.r.t. w_j and $\tilde{\mathbf{p}}_j$. That $Y'_f = w'_f L_f = w_f L_f = Y_f$, holds by construction. The claim that $Y'_h = aY_h$, meanwhile, can be shown along the following steps:

$$\begin{aligned} Y'_h &= w'_h L'_h + \sum_k \left[\frac{t'_k}{1 + t'_k} \tilde{p}'_{fh,k} q'_{fh,k} + \frac{x'_k}{1 + x'_k} \tilde{p}'_{hf,k} q'_{hf,k} \right] \\ &= aw_h L_h + \sum_k \left[\left(1 - \frac{a}{1 + t'_k}\right) \tilde{p}'_{fh,k} q_{fh,k} + \left[a - \frac{1}{1 + x'_k}\right] \tilde{p}_{hf,k} q_{hf,k} \right] \\ &= aw_h L_h + \sum_k \left[\left(1 - \frac{a}{a(1 + t_k)}\right) a\tilde{p}_{fh,k} q_{fh,k} + \left[a - \frac{1}{(1 + x_k)/a}\right] \tilde{p}_{hf,k} q_{hf,k} \right] \\ &= aw_h L_h + a \sum_k \left[\frac{t_k}{1 + t_k} \tilde{p}_{fh,k} q_{fh,k} + \frac{x_k}{1 + x_k} \tilde{p}_{hf,k} q_{hf,k} \right] = a(w_h L_h + \mathcal{R}_h), \end{aligned}$$

where the second line follows from the balanced trade condition, i.e., $\sum_k \left[\frac{\tilde{p}_{fh,k} q_{fh,k}}{1+t_k} - \tilde{p}_{hf,k} q_{hf,k} \right] = 0$. Together, Claims C1 and C2 establish Lemma 1.

B Proof of Theorem 1

Note that by definition $1 + t_k \equiv \tilde{p}_{fh,k}/p_{fh,k}$ and $1 + x_k \equiv \tilde{p}_{hf,k}/p_{hf,k}$. Hence, the optimal policy problem can be reformulated as a problem of directly choosing consumer prices, $\tilde{p}_{fh,k}$, and, $\tilde{p}_{hf,k}$, rather than taxes, t_k , and, x_k . In that case, the optimal policy problem can alternatively be expressed as

$$\max_{(\tilde{\mathbf{p}}; \mathbf{w}) \in \tilde{\mathcal{A}}} W_h(\tilde{\mathbf{p}}; \mathbf{w}),$$

where $\tilde{\mathcal{A}}$ denotes the set of feasible wage-price combinations that is defined analogous to D2. Note that in the presence of input-output linkages, equilibrium is conventionally defined in terms three sets of equilibrium outcomes: (a) wages, \mathbf{w} , (b) net income \mathbf{Y} , and (c) gross revenue, \mathbf{Y} . However, combining equilibrium conditions (iv) and (v) in D2, we can uniquely solve $\mathbf{Y} \equiv \mathbf{Y}(\tilde{\mathbf{p}}; \mathbf{w})$ and $\mathbf{Y} \equiv \mathbf{Y}(\tilde{\mathbf{p}}; \mathbf{w})$ in terms of prices and wages. Considering this we can uniquely solve all equilibrium variables, including welfare, in terms of prices and wages.

To handle the complex nature of the above problem, we use two technical tricks. First, we assume that the government can choose \tilde{p}_{hh} in addition to \tilde{p}_{fh} and \tilde{p}_{hf} . In that case, Home's tax revenues include revenue's from domestic taxation, which we denote by

$$\mathcal{R}_h^D \equiv \sum_k (\tilde{p}_{hh,k} - p_{hh,k}) q_{hh,k}.$$

Including the aforementioned instrument is innocuous since (i) markets are perfectly competitive, and (ii) governments have not political economy motives for intra-national redistribution. As a result, there is no rationale for domestic taxation (i.e., $\tilde{p}_{hh,k}^* = p_{hh,k}$) and the problem with extended instruments becomes isomorphic to our main problem of interest. The second trick that greatly simplifies our analysis is the observation that $dW_h(\tilde{\mathbf{p}}, \mathbf{w})/d\tilde{p}_{ff,k} = 0$. In light of the above result, we are simply solving a problem where the government can directly choose all price variables that matter for Home's welfare. So, we should not track how a change in, say price variable $\tilde{p}_{fh,k}$, affects any other consumer price variables.

Before proceeding to the first step of the proof, let us alert the readers of a slight change in notation, which is meant to streamline the presentation: We, hereafter, use $\tilde{p}_{hf,k} \equiv \tilde{p}_{hf,k}/(1 + t_{hf,k})$ to denote the consumer price of good hf,k net of Foreign's im-

port taxes, and $p_{fh,k} \equiv (1 + x_{fh,k})p_{fh,k}$ to denote the producer price of good fh,k inclusive of Foreign's export taxes.

Step 1: Characterizing the F.O.C. w.r.t. Import Taxes. The F.O.C. with respect to sector k 's tariff can be expressed as

$$\frac{dW_h(\tilde{p}; w)}{d \ln \tilde{p}_{fh,k}} = \frac{\partial V_h(\cdot)}{\partial Y_h} \left[\frac{\partial Y_h}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial Y_h}{\partial w_h} \frac{dw_h}{d \ln \tilde{p}_{fh,k}} \right] + \frac{\partial V_h(\cdot)}{\partial \tilde{p}_{fh,k}}$$

It should be noted upfront that the difference between the present setup and the pure Ricardian case, is that the (conditional) tariff pass-through $\sigma_{jh,g}^{fh,k} = \partial \ln \tilde{p}_{jh,g} / \partial \ln (1 + t_k)$ can be non-zero (even if $g \neq k$) due to (i) the upward sloping supply curve in industry g , plus (ii) the cross-substitutability between industry k and industry g goods. Plugging $Y_h = w_h L_h + \Pi_h + \mathcal{R}_h^X + \mathcal{R}_h^M + \mathcal{R}_h^D$, into the F.O.C. yields the following:

$$\begin{aligned} \frac{dW_h(\tilde{p}; w)}{d \ln \tilde{p}_{fh,k}} = & \frac{\partial V_h}{\partial Y_h} \left\{ \frac{\partial \Pi_h}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial \mathcal{R}_h^X}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial \mathcal{R}_h^M}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial \mathcal{R}_h^D}{\partial \ln \tilde{p}_{fh,k}} \right. \\ & \left. + \frac{\partial V_h / \partial \tilde{p}_{fh,k}}{\partial V_h / \partial Y_h} + \frac{\partial V_h / \partial w_h}{\partial V_h / \partial Y_h} \frac{dw_h}{d \ln \tilde{p}_{fh,k}} \right\} = 0, \end{aligned}$$

where $\partial V_h / \partial w_h = \partial Y_h / \partial w_h + \sum_g (\partial V_h / \partial \tilde{p}_{jh,g}) (\partial \tilde{p}_{jh,g} / \partial w_h)$. The above F.O.C. is characterized by six different elements that can be characterized as follows. First, The effect of tariffs on producer surplus, $\partial \Pi_h / \partial \ln \tilde{p}_{fh,k}$, can be expressed as

$$\begin{aligned} \frac{\partial \Pi_h(\tilde{p}; w)}{\partial \ln \tilde{p}_{fh,k}} &= \sum_g \sum_{i=h,f} \left(\frac{\partial \Pi_h}{\partial \ln p_{hi,g}} \left[\frac{\partial \ln p_{hi,g}}{\partial \ln \tilde{p}_{fh,k}^{\mathcal{I}}} + \frac{\partial p_{hi,g}}{\partial \ln q_{hh,g}} \frac{\partial \ln q_{hh,g}}{\partial \ln \tilde{p}_{fh,k}} \right] \right) + \frac{\partial \Pi_h}{\partial \ln \tilde{p}_{fh,k}^{\mathcal{I}}} \\ &= \sum_g \sum_{i=h,f} \left(p_{hi,g} q_{hi,g} \left[\frac{\partial \ln p_{hi,g}}{\partial \ln \tilde{p}_{fh,k}^{\mathcal{I}}} + \frac{\partial p_{hi,g}}{\partial \ln q_{hh,g}} \frac{\partial \ln q_{hh,g}}{\partial \ln \tilde{p}_{fh,k}} \right] \right) - p_{fh,k} q_{fh,k}^{\mathcal{I}}, \end{aligned}$$

where the second line follows from Hotelling's lemma that $\partial \Pi_{hi,g} / \partial p_{hi,g} = q_{hi,g}$ and $\partial \Pi_h / \partial \tilde{p}_{fh,k}^{\mathcal{I}} = q_{fh,k}^{\mathcal{I}}$. Second, noting that $dW_h / d \tilde{p}_{ff,g} = 0$, the effect of import taxes on

import tax revenues, $\mathcal{R}_h^M \equiv \sum_g (\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g}$, can be expressed as

$$\begin{aligned} \frac{\partial \mathcal{R}^M(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{fh,k}} &= \frac{\partial}{\partial \ln \tilde{p}_{fh,k}} \left\{ \sum_g (\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g} \right\} \\ &= \tilde{p}_{fh,q} q_{fh,k} + \sum_g \left[(\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g} \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} - p_{fh,g} q_{fh,g} \frac{\partial \ln p_{fh,g}}{\partial \ln q_{fh,g}} \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} \right], \end{aligned}$$

where $\partial \ln q_{fh,g} / \partial \ln \tilde{p}_{fh,k}$ accounts for the overall effect $\tilde{p}_{fh,k}$ on final and intermediate input demand. More specifically, letting $\tilde{\mathcal{D}}_{fh,gs}(\cdot)$ denote the demand for intermediate input variety fh, g from industry s ,

$$\begin{aligned} \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} &= \frac{q_{fh,g}^C}{q_{fh,g}} \left(\frac{\partial \ln \mathcal{D}_{fh,g}(\cdot)}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial \ln \mathcal{D}_{fh,g}(\cdot)}{\partial \ln Y_h} \frac{d \ln Y_h}{d \ln \tilde{p}_{fh,k}} \right) \\ &+ \frac{q_{fh,g}^I}{q_{fh,g}} \left(\frac{\partial \ln \tilde{\mathcal{D}}_{fh,gs}(\cdot)}{\partial \ln \tilde{p}_{fh,k}} + \sum_s \alpha_{h,gs} \frac{\partial \ln \tilde{\mathcal{D}}_{fh,gs}(\cdot)}{\partial \ln \mathcal{Y}_{h,s}} \frac{d \ln \mathcal{Y}_{h,s}}{d \ln \tilde{p}_{fh,k}} \right) = \varepsilon_{fh,g} + \eta_{fh,g} \frac{d \ln E_{h,g}}{d \ln \tilde{p}_{fh,k}} \end{aligned}$$

where, $\alpha_{h,gs} \equiv q_{fh,gs}^I / q_{fh,g}^I$, and $E_{i,g} \equiv \omega_{i,g}^C Y_i + \sum_s \omega_{i,gs}^I \mathcal{Y}_{i,s}$ denotes demand-weighted gross expenditure on final plus intermediate inputs from industry g , with the last line following for the implicit assumption that $\frac{\partial \ln \tilde{\mathcal{D}}_{fh,gs}(\cdot)}{\partial \ln \mathcal{Y}_{h,k}} = \frac{\partial \ln \mathcal{D}_{fh,g}(\cdot)}{\partial \ln Y_h}$ and $\frac{\partial \ln \tilde{\mathcal{D}}_{fh,g}(\cdot)}{\partial \ln \tilde{p}_{fh,k}} = \frac{\partial \ln \mathcal{D}_{fh,g}(\cdot)}{\partial \ln \tilde{p}_{fh,k}}$ for all s . Third, the effect of import taxes on export tax revenues, $\mathcal{R}_h^X \equiv \sum_g (\tilde{p}_{hf,g} - p_{hf,g}) q_{hf,g}$, can be expressed as

$$\begin{aligned} \frac{\partial \mathcal{R}^X(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{fh,k}} &= \frac{\partial}{\partial \ln \tilde{p}_{fh,k}} \left\{ \sum_g (\tilde{p}_{hf,g} - p_{hf,g}) q_{hf,g} \right\} \\ &= \sum_g \left[(\tilde{p}_{hf,g} - p_{hf,g}) q_{hf,g} \eta_{hf,g} \frac{\partial \ln E_{f,g}}{\partial \ln \tilde{p}_{fh,k}} - p_{hf,g} q_{hf,g} \left(\frac{\partial \ln p_{hf,g}}{\partial \ln \tilde{p}_{fh,k}^I} + \frac{\partial \ln p_{hf,g}}{\partial \ln q_{hh,g}} \frac{\partial \ln q_{hh,g}}{\partial \ln \tilde{p}_{fh,k}} \right) \right]. \end{aligned}$$

Fourth, the effect of import taxes on domestic tax revenues, $\mathcal{R}_h^D \equiv \sum_g (\tilde{p}_{hh,g} - p_{hh,g}) q_{hh,g}$, can be expressed as

$$\begin{aligned} \frac{\partial \mathcal{R}^D(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{fh,k}} &= \frac{\partial}{\partial \ln \tilde{p}_{fh,k}} \left\{ \sum_g (\tilde{p}_{hh,g} - p_{hh,g}) q_{hh,g} \right\} \\ &= - \sum_g \left[p_{hh,g} q_{hh,g} \left(\frac{\partial \ln p_{hh,g}}{\partial \ln \tilde{p}_{fh,k}^I} + \frac{\partial \ln p_{hh,g}}{\partial \ln q_{hh,g}} \frac{\partial \ln q_{hh,g}}{\partial \ln \tilde{p}_{fh,k}} \right) \right] \end{aligned}$$

where the second line implicitly assumes that by the second welfare theorem, $\tilde{p}_{hh,k}^* = p_{hh,k}$. Fifth, the effect of taxes on the consumer prices can be simplified using Roy's identity, $\frac{\partial V_i / \partial \tilde{p}_{ji,g}}{\partial V_i / \partial Y_i} = -q_{ji,g}$, as follows

$$\frac{\partial V_h / \partial \ln \tilde{p}_{fh,k}}{\partial V_h / \partial Y_h} = -\tilde{p}_{fh,k} q_{fh,k}^C.$$

Finally, the effect of tariffs on wages can be determined by applying the implicit function theorem to the balanced trade condition, $D_h(\tilde{p}, w) = \sum_g (p_{fh,g} q_{fh,g} - \tilde{p}_{hf,g} q_{hf,g})$. Doing so, implies $\frac{dw_h}{d\tilde{p}_{fh,k}} = -\frac{\partial D_h(\cdot)}{\partial \tilde{p}_{fh,k}} / \frac{\partial D_h(\cdot)}{\partial w_h}$. Hence, defining $\bar{\tau} \equiv \left(\frac{\partial V_h / \partial w_h}{\partial V_h / \partial Y_h} \right) / \frac{\partial D_h}{\partial w_h}$,

$$\begin{aligned} \frac{\partial V_h / \partial w_h}{\partial V_h / \partial Y_h} \frac{dw_h}{d \ln \tilde{p}_{fh,k}} &= -\bar{\tau} \sum_g \left\{ p_{fh,g} q_{fh,g} \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} + \right. \\ &\quad \left. + p_{fh,g} q_{fh,g} \frac{\partial \ln p_{fh,g}}{\partial \ln q_{hh,g}} \frac{\partial \ln q_{hh,g}}{\partial \ln \tilde{p}_{fh,k}} - \tilde{p}_{hf,g} q_{hf,g} \eta_{hf,g} \frac{\partial \ln E_f}{\partial \ln \tilde{p}_{fh,k}} \right\} \end{aligned}$$

Combing the above expression as well as noting that $\partial V_h / \partial Y_h > 0$ and $q_{ji,k} = q_{ji,k}^I + q_{ji,k}^C$, the F.O.C. can be simplified expressed as follows:

$$\begin{aligned} &\sum_g \left(\left[\frac{\tilde{p}_{fh,g}}{p_{fh,g}} - (1 + \bar{\tau}) \right] p_{fh,g} q_{fh,g} \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} - (1 + \bar{\tau}) \left(\frac{\partial \ln p_{fh,g}}{\partial \ln q_{fh,g}} \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} \right) p_{fh,g} q_{fh,g} \right) \\ &+ \sum_g \left[\left(1 - (1 + \bar{\tau}) \frac{p_{hf,g}}{\tilde{p}_{hf,g}} \right) \tilde{p}_{hf,g} q_{hf,g} \eta_{hf,g} \right] \frac{\partial \ln E_{f,g}}{\partial \ln \tilde{p}_{fh,k}} = 0 \end{aligned}$$

Given that $1 + t_{fh,g} \equiv \frac{\tilde{p}_{fh,g}}{p_{fh,g}}$ and $1 + x_{hf,g} \equiv \frac{\tilde{p}_{hf,g}}{p_{hf,g}}$, we can further simplify the above expression if we divide it by $1 + \bar{\tau}$ and $\sum_{g,i} p_{fi,g} q_{fi,g}$.²⁸

$$\sum_g \left[\left(\frac{1 + t_{fh,g}}{1 + \bar{\tau}} - 1 - \frac{\partial \ln p_{fh,g}}{\partial \ln q_{fh,g}} \right) \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} \hat{r}_{fh,g} + \Delta(x_h) \frac{\partial \ln E_{h,g}}{\partial \ln \tilde{p}_{fh,k}} \right] = 0$$

where $\Delta(x_h) \equiv \sum_g \left[\left(1 - \frac{1}{(1 + \bar{\tau})(1 + x_{hf,g})} \right) \lambda_{hf,g} \eta_{hf,g} \right]$. Since (a) Foreign labor is assigned as the numeraire, $w_f = 1$, (b) Foreign does not collect tax revenue, and (c) the share of intermediate inputs in Foreign's production is invariant to $\tilde{p}_{fh,k}$, then $\partial \ln E_{f,g} / \partial \ln \tilde{p}_{fh,k} = 0$. Finally, we can determine $\partial \ln p_{fh,k} / \partial \ln q_{fh,k}$ by applying the Implicit Function theorem to $p_{fh,k} = \tau_{fh,k} C_{f,k}(\dots; \sum_i \tau_{fi,k} q_{fi,k})$, which yields

²⁸Since Foreign does not collect tax revenue, gross expenditure in Foreign, E_f , also equals gross output revenue.

$\partial \ln p_{fh,k} / \partial \ln q_{fh,k} = \gamma_{f,k} r_{fh,k} / (1 - \gamma_{f,k} r_{ff,k} \varepsilon_{ff,k})$. Combining these results reduces the F.O.C. to

$$\sum_g \left[\left(\frac{1 + t_{fh,g}}{1 + \bar{\tau}} - 1 - \frac{\gamma_{f,k} r_{fh,k}}{1 - \gamma_{f,k} r_{ff,k} \varepsilon_{ff,k}} \right) \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{hf,k}} \dot{r}_{fh,g} \right] = 0. \quad (12)$$

Step 2. Characterizing the F.O.C. w.r.t. Export Taxes. The F.O.C. with respect to sector k 's export tax can be stated as

$$\begin{aligned} \frac{dW_h(\tilde{\mathbf{p}}; \mathbf{w})}{d \ln \tilde{p}_{hf,k}} &= \frac{\partial V_h}{\partial Y_h} \left\{ \frac{\partial \Pi_h}{\partial \ln \tilde{p}_{hf,k}} + \frac{\partial \mathcal{R}_h^X}{\partial \ln \tilde{p}_{hf,k}} + \frac{\partial \mathcal{R}_h^M}{\partial \ln \tilde{p}_{hf,k}} + \frac{\partial \mathcal{R}_h^D}{\partial \ln \tilde{p}_{hf,k}} \right. \\ &\quad \left. + \frac{\partial V_h / \partial \ln \tilde{p}_{hf,k}}{\partial V_h / \partial Y_h} + \frac{\partial V_h / \partial w_h}{\partial V_h / \partial Y_h} \frac{dw_h}{d \ln \tilde{p}_{hf,k}} \right\} = 0, \end{aligned}$$

In the above expression, $\frac{\partial V_h / \partial \ln \tilde{p}_{hf,k}}{\partial V_h / \partial Y_h} = 0$. The remaining non-zero elements can be expressed as follows. First, The effect of export taxes on producer surplus, $\partial \Pi_h / \partial \tilde{p}_{hf,k}$, can be expressed as

$$\begin{aligned} \frac{\partial \Pi_h(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{hf,k}} &= \sum_g \sum_{i=h,f} \left(\frac{\partial \Pi_{hi,g}}{\partial p_{hi,g}} \frac{\partial p_{hi,g}}{\partial \ln \tilde{p}_{hf,k}} \right) + \frac{\partial \Pi_h}{\partial \ln \tilde{p}_{hf,k}^{\mathcal{I}}} \\ &= \sum_g \sum_{i=h,f} \left(p_{hi,g} q_{hi,g} \frac{\partial p_{hi,g}}{\partial \ln q_{hf,g}} \frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,k}} \right), \end{aligned}$$

where the last line follows from Hotelling's lemma that $\partial \Pi_{hi,g} / \partial p_{hi,g} = q_{hi,g}$ and $\partial \Pi_h / \partial p_{hf,g}^{\mathcal{I}} = 0$. Second, the effect of export taxes on import tax revenues, $\mathcal{R}_h^M \equiv \sum_g (\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g}$, can be expressed as

$$\begin{aligned} \frac{\partial \mathcal{R}_h^M(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{hf,k}} &= \frac{\partial}{\partial \ln \tilde{p}_{hf,k}} \left\{ \sum_g (\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g} \right\} \\ &= - \sum_g \left[(\tilde{p}_{fh,g} - p_{fh,g}) \eta_{fh,g} \frac{\partial \ln E_h}{\partial \ln \tilde{p}_{hf,k}} + p_{fh,g} q_{fh,g} \frac{\partial \ln p_{fh,g}}{\partial \ln \tilde{p}_{hf,k}^{\mathcal{I}}} \right] \end{aligned}$$

Third, the effect of export taxes on export tax revenues, $\mathcal{R}_h^X \equiv \sum_g (\tilde{p}_{hf,g} - p_{hf,g}) q_{hf,g}$, can be expressed as

$$\begin{aligned} \frac{\partial \mathcal{R}_h^X(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{hf,k}} &= \frac{\partial}{\partial \ln \tilde{p}_{hf,k}} \left\{ \sum_g (\tilde{p}_{hf,g} - p_{hf,g}) q_{hf,g} \right\} \\ &= \sum_g \left[(\tilde{p}_{hf,g} - p_{hf,g}) q_{hf,g} \frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,k}} - p_{hf,g} q_{hf,g} \frac{\partial \ln p_{hf,g}}{\partial \ln q_{hf,g}} \frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,k}} \right] \end{aligned}$$

Fourth, the effect of import taxes on domestic tax revenues, $\mathcal{R}_h^X \equiv \sum_g (\tilde{p}_{hh,g} - p_{hh,g}) q_{hh,g}$, can be expressed as

$$\begin{aligned} \frac{\partial \mathcal{R}_h^D(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{hf,k}} &= \frac{\partial}{\partial \ln \tilde{p}_{hf,k}} \left\{ \sum_g (\tilde{p}_{hh,g} - p_{hh,g}) q_{fh,g} \right\} \\ &= - \sum_g \left[p_{hh,g} q_{hh,g} \frac{\partial \ln p_{hh,g}}{\partial \ln q_{hf,g}} \frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,k}} \right] \end{aligned}$$

where the second line implicitly assumes that by the second welfare theorem, $\tilde{p}_{hh,k}^* = p_{hh,k}$. Finally, the effect of export taxes on wages can be determined by applying the implicit function theorem to the balanced trade condition, $D_h = \sum_g p_{fh,g} q_{fh,g} - \tilde{p}_{hf,g} q_{hf,g}$. Doing so, implies $\frac{d\omega_h}{d\tilde{p}_{hf,k}} = -\frac{\partial D_h}{\partial \tilde{p}_{hf,k}} / \frac{\partial D_h}{\partial \omega_h}$. Hence, adopting our earlier definition, $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial \omega_h} / \frac{\partial V_h}{\partial Y_h} \right) / \frac{\partial D_h}{\partial \omega_h}$,

$$\begin{aligned} \frac{\partial V_h / \partial \omega_h}{\partial V_h / \partial Y_h} \frac{d\omega_h}{d \ln \tilde{p}_{hf,k}} &= -\bar{\tau} \left\{ \sum_g \left[p_{fh,g} q_{fh,g} \left(\frac{\partial \ln p_{fh,g}}{\partial \ln \tilde{p}_{hf,k}^T} + \eta_{fh,g} \frac{\partial \ln E_h}{\partial \ln \tilde{p}_{hf,k}} \right) \right] \right. \\ &\quad \left. - \tilde{p}_{hf,k} q_{hf,k} - \sum_g \tilde{p}_{hf,g} q_{hf,g} \frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,k}} \right\} \end{aligned}$$

Combing the above expression as well as noting that $\partial V_h / \partial Y_h > 0$ and $q_{ji,k} = q_{ji,k}^T + q_{ji,k}^C$, the F.O.C. can be simplified expressed as follows:

$$\begin{aligned} &(1 + \bar{\tau}) \sum_g \left(p_{fh,g} q_{fh,g} \alpha_{f,g}^{h,k} \right) + \sum_g \left(\left[(1 + \bar{\tau}) - \frac{p_{hf,g}}{\tilde{p}_{hf,g}} \right] \tilde{p}_{hf,g} q_{hf,g} \frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,k}} \right) \\ &+ \left(\sum_g \left[\frac{\tilde{p}_{fh,g}}{p_{fh,g}} - (1 + \bar{\tau}) \right] p_{fh,g} q_{fh,g} \eta_{fh,g} \right) \frac{\partial \ln E_h}{\partial \ln \tilde{p}_{hf,k}}, \end{aligned}$$

where recall that $\frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,k}} = \varepsilon_{hf,g}^{hf,k} + \eta_{hf,g} \frac{\partial \ln E_f}{\partial \ln \tilde{p}_{hf,k}}$. Given that $1 + t_{fh,g} \equiv \frac{\tilde{p}_{fh,g}}{p_{fh,g}}$ and $1 + x_{hf,g} \equiv \frac{\tilde{p}_{hf,g}}{p_{hf,g}}$, we can further simplify the above expression if we divide it by $1 + \bar{\tau}$ and E_f (noting that $p_{fh,k} q_{fh,k} = \hat{r}_{fh,k} E_f$ and $\tilde{p}_{hf,g} q_{hf,g} = \hat{\lambda}_{hf,g} E_f$):

$$\begin{aligned} \hat{\lambda}_{hf,k} + \sum_g \left(\left[1 - \frac{1}{(1 + x_{hf,g})(1 + \bar{\tau})} \right] \hat{\lambda}_{hf,g} \varepsilon_{hf,g}^{hf,k} \right) \\ - \sum_g \left(\hat{r}_{fh,g} \alpha_{f,g}^{h,k} \right) \frac{Y_f}{E_f} + \Delta_h(\mathbf{t}_h) \frac{\partial \ln E_h}{\partial \ln \tilde{p}_{hf,k}} + \Delta_f(\mathbf{x}) \frac{\partial \ln E_f}{\partial \ln \tilde{p}_{hf,k}} = 0, \end{aligned}$$

where $\Delta_h(\mathbf{t}_h) \equiv \sum_g \left(\left[\frac{1+t_{fh,g}}{1+\bar{\tau}} - 1 \right] \hat{r}_{fh,g} \eta_{fh,g} \right)$ is a uniform term. As noted before, given our choice of numeraire, $\partial \ln E_f / \partial \ln \tilde{p}_{hf,k} = 0$. Given the Lerner symmetry and the multiplicity of the optimal trade tax, there always exists a solution to the above problem where $\Delta_h(\mathbf{t}) = 0$. Henceforth, we restrict our attention to solving for this particular solution. Once we do that, the remaining solutions can be identified with a basic multiplicative transformation of the import and export tax vectors. yields the following optimality condition:

$$\hat{\lambda}_{hf,k} + \sum_g \left(\left[1 - \frac{1}{(1 + x_g)(1 + \bar{\tau})} \right] \hat{\lambda}_{hf,g} \varepsilon_{hf,g}^{hf,k} \right) - \sum_g \left(\hat{r}_{fh,g} \alpha_{f,g}^{h,k} \right) = 0. \quad (13)$$

Step 3: Simultaneously Solving the System of F.O.C. As a final step, we simultaneously solve the system of F.O.C.s for all tax instruments:

$$\begin{aligned} \sum_g \left[\left(\frac{1 + t_g}{1 + \bar{\tau}} - 1 - \gamma_{fh,g} \right) \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} r_{fh,g} \right], \quad \forall k \in \mathbb{K} \\ \sum_g \left(\left[\frac{1}{(1 + x_g)(1 + \bar{\tau})} - 1 \right] \frac{\hat{\lambda}_{hf,g}}{\hat{\lambda}_{hf,k}} \varepsilon_{hf,g} \right) = 1 - \sum_g \frac{\hat{r}_{fh,g}}{\hat{\lambda}_{hf,k}} \alpha_{f,g}^{h,k}, \quad \forall k \in \mathbb{K} \end{aligned}$$

First, note that by the Lerner symmetry, the value of $\bar{\tau}$ is redundant. In other words, replacing $1 + \bar{\tau}$ with any $1 + \bar{t} \in \mathbb{R}_+$ identifies an optimal tax schedule. To derive a simplified expression for optimal export taxes, define $\Omega_{hf,k} \equiv 1 - \sum_g \left(\frac{r_{fh,g}}{\lambda_{hf,k}} \alpha_{f,g}^{h,k} \right)$ and suppose the optimal export tax has the following formulation:

$$1 + x_{hf,k}^* = \frac{\varepsilon_{hf,k}}{\varepsilon_{hf,k} + \bar{\zeta}_{hf,k} + \Omega_{hf,k}}.$$

Plugging the above expression back into the F.O.C. implies that that $\sum_g (\xi_{hf,g} + \Omega_{hf,g}) \frac{\lambda_{hf,g} \varepsilon_{hf,g}^{hf,k}}{\lambda_{hf,k} \varepsilon_{hf,k}} = \Omega_{hf,k}$. This equation can be written in matrix notation as $\Xi (\xi + \Omega) = \Omega$, where $\xi \equiv [\xi_{hf,k}]_k$ and $\Omega \equiv [\Omega_{hf,k}]_k$ are $K \times 1$ vectors and $\Xi \equiv \left[\frac{\lambda_{hf,g} \varepsilon_{hf,g}^{hf,k}}{\lambda_{hf,k} \varepsilon_{hf,k}} \right]_{k,g}$ and I_K are $K \times K$ matrixes. Inverting the above system implies that

$$[\xi_{hf,k}]_k = \left(\Xi^{-1} - I_K \right) \Omega. \quad (14)$$

So, altogether the following tax schedule corresponds to an optimal policy for an arbitrary choice of $1 + \bar{t} \in \mathbb{R}_+$:

$$1 + t_{fh,k}^* = \left(1 + \frac{\gamma_{f,k} r_{fh,k}}{1 - \gamma_{f,k} r_{ff,k} \varepsilon_{ff,k}} \right) (1 + \bar{t})$$

$$1 + x_{hf,k}^* = \frac{\varepsilon_{hf,k}}{1 + \varepsilon_{hf,k} + \xi_{hf,k} - \sum_g \frac{r_{fh,g}}{\lambda_{hf,k}} \alpha_{fh,g}^{hf,k}} (1 + \bar{t})^{-1},$$

where $\xi_{hf,k}$ is given by Equation 14.

C Proof of Theorem 2

The F.O.C. with respect to sector k 's tariff can be expressed as

$$\frac{dW_h(\tilde{p}; w)}{d \ln \tilde{p}_{fh,k}} = \frac{\partial V_h(\cdot)}{\partial Y_h} \left[\frac{\partial Y_h}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial Y_h}{\partial w_h} \frac{dw_h}{d \ln \tilde{p}_{fh,k}} \right] + \frac{\partial V_h(\cdot)}{\partial \tilde{p}_{fh,k}}$$

It should be noted upfront that the difference between the present setup and the pure Ricardian case, is that the (conditional) tariff pass-through $\sigma_{jh,g}^{fh,k} = \partial \ln \tilde{p}_{jh,g} / \partial \ln (1 + t_k)$ can be non-zero (even if $g \neq k$) due to (i) the upward sloping supply curve in industry g , plus (ii) the cross-substitutability between industry k and industry g goods. Plugging $Y_h = w_h L_h + \mathcal{R}_h^M$, into the F.O.C. yields the following:

$$\frac{dW_h(\tilde{p}; w)}{d \ln \tilde{p}_{fh,k}} = \frac{\partial V_h}{\partial Y_h} \left\{ \frac{\partial \mathcal{R}_h^M}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial V_h / \partial \tilde{p}_{fh,k}}{\partial V_h / \partial Y_h} + \frac{\partial V_h / \partial w_h}{\partial V_h / \partial Y_h} \frac{dw_h}{d \ln \tilde{p}_{fh,k}} \right\} = 0,$$

By Roy's identity

$$\begin{aligned}\frac{\partial V_h / \partial \ln \tilde{p}_{fh,k}}{\partial V_h / \partial Y_h} &= -\tilde{p}_{fh,k} q_{fh,k}^C - \sum_g \tilde{p}_{hh,g} q_{hh,g}^C \frac{\partial \ln \tilde{p}_{hh,g}^C}{\partial \ln \tilde{p}_{fh,k}^T} \\ &= -\tilde{p}_{fh,k} q_{fh,k}^C - \sum_g \tilde{p}_{hh,g} q_{hh,g}^C \tilde{\alpha}_{h,g}^{f,k}\end{aligned}$$

where $\tilde{\alpha}_{i,k}^{f,k} \equiv \partial \ln \tilde{p}_{ij,g} / \partial \ln \tilde{p}_{fh,k}^T$ denotes the general equilibrium effect of a change in $\tilde{p}_{fh,k}$ on the price industry g 's output in country i .²⁹ The $K \times K$ matrix for $\tilde{\alpha}_{f,k}^{f,k}$ can be calculated by applying the Implicit Function Theorem to $\tilde{p}_{fi,k} = (1 + t_{fi,k}) \tau_{fi,k} C_{f,k}(w_f, \tilde{\mathbf{p}}_{ff}^T, \tilde{\mathbf{p}}_{hf}^T)$, which implies that

$$\left[\tilde{\alpha}_{f,g}^{f,k} \right]_{k,g} = \left(\mathbf{I}_K - \left[\alpha_{f,g}^{f,k} \right]_{k,g} \right) \left[\alpha_{f,g}^{h,k} \right]_{k,g} \left[\alpha_{h,g}^{f,k} \right]_{k,g},$$

Likewise, $\tilde{\alpha}_{f,k}^{f,k}$ can be calculated by applying the Implicit Function Theorem to $\tilde{p}_{hi,k} = \tau_{hi,k} C_{h,k}(w_f, \tilde{\mathbf{p}}_{fh}^T, \tilde{\mathbf{p}}_{hh}^T)$, which yields the following:

$$\left[\tilde{\alpha}_{h,g}^{f,k} \right]_{k,g} = \left(\mathbf{I}_K - \left[\alpha_{h,g}^{h,k} \right]_{k,g} \right) \left[\alpha_{h,g}^{f,k} \right]_{k,g}.$$

With the above definition in mind, the effect of import taxes on import tax revenues, $\mathcal{R}_h^M \equiv \sum_g (\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g}$, can be expressed as

$$\begin{aligned}\frac{\partial \mathcal{R}_h^M(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{fh,k}} &= \frac{\partial}{\partial \ln \tilde{p}_{fh,k}} \left\{ \sum_g (\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g} \right\} \\ &= \tilde{p}_{fh,k} q_{fh,k} + \sum_g \left[(\tilde{p}_{fh,g} - p_{fh,g}) q_{fh,g} \left(\frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{hh,g}} \tilde{\alpha}_{h,s}^{f,k} \right) - p_{fh,g} q_{fh,g} \tilde{\alpha}_{f,s}^{f,k} \right],\end{aligned}$$

where as before (with a slight abuse of notation) $\frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{ih,k}} = \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{ih,k}} + \eta_{fh,g} \frac{\partial \ln E_{h,g}}{\partial \ln \tilde{p}_{ih,k}}$. The income driven term, however, can be dropped as it has a second-order effect, since

²⁹To keep things simple, we are exercising a slight abuse of notation here. A more elaborate choice of notation would be $\tilde{\alpha}_{i,k}^{fh,k} \equiv \partial \ln \tilde{p}_{ij,g} / \partial \ln \tilde{p}_{fh,k}^T$, which clarifies that $\tilde{\alpha}_{i,k}^{fh,k} \neq \tilde{\alpha}_{i,k}^{ff,k}$.

$\frac{\partial \ln E_{h,g}}{\partial \ln \tilde{p}_{ih,k}} \propto (\mathcal{R}_h^M / Y_h) \lambda_{ih,k}$, where $\delta_h \approx \lambda_{ih,k} \approx 0$.

$$\begin{aligned} \frac{\partial V_h / \partial w_h}{\partial V_h / \partial Y_h} \frac{dw_h}{d \ln \tilde{p}_{fh,k}} = & -\bar{\tau} \sum_g \left\{ p_{fh,g} q_{fh,g} \left(\frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{fh,k}} + \frac{\partial \ln q_{fh,g}}{\partial \ln \tilde{p}_{hh,g}} \tilde{\alpha}_{h,s}^{f,k} \right) + \right. \\ & \left. + p_{fh,g} q_{fh,g} \tilde{\alpha}_{f,g}^{f,k} - \tilde{p}_{hf,g} q_{hf,g} \frac{\partial \ln q_{hf,g}}{\partial \ln \tilde{p}_{hf,g}} \tilde{\alpha}_{h,g}^{f,k} \right\} \end{aligned}$$

Finally, we have to plug Equations back into the F.O.C.. To simplify some terms we can use the zero profits condition, which entails that

$$\begin{aligned} \frac{\partial \Pi_h(\tilde{\mathbf{p}}; \mathbf{w})}{\partial \ln \tilde{p}_{fh,k}} &= \sum_g \sum_{i=h,f} \left(\frac{\partial \Pi_h}{\partial \ln p_{hi,g}} \frac{\partial \ln p_{hi,g}}{\partial \ln \tilde{p}_{fh,k}^{\mathcal{I}}} \right) + \frac{\partial \Pi_h}{\partial \ln \tilde{p}_{fh,k}^{\mathcal{I}}} + \sum_g \left(\frac{\partial \Pi_h}{\partial \ln \tilde{p}_{hh,k}^{\mathcal{I}}} \frac{\partial \ln \tilde{p}_{hh,g}^{\mathcal{I}}}{\partial \ln \tilde{p}_{fh,k}^{\mathcal{I}}} \right) \\ &= \sum_g \sum_{i=h,f} \left(p_{hi,g} q_{hi,g} \tilde{\alpha}_{h,g}^{f,k} \right) - p_{fh,k} q_{fh,k}^{\mathcal{I}} - \sum_g \left(p_{hh,g} q_{hh,g}^{\mathcal{I}} \tilde{\alpha}_{h,g}^{f,k} \right) = 0, \end{aligned}$$

Combining the above expressions, the F.O.C. can be stated as

$$\begin{aligned} & \sum_g \left(\left[\frac{\tilde{p}_{fh,g}}{p_{fh,g}} - (1 + \bar{\tau}) \right] \left(\varepsilon_{fh,g} + \varepsilon_{fh,g}^{hh,g} \tilde{\alpha}_{h,g}^{f,k} \right) - (1 + \bar{\tau}) \tilde{\alpha}_{f,g}^{f,k} \right) p_{fh,g} q_{fh,g} \\ & + \sum_g \left(p_{hf,g} q_{hf,g} \left[(1 + \bar{\tau}) \tilde{\alpha}_{h,g}^{f,k} + \bar{\tau} \varepsilon_{hf,g} \tilde{\alpha}_{h,g}^{f,k} \right] \right) = 0 \end{aligned}$$

where note that $\tilde{p}_{fh,g} / p_{fh,g} = 1 + t_{fh,k}$. To further simplify the above expression, we can divide by $1 + \bar{\tau}$ and gross expenditure in Foreign (which equal gross production revenue since Foreign collects no tax revenue, i.e., $\sum_k E_{f,k} = \sum_k \mathcal{Y}_{f,k}$), which yields

$$\begin{aligned} \sum_g \left(\left[\left(\frac{1 + t_{fh,g}}{1 + \bar{\tau}} - 1 \right) \left(\varepsilon_{fh,g} + \varepsilon_{fh,g}^{hh,g} \tilde{\alpha}_{h,g}^{f,k} \right) - \tilde{\alpha}_{f,g}^{f,k} \right] r_{fh,g} \right. \\ \left. + \left[\left(1 + \frac{\bar{\tau}}{1 + \bar{\tau}} \varepsilon_{hf,g} \right) \tilde{\alpha}_{h,g}^{f,k} \right] \lambda_{hf,k} \right) = 0. \end{aligned} \quad (15)$$

The next step is to characterize $\bar{\tau} \equiv \frac{-\partial(W_f/\partial w_f)}{\partial \sum_g (p_{fh,g}q_{fh,g} - p_{hf,g}q_{hf,g})/\partial w_f}$, where the partial derivative denotes the derivative with respect to w_f holding $\bar{p}_{fh,k}$ fixed.

$$\begin{aligned}\bar{\tau} &= \frac{-\sum_g p_{fh,k}q_{fh,g} \frac{\partial \ln p_{fh,g}}{\partial \ln w_f}}{\sum_g p_{fh,g}q_{fh,g} \frac{\partial \ln p_{fh,g}}{\partial \ln w_f} - p_{hf,g}q_{hf,g} \left(\frac{\partial \ln q_{hf,g}}{\partial \ln \bar{p}_{ff,g}} \frac{\partial \ln p_{ff,k}}{\partial \ln w_f} + \frac{\partial \ln q_{hf,g}}{\partial \ln E_{f,g}} \frac{\partial \ln E_{f,g}}{\partial \ln w_f} \right)} \\ &= \frac{-\sum_g p_{fh,k}q_{fh,g} \frac{\partial \ln p_{fh,g}}{\partial \ln w_f}}{\sum_g p_{fh,g}q_{fh,g} \frac{\partial \ln p_{fh,g}}{\partial \ln w_f} - \sum_g p_{hf,g}q_{hf,g} (\varepsilon_{hf,g} + \eta_{hf,g}) \delta_{f,g}} \\ &= \frac{-\sum_g p_{fh,k}q_{fh,g} \delta_{f,k}}{\sum_g p_{fh,g}q_{fh,g} \delta_{f,k} + \sum_g p_{hf,g}q_{hf,g} \varepsilon_{hf,k} \delta_{f,k}} = -\frac{1}{1 + \sum_g \omega_{hf,g} \varepsilon_{hf,g}} = -\frac{1}{1 + \bar{\varepsilon}_{hf}},\end{aligned}$$

where the second line follows from the well-known result in consumer theory that $\sum_i (\varepsilon_{hf,g}^{f,i}) + \eta_{hf,g} = 0$; $\bar{\varepsilon}_{hf}$ denotes the elasticity of foreign demand for labor as defined under Definition D5; and $\omega_{hf,g} \equiv \delta_{f,k} \hat{\lambda}_{hf,k} / \sum_k \delta_{f,k} \hat{r}_{fh,k}$ is the weight assigned to industry g when calculating the elasticity of labor demand. $\delta_{f,g}$, meanwhile, denotes the total contribution of Foreign labor to industry-level output. As discussed in Section 3.1, the vector $\delta_f \equiv [\delta_{i,k}]_k$ can be calculated as

$$\delta_f = (\mathbf{I}_K - \alpha_{ff})^{-1} \bar{\alpha}_f,$$

where recall that $\bar{\alpha}_f \equiv [\bar{\alpha}_{f,k}]_k$ is a $K \times 1$ vector and $\alpha_{ff} \equiv [\alpha_{f,g}^{f,k}]_{k,g}$ is a $K \times K$ input-output matrix. Also, note that if Home is sufficiently small compared to Foreign (which is often the case since Foreign is an aggregate of the rest of the world), then $\delta_{f,k} \approx 1$ in which case $\omega_{hf,g} \approx \hat{\lambda}_{hf,g} / \hat{\lambda}_{hf}$. Finally, plugging the all the above expressions back into the Equation 15, and inverting the system specified by this equation implies that

$$1 + t_{fh,k}^* = \frac{\bar{\varepsilon}_{hf}}{1 + \bar{\varepsilon}_{hf}} \left(1 + \frac{\tau_k}{\varepsilon_{fh,k}} \right),$$

where $\tau \equiv [\tau_k]$ is given by

$$\tau = \left[\mathbb{1}_{k=g} + \frac{\varepsilon_{fh,g}^{hh,g}}{\varepsilon_{fh,k}} \tilde{\alpha}_{h,g}^{f,k} \right]_{k,g}^{-1} \left[\sum_{g \in \mathbb{K}} \left(\tilde{\alpha}_{f,g}^{f,k} - \left[1 + \frac{\varepsilon_{hf,g}}{\bar{\varepsilon}_{hf}} \right] \frac{\lambda_{hf,g}}{r_{fh,g}} \tilde{\alpha}_{h,g}^{f,k} \right) \right]_k.$$

D The Multiple Country Case

Suppose there are arbitrarily many countries. Then, the effect of country i 's import tax on own welfare can be expressed as

$$\frac{dW_i(\cdot)}{d \ln(1 + t_{ji,k})} = \frac{\partial W_i(\cdot)}{\partial \ln(1 + t_{ji,k})} + \sum_{j \neq i} \frac{\partial W_i(\cdot)}{\partial \ln w_j} \frac{d \ln w_j}{d \ln(1 + t_{ji,k})}.$$

It is straightforward to show that $\frac{\partial W_i(\cdot)}{\partial \ln w_j} \frac{d \ln w_j}{d \ln(1 + t_{ji,k})} \propto \lambda_{ji} r_{ji,k} \lambda_{ji,k}$, and that based on actual trade data, $\lambda_{ji} r_{ji,k} / \lambda_{ii} r_{ii,k} \approx 0$ for $j \neq i$. So, treating labor in country j as the numeraire, changes in welfare can be approximated to a first-order as

$$\frac{dW_i(\cdot)}{d \ln(1 + t_{ji,k})} \approx \frac{\partial W_i(\cdot)}{\partial \ln(1 + t_{ji,k})} + \frac{\partial W_i(\cdot)}{\partial \ln w_i} \frac{d \ln w_i}{d \ln(1 + t_{ji,k})}.$$

The same applies to export taxes. We can, thus, cast the Country i 's optimal policy problem as one that maximizes $W_i(\tilde{p}, w_i)$ by choosing $\{\tilde{p}_{ij,k}\}$ and $\{\tilde{p}_{ji,k}\}$ subject to feasibility constraints. Following the same steps taken in Appendix B, the F.O.C. corresponding to price $\tilde{p}_{ji,k}$ (i.e., import tax, $t_{ji,k}$) can be expressed as:

$$\sum_g \sum_j \left[\left(\frac{1 + t_{ji,g}}{1 + \bar{\tau}_i} - 1 - \gamma_{ji,g} \right) \frac{\partial \ln q_{ji,g}}{\partial \ln \tilde{p}_{ji,k}} p_{ji,g} q_{ji,g} \right] = 0.$$

$\gamma_{ji,k}$ can be derived by applying the Implicit Function Theorem to $p_{ji,k} = \tau_{ji,k} C_i(\dots, \sum_g \sum_l \tau_{jl,g} q_{jl,g})$, which implies the following:

$$\gamma_{ji,k} = \frac{\gamma_{j,k} r_{ji,k}}{1 - \sum_{l \neq i} \gamma_{j,k} r_{jl,k} \varepsilon_{jl,k}}.$$

Likewise, the F.O.C. corresponding price $\tilde{p}_{ij,k}$ (i.e., export tax, $x_{ij,k}$) can be expressed as:

$$\begin{aligned} & \sum_g \sum_j \left[\left(\frac{1 + t_{ji,g}}{1 + \bar{\tau}} - 1 - \gamma_{ji,g} \right) \frac{\partial \ln q_{ji,g}}{\partial \ln \tilde{p}_{ij,k}} \frac{p_{ji,g} q_{ji,g}}{\tilde{p}_{ij,k}} \right], \quad \forall k \in \mathbb{K} \\ & \sum_g \left(\left[\frac{1}{(1 + x_{ij,g})(1 + \bar{\tau}_i)} - 1 \right] \frac{\hat{\lambda}_{ij,g}}{\hat{\lambda}_{ij,k}} \varepsilon_{ij,g} \right) = 1 - \sum_g \sum_j \frac{\Lambda_{ji,g}}{\tilde{\Lambda}_{ij,k}} \tilde{\alpha}_{jj,g}^{i,k}, \quad \forall k \in \mathbb{K} \end{aligned}$$

where $\Lambda_{j,i,g} \equiv p_{j,i,g}q_{j,i,g} / \sum_{j \neq i} \sum_k p_{j,i,k}q_{j,i,k}$; $\tilde{\Lambda}_{i,j,g} \equiv \tilde{p}_{i,j,g}q_{i,j,g} / \sum_{j \neq i} \sum_k \tilde{p}_{i,j,k}q_{i,j,k}$; and $[\tilde{\alpha}_{jj,g}^{i,k}] = \alpha_j^i \alpha_j^j$ if $j \neq j$, with $[\tilde{\alpha}_{jj,g}^{i,k}] = \alpha_j^i$. Combining the above expressions implies

$$1 + t_{ji,k}^* = \left(1 + \frac{\gamma_{j,k} r_{ji,k}}{1 - \sum_{l \neq i} \gamma_{j,k} r_{jl,k} \varepsilon_{jl,k}} \right) (1 + \bar{t})$$

$$1 + x_{ij,k}^* = \frac{\varepsilon_{ij,k}}{1 + \varepsilon_{ij,k} + \zeta_{ij,k} - \sum_g \sum_j \frac{\Lambda_{j,i,g}}{\tilde{\Lambda}_{i,j,k}} \tilde{\alpha}_{jj,g}^{i,k}} (1 + \bar{t})^{-1}.$$

Uniformity of Import Taxes. Here, we establish the uniformity of import taxes (when $\gamma_{i,k} = 0$) without invoking any first-order approximation. Analogous to the two-country model, welfare in country i can be expressed as $W_i = \partial V_i(Y_i, \tilde{\mathbf{p}}_i)$, where $Y_i = w_i L_i + \sum_k (t_{ji,k} p_{ji,k} q_{ji,k} + x_{ij,k} p_{ij,k} q_{ij,k})$. Correspondingly, W_i is uniquely determined by the vector of import and export taxes, $\mathbf{t}_i = \{t_{ji,k}\}$ and $\mathbf{x}_i = \{x_{ij,k}\}$, plus the vector of country-level wages, $\mathbf{w} = \{w_j\}$:

$$W_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{w}) = V_i(Y_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{w}), \tilde{\mathbf{p}}_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{w}))$$

Defining $D_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{w}) = w_i L_i - \sum_k \sum_\ell (p_{i\ell,k} q_{i\ell,k})$, the equilibrium vector of aggregate wages, \mathbf{w} , solves the following system of equations:

$$\begin{cases} D_1(\mathbf{t}_1, \mathbf{x}_1; \mathbf{w}) = 0 \\ \vdots \\ D_N(\mathbf{t}_N, \mathbf{x}_N; \mathbf{w}) = 0 \end{cases} \quad (16)$$

Keeping the above observation in mind, we can write the F.O.C. with respect to $t_{ji,k}$ as

$$\begin{aligned} \frac{dW_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{w})}{d(1 + t_{ji,k})} &= \frac{\partial V_i(Y_i, \tilde{\mathbf{p}}_i)}{\partial Y_i} \left[\frac{\partial Y_i}{\partial (1 + t_{ji,k})} + \frac{\partial Y_i}{\partial w_i} \frac{dw_i}{d(1 + t_{ji,k})} \right] \\ &+ \sum_{g \in \mathbb{K}} \sum_{\ell \in \mathbb{C}} \left(\frac{\partial V_i(Y_i, \tilde{\mathbf{p}}_i)}{\partial \tilde{p}_{li,g}} \left[\frac{\partial \tilde{p}_{li,g}}{\partial (1 + t_{ji,k})} + \frac{\partial \tilde{p}_{li,g}}{\partial w_\ell} \frac{dw_\ell}{d(1 + t_{ji,k})} \right] \right) \\ &= \frac{\partial V_i(Y_i, \tilde{\mathbf{p}}_i)}{\partial Y_i} \left\{ \frac{\partial Y_i}{\partial (1 + t_{ji,k})} + \frac{\partial V_i(Y_i, \tilde{\mathbf{p}}_i) / \partial \tilde{p}_{ji,k}}{\partial V_i(Y_i, \tilde{\mathbf{p}}_i) / \partial Y_i} \frac{\partial \tilde{p}_{ji,g}}{\partial (1 + t_{ji,k})} + \sum_\ell \left(\frac{\partial W_i(\cdot)}{\partial w_\ell} \frac{dw_\ell}{d(1 + t_{ji,k})} \right) \right\} \end{aligned}$$

Invoking Roy's identity, $\frac{\partial V_i(Y_i, \tilde{p}_i)/\partial \tilde{p}_{ji,k}}{\partial V_i(Y_i, \tilde{p}_i)/\partial Y_i} = q_{ji,k}$, and noting that

$$\frac{\partial Y_i}{\partial \ln(1+t_{ji,k})} = \tilde{p}_{ji,k} q_{ji,k} + \sum_{\ell \in \mathbb{C}} \sum_{g \in \mathbb{K}} \left[t_{\ell i,k} p_{\ell i,k} q_{\ell i,k} \left(\varepsilon_{\ell i,g}^{ji,k} + \eta_{\ell i,g} \frac{\partial \ln Y_i}{\partial \ln(1+t_{ji,k})} \right) \right]$$

Plugging the above equation back into the F.O.C. and defining $\Delta_{\ell i,g}^{ji,k} \equiv \varepsilon_{\ell i,g}^{ji,k} + \eta_{\ell i,g} \frac{\partial \ln Y_i}{\partial \ln(1+t_{ji,k})}$, will yield the following optimality condition

$$\frac{dW_i(\mathbf{t}_i, \mathbf{x}_i; \mathbf{w})}{d \ln(1+t_{ji,k})} = \frac{\partial V_i(Y_i, \tilde{p}_i)}{\partial Y_i} \left[\tilde{p}_{ji,k} q_{ji,k} - \tilde{p}_{ji,k} q_{ji,k} + \sum_{\ell \in \mathbb{C}} \sum_{g \in \mathbb{K}} \left(t_{\ell i,k} p_{\ell i,k} q_{\ell i,k} \Delta_{\ell i,g}^{ji,k} \right) - \sum_{\ell \in \mathbb{C}} \left(v_{i\ell} \frac{d \ln w_\ell}{d \ln(1+t_{ji,k})} \right) \right]$$

where $v_{i\ell} \equiv \partial W_i / \partial \ln w_\ell$. Applying the implicit function theorem to the System of Equations 16, we can solve for $d \ln \mathbf{w} / d \ln 1 + \mathbf{t}_i$ as follows:

$$\begin{bmatrix} \frac{d \ln w_1}{d \ln(1+t_{1i,k})} & \cdots & \frac{d \ln w_1}{d \ln(1+t_{Ni,k})} \\ \vdots & \ddots & \vdots \\ \frac{d \ln w_N}{d \ln(1+t_{1i,k})} & \cdots & \frac{d \ln w_N}{d \ln(1+t_{Ni,k})} \end{bmatrix} = \left(\frac{\partial \ln \mathbf{D}}{\partial \ln \mathbf{w}} \right)^{-1} \begin{bmatrix} \frac{\partial \ln D_1}{\partial \ln(1+t_{1i,k})} & \cdots & \frac{\partial \ln D_1}{\partial \ln(1+t_{Ni,k})} \\ \vdots & \ddots & \vdots \\ \frac{\partial \ln D_N}{\partial \ln(1+t_{1i,k})} & \cdots & \frac{\partial \ln D_N}{\partial \ln(1+t_{Ni,k})} \end{bmatrix},$$

Letting $\tau_{\ell i}$ denotes element ℓi of matrix $(\partial \ln \mathbf{D} / \partial \ln \mathbf{w})^{-1}$, the above system implies that for every $\ell \in \mathbb{C}$

$$\frac{d \ln w_\ell}{d \ln(1+t_{ji,k})} = \sum_{\ell \in \mathbb{C}} \left(\tau_{\ell n} \sum_{g \in \mathbb{K}} \left(p_{ni,g} q_{ni,g} \Delta_{ni,g}^{ji,k} \right) \right).$$

Plugging the above expression back into the F.O.C. implies the following

$$\sum_{n \in \mathbb{C}} \sum_{g \in \mathbb{K}} \left(t_{ni,g} p_{ni,g} q_{ni,g} \Delta_{ni,g}^{ji,k} \right) - \sum_{\ell \in \mathbb{C}} \left(v_\ell \sum_{n \in \mathbb{C}} \left(\tau_{\ell n} \sum_{g \in \mathbb{K}} p_{ni,g} q_{ni,g} \Delta_{ni,g}^{ji,k} \right) \right) = 0.$$

The above expression can in turn be rearranged as

$$\begin{aligned} & \sum_{n \in \mathbb{C}} \sum_{g \in \mathbb{K}} \left[\left(t_{ni,g} - \sum_{\ell \in \mathbb{C}} \tau_{\ell n} v_{i\ell} \right) p_{ni,g} q_{ni,g} \Delta_{ni,g}^{ji,k} \right] \\ & = \sum_{n \in \mathbb{C}} \sum_{g \in \mathbb{K}} \left[(t_{ni,g} - \bar{\tau}_{ni}) p_{ni,g} q_{ni,g} \Delta_{ni,g}^{ji,k} \right] = 0, \end{aligned}$$

where $\bar{\tau}_{ni} \equiv \sum_{\ell \in \mathbb{C}} \tau_{\ell n} v_{i\ell}$, or, equivalently

$$\mathbf{T}_i \Delta_i = \mathbf{0},$$

where $\mathbf{T}_i = [t_{ni,g} - \bar{\tau}_{ni}]_{ng \in \mathbb{C} \times \mathbb{K}}$ and $\Delta_i = [p_{ni,g} q_{ni,g} \Delta_{ni,g}^{j,k}]_{ng, jk \in \mathbb{C} \times \mathbb{K}}$ are respectively $1 \times N \cdot K$ and $N \cdot K \times N \cdot K$ matrixes. If $\det \Delta_i \neq 0$, then $\mathbf{T}_i = \mathbf{0}$ is the unique solution to the above system, which implies that the optimal import tax is uniform across products originating from the same exporting country:

$$t_{ji,k}^* = \bar{\tau}_{ji}, \quad \forall j, k.$$

E Quantitative Implementation with Many Countries

This appendix extends the quantitative procedure outlined in Section 6 to environments that feature arbitrarily many countries.

Baseline Model without IO Networks. Without input-output network (i.e., $\tilde{\alpha}_{jj,g}^{i,k} = 0$) and with CES-Cobb-Douglas preferences (i.e., $\varepsilon_{ij,k} = -1 - \epsilon_k(1 - \lambda_{ij,k})$), Equation 8 implies the an optimal export tax on good ji, k that is equal to

$$1 + x_{ij,k}^* = 1 + \frac{1}{\epsilon_k(1 - \hat{\lambda}_{ij,k} \lambda_{ij,k})},$$

where for each variable, \hat{x} denotes the counterfactual value in the non-cooperative equilibrium. Combining the above formula with the equilibrium conditions specified in Section 6 yields the following system of equations:

$$\begin{cases} 1 + x_{ij,k}^* = 1 + \frac{1}{\epsilon_k(1 - \hat{\lambda}_{ij,k} \lambda_{ij,k})}; & 1 + t_{ji,k}^* = 1 \\ \hat{\lambda}_{ji,k} = \left[\frac{(1+t_{ji,k}^*)(1+x_{ji,k}^*)}{(1+t_{ji,k})(1+x_{ji,k})} \hat{w}_j \right]^{-\epsilon_k} \hat{P}_{i,k}^{\epsilon_k} \\ \hat{P}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[\frac{(1+t_{ji,k}^*)(1+x_{ji,k}^*)}{(1+t_{ji,k})(1+x_{ji,k})} \hat{w}_j \right]^{-\epsilon_k} \lambda_{ji,k} \right) \\ \hat{w}_i w_i L_i = \sum_k \sum_j \left[\frac{\hat{\lambda}_{ij,k} \lambda_{ij,k}}{(1+x_{ij,k}^*)(1+t_{ij,k}^*)} e_{j,k} \hat{Y}_j Y_j \right] \\ \hat{Y}_i Y_i = \hat{w}_i w_i L_i + \sum_k \sum_j \left(\frac{t_{ji,k}^*}{1+t_{ji,k}^*} \hat{\lambda}_{ji,k} \lambda_{ji,k} e_{i,k} \hat{Y}_i Y_i + \frac{x_{ij,k}^*}{1+x_{ij,k}^*} \hat{\lambda}_{ij,k} \lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j \right) \end{cases}.$$

Solving the above system determines the welfare consequences of dissolving the existing trade agreement: $\hat{W}_i = \hat{Y}_i / \prod_k \hat{P}_{i,k}$.

Main Model with IO Networks. Now, consider the main model with input-output networks. By Equation 8, the optimal export tax on good ji, k can be expressed as

$$1 + x_{ij,k}^* = \frac{1 + \epsilon_k(1 - \hat{\lambda}_{ij,k}\lambda_{ij,k})}{\epsilon_k(1 - \hat{\lambda}_{ij,k}\lambda_{ij,k}) - \sum_g \sum_{l \neq i} \frac{\hat{\lambda}_{ij,g}\lambda_{ij,g}\hat{E}_{j,g}E_{j,g}}{(\hat{\lambda}_{ij,k}\lambda_{ij,k}\hat{E}_{j,k}E_{j,k})(1+t_{ij,g})(1+x_{ij,g})} \tilde{\alpha}_{jj,g}^{i,k}},$$

where for each variable, \hat{x} denotes the counterfactual value in the non-cooperative equilibrium; while $\tilde{\alpha}_{jj,g}^{i,k} = \hat{\lambda}_{ij,k}\lambda_{ij,k}\alpha_{j,k,g}$ and

$$\left[\tilde{\alpha}_{jj,g}^{i,k} \right] = \left[\hat{\lambda}_{ij,k}\lambda_{ij,k}\alpha_{j,k,g} \right]_{k,g} \left[\hat{\lambda}_{jj,g}\lambda_{ij,g}\alpha_{j,g,k} \right]_{k,g},$$

if $j \neq j$. Combining the above formula with the equilibrium conditions specified in Appendix 6 yields the following system of equations:

$$\left\{ \begin{array}{l} 1 + x_{ij,k}^* = \frac{1 + \epsilon_k(1 - \hat{\lambda}_{ij,k}\lambda_{ij,k})}{\epsilon_k(1 - \hat{\lambda}_{ij,k}\lambda_{ij,k}) - \sum_g \sum_{l \neq i} \frac{\hat{\lambda}_{ij,g}\lambda_{ij,g}\hat{E}_{j,g}E_{j,g}}{(\hat{\lambda}_{ij,k}\lambda_{ij,k}\hat{E}_{j,k}E_{j,k})(1+t_{ij,g})(1+x_{ij,g})} \tilde{\alpha}_{jj,g}^{i,k}}; \quad 1 + t_{ji,k}^* = 1 \\ \hat{\lambda}_{ji,k} = \left[\frac{(1+t_{ji,k}^*)(1+x_{ji,k}^*)}{(1+t_{ji,k})(1+x_{ji,k})} \hat{w}_j \bar{\alpha}_{i,k} \prod_g \hat{P}_{i,g}^{\alpha_{i,gk}} \right]^{-\epsilon_k} \hat{P}_{i,k}^{\epsilon_k} \\ \hat{P}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[\frac{(1+t_{ji,k}^*)(1+x_{ji,k}^*)}{(1+t_{ji,k})(1+x_{ji,k})} \hat{w}_j \bar{\alpha}_{i,k} \prod_g \hat{P}_{i,g}^{\alpha_{i,gk}} \right]^{-\epsilon_k} \lambda_{ji,k} \right) \\ \hat{r}_{ji,k} r_{ji,k} = \frac{\hat{\lambda}_{ji,k}\lambda_{ji,k}\hat{E}_{i,k}E_{i,k}}{(1+x_{ji,k}^*)(1+t_{ji,k}^*)} / \sum_n \left(\frac{\hat{\lambda}_{jn,k}\lambda_{jn,k}\hat{E}_{n,g}E_{n,g}}{(1+x_{jn,k}^*)(1+t_{jn,k}^*)} \right) \\ \hat{E}_{i,k}E_{i,k} = e_{i,k}\hat{Y}_iY_i + \sum_g \sum_j \alpha_{i,gk} \frac{\hat{\lambda}_{ij,g}\lambda_{ij,g}\hat{Y}_{j,g}Y_{j,g}}{(1+x_{ij,g}^*)(1+t_{ij,g}^*)} \\ \hat{w}_i w_i L_i = \sum_k \sum_j \bar{\alpha}_{i,k} \frac{\hat{\lambda}_{ij,k}\lambda_{ij,k}\hat{Y}_{j,k}Y_{j,k}}{(1+x_{ij,k}^*)(1+t_{ij,k}^*)} \\ \hat{Y}_i Y_i = \hat{w}_i w_i L_i + \sum_k \sum_j \left(\frac{t_{ji,k}^*}{1+t_{ji,k}^*} \hat{\lambda}_{ji,k}\lambda_{ji,k}\hat{E}_{i,k}E_{i,k} + \frac{x_{ij,k}^*}{1+x_{ij,k}^*} \hat{\lambda}_{ij,k}\lambda_{ij,k}\hat{E}_{j,k}E_{j,k} \right) \end{array} \right.$$

Solving the above system determines the welfare consequences of dissolving the existing trade agreement: $\hat{W}_i = \hat{Y}_i / \prod_k \hat{P}_{i,k}$

Table 4: List of countries in quantitative analysis

Country name	WIOD code	Basic aggregation
Australia	AUS	Australia
Brazil	BRA	Brazil
Canada	CAN	Canada
China	CHN	China
Indonesia	IDN	Indonesia
India	IND	India
Japan	JPN	Japan
Korea	KOR	Korea
Mexico	MEX	Mexico
Russia	RUS	Russia
Turkey	TUR	Turkey
Taiwan	TWN	Taiwan
United States	USA	United States
Austria	AUT	
Belgium	BEL	
Bulgaria	BGR	
Cyprus	CYP	
Czech Republic	CZE	
Germany	DEU	
Denmark	DNK	
Spain	ESP	
Finland	FIN	
France	FRA	
United Kingdom	GBR	
Greece	GRC	
Hungary	HUN	
Ireland	IRL	
Italy	ITA	European Union
Netherlands	NLD	
Poland	POL	
Portugal	PRT	
Romania	ROM	
Slovakia	SVK	
Slovenia	SVN	
Sweden	SWE	
Estonia	EST	
Latvia	LVA	
Lithuania	LTU	
Luximburg	LUX	
Malta	MLT	
Rest of the World	RoW	Rest of the World