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

Ahmad Lashkaripour
Indiana University

Volodymyr Lugovskyy
Indiana University

First version: August, 2016
This version: December, 2019

Abstract

Trade taxes are often used as (a) a first-best instrument to manipulate the terms-of-trade, or (b) a second-best instrument to correct pre-existing market distortions. We analyze the (in)effectiveness of trade taxes at achieving these two policy goals. To this end, we derive sufficient statistics formulas for *first-best* and *second-best* trade/domestic taxes in an important class of multi-industry quantitative trade models featuring market distortions. Guided by these formulas, we estimate the key parameters that govern the gains from policy in these frameworks. Our estimates indicate that (i) the gains from terms-of-trade manipulation are small; (ii) trade taxes are remarkably ineffective at correcting domestic distortions; and (iii) the gains from a "deep" trade agreement that corrects distortions at the global level dominate those of any non-cooperative policy alternative.

1 Introduction

Over the past few years, trade and industrial policy have reemerged at the center stage of heated political debates. The Trump administration recently initiated the *National Trade Council* with the mission to restructure the United States' trade and industrial policy. This initiative itself was marketed as an effort to counter the Chinese trade and industrial policy initiative known as "Made in China 2025."[1]

*We are grateful to Adina Ardelean, Dominick Bartelme, Kerem Cosar, Arnaud Costinot, Farid Farrokhi, Harald Fadinger, Fabio Ghironi, David Hummels, Kala Krishna, Konstantin Kucheryavyy, Nuno Limao, Gary Lyn, Ralph Ossa, James Rauch, Andrés Rodríguez-Clare, Kadee Russ, Peter Schott, Alexandre Skiba, Anson Soderbery, Robert Staiger, Jonathan Vogel and conference participants at the Midwest Trade Meetings, Chicago Fed WIE, UECE Lisbon Meetings, 2017 NBER ITI Summer Institute, 2019 WCTW, Indiana University, Purdue University, SIU, and University of Mannheim for helpful comments and suggestions. We thank Nicolas de Roux and Santiago Tabares for providing us with the data on the Colombian HS10 product code changes over time. We are grateful to Fabio Gomez for research assistance. Lugovskyy thanks Indiana University SSRC for financial support. All remaining errors are our own.

[1] This level of interest in trade and industrial policy is not unprecedented. In 1791 Alexander Hamilton approached Congress with a report (namely, *"the Report on the Subject of Manufactures"*) that encouraged the implementation of protective tariffs and targeted subsidies. At that time, these policies were intended to help the US economy catch up with Britain's economy.

The revival of real-world interest in trade and industrial policy has re-ignited academic interest in these topics, resurfacing some old but unresolved questions:

i. How large are the gains from *terms-of-trade (ToT)* manipulation once we account for market distortions and general equilibrium linkages?

ii. How effective are trade taxes as a second-best instrument in correcting pre-existing market distortions?

iii. How large are the gains from "deep" trade agreements that restrict trade policies but allow for cooperation in corrective industrial policies?

Despite all the recent advances in quantitative trade theory, addressing the above questions remains challenging for two reasons:

First, we lack a sharp analytical understanding of optimal *first-best* and *second-best* policies in general equilibrium models with pre-existing market distortions. The prevailing views on optimal trade/industrial policy are mostly based on two-good or partial equilibrium trade models.[2]

Second, we lack appropriate estimates for the parameters that govern the *corrective* and *ToT* gains from policy. We possess multiple estimates for the trade elasticity that governs the gains from ToT manipulation. Likewise, we possess multiple estimates for the degree of firm-level market power that governs the corrective gains from policy. However, we lack a unified technique that can estimate both of these parameters in a mutually consistent way.

In this paper, we aim to fill this gap along both dimensions. We first present an analytical characterization of optimal trade/industrial policy in an important class of multi-industry quantitative trade models where market distortions arise from scale economies or profits. Our characterization delivers simple, sufficient statistics formulas for *first-best* and *second-best* trade and domestic taxes.

We then employ micro-level trade data to estimate the structural parameters that govern the gains from trade/industrial policy in our framework. We finally use these estimates and our sufficient statistics formulas to quantify the gains from trade/industrial policy under various scenarios. Our findings suggest that (*i*) the gains from ToT manipulation are small; (*ii*) trade taxes are remarkably ineffective at correcting domestic distortions; (*iii*) most countries lose from implementing corrective industrial subsidies unilaterally; but (*iv*) the gains from a "deep" trade agreement that corrects market distortions multilaterally can dominate those of any non-cooperative policy alternative.

[2]Recently, several studies have taken notable strides towards characterizing the optimal trade and industrial policies in modern quantitative trade models with distortions. Two prominent examples are Campolmi, Fadinger, and Forlati (2014) and Costinot, Rodríguez-Clare, and Werning (2016) who respectively characterize optimal policy in the two-sector Krugman and Melitz models featuring (*i*) a differentiated and a costlessly traded homogeneous sector, (*ii*) fixed wages, and (*iii*) zero cross-substitutability between sectors. As noted by Costinot et al. (2016), given the state of the literature and beyond these special cases, *"there is little that can be said, in general, about the optimal structure of macro-level taxes"* when dealing with *"arbitrarily many sectors."*

Section 2 presents our theoretical framework. As our baseline model, we adopt a generalized multi-industry Krugman (1980) model that features (i) a non-parametric utility aggregator across industries, and (ii) a nested CES utility aggregator within industries that allows for the degrees of firm-level and country-level market power to diverge. We show that this framework also nests (a) a multi-industry Melitz (2003) model with a Pareto productivity distribution, and (b) a multi-industry Eaton and Kortum (2002) model with industry-level Marshallian externalities.

Section 3 derives sufficient statistics formulas for *first-best* and *second-best* trade and domestic taxes. To summarize these formulas, note that a non-cooperative government can use taxes to correct two types of inefficiency in our framework: (i) a "ToT inefficiency" that concerns unexploited market power with respect to the rest of the world, and (ii) an "allocative inefficiency" that concerns sub-optimal production in high-profit/high-returns-to-scale industries.

Our formulas indicate that *first-best* non-cooperative trade taxes are targeted solely at correcting the ToT inefficiency. Specifically, first-best export taxes equal the optimal markup of a country that acts as a multi-product monopolist, while the first-best import taxes equal the optimal mark-down of a country that acts as a multi-product monopsonist. The first-best domestic taxes, meanwhile, only correct the allocative inefficiency. They are Pigouvian taxes that equal the inverse of the industry-level markup wedge under restricted entry and the inverse of the scale elasticity under free entry.[3]

Our *second-best* trade tax formulas concern a case where governments cannot use corrective domestic subsidies. In that case, optimal trade taxes feature a standard ToT-correcting component and an additional component that corrects allocative inefficiency. This additional component restricts imports and subsidizes exports in high-markup industries in an attempt to mimic the restricted Pigouvian taxes.

Our formulas highlight a lesser unknown trade-off facing second-best trade taxes: Correcting allocative inefficiency with trade taxes often worsens the ToT inefficiency. This trade-off limits the effectiveness of trade taxes as a second-best corrective policy measure; beyond what is implied by the well-known *targeting principle*.

We finally derive formulas for *second-best import taxes* when the use of both export taxes and domestic subsidies is restricted–this scenario has been a focal point in the recent trade policy literature. We show that second-best optimal import taxes feature an additional uniform component that accounts for ToT manipulation through general equilibrium wage effects. This uniform term is redundant when export taxes are available because of the Lerner symmetry, but plays a key role otherwise.

[3] The link between the degree of scale economies (i.e., the scale elasticity) and firm-level market power has deep roots in the literature. Under free entry and monopolistically competitive firms, the following relationship holds (Hanoch (1975) and Helpman (1984)):

$$\text{elasticity of firm-level ouput w.r.t. input cost} = \text{firm-level markup}.$$

The other component of the scale elasticity concerns Marshallian externalities, which operate orthogonally to "firm-level" market power.

Section 5 estimates the structural parameters that govern the gains from policy in our framework. Our sufficient statistics formulas indicate that gains from policy depend on (i) industry-level *trade elasticities* that govern the extent of the ToT inefficiency, and (ii) industry-level *scale elasticities* (or markups) that govern the extent of allocative inefficiency. We develop a new methodology that *simultaneously* estimates these parameters using transaction-level trade data.

Our approach involves fitting a structural import demand function to the universe of Colombian import transactions from 2007 to 2013. Our data covers 226,288 exporting firms from 251 different countries. The main advantage of our approach is its unique ability to *separately* identify the degree of firm-level market power (that determines the scale elasticity) from the degree of national-level market power (that determines the trade elasticity).

The firm-level nature of our empirical strategy exposes us to a non-trivial identification challenge. Standard estimations of import demand are often conducted at the country level and use tariffs as an exogenous instrument to identify the underlying parameters. This identification strategy is, however, not fully-applicable to our firm-level estimation. To attain identification, we compile a comprehensive database on monthly exchange rates. We then interact aggregate movements in monthly bilateral exchange rates with the (lagged) monthly firm-level export sales to construct a *shift-share* instrument that measures exposure to exchange rate shocks at the firm-product-year level.

Section 6 combines our micro-level estimates, our sufficient statistics formulas, and macro-level data from the 2012 World Input-Output Database to quantify the gains from policy for 32 major economies. Our main findings can be summarized as follows:

i. The pure gains from ToT manipulation are relatively small. Suppose a country applies its first-best non-cooperative trade and domestic taxes and the rest of the world does not retaliate. In that case, the average country can raise its real GDP by 1.7% under restricted entry and by 3% under free entry. However, only one-third of these gains are driven by ToT considerations. The rest is driven by restoring allocative efficiency in the domestic economy.

ii. Trade taxes are remarkably ineffective as a second-best measure at tackling allocative inefficiency. Under restricted entry, second-best trade taxes can raise the real GDP by only 0.5% for the average country. That accounts for only 29% of the gains attainable under the fist-best policy schedule. Second-best import taxes are even less effective, raising the real GDP by a mere 0.4%. As noted earlier, using trade taxes to improve allocative inefficiency worsens the ToT. This tension renders trade taxes ineffective at simultaneously targeting and correcting both the ToT and allocative inefficiencies.

iii. For most economies, the gains from "deep" cooperation are larger than the gains from the first-best non-cooperative policy; even if non-cooperation does not trigger retaliation. Deep cooperation in our analysis corresponds to a case where governments abolish trade taxes, put aside political economy motives for intra-national redistribution, and commit to corrective subsidies that restore efficiency at the global level. On average, the gains from deep cooperation are 0.2% percentage points higher than the gains from the first-best non-cooperative policy under free entry. That is, governments are better off forgoing the ToT gains from policy in return for importing subsidized goods from the rest of the world.

iv. While deep cooperation could deliver large welfare gains, most countries can lose from a unilateral application of corrective policies. When countries cannot apply trade taxes, applying corrective Pigouvian taxes can worsen their ToT. The ToT-worsening effect can be so large that it offsets any possible gains from restoring allocative efficiency in the local economy.[4]

Related Literature

Our paper contributes to the existing literature in a number of areas. Our *first-best* tax formulas generalize those in Campolmi et al. (2014), Costinot et al. (2016), Campolmi, Fadinger, and Forlati (2018), and Bartelme, Costinot, Donaldson, and Rodriguez-Clare (2018) by allowing for a large economy, arbitrarily many industries, and arbitrary cross-demand effects.[5] We also present a novel characterization of optimal (first-best) policy in the presence of political economy pressures and input-output linkages.

Our *second-best* tax formulas, to the best of our knowledge, are the first to analytically characterize second-best export and import taxes in general equilibrium quantitative trade models with market distortions. Previous studies on second-best trade taxes either neglect pre-existing market distortions, are partial equilibrium in nature, or focus on the local effects of piecemeal trade tax reforms (e.g., Bond (1990); Bagwell and Staiger (2012); Bagwell and Lee (2018); Caliendo, Feenstra, Romalis, and Taylor (2015); Demidova (2017); Beshkar and Lashkaripour (2019); Costinot and Rodríguez-Clare (2014); Costinot, Donaldson, Vogel, and Werning (2015)).

We build heavily on Kucheryavyy, Lyn, and Rodríguez-Clare (2016) when establishing isomorphism between our baseline model and other workhorse models in the literature. Kucheryavyy et al. (2016) identify the trade and scale elasticities as key determinants of the gains from trade. We estimate these two elasticities and demonstrate that the same elasticities govern the gains from optimal trade and domestic policies.

[4]This result highlights a previously-overlooked benefit of international coordination that is typically overlooked in critiques of global governance (e.g., Rodrik (2019)).

[5]To be clear, we generalize the *macro-level* optimal tax formulas in Costinot et al. (2016). The aforementioned study also derives optimal *firm-level* tax formulas, which are not accommodated by our paper. Campolmi et al. (2018) allow for an arbitrary firm-productivity distribution but abstract from general equilibrium wage, income, and cross-industry demand effects Relatedly, Demidova and Rodriguez-Clare (2009) and Felbermayr, Jung, and Larch (2013) characterize optimal import taxes in a single industry, constant markup framework. In their setups, the market outcome is efficient (Dhingra and Morrow (2019)). Moreover, since the economy is modeled as one industry, import taxes and export taxes are equivalent by the Lerner symmetry. So, the first-best allocation can be attained with only import taxes.

Our estimation of the scale and trade elasticities exhibits key differences from the prior literature. Our approach separately identifies the degree of firm-level power from country-level market power. In comparison, the prior literature typically estimates the trade elasticity with macro-level trade data.[6] The scale elasticity is then normalized to either *zero* or to the inverse of the trade elasticity, which can create an arbitrary link between firm-level and country-level market power (Benassy (1996)).[7]

A notable exception is Bartelme et al. (2018), who (concurrent with us) have developed a strategy to estimate the *product* of scale and trade elasticities. Their approach has the advantage of detecting industry-level Marshallian externalities but relies on the assumption that there are no diseconomies of scale due to industry-specific factors of production. Our approach cannot detect Marshallian externalities but has the advantage of separately identifying the trade elasticity from the scale elasticity and is robust to the presence of industry-specific factors of production.

Several studies have used the exact hat-algebra methodology to study the consequences of counterfactual tariff reductions (Costinot and Rodríguez-Clare (2014); Caliendo and Parro (2015); Ossa (2014, 2016); Spearot (2016)). We contribute to these studies by combining the exact hat-algebra technique with sufficient statistics tax formulas. Doing so simplifies the analysis of optimal policy, allowing us to bypass some of the most important computational challenges facing this literature.

Our quantitative analysis of deep trade agreements has little precedent in the literature. Our closest counterpart is Ossa (2014), who quantifies the gains from an agreement whereby countries cooperate in setting second-best corrective import tariffs. We contribute to Ossa (2014) by analyzing cooperation in first-best corrective policies. Doing so enables us to highlight a previously-neglected tension between the corrective and ToT gains from taxation.

Finally, our paper is related to a vibrant literature that studies the effects of trade openness on allocative efficiency (e.g., Epifani and Gancia (2011); Bond (2013); Holmes, Hsu, and Lee (2014); De Blas and Russ (2015); Edmond, Midrigan, and Xu (2015); Baqaee and Farhi (2019)). We contribute to this literature by highlighting a previously-unknown tension between the allocative and ToT efficiencies. We argue that this tension diminishes the extent to which trade restriction can restore allocative inefficiency.

2 Theoretical Framework

We consider a world economy consisting of a Home country (h) and an aggregate of the rest of the world, which we call Foreign (f). $\mathbb{C} = \{h, f\}$ denotes the set of countries. There are $k = 1, ..., K$ industries in each country, which can differ in fundamentals such as the degree of scale economies and/or market power. $\mathbb{K}$ denotes the set of industries.

[6] See Broda and Weinstein (2006), Simonovska and Waugh (2014), Caliendo and Parro (2015), Soderbery (2015), and Feenstra, Luck, Obstfeld, and Russ (2018) for different estimation approaches.

[7] See Kucheryavyy et al. (2016) for a review on how the scale and trade elasticities are treated across different workhorse trade models.

Each country $i \in \mathbb{C}$ is populated by L_i individuals who are each endowed with one unit of labor, which is the sole factor of production. All individuals are perfectly mobile across the production of different goods and are paid a country-wide wage, w_i, but are immobile across countries.

2.1 Preferences

Cross-Industry Demand. The representative consumer in country $i \in \mathbb{C}$ maximizes a non-parametric utility function subject to their budget constraint. The solution to this problem delivers the following indirect utility as a function of income, Y_i, and industry-level *"consumer"* prices indexes in market i, $\tilde{\boldsymbol{P}}_i \equiv \{\tilde{P}_{i,k}\}$:

$$V_i\left(Y_i, \tilde{\boldsymbol{P}}_i\right) = \max_{\boldsymbol{Q}_i} \;\; U_i(\boldsymbol{Q}_i)$$
$$s.t. \;\; \sum_{k \in \mathbb{K}} \sum_{j \in \mathbb{C}} \tilde{P}_{i,k} Q_{i,k} = Y_i. \tag{1}$$

$\boldsymbol{Q}_i \equiv \{Q_{i,k}\}$, in the above problem, denotes the vector of composite industry-level consumption bundles, with $\tilde{\boldsymbol{P}}_i \equiv \{\tilde{P}_{i,k}\}$ denoting the corresponding consumer price indexes. To avoid any confusion later on, we should emphasize that the *tilde* notation distinguishes between "consumer" and "producer" prices. To keep track of Problem 1, we denote the *industry-level* Marshallian demand function implied by the above problem as $\boldsymbol{Q}_i = \mathcal{D}_i\left(Y_i, \tilde{\boldsymbol{P}}_i\right)$.

Within-Industry Demand. Each industry-level bundle aggregates over various country-level composite varieties: $Q_{i,k} = \mathcal{U}_{i,k}\left(Q_{hi,k}, Q_{fi,k}\right)$. Each country-level composite variety itself aggregates over multiple firm-level varieties: $Q_{ji,k} = \mathcal{U}_{ji,k}(\boldsymbol{q}_{ji,k})$ where $\boldsymbol{q}_{ji,k} = \left\{q_{ji,k}\left(\omega\right)\right\}_{\omega \in \Omega_{j,k}}$ is a vector with each element $q_{ji,k}\left(\omega\right)$ denoting the quantity consumed of firm ω's output. $\Omega_{j,k}$ denotes the set of all firms serving industry k from country j.[8] As a key restriction, we assume that the within-industry utility aggregator, $\mathcal{U}_{i,k}\left(.\right)$, has a nested-CES structure. This assumption enables us to abstract from variable markups, thereby directing attention to the scale-driven or profit-shifting effects of policy.

A1. *The within-industry utility aggregator is nested-CES ($\rho_k \in (0,1)$):*

$$Q_{i,k} = \left(\sum_{j \in \mathbb{C}} Q_{ji,k}^{\rho_k} \right)^{1/\rho_k},$$

where $Q_{ji,k}$ is a CES-composite bundle of firm-level varieties from country j ($\varrho_k \in (0,1)$):

$$Q_{ji,k} = \left(\int_{\omega \in \Omega_{j,k}} q_{ji,k}\left(\omega\right)^{\varrho_k} d\omega \right)^{1/\varrho_k}.$$

[8] As we elaborate later, in the presence of firm-selection effects, only a sub-set of firms in set $\Omega_{j,k}$ serve each market. In that case, $\Omega_{ji,k} \subset \Omega_{j,k}$ will denote the set of firms serving market i from country j.

Defining $\epsilon_k \equiv \rho_k/(1-\rho_k)$ and $\xi_k \equiv \varrho_k/(1-\varrho_k)$, the *national-level* demand facing country j in industry k is given by

$$Q_{ji,k} = \left(\tilde{P}_{ji,k}/\tilde{P}_{i,k}\right)^{-\epsilon_k - 1} Q_{i,k}, \tag{2}$$

where $\tilde{P}_{ji,k} = \left(\sum_{\omega\in\Omega_{ji,k}} \tilde{p}_{ji,k}(\omega)^{-\xi_k}\right)^{-1/\xi_k}$ and $\tilde{P}_{i,k} = \left(\sum_{j\in\mathbb{C}} \tilde{P}_{ji,k}^{-\epsilon_k}\right)^{-1/\epsilon_k}$ denote the origin-specific and industry-level price indexes; while $Q_{i,k}$ denotes the industry-level demand that is given by $Q_{i,k} = \mathcal{D}_{i,k}\left(Y_i, \tilde{\boldsymbol{P}}_i\right)$. The demand facing individual firms from country j is also given by

$$q_{ji,k}(\omega) = \left(\frac{\tilde{p}_{ji,k}(\omega)}{\tilde{P}_{ji,k}}\right)^{-1-\xi_k} \left(\frac{\tilde{P}_{ji,k}}{\tilde{P}_{i,k}}\right)^{-1-\epsilon_k} \mathcal{D}_{i,k}\left(Y_i, \tilde{\boldsymbol{P}}_i\right). \tag{3}$$

Importantly, the above demand structure allows for the *firm-level* and *national-level* degrees of market power to diverge, with ξ_k reflecting the degree of firm-level market power and ϵ_k reflecting the degree of national-level market power.

National-Level Demand Elasticities. The aggregate demand facing country j varieties in market i in industry k, $Q_{ji,k} = Q_{ji,k}\left(Y_i, \tilde{\boldsymbol{P}}_{i,k}\right)$, is a function of total income in market i, as well as the vector of all origin-specific consumer price indexes in that market, $\tilde{\boldsymbol{P}}_{i,k} \equiv \{\tilde{P}_{ji,g}\}$. Considering this and for future reference, we formally define the price elasticity of demand facing national-level exports as follows.

Definition D1. (i) [own price elasticity] $\varepsilon_{ji,k} \equiv \partial \ln Q_{ji,k}(.)/\partial \ln \tilde{P}_{ji,k}$;
(ii) [cross-price elasticity] $\varepsilon_{ji,k}^{\jmath\iota,g} \equiv \partial \ln Q_{ji,k}(.)/\partial \ln \tilde{P}_{\jmath\iota,g}$ *for* $\jmath\iota, g \neq ji, k$.

In the special case with *zero* cross-substitutability between industries, the national-level demand elasticity is fully determined by parameter ϵ_k and national-level expenditure shares. Namely, $\varepsilon_{ji,k}^{\jmath i,g} = 0$ if $g \neq k$, while

$$\varepsilon_{ji,k}^{\jmath i,k} = \begin{cases} -1 - \epsilon_k(1-\lambda_{ji,k}) & \jmath = j \\ \epsilon_k \lambda_{\jmath i,g} & \jmath \neq j \end{cases},$$

where $\lambda_{ji,k} \equiv \tilde{P}_{ji,k}Q_{ji,k}/\sum_{\jmath} \tilde{P}_{\jmath i,k}Q_{\jmath i,k}$ denotes the share of country i's expenditure on country j varieties within industry k. Present cross-substitutability between industries, the demand elasticity features an additional term that accounts for such cross-industry demand effects.

2.2 Production and Firms

Economy $j \in \mathbb{C}$ is populated with a mass, $M_{j,k}$, of single-product firms in industry k that compete under monopolistic competition. Labor is the only factor of production. Firm entry into industry k is either free or restricted. Under restricted entry, $M_{j,k} = \overline{M}_{j,k}$

is invariant to policy. Under free entry, a pool of ex-ante identical firms can pay an entry cost $w_j f_k^e$ to serve industry k from country j.

After paying the entry cost, each firm $\omega \in \Omega_{j,k}$ draws a productivity $z(\omega) \geq 1$ from distribution $G_{j,k}(z)$. The marginal cost of producing and delivering goods to market i by firm ω is, thus, given by $\tau_{ji,k} w_j / z(\omega)$, where $\tau_{ji,k}$ denotes a flat iceberg transport cost. We summarize these production-side assumptions under the following assumption regarding *national-level* producer price indexes.

A2. *The "producer" price index of composite good ji,k is given by*

$$P_{ji,k} = (1+\mu_k)\, \bar{a}_{ji,k} w_j M_{j,k}^{-\psi_k},$$

where $\bar{a}_{ji,k} = \tau_{ji,k} \left[\int_1^\infty z^{\xi_k} dG_{j,k}(z)\right]^{-\frac{1}{\xi_k}}$ (average unit labor cost of production/delivery), μ_k (markup margin), and ψ_k (scale elasticity) are invariant to policy.

To elaborate, the fact that μ_k or ψ_k are invariant to policy follows from our previous assumption, A1. The fact that $\bar{a}_{ji,k}$ is invariant to policy follows from the implicit assumption that there are no fixed costs associated with serving individual markets. This assumption rules out selection effects but is not as restrictive as it may appear. As elaborated in Subsection 2.5, the above setup is isomorphic to a model that admits selection effects as long as the productivity distribution, $G_{j,k}(z)$, is Pareto.

In our baseline model, the scale elasticity and the markup wedge are identical: $\psi_k = \mu_k = 1/\xi_k$. We, however, maintain the notational distinction between μ_k and ψ_k for two reasons. First, markups and scale effects influence welfare in distinct ways. Second and more importantly, our theory nests models featuring Marshallian externalities and firm-selection effects (à la Melitz (2003)) in which $\mu_k \neq \psi_k$. We elaborate more on these models in Section 2.5.

We can summarize the supply-side of the economy in terms of a set producer prices and aggregate profits. Specifically, given the constancy of markups, producer prices are given by

$$P_{ji,k} = \begin{cases} \bar{\rho}_{ji,k} w_j & \text{if entry is restricted} \\ \bar{\varrho}_{ji,k} w_j \mathcal{Q}_{j,k}^{-\psi_k/(1+\psi_k)} & \text{if entry is free} \end{cases} \tag{4}$$

where $\bar{\varrho}_{ji,k} \equiv \bar{a}_{ji,k}(1+\mu_k)/\mu_k f_k^e$ and $\bar{\rho}_{ji,k} \equiv \bar{a}_{ji,k}(1+\mu_k)\, \overline{M}_{j,k}^{-\psi_k}$ are invariant to policy; with $\mathcal{Q}_{j,k} \equiv \sum_i \bar{a}_{ji,k} Q_{ji,k}$ denoting the effective output of country j in industry k. Given A1 and A2, profits (gross of entry costs) are a constant faction of total revenue, with total net profits in country i given by

$$\Pi_i = \begin{cases} \sum_k \sum_n \left(\frac{\mu_k}{1+\mu_k} P_{in,k} Q_{in,k}\right) & \text{if entry is restricted} \\ 0 & \text{if entry is free} \end{cases}. \tag{5}$$

2.3 The Instruments of Policy

The government in country i has access to a full set revenue-raising trade and domestic taxes. Namely,

i. industry-level ***import taxes***, $\{t_{ji,k}\}$ that are applied by country i to all varieties imported from country $j \neq i$ in industry k (with $t_{ii,k} = 0$);

ii. industry-level ***export taxes***, $\{x_{ji,k}\}$ that are applied by country j to all export varieties sold to market $i \neq j$ in industry k (with $x_{jj,k} = 0$);

iii. industry-level ***production taxes***, $\{s_{j,k}\}$, that are applied by country j on industry k's output irrespective of the location of final sale; and

iv. industry-level ***consumption tax***es, which are redundant given the availability of the other tax instruments and are henceforth normalized to zero.

The above specification of policy is quite flexible as it accommodates import or export subsidies ($t < 0$ or $x < 0$) as well as production subsidies ($s < 0$). A formal proof for the redundancy of consumption taxes is provided in Appendix A. There is a simple intuition behind this result: The Home country, h, has access to 4 different tax instruments in each industry. These 4 tax instruments can directly affect 3 consumer price variables: $\tilde{P}_{hf,k}$, $\tilde{P}_{fh,k}$, and $\tilde{P}_{hh.k}$. So, by construction 1 of the 4 tax instruments, in each industry, is redundant. Here, we treat the consumption tax as a redundant instrument.

The above tax instruments create a wedge between consumer price indexes, $\{\tilde{P}_{ji,k}\}$ and producer price indexes, $\{P_{ji,k}\}$, as follows

$$\tilde{P}_{ji,k} = \left(1 + t_{ji,k}\right)\left(1 + x_{ji,k}\right)\left(1 + s_{j,k}\right) P_{ji.k}, \quad \forall j, i \in \mathbb{C},\ k \in \mathbb{K}. \tag{6}$$

Each of the tax instruments also generates/exhausts revenue for the imposing country, with the combination of all taxes imposed by country $i \in \mathbb{C}$ producing a revenue equal to

$$\mathcal{R}_i = \sum_{k\in\mathbb{K}} \sum_{j\in\mathbb{C}} \left(t_{ji,k} P_{ji,k} Q_{ji,k} + \left[(1 + x_{ij,k})(1 + s_{i,k}) - 1\right] P_{ij,k} Q_{ij,k}\right), \tag{7}$$

Accounting for tax revenues, total income in country i equals wage income, $w_i L_i$, plus total profits, Π_i, plus total tax revenues: $Y_i = w_i L_i + \Pi_i + \mathcal{R}_i$, with $\mathcal{R}_i$ being positive or negative depending on whether country i's policy consists of net taxes or subsidies.

2.4 Equilibrium

We assume throughout the paper that the necessary and sufficient conditions for the uniqueness of equilibrium are satisfied.

A3. *For any combination of tax instruments, equilibrium is unique.*

Following Kucheryavyy et al. (2016), Assumption A3 is equivalent to assuming that $\psi_k \epsilon_k \leq 1$.[9] Given the above assumption, any combination of taxes ($\boldsymbol{x}$, $\boldsymbol{t}$, and $\boldsymbol{s}$) is consistent with only one vector of wages, $\boldsymbol{w}$. Accordingly, the set of feasible tax-wage combinations can be defined as follows.

Definition. [Feasible Tax-Wage Combinations] *The tax-wage combination $A \equiv (\boldsymbol{x}, \boldsymbol{t}, \boldsymbol{s}; \boldsymbol{w})$ is feasible if (i) Producer prices for any good ji, k are characterized by 4; (ii) Consumer prices for any good ji, k are given by 6; (iii) Industry-level consumption choices are a solution to 1 and the variety-level demand, $Q_{ji,k}$, is characterized by 2; (iv) Factor markets clear, $w_i L_i + \Pi_i = \sum_n \sum_k P_{in,k} Q_{in,k}$, with profits given by 5; and (v) Total income equals factor income plus tax revenue, $Y_i = w_i L_i + \Pi_i + \mathcal{R}_i$, where $\mathcal{R}_i$ is given by 7.*

Following the above definition, we let $\mathbb{F}$ denote the set of *all* feasible tax-wage combinations. For any feasible outcome, $(\boldsymbol{x}, \boldsymbol{t}, \boldsymbol{s}; \boldsymbol{w}) \in \mathbb{F}$, we can express all equilibrium variables including welfare as a function of taxes and wages:

$$W_i(\boldsymbol{x}, \boldsymbol{t}, \boldsymbol{s}; \boldsymbol{w}) \equiv V_i\left(Y_i(\boldsymbol{x}, \boldsymbol{t}, \boldsymbol{s}; \boldsymbol{w}), \tilde{\boldsymbol{P}}_i(\boldsymbol{x}, \boldsymbol{t}, \boldsymbol{s}; \boldsymbol{w})\right).$$

Noting this choice of notation, we present an intermediate result on the equivalence of policy equilibria, which simplifies our subsequent analysis.

Lemma 1. *[**Equivalence of Policy Equilibria**] For any a and $\tilde{a} \in \mathbb{R}_+$ (i) if $A \equiv (\boldsymbol{1} + \boldsymbol{t}_i, \boldsymbol{t}_{-i}, \boldsymbol{1} + \boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{1} + \boldsymbol{s}_i, \boldsymbol{s}_{-i}; w_i, \boldsymbol{w}_{-i}) \in \mathbb{F}$, then $A' \equiv (a(\boldsymbol{1} + \boldsymbol{t}_i), \boldsymbol{t}_{-i}, (\boldsymbol{1} + \boldsymbol{x}_i)/a, \boldsymbol{x}_{-i}, \tilde{a}(\boldsymbol{1} + \boldsymbol{s}_i), \boldsymbol{s}_{-i}; a w_i/\tilde{a}, \boldsymbol{w}_{-i}) \in \mathbb{F}$. Moreover, (ii) the real welfare is preserved under allocations A and A': $W_i(A) = W_i(A')$.*

The above lemma, which is proven in Appendix B, implies that there are multiple optimal tax combinations. This result is a simple extension of the Lerner symmetry to an environment where domestic taxes are also applied. It is easy to verify that "one" of the multiple optimal tax combinations exhibits a *nominal* wage rate equal to that implied by the no-tax equilibrium. This result allows us to derive the optimal tax schedule while netting out *either* the general equilibrium wage effects or the general equilibrium income effects.

2.5 Other Micro-Founded Models Nested by our Framework

Before moving forward, it is useful to outline the breadth of our theoretical framework. When presenting our framework, we used the generalized multi-industry Krugman (1980) model as our choice of micro-foundation. However, in terms of macro-level representation, our framework nests two other widely-used models in the literature:

[9] Two remarks regarding the above claim are in order. First, Kucheryavyy et al. (2016) prove uniqueness under zero cross-substitutability between industries. Second, they prove their uniqueness result in the presence of iceberg trade barriers. This is, however, not problematic, as there is a one-to-one correspondence between the present model, which features revenue-raising taxes, and the Kucheryavyy et al. (2016) model, which features iceberg trade barriers.

Table 1: Summary of Key Variables

Variable	Description
$\tilde{P}_{ji,k}$	Consumer price index (exporter j–importer i–industry k)
$P_{ji,k}$	Producer price index (exporter j–importer i–industry k)
Y_i	Total income in country i
Π_i	Total profits in country i
$\mathcal{R}_i$	Total tax revenue in country i
$w_i L_i$	Factor income in country i (wage$\times$population size)
$x_k \equiv x_{hf,k}$	Home's export tax in industry k (corresponding vector: $\boldsymbol{x} \equiv \{x_k\}$)
$t_k \equiv t_{fh,k}$	Home's import tax in industry k (corresponding vector: $\boldsymbol{t} \equiv \{t_k\}$)
$s_k \equiv s_{h,k}$	Home's production tax in industry k (corresponding vector: $\boldsymbol{s} \equiv \{s_k\}$)
$\lambda_{ji,k}$	Within-industry expenditure share (good ji,k): $\tilde{P}_{ji,k}Q_{ji,k}/\sum_\jmath \tilde{P}_{\jmath i,k}Q_{\jmath i,k}$
$r_{ji,k}$	Within-industry revenue share (good ji,k): $P_{ji,k}Q_{ji,k}/\sum_\iota P_{j\iota,k}Q_{j\iota,k}$
$\acute{\lambda}_{ji,k}$	Overall expenditure share (good ji,k): $\tilde{P}_{ji,k}Q_{ji,k}/Y_i$
$\acute{r}_{ji,k}$	Overall revenue share (good ji,k): $P_{ji,k}Q_{ji,k}/w_jL_j$
$e_{i,k}$	Industry-level expenditure share (country i–industry k)
ξ_k	Degree of firm-level market power in industry k ~ sub-national CES parameter
ϵ_k	Trade elasticity in industry k ~ cross-national CES parameter
μ_k	Markup/Profit margin in industry k
ψ_k	Scale elasticity in industry k
$\varepsilon_{ji,k}$	Own price elasticity of national-level demand: $\partial \ln Q_{ji,k}/\partial \ln \tilde{P}_{ji,k}$
$\varepsilon_{ji,k}^{ji,g}$	Cross price elasticity of national-level demand: $\partial \ln Q_{ji,k}/\partial \ln \tilde{P}_{ji,g}$

i. The multi-industry Eaton and Kortum (2002) model with Marshallian externalities as a special case where $\mu_k = 0$ and ψ_k equals the elasticity of industry-level productivity with respect to industry-level employment. A version of this isomorphism has been already established in Kucheryavyy et al. (2016). In Appendix C, though, we establish a deeper version of this isomorphism by demonstrating that the nested-CES demand function implied by A1 may analogously arise from within-industry specialization á la Eaton and Kortum (2002). Provided that the firm-level productivity distribution is nested Fréchet. Under this interpretation, structural parameters ϵ_k and ξ_k represent the degrees of cross-national and sub-national productivity dispersion.

ii. The Melitz-Pareto model as a special case where $\mu_k = (\xi_k - 1)/\gamma_k \xi_k$ and $\psi_k =$

$1/\gamma_k$ with γ_k being the shape of the Pareto productivity distribution. Appendix D establishes the isomorphism between our benchmark model and the Melitz-Pareto model featuring firm-selection effects.[10] In the Melitz-Pareto interpretation of the model, the national-level demand elasticities also assume a different specification, which we fully elaborate on in Appendix D.

The analytic optimal policy formulas presented next, therefore, apply to the above two models with no further qualifications. When mapping the formulas to data, however, the choice of micro-foundation is consequential since the structural parameters of the model assume different interpretations under different micro-foundations. We elaborate more on this issue in Section 6.

3 Sufficient Statistics Formulas for Optimal Policy

In this section, we derive sufficient statistics formulas for optimal trade and domestic taxes. These formulas are later employed to quantify the gains from policy. Before deriving our formulas, let us highlight the inefficiencies taxes can target in our setup. From the perspective of the Home government, taxes can correct two distinct types of inefficiency:

a) *Terms-of-trade (ToT) inefficiency* due to unexploited export/import market power with respect to the rest of the world, and

b) *Allocative inefficiency* due to differential profit margins or scale elasticities across industries.

Correcting the ToT inefficiency will always lead to a Pareto sub-optimal outcome. That is, by manipulating the ToT, non-cooperative governments transfer consumer/producer surplus from the rest of the world to their own economy. Improving allocative inefficiency, however, is Pareto improving. Specifically, there is sub-optimal global output in high-profit (high-μ) industries under restricted entry; or, analogously, there is sub-optimal global firm-entry in high-returns-to–scale (high-ψ) industries under free entry. When these distortions are corrected, consumers all around the world will reap the benefits.

Optimal Cooperative Policy. As a useful benchmark, we characterize the optimal cooperative policy in our model. Such a policy maximizes global welfare, as defined by a population-weighted sum of welfare across all countries:

$$\max_{(\boldsymbol{x},\boldsymbol{t},\boldsymbol{s};\boldsymbol{w})\in\mathbb{F}} \quad \alpha_h W_h\left(\boldsymbol{x},\boldsymbol{t},\boldsymbol{s};\boldsymbol{w}\right)+\alpha_f W_f\left(\boldsymbol{x},\boldsymbol{t},\boldsymbol{s};\boldsymbol{w}\right),$$

[10] In this process, we borrow heavily from Kucheryavyy et al. (2016).

It is straightforward to show that the solution to above problem involves *zero* trade taxes and corrective Pigouvian subsidies in all countries:

$$\begin{aligned} t^*_{ji,k} = & x^*_{ji,k} = 0 \quad \forall ji,k \\ 1 + s^*_{i,k} = & \frac{1}{1+\mu_k} \quad \forall i,k \end{aligned} \tag{8}$$

The above characterization applies to both the case of free and restricted entries and is analogous to the cooperative policy structure in Bagwell and Staiger (2001). Under free entry, though, we should interpret the corrective subsidies as being proportional to the industry-level scale elasticity, i.e., $1 + s^*_{i,k} = 1/(1+\psi_k)$.

The fact that international cooperation restores allocative efficiency at the global level is an artifact of assuming non-politically-weighted welfare functions. We can, thus, think of the policy schedule highlighted above as the outcome of deep cooperation or a ***deep trade agreement***. As we discuss in Section 4, domestic taxes applied by politically-motivated governments can worsen allocative efficiency without necessarily worsening the ToT.

3.1 First-Best Optimal Trade/Domestic Taxes

Now we turn to characterizing Home's optimal non-cooperative policy. We consider a case where Home applies taxes $t_k \equiv t_{fh,k}$, $x_k \equiv x_{hf,k}$, and $s_k \equiv s_k$, but the rest of the world is passive, $t_{hf,k} = x_{fh,k} = s_{f,k} = 0$. So, the vectors $\boldsymbol{t} \equiv \{t_k\}$, $\boldsymbol{x} \equiv \{x_k\}$, and $\boldsymbol{s} \equiv \{s_k\}$ correspond to only the Home government's policy instruments, hereafter. With this choice of notation, Home's optimal (first-best) non-cooperative taxes solve the following problem:

$$\max_{(\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w})\in\mathbb{F}} \quad W_h\left(\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w}\right) \qquad (\text{P1}).$$

Below, we analytically solve Problem (P1) under both restricted and free entry. The two cases bear some resemblance but also exhibit key differences.

3.1.1 The Restricted Entry Case

Under restricted entry, the *direct* passthrough of a tax onto the own consumer price is complete and the cross-passthrough is zero:

$$\frac{\partial \ln \tilde{P}_{ji,k}\left(\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w}\right)}{\partial \ln(1+t_{\jmath\imath,g})} = \begin{cases} 1 & \text{if } ji,k = \jmath\imath,g \\ 0 & \text{if } ji,k \neq \jmath\imath,g \end{cases}. \tag{9}$$

So, aside from general equilibrium wage and profit-shifting effects, the burden of t_k falls on Home consumers, while the burden of x_k falls on Foreign consumers. Export taxes (x_k) can, therefore, extract monopoly markups from Foreign consumers on an industry-specific basis. Optimal import taxes (t_k), however, can impose markups/markdowns on foreign prices ($\tilde{P}_{hf,g}$ or $P_{fh,g}$) only through a flat general equilibrium wage effect. Production taxes can serve as a first-best instrument to eliminate allocative inefficiency arising from markup heterogeneity. The following theorem summarizes these results, providing sufficient statistics formulas for optimal tax rates.

Theorem 1.A. *[**Optimal Policy Under Restricted Entry**] With restricted entry, the optimal policy is unique up to two uniform tax shifters $\bar{s}$ and $\bar{t} \in \mathbb{R}_+$, and is given by*

$$[(1+s_k^*)\,(1+\mu_k)\,/\,(1+\bar{s})]_k = \mathbf{1}$$
$$[(1+t_k^*)\,/\,(1+\bar{t})]_k = \mathbf{1}$$
$$[1/\,(1+x_k^*)\,(1+\bar{t})]_k = \boldsymbol{\mathcal{E}}_{hf}^{-1}\left(\boldsymbol{I}+\boldsymbol{\mathcal{E}}_{hf}\right)\mathbf{1}$$

where $\boldsymbol{\mathcal{E}}_{hf} \equiv \left[\varepsilon_{hf,k}^{hf,g}\right]_{k,g}$ is a $K \times K$ matrix composed of reduced-form demand elasticities.[11]

The above theorem, which is proven in Appendix E, states that the optimal production tax is a Pigouvian tax that solely corrects allocative inefficiency. The optimal import tax can be normalized to zero by choice of $\bar{t} = 0$ and is essentially redundant. The optimal export tax is equal to the optimal markup of a *multi-product* monopolist, which takes into account all the cross-demand effects. The fact that the uniform tax shifters, $\bar{s}$, and $\bar{t}$, can be assigned any arbitrary value corresponds to the multiplicity of the optimal policy equilibria as highlighted by Lemma 1.

To gain a deeper intuition, consider a special case where there is zero cross-substitutability between industries. In that case, the formulas specified by Theorem 1.A reduces to[12]

$$1+s_k^* = \frac{1+\bar{s}}{1+\mu_k}, \qquad \forall k$$
$$1+t_k^* = 1+\bar{t}, \qquad \forall k$$
$$1+x_k^* = (1+\bar{t})^{-1}\left(1+\frac{1}{\epsilon_k \lambda_{ff,k}}\right), \quad \forall k$$

with the export tax now corresponding to the optimal markup of a single product monopolist facing a CES demand. Based on the above, we can express the optimal tax schedule as a function of three *sufficient statistics*: (i) observable expenditure shares, $\lambda_{ff,k}$, and estimable industry-level parameters(ii) ϵ_k, and (iii) $\mu_k = 1/\xi_k$. This feature greatly simplifies our quantitative analysis of optimal policy in Section 6.

The above formulas indicate that optimal trade taxes are independent of the degree of allocative inefficiency, which is reflected in the μ_k's. The optimal trade tax formulas are in fact identical to that arising from a perfectly competitive model à la Beshkar and Lashkaripour (2019). This outcome corresponds to the result in Kopczuk (2003),

[11] More generally, $\boldsymbol{\mathcal{E}}_{hf} \equiv [\lambda_{hf,g}/\lambda_{hf,k}\varepsilon_{hf,g}^{hf,k}]_{k,g}$. But when preferences are homothetic, the symmetry of the Slutsky matrix entails that $\lambda_{hf,g}\varepsilon_{hf,g}^{hf,k} = \lambda_{hf,k}\varepsilon_{hf,k}^{hf,g}$, which delivers the above specification.

[12] Recall that with zero cross-substitutability, $\varepsilon_{hf,k} = -\epsilon_k\lambda_{ff,k} - 1$ and $\varepsilon_{hf,g}^{hf,k} = 0$ if $k \neq g$.

whereby we can first eliminate the underlying market inefficiency with the appropriate policy instruments and then solve Problem (P1) as if markets where efficient.

3.1.2 The Free Entry Case

Under free entry, taxes can affect Foreign prices through scale or *firm delocation* effects. Specifically, a tax on good ji, k, can affect the scale of production for that good as well as all other goods through cross-demand and firm-entry effects. As a result, the passthrough of taxes onto own price is incomplete (i.e., $\partial \ln \tilde{P}_{ji,k}\left(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}; \boldsymbol{w}\right) / \partial \ln(1 + t_{ji,k}) \neq 1$) while the passthrough onto other international prices is possibly non-zero (i.e., $\partial \ln \tilde{P}_{ji,k}\left(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}; \boldsymbol{w}\right) / \partial \ln(1 + t_{\jmath\imath,g}) \neq 0$).

These considerations present the Home economy with an ability to manipulate its import market beyond what is possible under restricted entry. The optimal tax schedule, characterized by the following theorem, reflects this additional degree of import market power.

Theorem 1.B. ***[Optimal Policy Under Free Entry]*** *Under free entry, Home's optimal first-best policy is unique up to two uniform tax shifters $\bar{s}$ and $\bar{t} \in \mathbb{R}_+$, and is given by*

$$\begin{cases} \left[(1+s_k^*)(1+\psi_k)/(1+\bar{s})\right]_k = \mathbf{1} \\ \left[(1+t_k^*)/\left(\frac{1+\psi_k r_{ff,k}}{1+\psi_k}\right)(1+\bar{t})\right]_k = \mathbf{1} \\ \left[\left(\frac{1+\psi_k}{1+\psi_k r_{ff,k}}\right)/(1+x_k^*)(1+\bar{t})\right]_k = \tilde{\boldsymbol{\mathcal{E}}}_{hf}^{-1}\left(\boldsymbol{I} + \tilde{\boldsymbol{\mathcal{E}}}_{hf}\right)\mathbf{1} \end{cases},$$

where $\tilde{\boldsymbol{\mathcal{E}}}_{hf} \equiv \left[\frac{(1+\psi_g)(1+\psi_k r_{ff,k})}{(1+\psi_k)(1+\psi_g r_{ff,g})}\varepsilon_{hf,k}^{hf,g}\right]_{k,g}$ is a $K \times K$ matrix composed of reduced-form demand elasticities, scale elasticities, and revenue shares.[13]

As with Theorem 1.A, the above theorem characterizes optimal taxes in terms of a set of sufficient statistics: (i) observable expenditure and revenue shares, $\lambda_{ji,k}$, and $r_{ji,k}$, as well as (ii) estimable parameters, ϵ_k and ψ_k. Also, similar to the restricted entry case, optimal domestic taxes are solely corrective Pigouvian taxes, while optimal trade taxes target only the ToT inefficiency.

Theorems 1.A and 1.B, however, exhibit a key difference: import taxes are neither redundant nor uniform under free entry. That is because, under free entry, import taxes can affect producer prices in Foreign ($P_{fh,k}$) through both *firm delocation* and a wage effect. The non-uniform import taxes reflect Home's ability to manipulate its industry-specific import market power to lower $P_{fh,k}$.[14]

As before, we can gain extra intuition by focusing on the special case where there is

[13] More generally, $\boldsymbol{\mathcal{E}}_{hf} \equiv [\frac{(1+\psi_g)(1+\psi_k r_{ff,k})}{(1+\psi_k)(1+\psi_g r_{ff,g})}\frac{\acute{\lambda}_{hf,g}}{\acute{\lambda}_{hf,k}}\varepsilon_{hf,g}^{hf,k}]_{k,g}$. But when preferences are homothetic, the symmetry of the Slutsky matrix entails that $\acute{\lambda}_{hf,g}\varepsilon_{hf,g}^{hf,k} = \acute{\lambda}_{hf,k}\varepsilon_{hf,k}^{hf,g}$, which delivers the above specification.

[14] Note that, if we assume that trade taxes generate no revenue à la Ossa (2011), the optimal import and export tax will be zero. That is, non-revenue-raising import barriers are optimal as a second-best policy, only if the government does not have access to production taxes.

zero cross-substitutability between industries. In this special case, Theorem 1.B yields the following formulas:

$$
\begin{aligned}
1+s_k^* &= \frac{1+\bar{s}}{1+\psi_k}, \quad \forall k \\
1+x_k^* &= \left(1+\frac{1}{\epsilon_k \lambda_{ff,k}}\right)\left(\frac{1+\psi_k}{1+\psi_k r_{ff,k}}\right)(1+\bar{t})^{-1}, \quad \forall k \\
1+t_k^* &= (1+\bar{t})\left(\frac{1+\psi_k r_{ff,k}}{1+\psi_k}\right), \quad \forall k.
\end{aligned}
$$

The above special-case formula is itself a strict generalization of the formula derived by Bartelme et al. (2018) for a small open economy (for which $r_{ff,k} \approx \lambda_{ff,k} \approx 1$). The characterization of t_k^* also bears resemblance to the classic optimal tariff formula, as it essentially indicates that the optimal tariff is related to the inverse of Foreign's *backward-falling* supply curve.

As Home's size relative to the rest of the world diminishes (i.e., $r_{ff,k} \rightarrow 1$), the above formulas converge to those derived under restricted entry. So, even though the free and restricted entry cases exhibit key differences, these differences are quantitatively important *only* if Home is sufficiently large relative to the rest of the world.[15] Otherwise, the two cases are effectively identical insofar as optimal tax schedules are concerned.

3.2 Second-Best Optimal Export and Import Taxes

Suppose the Home government cannot raise domestic taxes due to institutional or political constraints. In that case, it is optimal to use trade taxes as a second-best instrument to tackle allocative inefficiency. In this section, we derive sufficient statistics formulas for second-best optimal trade taxes in such instances.

We have already highlighted the role of cross-demand effects as well as the subtle distinctions that separate the restricted and free entry cases. To streamline the presentation, though, we hereafter present results corresponding to the case of restricted entry and zero cross-industry demand effects. Formulas corresponding to free entry or arbitrary cross-demand effects are nonetheless presented in the appendix.

When domestic taxes are unavailable, the Home government's optimal policy problem includes an added constraint that $\boldsymbol{s}=\mathbf{0}$:

$$
\max_{(\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w}|\boldsymbol{s}=\mathbf{0})\in\mathbb{F}} \quad W_h\left(\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w}\right) \qquad \text{(P2)}.
$$

By analytically solving Problem (P2), we can derive sufficient statistics formulas for second-best optimal trade taxes. The following proposition presents these formulas,

[15] Another difference between the free and restricted entry cases is that the first-best tax schedule can be replicated with a combination of trade taxes and *entry subsidies*.

with a formal proof provided in Appendix G.[16]

Theorem 2. *Suppose the use of both domestic taxes is restricted, the second-best optimal trade taxes are unique up to a uniform tax shifters $\bar{t} \in \mathbb{R}_+$, and are given by:*

$$1 + t_k^* = \frac{(1+\mu_k)\left[1+\epsilon_k \lambda_{hh,k}\right]}{1+\mu_k+\epsilon_k \lambda_{hh,k}}(1+\bar{t}), \quad \forall k$$
$$1 + x_k^* = \frac{1+1/\epsilon_k \lambda_{hh,k}}{1+\mu_k}(1+\bar{t})^{-1}. \quad \forall k \tag{10}$$

The above theorem indicates that, when the use of domestic taxes is restricted, (i) the optimal export tax equals the wedge between the firm-level and country-level market powers in industry k, and (ii) the optimal import tax is non-uniform and positively related to the industry-level profit margin. However, due to a lack of sufficient policy instruments, the above tax schedule cannot perfectly reproduce the first-best outcome.[17]

Intuitively, the government's objective when solving (P2) is to mimic domestic Pigouvian subsidies with trade taxes. It is, thus, optimal to raise import taxes and lower export taxes in high-profit industries relative to the first-best benchmark. Doing so indirectly promotes more production in high-μ industries.

This adjustment, however, comes with a lesser-known side-effect: if ϵ_k and $\xi_k = 1/\mu_k$ are positively correlated, such an adjustment in trade taxes worsens Home's ToT. The intuition being that correcting allocative inefficiency requires subsidizing the high-ϵ industries, whereas the correcting the ToT inefficiency requires the opposite.

There is, thus, an inherent tension between the ToT-enhancing and efficiency-enhancing motives for trade taxation when $\text{cov}(1/\mu_g, \epsilon_g) > 0$. This tension can diminish the effectiveness of trade taxes as a second-best corrective instrument, beyond what is implied by the *targeting principle* (Bhagwati and Ramaswami (1963)).[18] we summaries these arguments in the following remark.

Remark 1. *If the industry-level trade and scale elasticities are negatively correlated, then correcting allocative inefficiency with "trade taxes" worsens the ToT inefficiency. This tension diminishes the effectiveness of trade taxes as a second-best corrective policy measure, beyond what is already implied by the targeting principle.*

3.3 Second-Best Optimal Import Taxes

Now suppose that in addition to domestic taxes, the use of export taxes is also restricted. The Home government's optimal policy problem in this case features two

[16] See Appendix E for a more general version of the formulas under arbitrary cross-demand effects.

[17] As noted earlier, the government needs at least three tax instruments (per industry) to achieve the first-best outcome. In *homogeneous* good models, though, trade taxes can fully mimic domestic taxes, because two instruments suffice to achieve the first-best.

[18] On the flip side, when trade taxes are restricted, governments will use domestic taxes as a second-best instrument to improve their ToT. Moving from the *first-best* to the *second-best* production tax-cum-subsidy schedule, in that case, improves the ToT inefficiency at the expense of worsening allocative efficiency.

additional constraints, $\boldsymbol{x} = \boldsymbol{s} = \mathbf{0}$:

$$\max_{(\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w}|\boldsymbol{s},\boldsymbol{x}=\mathbf{0})\in\mathbb{F}} \quad W_h\left(\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w}\right) \quad (\text{P3}).$$

Some variation of the above problem, which concerns second-best optimal *import* taxes, has been the focus of expansive literature on trade policy. As before, we can solve Problem (P3) analytically to produce sufficient statistics formulas for the optimal second-best import taxes. The following proposition presents these formulas, with a formal proof provided in Appendix H.

Theorem 3. *Suppose the use of both domestic and export taxes is restricted, the second-best optimal import taxes are uniquely given by:*

$$1 + t_k^* = \frac{(1+\mu_k)\left[1+\epsilon_k\lambda_{hh,k}\right]}{1+\mu_k+\epsilon_k\lambda_{hh,k}}\left(\frac{1+\sum_g\left[\frac{1}{1+\mu_g}\frac{\acute{\lambda}_{hf,g}}{\acute{\lambda}_{hf}}\epsilon_g\lambda_{ff,g}\right]}{\sum_g\frac{\acute{\lambda}_{hf,g}}{\acute{\lambda}_{hf}}\epsilon_g\lambda_{ff,g}}\right), \quad \forall k$$

where $\acute{\lambda}_{hf,g} \equiv \lambda_{hf,g}e_{f,g}$ *and* $\acute{\lambda}_{hf} = \sum_g \acute{\lambda}_{hf,g}$.

The above optimal tariff formula is composed of and industry-specific component and a uniform component in parenthesis. The industry-specific component bears the same intuition as that provided under Theorem 2.

The uniform component accounts for the ability of import taxes to manipulate the ToT through general equilibrium wage effects. These effects where previously redundant due to the availability of export taxes. The intuition being that, previously, the uniform tariff component could be perfectly mimicked with an across-the-board shift in export taxes (see Lemma 1). Since export taxes are now restricted, the uniform tariff component becomes consequential to ToT manipulation.

3.4 Decomposing the Gains from Policy

One of our main objectives in this paper is to quantify the gains from trade taxation. As noted before, these gains come from (i) correcting the ToT inefficiency as a (non-cooperative) first-best instrument, and/or (ii) correcting allocative inefficiency as a second-best instrument. Below, we briefly discuss how large these gains are as a function of the model's underlying parameters.

Let $W_{h,0} \equiv W_h(\mathbf{0},\mathbf{0},\mathbf{0})$ denote Home's under status quo. Also, let $W_h(\boldsymbol{t}^*,\boldsymbol{x}^*,\boldsymbol{s}^*)$ denote Home's welfare under the optimal non-cooperative policy characterized by Theorems 1.A or 1.B; and let $W_h(\boldsymbol{t}^*,\boldsymbol{x}^*,\mathbf{0})$ denote Home's welfare under first-best non-cooperative trade taxes but in the absence of domestic taxes, i.e., $\boldsymbol{s} = \mathbf{0}$. The gains from the (first-best) optimal policy can be decomposed as follows:

$$\frac{W_h(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*)}{W_{h,0}} = \underbrace{\frac{W_h(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*)}{W_h(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{0})}}_{\text{corrective gains}} \times \underbrace{\frac{W_h(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{0})}{W_{h,0}}}_{\text{ToT gains}}.$$

The first component, labeled "corrective gains," accounts for the pure gains from eliminating allocative inefficiency. The second component, labeled "ToT gains," accounts for the pure gains from ToT (terms-of-trade) improvement. The above decomposition is theoretically grounded in the fact that $\boldsymbol{t}^*$ and $\boldsymbol{x}^*$ only target the ToT inefficiency, whereas $\boldsymbol{s}^*$ only targets allocative inefficiency—both from the perspective of the Home country.

Noting the above decomposition, consider two extreme cases. First, suppose $\text{Var}_k(\mu_k) \approx 0$. In that case, Theorems 1.A and 1.B indicate that $\boldsymbol{s}^* = 0$, by choice of shifter $\bar{s}$. In other words, there are no "corrective gains" from policy in this case (i.e., $W_h(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*)/W_h(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{0}) \approx 1$). Second, consider the case where $\epsilon_k \to \infty$. In this other extreme, optimal export and import taxes approach zero (i.e., $\boldsymbol{t}^*$=$\boldsymbol{x}^*$=$\boldsymbol{0}$) by choice of shifter $\bar{t}$. As a result, there are no "ToT gains" from policy (i.e., $W_h(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{0})/W_{h,0} \approx 1$). Comparing and combining these two extreme cases, we can deduce the following remark.

Remark 2. *(a) the ToT gains from policy depend primarily on the mean of ϵ_k's, whereas (b) the corrective gains from policy depend primarily on the dispersion in μ_k or ψ_k's.*

Considering Remarks 1 and 2, questions such as "how large are gains from ToT manipulation?" or "how effective are trade taxes as a second-best instrument?" depend critically on our estimates for ϵ_k and $\mu_k = \psi_k = 1/\xi_k$. The next section develops a new methodology to estimate these key policy parameters. But before moving on to the estimation, let us briefly discuss the robustness of our theory to three basic extensions.

4 Extensions of Optimal Tax Formulas

In our baseline model, we abstracted from political economy considerations, input-output linkages, and multilateral considerations. Below, we introduce these previously-overlooked margins and discuss how they modify our baseline optimal tax formulas.

4.1 Accounting for Political Economy Motives

When governments are politically-motivated, their policy objectives may no longer align insofar as domestic policies are concerned. In particular, some governments may assign a greater weight to low-profit industries such as agriculture. These political economy considerations can prompt governments to subsidize and promote the wrong domestic industries from the perspective of the rest of the world. This concern is at the root of some existing opposition to industrial policy.

To account for such political economy motives, we follow Ossa's (2014) adaptation of Grossman and Helpman (1994). Specifically, we consider the Cobb-Douglas-CES case of our model where social welfare in country i is given by $W_i \equiv V_i(.) =$

$\frac{w_i L_i + \mathcal{R}_i}{\tilde{P}_i} + \sum_k \frac{\Pi_{i,k}}{\tilde{P}_i}$. Instead of maximizing W_i, we assume that the government in country i maximizes a politically-weighted welfare function, $W_i^P = \frac{w_i L_i + \mathcal{R}_i}{\tilde{P}_i} + \sum_k \theta_{i,k} \frac{\Pi_{i,k}}{\tilde{P}_i}$; where $\theta_{i,k}$ is the political economy weight assigned to industry k's profits (with $\sum_k \theta_{i,k}/K = 1$). It follows trivially from Theorem 1.A that the optimal (first-best) policy in this setup is given by

$$1 + s_k^* = \frac{1+\bar{s}}{1+\mu_{i,k}^P}, \qquad \forall k$$
$$1 + t_k^* = 1 + \bar{t}, \qquad \forall k$$
$$1 + x_k^* = (1+\bar{t})^{-1}\left(1 + \frac{1}{\epsilon_k \lambda_{ff,k}}\right), \quad \forall k$$

where $\mu_{i,k}^P = \frac{\theta_{i,k}\mu_k}{1+\mu_k}$ is the *political economy-adjusted* markup of industry k. Analogously, second-best optimal trade taxes are also characterized by a variation of Theorems 2 and 3, where μ_k is replaced with the *political economy-adjusted* markup, $\mu_{i,k}^P$. Considering the above formulas, if $\text{cov}(\theta_{i,k}, \mu_k) < 0$, then the gains from global cooperation will be strictly smaller than those predicted by our non-political baseline model. For this reason, we label cooperation by non-politicly-motivated governments as *deep* cooperation.

4.2 Accounting for Input-Output Linkages

Next, consider an extension of our baseline model where production employs both labor and intermediate inputs that are denoted by superscript $\mathcal{I}$. Supposing entry is restricted, the producer price of good ji,k depends on (i) the local wage rate and (ii) the price of all intermediate inputs available to suppliers in country j: $P_{ji,k} = (1+\mu_k)\mathcal{C}_{ji,k}(w_j, \tilde{\boldsymbol{P}}_j^{\mathcal{I}})$ where $\tilde{\boldsymbol{P}}_j^{\mathcal{I}} \equiv \{P_{ij,k}^{\mathcal{I}}\}$. Without loss of generality, assume that each good ji,k can be used as either an intermediate input or a final consumption good and that taxes are applied on a good irrespective of the intended final use, i.e., $\tilde{P}_{ij,k}^{\mathcal{I}} = \tilde{P}_{ij,k}$.[19]

By appealing to supply-side envelop conditions, we can derive the optimal policy under restricted entry as follows (see Appendix I):

$$1 + s_k^* = \frac{1+\bar{s}}{1+\mu_k}$$
$$1 + t_k^* = 1 + \bar{t}$$
$$1 + x_k^* = \left(1 + \frac{1}{\lambda_{ff,k}\epsilon_k}[1 - \frac{\acute{\lambda}_{hf}}{\acute{\lambda}_{hf,k}}\alpha_{fh}^{hf,k}]\right)(1+\bar{t})^{-1}.$$

In the above formula $\alpha_{fh}^{hf,k} \equiv \sum_g \left(\tilde{P}_{hf,g}^{\mathcal{I}} Q_{hf,k}^{fh,g}\right) / \sum_g \left(P_{fh,g} Q_{fh,g}\right)$ denotes the total share of

[19] This assumption is just a matter of labeling the goods. We can always break up every industry k into a final good version k' and an intermediate good version k''. Since we do not impose any restrictions on the number of industries, all the results carry over to this economy, which allows for differential taxes on the *final* versus *intermediate-input* version of industry k goods.

input hf,k in Foreign's exports to Home ($Q_{hf,k}^{fh,g}$ denotes the amount of input hf,k used in the production of output fh,g). The first-best domestic and import tax formulas are,thus, unaffected by the presence input-output linkages. The optimal export tax includes an extra term that accounts for the fact that export taxes are partially passed on to Home consumers through the input-output structure. The gains from policy will also multiply through the input-output structure.[20]

That the optimal import tariff is uniform even with input-output linkages is an artifact of governments having access to export taxes. If the use of export taxes is restricted, optimal import tariffs will vary across industries. This variation is due to the ability of import taxes to manipulate Home's export market power *indirectly* through input-output linkages.

4.3 Many Countries and Strategic Considerations

With arbitrarily many countries, we can follow the same steps to derive sufficient statistics formulas for optimal taxes up-to a first-order welfare approximation. Doing so, as shown in Appendix J, yields the following optimal tax formulas under restricted entry:

$$\begin{aligned} 1+s_{i,k}^{*} &= \frac{1+\bar{s}}{1+\mu_k}, \quad \forall k \\ 1+t_{ji,k}^{*} &= 1+\bar{t}, \quad \forall j,k \\ 1+x_{ij,k}^{*} &= \left(1+\frac{1}{\epsilon_k(1-\lambda_{ij,k})}\right)(1+\bar{t})^{-1}, \quad \forall j,k. \end{aligned}$$

The above formulas differ from the baseline formula in that the export taxes now depends on the destination-specific import demand elasticity. Analogously, the optimal policy under free entry is the same for a sufficiently small economy expect that $1+s_{i,k}^{*} = (1+\bar{s})/(1+\psi_k)$ for all k.

Theorems 1-3 also apply to cases both Home and Foreign apply taxes, provided that income effects are absent. To elaborate, suppose the cross-industry utility aggregator is quasi-linear, with ℓ denoting the *non-traded* quasi-linear sector: $U_i(\boldsymbol{Q}_i) = Q_{i,\ell} + V(\boldsymbol{Q}_{i,-\ell})$.[21] In that case, $\partial \ln \mathcal{D}_{i,k}\left(\tilde{\boldsymbol{P}}_i, Y_i\right)/\partial \ln Y_i = 0$ for every traded industry k. It is then straightforward to show that the optimal tax formulas presented under Theorems 1-3 remain valid even when Foreign applies taxes. The only qualification being that the expenditure shares (λ), revenue shares (r), and reduced-form demand elasticities (ε) now also depend on Foreign's applied taxes.[22]

[20]See Beshkar and Lashkaripour (2019) for a discussion on the magnification of the gain from trade taxation in the presence of input-output linkages. See Liu (2018) for discussion on the magnification of the gain from corrective subsidies in the presence of input-output linkages.

[21]Being non-traded simply means that $\tau_{ji,\ell}$ or ϵ_ℓ are prohibitively large. Otherwise, industry ℓ can be modeled like any other industry in the economy.

[22]With income effects, the optimal tax formulas characterized by Theorems 1 and 2 require slight

5 Estimating the Key Policy Parameters

To compute the gains from policy, we need credible estimates for (i) ϵ_k, which is generally labeled the *trade elasticity* and reflects the degree of national-level market power, and (ii) ξ_k that reflects the degree of firm-level market power and pins down the industry-level scale elasticity under free entry ($\psi_k = 1/\xi_k$) or the markup margin under restricted entry ($\mu_k = 1/\xi_k$). The trade literature has paid considerable attention to the estimating ϵ_k, but less attention has been devoted to estimating ξ_k and its byproducts ψ_k and μ_k. So, before delving into the actual estimation, let us provide a brief background on how the existing literature typically handles parameters ψ_k or μ_k.

Background. The trade policy literature typically handles parameters μ_k and ψ_k by adopting one of the following two normalizations: (i) $\mu_k = \psi_k = 1/\epsilon_k$ in monopolistically competitive models like Ossa (2016), and (ii) $\mu_k = \psi_k = 0$ in perfectly competitive models like Caliendo and Parro (2015). These normalizations allow researchers to pin down ψ_k and μ_k based on existing estimates for the trade elasticity but also create an arbitrary link between the market power of individual firms and the national-level market power (see Benassy (1996)).

Outside of the trade policy literature, several studies, including De Loecker and Warzynski (2012), Edmond et al. (2015), and De Loecker, Goldberg, Khandelwal, and Pavcnik (2016), have used firm-level production data to estimate the degree of firm-level market power within industries. These production-side approaches, however, pin down μ_k (or equivalently ψ_k) in isolation; requiring the trade elasticity to be estimated with separate data.

To draw credible comparisons between the *ToT gains* and *corrective gains* from policy, we (a) need to separately identify ϵ_k from μ_k or ψ_k, and (b) we want our estimates for ϵ_k and $\psi_k = \mu_k = 1/\xi_k$ to be mutually-consistent. We, thus, propose a new demand-based methodology that simultaneously estimates ϵ_k and ξ_k, thereby satisfying requirements (a) and (b).[23] Our approach involves fitting the structural firm-level import demand function implied by A1 to the universe of Colombian import transactions from 2007–2013. We outline this approach below, starting with a description of the data used in our estimation.

Data Description. Our estimation uses import transactions data from the Colombian Customs Office for the 2007–2013 period.[24] The data include detailed informa-

amendments if Foreign imposes retaliatory taxes. Specifically, if Foreign applies taxes and collects tax revenue, Home's taxes can alter Foreign's income through their effect on Foreign's tax revenues. Taking these effects into consideration will modify Home's optimal policy, though slightly.

[23]In the presence of selection effects, our estimated parameters are necessary but not sufficient to pin down the trade and scale elasticities. In addition to these parameters, we also need information on the shape of the Pareto productivity distribution—see Appendix D.

[24]The data is obtained from Datamyne, a company that specializes in documenting import and export transactions in Americas. For more detail, please see www.datamyne.com.

tion about each transaction, such as the Harmonized System 10-digit product category (HS10), importing and exporting firms,[25] f.o.b. (free on board) and c.i.f. (customs, insurance, and freight) values of shipments in US dollars, quantity, unit of measurement (of quantity), freight in US dollars, insurance in US dollars, value-added tax in US dollars, country of origin, and weight. A unique feature of this data set is that it reports the identities of all foreign firms exporting to Colombia, allowing us to define import varieties as firm-product combinations—in comparison, most papers focusing on international exports to a given location typically treat varieties as more aggregate country-product combinations. Table 6 (in the appendix) reports a summary of basic trade statistics in our data.[26]

When working with the above data set, we face the challenge that, for some products, Colombia has been changing the HS10 classification between 2007 and 2013. Fortunately, the Colombian Statistical Agency, DANE, has kept track of these changes,[27] and we utilized this information to concord the Colombian HS10 codes over time. In the process, we followed the guidelines outlined by Pierce and Schott (2012) for the concordance of the U.S. HS10 codes over time.[28] Overall, changes in HS10 codes between 2007 and 2013 affect a very small portion (less than 0.1%) of our dataset.

5.1 Estimation Strategy

Since we are focusing on one importer, we hereafter drop the importer's subscript i and add a year subscript t to account for the time dimension of our data. With this slight change of notation, the demand facing firm ω located in country j and supplying product k in year t is given by (see A1):

$$q_{j,kt}(\omega) = \left(\frac{\tilde{p}_{j,kt}(\omega)}{\tilde{P}_{j,kt}}\right)^{-1-\xi_k} \left(\frac{\tilde{P}_{j,kt}}{\tilde{P}_{kt}}\right)^{-1-\epsilon_k} Q_{kt}, \tag{11}$$

Note that subscript k has been used thus far to reference industries. But in our empirical analysis, k will designate the most disaggregated industry/product category in our dataset, which is an HS10 product category. The quadruplet "ωjkt" therefore denotes a unique imported variety corresponding to *firm ω–country of origin j–HS10 product k–year t*.

[25] The identification of the Colombian importing firms is standardized by the national tax ID number. For the foreign exporting firms, the data provide the name of the firm, phone number, and address. The names of the firms are not standardized, and thus there are instances in which the name of a firm and its address are recorded differently (e.g., using abbreviations, capital and lower-case letters, dashes, etc.). We deal with this problem by standardizing the spelling and the length of the names along with utilizing the data on firms' phone numbers. The detailed description of cleaning the exporters' names is provided in the Technical Appendix.

[26] Our estimation also employs data on monthly average exchange rates, which are taken from the Bank of Canada: *http://www.bankofcanada.ca/rates/exchange/monthly-average-lookup/*.

[27] We thank Nicolas de Roux and Santiago Tabares for providing us with this information.

[28] To preserve the industry identifier of the product codes, and in contrast to Pierce and Schott (2012), we try to minimize the number of the synthetic codes. The concordance data and do files are provided in the data appendix.

Rearranging Equation 11, we can produce the following equation

$$q_{j,kt}(\omega) = \tilde{p}_{j,kt}(\omega)^{-\epsilon_k - 1} \left(\frac{\tilde{p}_{j,kt}(\omega)}{\tilde{P}_{j,kt}}\right)^{\epsilon_k - \xi_k} \tilde{P}_{kt}^{\epsilon_k+1} Q_{kt} = \tilde{p}_{j,kt}(\omega)^{-\epsilon_k - 1} \lambda_{j,kt}(\omega)^{1-\alpha_k} \tilde{P}_{kt}^{\epsilon_k+1} Q_{kt}, \tag{12}$$

where α_k denotes the wedge between the degree of national-level and firm-level market powers,

$$\alpha_k \equiv \epsilon_k / \xi_k;$$

while $\lambda_{j,kt}(\omega)$ denotes the share of expenditure on firm ω *"conditional"* on buying from country j,

$$\lambda_{j,kt}(\omega) \equiv \left(\frac{\tilde{p}_{j,kt}(\omega)}{\tilde{P}_{j,kt}}\right)^{-\xi_k} = \frac{\tilde{p}_{j,kt}(\omega) q_{j,kt}(\omega)}{\sum_{\omega' \in \Omega_{j,kt}} \tilde{p}_{j,kt}(\omega') q_{j,kt}(\omega')}.$$

Taking logs from Equation 12, writing the demand function in stochastic form, and letting $\chi(\omega) \equiv \tilde{p}(\omega) q(\omega)$ denote sales, yields the following firm-level import demand function:

$$\ln \chi_{j,kt}(\omega) = -\epsilon_k \ln \tilde{p}_{j,kt}(\omega) + (1 - \alpha_k) \ln \lambda_{j,kt}(\omega) + \delta_{kt} + v_{\omega jkt}, \tag{13}$$

where $\delta_{kt} \equiv \ln P_{kt}^{-\epsilon_k - 1} Q_{kt}$ is a product-year fixed effect, while $v_{\omega jkt}$ encapsulates all non-price-related demand shifters. We can decompose $v_{\omega jkt} = \delta_{\omega k} + \nu_{\omega jkt}$ into a (i) systematic term, $\delta_{\omega k}$, which is a firm-product fixed effect that reflects firm-level quality among other things, and (ii) idiosyncratic terms $\nu_{\omega jkt}$, which reflects idiosyncratic variations in quality, measurement errors, or non-technological demand shifters specific to variety ωjkt.

In the estimating Equation 13, price and sales data ($\tilde{p}_{j,kt}(\omega)$ and $\chi_{j,kt}(\omega)$) are directly observable per import variety and the within-national market share, $\lambda_{j,kt}(\omega)$, can be calculated using the universe of firm-level sales to Colombia. Some readers may notice parallels between our estimating equation and the nested demand function analyzed in Berry (1994); but given the structure of our data, we adopt a distinct identification strategy.

Before presenting our estimation strategy, a discussion of what identifies $\alpha_k \equiv \epsilon_k / \xi_k$ is in order. Parameter α_k reflects the wedge between the national-level and firm-level degrees of market power. Broadly speaking, variety ωjkt is (i) either imported from a thick market like China in which case it competes with many other Chinese varieties, hence a low $\lambda_{j,kt}(\omega)$, or (ii) it is imported from a thin market like Taiwan where it competes with a few other Taiwanese varieties, hence a high $\lambda_{j,kt}(\omega)$. Controlling for prices, if varieties originating from thick markets generate lower sales, our import demand function identifies this as a case where $1 > \alpha_k > 0$. That is, $1 > \alpha_k > 0$ corresponds to a situation where firm-level market power is less than the national-level market power due to sub-national competition.[29]

As a first step in our analysis, we estimate Equation 13 under a pooled specification where α_k and ϵ_k are assumed to be uniform across product categories. To this end, we employ a first-difference estimator that eliminates the firm-product fixed effect, $\delta_{\omega k}$, and also observations pertaining to one-time exporters. We deem the first-difference approach appropriate given the possibility that $\nu_{\omega jkt}$'s are sequentially correlated. Stated in terms of first-differences, our estimating equation takes the following form

$$\Delta \ln \chi_{j,kt}(\omega) = -\epsilon \Delta \ln \tilde{p}_{j,kt}(\omega) + (1-\alpha)\, \Delta \ln \lambda_{j,kt}(\omega) + \Delta\delta_{kt} + \zeta_{\omega jkt}, \qquad (14)$$

where $\zeta_{\omega jkt} \equiv \Delta \nu_{\omega jkt}$ represents a variety-specific demand shock, while $\Delta\delta_{kt}$ can again be treated as a product-year fixed effect.[30] The identification challenge here is that price and within-national market share are endogenous and may co-move with the demand shock, $\zeta_{\omega jkt}$.[31]

While the first-difference transformation is a partial remedy to the aforementioned identification problem, we also employ an instrumental variable strategy to recover the price and within-national market share coefficients. To this end, we construct a variety-specific cost shifter that is arguably uncorrelated with the variety-specific demand shock, $\zeta_{\omega jkt}$. Our identification strategy capitalizes on the monthly frequency of import transactions in the data. Specifically, we compile an external database on aggregate monthly exchange rates and interact the monthly variation in aggregate exchange rates with the (lagged) monthly composition of firm-level exports to construct a *shift-share* instrument, $z_{j,kt}(\omega)$, as follows:

$$z_{j,kt}(\omega) = \sum_{m=1}^{12} \frac{\chi_{j,kt-1}(\omega; m)}{\chi_{j,kt-1}(\omega)} \Delta E_{jt}(m).$$

In the above expression, $\Delta E_{jt}(m)$ denotes the change in country j's exchange rate with Colombia in month m of year t; and $x_{j,kt-1}(\omega; m)$ represents month m sales of firm ω (from country j, within product category k) in the prior year, $t-1$, with $\chi_{j,kt-1}(\omega) = \sum_{m=1}^{12} \chi_{j,kt-1}(\omega; m)$. So, for variety ωjkt, the term $z_{j,kt}(\omega)$ measures the variety-specific *exposure to exchange rate shocks* in year t. Put differently, depending on the monthly

[29] Note that the case $\alpha_k > 1$ undermines assumption A3 regarding the uniqueness of the equilibrium. As we will see shortly, our estimation always yields a value of $0 < \alpha_k < 1$.

[30] In the previous version of the paper, we included all observations that reported a non-missing $\Delta \ln \tilde{p}_{j,kt}$. Some of our industry-level estimates, however, display sensitivity to outlier observations. Considering this, we now trim the sample to exclude observations that report a price change, $\Delta \ln \tilde{p}_{j,kt}$, above the 99th percentile or below the 1st percentile of the industry.

[31] Note that the within-national market share, $\lambda_{\omega|j,ht}$, could be correlated with $\varepsilon_{\omega,ht}$ due to measurement errors in export sales. Our identification strategy takes care of this alternative source of endogeneity as long as our instruments are not correlated with variety-specific measurement errors. Similarly, unit prices could be contaminated with measurement errors, as they are averaged across transactions and consumers. This type of measurement error, however, is fairly innocuous given the log-linear structure of our demand function (see Berry (1994)),

composition of sales to Colombia, aggregate exchange rate movements have differential effects on individual firms. These differential exposures to exchange rate shocks are picked up by $z_{j,kt}(\omega)$—compare this with the widely-used *Bartik* instrument, which asserts that different regions are affected differentially by national labor market shocks depending on the industry composition of that region's production.

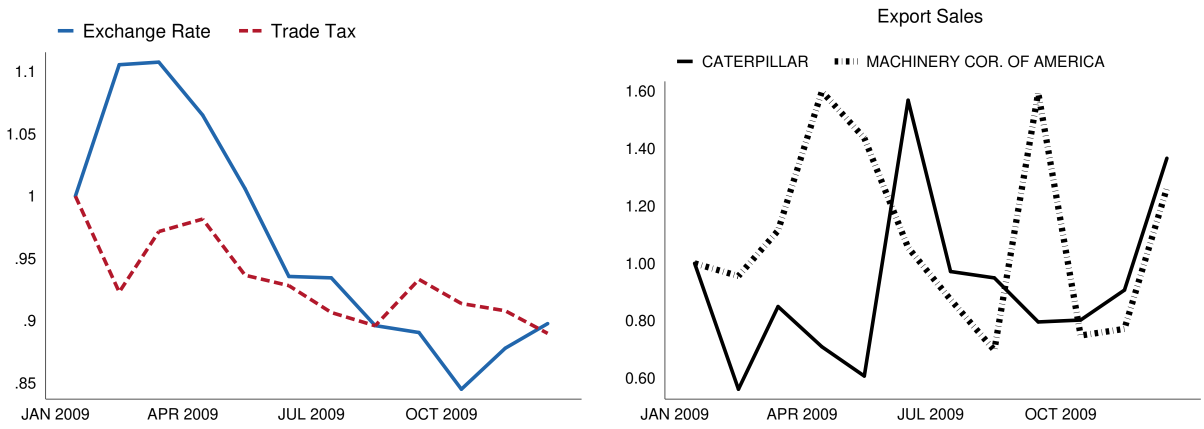
Figure 1 illustrates the workings of our instrument in further detail. It concerns U.S. exports in product category HS8431490000 (PARTS AND ATTACHMENTS OTHER FOR DERRIKS ETC.)—a product category that features one of the most frequently imported varieties: *machine parts* from "CATERPILLAR." The left panel of Figure 1 displays how both the exchange rate and the *average* import tax rate paid by U.S. based firms varied considerably on a monthly basis in 2009. The right panel plots the monthly variation in the export sales of the two largest U.S. based firms within category HS8431490000 (namely, "CATERPILLAR" and "MACHINERY CORP. OF AMERICA"). Given that the monthly composition of exports from "CATERPILLAR" and "MACHINERY CORP. OF AMERICA" are markedly different, the two firms are affected differently by aggregate movements in the monthly exchange rate.

In addition to our measure of exchange rate shock exposure, we adopt a set of standard instruments borrowed from the existing literature. In particular, we instrument for the c.i.f. price with variety-specific import tax rates (which include tariff and the Columbian value-added tax).[32] Several studies, most notably Caliendo and Parro (2015), treat import taxes as an exogenous cost shifter to identify the trade elasticity. Following Khandelwal (2010), we construct two additional instruments for the annual variation in the within-national market shares, $\Delta \ln \lambda_{j,kt}(\omega)$: (i) annual changes in the total number of country j firms serving the Colombian market in product category k, and (ii) changes in the total number of HS10 product categories actively served by firm ω in year t. These count measures will be correlated with $\Delta \ln \lambda_{j,kt}(\omega)$ but uncorrelated with $\zeta_{\omega jkt}$ if variety-level entry and exit occur prior to, or independent of, the demand shock realization of competing varieties. This assumption is rather standard in the literature that estimates discrete choice demands curves (see Berry, Levinsohn, and Pakes (1995) or Khandelwal (2010)).

In light of our theory, we are interested in estimating ϵ and ξ for individual industries that are comprised of multiple HS10 product categories. To distinguish between industries and HS10 product categories, we *temporarily* use κ to index industries and k to index HS10 product categories. To identify the industry-level elasticities, we simply pool all HS10 products in industry κ and assume that ϵ_k and ξ_k are uniform with that industry. Namely, $\alpha_k = \alpha_\kappa$ and $\epsilon_k = \epsilon_\kappa$ for all $k \in \mathbb{K}_\kappa$. Then, we estimate the following

[32] We do not include the value-added tax in the 'Transportation' and 'Petroleum' industries since the value-added tax in these industries discriminates by the method of delivery and level of luxury, both of which may be correlated with the demand shocks.

Figure 1: Monthly variations in national exchange rate, firm-specific exports, and import tax within a selected product-country-year category.



Notes: The left panel plots monthly variations in exchange rate and (value-weighted) average import tax for US-based firms within product category HS8431490000-year 2009. The right panel plots monthly movements in export sales for the two biggest US firms in product HS8431490000 in year 2009— namely, Caterpillar and Machinery Corp. of America.

import demand function separately for each industry κ:

$$\Delta \ln \chi_{j,kt}(\omega) = -\epsilon_\kappa \Delta \ln \tilde{p}_{j,kt}(\omega) + (1-\alpha_\kappa)\,\Delta \ln \lambda_{j,kt}(\omega) + \Delta \ln Q_{kt} + \tilde{\varepsilon}_{\omega jkt}\,, \quad k \in \mathbb{K}_\kappa \tag{15}$$

For the purpose of our quantitative analysis we estimate the above import demand function separately for 14 tradable industries based on the industry classification used in the World Input-Output Database.[33] In principle, we can also estimate the import demand function separately for each HS10 product category. Such disaggregated elasticities, however, will be of little use for the quantitative analysis of policy, given that multi-country data on trade, production, and expenditure shares are scarce at finer levels of disaggregation.

Estimation Results. To attain a basic overview of our results, we first present the results corresponding to the pooled estimation in Table 2. Here, we break down the sample into manufacturing and non-manufacturing goods and compare the 2SLS and OLS estimates. Across manufacturing industries, the average trade elasticity is 3.6 while the average scale elasticity (or profit margin) is $\psi = \mu = \alpha/\epsilon \approx 0.19$. Comparing the OLS and 2SLS estimates indicates that our IV strategy is working in the expected direction. Moreover, our estimation yields a heteroskedasticity-robust Kleibergen-Paap Wald rk F-statistic of 95. Hence, we can reject the null of weak instruments given the Stock-Yogo

[33]Sector groupings are based on the ISIC Rev. 3 classification to match the WIOD sectors—see `https://wits.worldbank.org/product_concordance.html` for concordance between HS10 codes and two-digit ISIC sectors.

Table 2: Pooled estimation results

	Manufacturing		Non-Manufacturing	
Variable (log)	2SLS	OLS	2SLS	OLS
Price, $-\epsilon$	-3.588***	0.202***	-4.658***	0.104***
	(0.221)	(0.002)	(0.504)	(0.004)
Within-national share, $1-\alpha$	0.333***	0.814***	0.190***	0.802***
	(0.012)	(0.001)	(0.028)	(0.003)
Weak Identification Test	94.84	...	24.16	...
Under-Identification P-value	0.00	...	0.00	...
Within-R^2	...	0.77	...	0.73
N of Product-Year Groups	23,683		10,461	
Observations	1,126,976		205,580	

Notes: *** denotes significant at the 1% level. The Estimating Equation is (14). Standard errors in brackets are robust to clustering within product-year. The estimation is conducted with HS10 product-year fixed effects. The reported R^2 in the OLS specifications correspond to within-group goodness of fit. Weak identification test statistics is the F statistics from the Kleibergen-Paap Wald test for weak identification of all instrumented variables. The p-value of the under-identification test of instrumented variables is based on the Kleibergen-Paap LM test. The test for over-identification is not reported due to the pitfalls of the standard over-identification Sargan-Hansen J test in the multi-dimensional large datasets pointed by Angrist, Imbens, and Rubin (1996).

critical values.

Given that $\alpha \approx 0.67$, our estimation decisively rejects the arbitrary link often assumed between the firm-level and national-level degrees of market power in the literature.[34] To elaborate, our estimated α lies in-between two standard but polar specifications: (i) perfectly competitive models that assume firms have zero market power (i.e., $\alpha = 0$; Eaton and Kortum (2002); Anderson and Van Wincoop (2003); Alvarez and Lucas (2007); Caliendo and Parro (2015)), and (ii) imperfectly competitive models that assume similar degrees of firm-level and country-level market power (i.e., $\alpha = 1$; Krugman (1980); Melitz (2003); Balistreri, Hillberry, and Rutherford (2011); Ossa (2014)).

The industry-level elasticity estimates based on Equation (15) are reported in Table 3. The estimated elasticities here display a considerable amount of variation across industries. The estimated scale elasticity (or profit margin), $\mu = \psi = \alpha/\epsilon$, is highest in the 'Electrical & Optical Equipment' ($\mu = \psi = 0.45$) and 'Petroleum' ($\mu = \psi = 1.7$) industries; both of which are associated with high R&D or fixed costs. Our estimated scale elasticity is lowest in 'Agricultural & Mining' ($\psi = 0.14$) and 'Machinery' ($\psi = 0.10$) industries. Furthermore, with the exception of 'Agriculture & Mining,' we cannot

[34]In other words, our estimation rejects the *independence of irrelevant alternatives* (IIA) because product varieties or technologies are less differentiated intra-nationally than inter-nationally. While the IIA assumption has garnered considerable attention in the industrial organization literature, the trade literature has only recently tested this assumption against data. Redding and Weinstein (2016) estimate an international demand system that relaxes the IIA assumption by accommodating heterogeneous taste across consumers. Adao, Costinot, and Donaldson (2017) estimate a trade model that (unlike standard CES models) permits varieties from certain countries to be closer substitutes. Our results contribute to this emerging literature by highlighting another aspect of the trade data that is at odds with the IIA assumption.

reject the prevalence of scale economies.[35]

Our industry-level trade elasticity estimates, ϵ_k, also display some novel properties. To our knowledge, our estimation is the first to identify the industry-level trade elasticities using (i) firm-level data, and (ii) without externally setting α to either 0 or 1. Qualitatively speaking, our estimated trade elasticities are similar in magnitude to those estimated by Simonovska and Waugh (2014) but slightly lower than those estimated by Caliendo and Parro (2015). Aside from the firm-level nature of our estimation, these differences may be driven by the fact that instead of controlling for f.o.b. prices with exporter fixed effects, we directly use data on f.o.b. price levels.

5.2 Discussion

Plausibility of Estimates. While our theoretical model imposed strong structure on the supply-side of the economy, the credibility of our estimated $\psi_k = \mu_k = 1/\xi_k$ is not necessarily contingent on these assumptions. As long as assumption A1 is satisfied, our estimates of ψ_k or μ_k are compatible with a non-parametric production function that admits multiple factors of production.[36] To elaborate, consider the standard definition of the scale elasticity, which is elasticity of firm-level output, $q(\omega)$, with respect to total input costs, $c(\omega)$ (Morrison (1992)). There is a well-known result that if (i) firms are profit-maximizing and monopolistically competitive, and (ii) there is free entry into industry k, i.e., $p(\omega) = AC(\omega)$,[37] then the scale elasticity equals the markup margin, which is itself equal to the inverse of ξ_k. Namely,

$$1 + \psi_k \equiv \frac{\partial \ln q(\omega)}{\partial \ln c(\omega)} = \frac{AC(\omega)}{MC(\omega)} = 1 + \mu_k = 1 + \frac{1}{\xi_k}.$$

Considering this, our estimate for ξ_k pin down μ_k and ψ_k in environments that are considerably more general than our baseline theoretical model.[38]

Despite the generality highlighted above, one may still be concerned about the plausibility of our estimates. Take, for example, our elasticity estimates for the 'Petroleum' industry, which appear somewhat low. First, our estimate for ϵ_k aligns with the general consensus in the Energy Economics literature that, at the national level, demand for petroleum products is price-inelastic.[39] Second, our estimated ψ_k and μ_k for the 'Petroleum' industry align closely with existing estimates in the Industrial Organization literature (e.g., Considine (2001)). Finally, our finding that the 'Petroleum' industry

[35]Our finding that returns-to-scale are negligible in the agricultural sector is not unprecedented. Starting from Sen (1962), a whole body of literature has been devoted to documenting and analyzing the inverse farm-size productivity (IFSP) relationship.

[36]Our optimal policy formulas may, however, change under a richer production structure.

[37]$AC(\omega)$ and $MC(\omega)$ are used here to denote the firm-level average and marginal cost.

[38]Also, as noted earlier, in the presence of firm-selection effects, our estimates for ξ_k and ϵ_k along with the Pareto shape parameter would be sufficient to pin down ψ_k and μ_k (provided that the firm-productivity distribution is Pareto—see Appendix D).

[39]See Pesaran, Smith, and Akiyama (1998), which is the main reference in this literature, as well as Fattouh (2007) for a survey of that literature

Table 3: Industry-level estimation results

Sector	ISIC4 codes	Estimated Parameter ϵ_k	$1-\alpha_k$	ψ_k	Obs.	Weak Ident. Test
Agriculture & Mining	100-1499	6.212 (2.112)	0.125 (0.142)	0.141 (0.167)	11,962	2.51
Food	100-1499	3.333 (0.815)	0.117 (0.050)	0.265 (0.131)	20.042	6.00
Textiles, Leather & Footwear	100-1499	3.413 (0.276)	0.293 (0.020)	0.207 (0.022)	126,483	63.63
Wood	100-1499	3.329 (1.331)	0.101 (0.181)	0.270 (0.497)	5,962	1.76
Paper	100-1499	2.046 (0.960)	0.187 (0.216)	0.397 (0.215)	37,815	2.65
Petroleum	100-1499	0.397 (0.342)	0.302 (0.081)	1.758 (1.584)	4,035	2.03
Chemicals	100-1499	4.320 (0.376)	0.085 (0.027)	0.212 (0.069)	134,413	42.11
Rubber & Plastic	100-1499	3.599 (0.802)	0.418 (0.041)	0.162 (0.039)	107,713	7.22
Minerals	100-1499	4.561 (1.347)	0.153 (0.096)	0.186 (0.129)	28,197	3.19
Basic & Fabricated Metals	100-1499	2.959 (0.468)	0.441 (0.024)	0.189 (0.032)	155,032	16.35
Machinery	100-1499	8.682 (1.765)	0.130 (0.080)	0.100 (0.065)	266,628	8.54
Electrical & Optical Equipment	100-1499	1.392 (0.300)	0.369 (0.015)	0.453 (0.099)	260,207	17.98
Transport Equipment	100-1499	2.173 (0.589)	0.711 (0.028)	0.133 (0.036)	86,853	5.09
N.E.C. & Recycling	100-1499	6.704 (1.133)	0.049 (0.100)	0.142 (0.289)	70,974	8.51

Notes. Estimation results of Equation (15). Standard errors in parentheses. The estimation is conducted with HS10 product-year fixed effects. All standard errors are simultaneously clustered by product-year and by product-'country of origin.' The weak identification test statistics is the F statistics from the Kleibergen-Paap Wald test for weak identification of all instrumented variables. The test for over-identification is not reported due to the pitfalls of the standard over-identification Sargan-Hansen J test in the multi-dimensional large datasets pointed by Angrist et al. (1996).

is the most scale-intensive industry is consistent with the finding in Antweiler and Trefler (2002), which is based on more aggregated data.

Likewise, consider the case of 'Transportation' or auto industry, where we estimate that $\psi_k = \mu_k = 0.13$. Our estimate aligns well with existing estimates from various industry-level studies. Recently, Coşar, Grieco, Li, and Tintelnot (2018) have estimated markups for the auto industry that range between roughly 6% to 13%. Previously, Berry et al. (1995) had estimated markups of around 20% in the U.S. auto industry using data from 1971-1990 (page 882).

Another way to assess the plausibility of our estimates is from the lens of the cross-national income-size elasticity. As pointed out by Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), the factual relationship between real per capita income and population size (i.e., the income-size elasticity) is negative and statistically insignificant. Quantitative trade models featuring the normalization $\alpha = \epsilon/\xi = 1$, however, predict a strong and positive income-size elasticity that remains significant even after the introduction of domestic trade frictions. Ramondo et al. (2016) call this observation the income-size elasticity puzzle. Considering this puzzle, in Appendix L we compute the income-size elasticity implied by $\alpha \approx 0.67$, as a way to assess the plausibility of our estimates. Encouragingly, we find that our estimated value for α completely resolves the aforementioned puzzle. In other words, our micro-estimated elasticities are consistent with the macro-level cross-national relationship between population size and real per capita income.

5.2.1 Comparison to Bartelme et al. (2018)

Concurrent with us, Bartelme et al., (2018, BCDR) have developed an alternative methodology to estimate the scale elasticity. To be specific, they estimate the elasticity of export sales with respect to industry-level employment, namely α_k^{BCDR}. From the lens of gravity trade models, this elasticity assumes the following interpretation:

$$\alpha_k^{BCDR} = \left(\psi_k^{\mathcal{X}} + \psi_k\right)\epsilon_k - \beta_k,$$

where $\psi_k^{\mathcal{X}}$ denotes the industry-level agglomeration elasticity and β_k denotes the share of industry-specific factors in total production.

BCDR's estimation implicitly assumes that there are *no* are industry-specific factors of production, which amounts to $\beta_k = 0$. They then take estimates for the trade elasticity, ϵ_k, from other sources in the literature and calculate the scale elasticity as $\psi_k^{\mathcal{X}} + \psi_k = \alpha_k^{BCDR}/\epsilon_k$. The advantage of their approach is that it detects Marshallian (or agglomeration) externalities. The advantages of our approach are two-fold: (a) we separately identify ϵ_k from ψ_k;[40] and (b) our estimates are robust to the presence of

[40] As noted in Appendix D, in the presence of selection effect, our approach can identify the scale elasticity only up-to an externally chosen trade elasticity. Second, our approach can identify the scale elasticity even when there are diseconomies of scale at the industry-level.

industry-specific factors of production (i.e., our identification does not rely on $\beta_k = 0$).

These differences can also explain why we estimate a larger scale elasticity than BCDR in some industries. Consider, for instance, the 'Petroleum' industry where we estimate a considerably larger ψ_k than BCDR. This difference can be driven by the 'Petroleum' industry employing a large amount of industry-specific factors of production, like natural resources or offshore platforms. This situation corresponds to a high β_k that can attenuate the estimates in BCDR relative to ours.

6 Quantifying the Gains from Policy

As a final step, we quantify the gains from trade and industrial policy across 32 major economies. To this end, we combine our micro-level estimates for ϵ_k and $\mu_k = \psi_k = 1/\xi_k$, our sufficient statistic tax formulas from Section 3, and macro-level data production and expenditure from the 2012 World Input-Output Database (WIOD, Timmer, Erumban, Gouma, Los, Temurshoev, de Vries, Arto, Genty, Neuwahl, Francois, et al. (2012)). The original WIOD database covers 35 industries and 40 countries, which 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 plus 13 other major economies, namely, Australia, Brazil, Canada, China, India, Indonesia, Japan, Mexico, Russia, South Korea, Taiwan, Turkey, and the United States. The 35 industries in the WIOD database include the 15 industries listed in Table 3, plus 20 service-related industries. Our analysis contracts the original WIOD data into a sample of 32 major economies (listed in Table 4, first column) plus an aggregate of the rest of the world and 16 industries, including 15 tradable industries plus an aggregate of all service-related industries.[41] As a first step, we closely follow the methodology outlined in Costinot and Rodríguez-Clare (2014) to purge the data from trade imbalances.

Our goal is to calculate the optimal policy and the corresponding gains for each of the 32 economies in our sample. So, in each iteration of our analysis, we treat one of the 32 economies as "Home," and aggregate the rest of the world into one country and treat it as "Foreign." We then compute the optimal policy for the designated Home country relative to the rest of the world. To simplify this task, we appeal to our sufficient statistics tax formulas derived in Section 3.

6.1 Mapping the Sufficient Statistics Formulas to Data

The sufficient statistics formulas, presented under Theorems 1-3, simplify the task of computing the gains from policy. So, we start by discussing how these formulas can be mapped to data. Our quantitative analysis assumes that the cross-industry utility aggregator is Cobb-Douglas:

$$U_i\left(\boldsymbol{Q}_i\right) = \prod_k Q_{i,k}^{e_{i,k}}. \tag{16}$$

[41] Our shrinking of the sample along these dimensions is akin to Costinot and Rodríguez-Clare (2014). See their appendix for a discussion of why shrinking the WIOD sample in this regard is appropriate.

With the above assumption, the reduced-form demand elasticities are given by 2. We also assume that the status-quo is a tax-free equilibrium.

To map our formulas to data, we adopt the conventional exact hat-algebra notation ($\hat{x} \equiv x'/x$). We then combine the optimal tax formulas specified by Theorems 1-3 with a set of equilibrium conditions, expressing them in terms of the hat-algebra notation. This procedure yields a system of equations, the solution of which determines the optimal tax rates, $(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*)$, and their corresponding welfare effects.

The propositions that follow correspond to the restricted entry case, with analogous propositions for the free case entry presented in Appendix M. Our first proposition solves for the first-best non-cooperative tax schedule characterized by Theorem 1.A.

Proposition 1. *Suppose the observed data is generated by a model that exhibits the cross-industry utility aggregator 16. Then, the optimal (first-best) taxes and their effects on wages, profits, and income can be fully characterized by solving the following system of equations:*

$$\begin{cases} x_{hf,k} = \frac{1}{\epsilon_k \hat{\lambda}_{ff,k} \lambda_{ff,k}}; \quad t_{fh,k} = 0; \quad 1 + s_{h,k} = \frac{1}{1+\mu_k} & \forall k \in \mathbb{K} \\ x_{fh,k} = 0, \quad t_{hf,k} = 0; \quad s_k = 0 & \forall k \in \mathbb{K} \\ \hat{\lambda}_{ji,k} = \left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k} \hat{P}_{i,k}^{\epsilon_k} & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \hat{P}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k} \lambda_{ji,k} \right) & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \widehat{w}_i w_i L_i = \sum_k \sum_j \left(\frac{\hat{\lambda}_{ij,k} \lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})} \right) & \forall i \in \mathbb{C} \\ \widehat{\Pi}_i \Pi_i = \sum_k \sum_j \left(\frac{\mu_k \hat{\lambda}_{ij,k} \lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})} \right) & \forall i \in \mathbb{C} \\ \widehat{\mathcal{R}}_i \mathcal{R}_i = \sum_k \sum_j \left(\frac{t_{ji,k}}{1+t_{ji,k}} \hat{\lambda}_{ji,k} \lambda_{ji,k} e_{i,k} \hat{Y}_i Y_i + \frac{(1+x_{ij,k})(1+s_{i,k})-1}{(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})} \hat{\lambda}_{ij,k} \lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j \right) & \forall i \in \mathbb{C} \\ \hat{Y}_i Y_i = \widehat{w}_i w_i L_i + \widehat{\Pi}_i \Pi_i + \widehat{\mathcal{R}}_i \mathcal{R}_i & \forall i \in \mathbb{C} \end{cases}.$$

The above system involves $3K$ independent unknowns, $\{x_{hf,k}\}$, $\{\widehat{w}_i\}$, and $\{\hat{Y}_i\}$, and $3K$ independent equations. Moreover, it can be solved with knowledge of only (a) observable expenditure shares $\lambda_{ji,k}$, and $e_{i,k}$, wage income, $w_i L_i$, aggregate profits, Π_i, and total income $Y_i = w_i L_i + \Pi_i$, all of which are readily available from the WIOD. Solving the system also requires knowledge of (b) industry-level trade elasticities and markup margins, ϵ_k and $\mu_k = 1/\xi_k$, which were estimated in the previous section.

The next proposition outlines a procedure to solve for the second-best optimal trade taxes characterized by Theorem 2. Recall that these are optimal non-cooperative export and import taxes when the government is prohibited from using domestic taxes.

Proposition 2. *Suppose the observed data is generated by a model that exhibits the cross-industry utility aggregator 16. Then, the second-best optimal trade taxes and their effects on wages, profits, and income can be fully characterized by solving the following system of equa-*

tions:

$$\begin{cases} x_{hf,k} = \left(1 + \frac{1}{\epsilon_k \hat{\lambda}_{ff,k}\lambda_{ff,k}}\right)(1+\mu_k)^{-1}; \quad t_{fh,k} = \frac{(1+\mu_k)\left[1+\epsilon_k\hat{\lambda}_{hh,k}\lambda_{hh,k}\right]}{1+\mu_k+\epsilon_k\hat{\lambda}_{hh,k}\lambda_{hh,k}}; \quad s_{h,k}=0 & \forall k \in \mathbb{K} \\ x_{fh,k} = 0, \quad t_{hf,k} = 0; \quad s_k = 0 & \forall k \in \mathbb{K} \\ \hat{\lambda}_{ji,k} = \left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k} \hat{\tilde{P}}_{i,k}^{\epsilon_k} & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \hat{\tilde{P}}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k}\lambda_{ji,k}\right) & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \widehat{w}_i w_i L_i = \sum_k \sum_j \left(\frac{\hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\right) & \forall i \in \mathbb{C} \\ \widehat{\Pi}_i \Pi_i = \sum_k \sum_j \left(\frac{\mu_k \hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\right) & \forall i \in \mathbb{C} \\ \widehat{\mathcal{R}}_i \mathcal{R}_i = \sum_k \sum_j \left(\frac{t_{ji,k}}{1+t_{ji,k}}\hat{\lambda}_{ji,k}\lambda_{ji,k}e_{i,k}\hat{Y}_i Y_i + \frac{(1+x_{ij,k})(1+s_{i,k})-1}{(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j\right) & \forall i \in \mathbb{C} \\ \hat{Y}_i Y_i = \widehat{w}_i w_i L_i + \widehat{\Pi}_i \Pi_i + \widehat{\mathcal{R}}_i \mathcal{R}_i & \forall i \in \mathbb{C} \end{cases},$$

The system specified by Proposition 2 involves $4K$ independent unknowns, $\{x_{hf,k}\}$, $\{t_{fh,k}\}$, $\{\widehat{w}_i\}$, and $\{\hat{Y}_i\}$, and $4K$ independent equations. Moreover, like the previous system, it can be solved with knowledge of only observables and industry-level parameters, ϵ_k, and $\mu_k = 1/\zeta_k$, which were estimated in the previous section.

Our last proposition outlines a procedure to solve for the second-best optimal import taxes characterized by Theorem 3. Recall that these are optimal non-cooperative import taxes when the government is prohibited from using both domestic and export taxes.

Proposition 3. *Suppose the observed data is generated by a model that exhibits the cross-industry utility aggregator 16. Then, the second-best optimal trade taxes and their effects on wages, profits, and income can be fully characterized by solving the following system of equations:*

$$\begin{cases} t_{fh,k} = \frac{(1+\mu_k)\left[1+\epsilon_k\hat{\lambda}_{hh,k}\lambda_{hh,k}\right]}{1+\mu_k+\epsilon_k\hat{\lambda}_{hh,k}\lambda_{hh,k}}(1+\bar{\tau}); x_{hf,k} = 0; \quad ; s_{h,k} = 0 & \forall k \in \mathbb{K} \\ x_{fh,k} = 0, \quad t_{hf,k} = 0; \quad s_k = 0 & \forall k \in \mathbb{K} \\ 1+\bar{\tau} = \sum_g \left[(1+\epsilon_g \frac{\hat{\lambda}_{hf,g}\lambda_{hf,g}e_{f,g}}{\hat{\lambda}_{hf}\lambda_{hf}}\hat{\lambda}_{ff,g}\lambda_{ff,g})(1+\mu_g)^{-1}\right] / \sum_g \left(\epsilon_g \frac{\hat{\lambda}_{hf,g}\lambda_{hf,g}e_{f,g}}{\hat{\lambda}_{hf}\lambda_{hf}}\hat{\lambda}_{ff,g}\lambda_{ff,g}\right) & \\ \hat{\lambda}_{ji,k} = \left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k} \hat{\tilde{P}}_{i,k}^{\epsilon_k} & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \hat{\tilde{P}}_{i,k}^{-\epsilon_k} = \sum_j \left(\left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k}\lambda_{ji,k}\right) & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \widehat{w}_i w_i L_i = \sum_k \sum_j \left(\frac{\hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\right) & \forall i \in \mathbb{C} \\ \widehat{\Pi}_i \Pi_i = \sum_k \sum_j \left(\frac{\mu_k \hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\right) & \forall i \in \mathbb{C} \\ \widehat{\mathcal{R}}_i \mathcal{R}_i = \sum_k \sum_j \left(\frac{t_{ji,k}}{1+t_{ji,k}}\hat{\lambda}_{ji,k}\lambda_{ji,k}e_{i,k}\hat{Y}_i Y_i + \frac{(1+x_{ij,k})(1+s_{i,k})-1}{(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\hat{\lambda}_{ij,k}\lambda_{ij,k}e_{j,k}\hat{Y}_j Y_j\right) & \forall i \in \mathbb{C} \\ \hat{Y}_i Y_i = \widehat{w}_i w_i L_i + \widehat{\Pi}_i \Pi_i + \widehat{\mathcal{R}}_i \mathcal{R}_i & \forall i \in \mathbb{C} \end{cases}.$$

The above system also involves $3K$ independent unknowns, $\{t_{fh,k}\}$, $\{\widehat{w}_i\}$, and $\{\hat{Y}_i\}$, and $3K$ independent equations, and can be solved with knowledge of only observables and industry-level parameters, ϵ_k, and $\mu_k = 1/\xi_k$. Once we solve the above three systems, the gains from optimal policy can be calculated as

$$\hat{W}_i = \hat{Y}_i / \prod_k \left(\hat{\tilde{P}}_{i,k}^{e_{i,k}} \right).$$

To repeat ourself, Propositions 1-3 take an important step in simplifying the task of computing the gains from policy. Existing analyses of optimal policy often solve for the optimal policy schedule using the constrained global optimization technique, also known as MPEC (Mathematical Programming with Equilibrium Constraint, see e.g., Ossa (2014)). Propositions 1-3 allow us to bypass the full-blown global optimization procedure, leading to a notable increase in both computational speed and accuracy.

6.2 The Gains from *First-* and *Second-Best* Non-Cooperative Policies

Table 4 reports the gains from non-cooperative policies both under free and restricted entry. In all cases, the welfare gains are computed assuming the rest of the world does not retaliate. The first column of each panel reports the gains from the first-best non-cooperative tax schedule (as characterized by Theorems 1.A and 1.B). The second column reports the gains from second-best optimal import and export taxes (as characterized by Theorem 2). The third column reports the gains from second-best optimal import taxes (as characterized by Theorem 3). We can draw three main conclusions from the results in Table 4:

a) The gains from ToT manipulation are *relatively* small. This is partly evident from the fact that when governments are restricted to only trade taxes, the resulting gains are significantly smaller than the first-best case. We can also directly verify *pure* ToT gains account for less than a third of the total gains from first-best taxes, which average around 1.7% under restricted entry and 3% under free entry.

b) Trade taxes are a poor *second-best* substitute for domestic Pigouvian taxes. This conclusion can be again drawn from the fact that trade taxes alone can barely replicate a third of the welfare gains attainable under the first-best policy. Under restricted entry, for instance, the fist-best policy increases welfare by 1.67% on average, whereas second-best trade taxes increase welfare by only 0.54%. Second-best import taxes are even less effective, with similar results applying to the free entry case.
   Intuitively, when governments try to mimic Pigouvian taxes with second-best trade taxes, they inevitably worsen the ToT. The reason being that high-markup industries, where exports need to be subsidized, are also industries where countries enjoy more national-level market power and exports need to be taxed for ToT motives. This tension renders second-best trade taxes as largely ineffective.

c) The gains from policy are larger under free entry than under restricted entry. This is an important nuance because, for a small country, the optimal policy schedules look identical in the two cases. But the same tax rates deliver larger gains under free entry, which suggests that *firm-delocation* gains from policy dominate the *profit-shifting* gains.

6.3 The Gains from "Deep" Cooperation

Next, we analyze the gains from deep cooperation. That is, a scenario where all governments cooperatively apply corrective domestic subsidies that are not contaminated with political economy considerations. Trade taxes are also set to zero all over the world. This scenario resembles what Bagwell, Bown, and Staiger (2016) label *deep integration*.

When committing to deep cooperation, governments face an important trade-off. On the one hand, each country forgoes the *ToT gains* from non-cooperative taxation. On the other hand, each country experiences a positive externality from importing subsidized Foreign varieties in high-ψ or high-μ industries. We want to quantify how these two trade-offs compare.

To compute the gains from deep cooperation, we use the closed-form formula specified by Equation 8 and apply the exact hat-algebra methodology. The results are displayed in Figure (2).[42] The *x-axis* in the aforementioned figure corresponds to the gains from deep cooperation. The *y-axis* corresponds to the gains from the optimal (first-best) non-cooperative policy in the absence of retaliation.

For most economies, the gains from deep cooperation dominate those of the optimal non-cooperative policy. What is perhaps surprising is that this outcome emerges even when there is no threat of retaliation from the rest of the world. It simply reflects the fact that the ToT gains from taxation are relatively small, whereas the extent of allocative inefficiency in the global economy is relatively large. So, for many countries, it is beneficial to forgo the ToT gains in exchange for importing more efficient varieties.

Also interestingly, the gains from deep cooperation favor countries that have a comparative disadvantage in high-scale-intensive (or high-profit) industries, e.g., Greece, Russia, and Mexico. The intuition being that these countries depend relatively more on imported varieties in high-ψ or high-μ industries, and, under deep cooperation, these industries are subsidized across the globe.

6.4 The Pitfalls of Unilateralism in Corrective Policies

As noted in Section 2, there is an inherent tension between the ToT and the corrective gains from taxation. Due to this tension, a unilateral application of Pigouvian subsidies will improve allocative inefficiency but can worsen the ToT inefficiency. Table 5 displays the extent of this trade-off for the average country. The unilateral case

[42] The graph excludes Ireland, which experiences negative gains from cooperation.

Table 4: The gains from optimal non-cooperative policies

Country	Restricted Entry			Free Entry		
	All Taxes (First-Best)	Only Exp/Imp Taxes	Only Import Taxes	All Taxes (First-Best)	Only Exp/Imp Taxes	Only Import Taxes
AUS	0.91%	0.29%	0.25%	1.97%	0.85%	0.14%
AUT	1.36%	0.79%	0.62%	2.07%	1.25%	0.72%
BEL	1.34%	0.68%	0.49%	1.77%	0.79%	0.49%
BRA	1.69%	0.31%	0.27%	3.47%	1.15%	0.37%
CAN	1.53%	0.62%	0.49%	2.61%	1.25%	0.57%
CHN	1.74%	0.57%	0.50%	2.82%	0.96%	0.56%
CZE	1.84%	1.03%	0.77%	3.04%	1.86%	0.99%
DEU	1.83%	0.86%	0.65%	3.28%	1.60%	0.80%
DNK	1.03%	0.54%	0.42%	1.84%	0.96%	0.45%
ESP	1.46%	0.55%	0.45%	2.55%	1.19%	0.50%
FIN	1.65%	0.60%	0.51%	2.69%	1.18%	0.57%
FRA	1.85%	0.57%	0.44%	3.98%	1.77%	0.64%
GBR	0.97%	0.53%	0.42%	1.35%	0.77%	0.46%
GRC	1.75%	0.60%	0.57%	3.66%	1.75%	0.89%
HUN	2.41%	1.15%	0.86%	4.74%	2.88%	1.46%
IDN	1.12%	0.42%	0.37%	1.37%	0.50%	0.41%
IND	2.08%	0.45%	0.39%	4.08%	1.41%	0.41%
IRL	1.75%	0.99%	0.85%	2.63%	1.52%	0.94%
ITA	1.43%	0.52%	0.43%	2.47%	1.12%	0.51%
JPN	1.60%	0.43%	0.37%	3.07%	0.94%	0.43%
KOR	2.79%	1.08%	0.92%	4.61%	2.13%	0.98%
MEX	2.04%	0.82%	0.63%	4.40%	2.21%	1.07%
NLD	1.14%	0.61%	0.44%	1.37%	0.75%	0.46%
POL	1.94%	0.83%	0.66%	3.38%	1.74%	0.85%
PRT	1.60%	0.62%	0.54%	2.91%	1.39%	0.66%
ROM	1.75%	0.74%	0.64%	3.18%	1.60%	0.88%
RUS	2.69%	0.51%	0.45%	5.43%	1.92%	0.60%
SVK	1.73%	0.93%	0.69%	2.65%	1.55%	0.86%
SWE	1.13%	0.72%	0.57%	1.35%	0.90%	0.66%
TUR	1.43%	0.61%	0.52%	2.24%	1.08%	0.61%
TWN	2.37%	0.99%	0.82%	4.79%	2.70%	1.43%
USA	1.55%	0.39%	0.33%	3.02%	1.13%	0.40%
Average	1.67%	0.54%	0.45%	3.02%	1.21%	0.54%

Figure 2: *Deep cooperation vs. first-best non-cooperative policy*

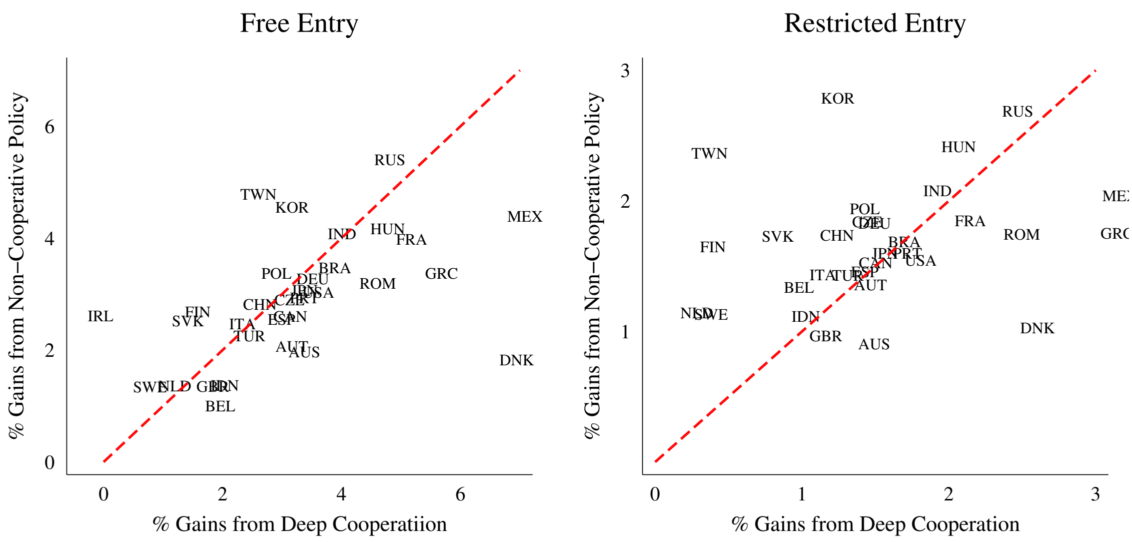


Table 5: *The gains from corrective subsidies: unilateral vs. multilateral adoption*

	Restricted Entry		Free Entry	
	Unilateral	Multilateral	Unilateral	Multilateral
% ΔW_{avg}	-0.36%	1.67%	-1.18%	3.02%

corresponds to a scenario where a country unilaterally imposes Pigouvian subsidies without erecting any trade taxes. The multilateral case corresponds to deep cooperation on Pigouvian subsidies, as outlined in the previous section.

Evidently, the ToT losses from applying unilateral Pigouvian subsidies outweigh the corrective gains for the average country. This result highlights the importance of international coordination in domestic policies. The failure of international coordination can deter most countries from undertaking corrective subsidies that can be ultimately beneficial. This point, though, has been largely overlooked by existing critiques of global governance (e.g., Rodrik (2019)).

6.5 Plausibility of Results

Our finding that the gains from deep cooperation are large sits well with the findings in Baqaee and Farhi (2017) that eliminating sectoral markup-heterogeneity in the U.S. economy can raise real GDP by 2.3%.[43] However, we find larger gains from corrective policies than Bartelme et al. (2018).

How can we reconcile these findings? As an intermediate step, it may be helpful to highlight the Hsieh and Klenow (2009) exact formula for distance from the efficient

[43] This number is the average of the numbers reported in the last column of Table 2 in Baqaee and Farhi (2017).

frontier. Specifically, suppose $\epsilon_k = \epsilon$ is uniform across industries, and that industry-wide productivity levels and μ_k's have a joint log-normal distribution. In that case, the distance from the frontier can be approximated to a first-order as

$$\mathcal{L} \approx \frac{1}{2}\epsilon Var(\ln \mu_k) = \frac{1}{2}\epsilon Var(\ln \psi_k).$$

Considering the above formula, the difference between our results and those in BDCR can be driven by two factors. On the one hand, like Baqaee and Farhi (2017), our analysis does not account for Marshallian externalities. This omission can lead to an overstatement of $\mathcal{L}$ if Marshallian externalities are negatively correlated with firm-level market power.[44] On the other hand, as noted in Section 5.2, the scale elasticity estimated by BDCR can understate the true scale elasticity (and therefore $\mathcal{L}$) when there are diseconomies of scale due to industry-specific factors of production.

7 Concluding Remarks

For centuries scale economies have served as the backbone of many arguments concerning trade and industrial policy. Despite the immense conceptual interest, we know surprisingly little about the empirical relevance of scale economies for the optimal design of and the gains from policy. Against the backdrop of this disconnect, we took a preliminary step toward identifying the force of industry-level scale economies using micro-level trade data. Our estimates indicated that, due to significant cross-industry heterogeneity in scale economies, corrective domestic taxes deliver greater gains than foreign trade taxes for most economies. Furthermore, the gains from a deep multilateral agreement that eliminates inefficiency at the global level dominate the gains from any unilateral policy.

While we highlighted several macro-level implications, our micro-estimates have an even broader reach. Two implications, which were left out in the interest of space, merit particularly close attention. First, our scale elasticity estimates can help disentangle the relative contribution of scale economies and Ricardian comparative advantage to industry-level specialization. This is an old question, which, from an empirical perspective, we know surprisingly little about.

Second, our estimates can shed fresh light on the puzzlingly large income gap between rich and poor countries. Economists have always hypothesized that a fraction of this income gap is driven by rich countries specializing in scale-intensive, high-return industries. An empirical assessment of these hypotheses, however, has been impeded by a lack of comprehensive estimates for industry-level scale elasticities. Our micro-level estimates pave the way for an empirical exploration in this direction.

[44] Another issue is that we are assuming away selection effects in our quantitative analysis. In the presence of selection effects, we can still use our estimates for ϵ_k and ξ_k to identify the scale elasticity up-to an externally chosen trade elasticity. Doing so, however, may lead to a lower or higher $\mathcal{L}$.

References

Adao, R., A. Costinot, and D. Donaldson (2017, March). Nonparametric Counterfactual Predictions in Neoclassical Models of International Trade. *The American Economic Review 107*(3), 633–89.

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.

Anderson, J. E. and E. Van Wincoop (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *The American Economic Review 93*(1), 170–192.

Angrist, J., G. Imbens, and D. Rubin (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association 91*, 444–455.

Antweiler, W. and D. Trefler (2002). Increasing returns and all that: a view from trade. *American Economic Review 92*(1), 93–119.

Bagwell, K., C. P. Bown, and R. W. Staiger (2016). Is the wto passé? *Journal of Economic Literature 54*(4), 1125–1231.

Bagwell, K. and S. H. Lee (2018). Trade Policy under Monopolistic Competition with Firm Selection. *Mimeo, Stanford University*.

Bagwell, K. and R. W. Staiger (2001). Domestic policies, national sovereignty, and international economic institutions. *The Quarterly Journal of Economics 116*(2), 519–562.

Bagwell, K. and R. W. Staiger (2012). Profit Shifting and Trade Agreements in Imperfectly Competitive Markets. *International Economic Review 53*(4), 1067–1104.

Balistreri, E. J., R. H. Hillberry, and T. F. Rutherford (2011). Structural estimation and solution of international trade models with heterogeneous firms. *Journal of International Economics 83*(2), 95–108.

Baqaee, D. and E. Farhi (2019). Networks, barriers, and trade. Technical report, National Bureau of Economic Research.

Baqaee, D. R. and E. Farhi (2017). Productivity and misallocation in general equilibrium. Technical report, National Bureau of Economic Research.

Bartelme, D., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2018). External Economies of Scale and Industrial Policy: A View from Trade. *Working Paper*.

Benassy, J.-P. (1996). Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters 52*(1), 41–47.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Beshkar, M. and A. Lashkaripour (2019). Interdependence of Trade Policies in General Equilibrium.

Bhagwati, J. and V. K. Ramaswami (1963). Domestic Distortions, Tariffs and the Theory of Optimum Subsidy. *Journal of Political Economy 71*(1), 44–50.

Bond, E. W. (1990). The optimal tariff structure in higher dimensions. *International Economic Review*, 103–116.

Bond, E. W. (2013). Trade policy and competition policy: conflict vs. mutual support. In *The International Handbook of Competition–Second Edition*. Edward Elgar Publishing.

Broda, C. and D. E. Weinstein (2006). Globalization and the Gains from Variety. *The Quarterly Journal of Economics 121*(2), 541–585.

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 (2015). Estimates of the Trade and Welfare Effects of NAFTA. *Review of Economic Studies 82*(1), 1–44.

Campolmi, A., H. Fadinger, and C. Forlati (2014). Trade policy: Home market effect versus terms-of-trade externality. *Journal of International Economics 93*(1), 92–107.

Campolmi, A., H. Fadinger, and C. Forlati (2018). Trade and domestic policies in models with monopolistic competition.

Considine, T. J. (2001). Markup pricing in petroleum refining:: A multiproduct framework. *International Journal of Industrial Organization 19*(10), 1499–1526.

Coşar, A. K., P. L. Grieco, S. Li, and F. Tintelnot (2018). What drives home market advantage? *Journal of international economics 110*, 135–150.

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. *Handbook of International Economics 4*, 197.

Costinot, A., A. Rodríguez-Clare, and I. Werning (2016). Micro to Macro: Optimal Trade Policy with Firm Heterogeneity. Technical report, National Bureau of Economic Research.

De Blas, B. and K. N. Russ (2015). Understanding markups in the open economy. *American Economic Journal: Macroeconomics 7*(2), 157–80.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, Markups, and Trade Reform. *Econometrica 84*(2), 445–510.

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

Demidova, S. (2017). Trade policies, firm heterogeneity, and variable markups. *Journal of International Economics 108*, 260–273.

Demidova, S. and A. Rodriguez-Clare (2009). Trade policy under firm-level heterogeneity in a small economy. *Journal of International Economics 78*(1), 100–112.

Dhingra, S. and J. Morrow (2019). Monopolistic competition and optimum product diversity under firm heterogeneity. *Journal of Political Economy 127*(1), 196–232.

Eaton, J. and S. Kortum (2001). Technology, trade, and growth: A unified framework. *European Economic Eeview 45*(4), 742–755.

Eaton, J. and S. Kortum (2002). Technology, Geography, and Trade. *Econometrica 70*(5), 1741–1779.

Edmond, C., V. Midrigan, and D. Y. Xu (2015). Competition, markups, and the gains from international trade. *American Economic Review 105*(10), 3183–3221.

Epifani, P. and G. Gancia (2011). Trade, markup heterogeneity and misallocations. *Journal of International Economics 83*(1), 1–13.

Fattouh, B. (2007). *The drivers of oil prices: the usefulness and limitations of non-structural model, the demand-supply framework and informal approaches*.

Feenstra, R. C., P. Luck, M. Obstfeld, and K. N. Russ (2018, March). In Search of the Armington Elasticity. *The Review of Economics and Statistics 100*(1), 135–150.

Felbermayr, G., B. Jung, and M. Larch (2013). Optimal tariffs, retaliation, and the welfare loss from tariff wars in the melitz model. *Journal of International Economics 89*(1), 13–25.

Grossman, G. and E. Helpman (1994). Protection for sale. *American Economic Review 84*(4), 833–50.

Hanoch, G. (1975). The elasticity of scale and the shape of average costs. *The American Economic Review 65*(3), 492–497.

Helpman, E. (1984). Increasing returns, imperfect markets, and trade theory. *Handbook of international economics 1*, 325–365.

Holmes, T. J., W.-T. Hsu, and S. Lee (2014). Allocative efficiency, mark-ups, and the welfare gains from trade. *Journal of International Economics 94*(2), 195–206.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Khandelwal, A. (2010). The Long and Short (of) Quality Ladders. *The Review of Economic Studies 77*(4), 1450–1476.

Kopczuk, W. (2003). A note on optimal taxation in the presence of externalities. *Economics Letters 80*(1), 81–86.

Kortum, S. S. (1997). Research, Patenting, and Technological Change. *Econometrica 65*(6), 1389–1420.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review 70*(5), 950–959.

Kucheryavyy, K., G. Lyn, and A. Rodríguez-Clare (2016). Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale. NBER Working Paper 22484.

Liu, E. (2018). Industrial policies in production networks. *Available at SSRN*.

Manski, C. F. and D. McFadden (1981). *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press Cambridge, MA.

Melitz, M. (2003). The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations. *Econometrica 71*(6), 1695–1725.

Morrison, C. J. (1992). Unraveling the Productivity Growth Slowdown in the United States, Canada and Japan: The Effects of Subequilibrium, Scale Economies and Markups. *The Review of Economics and Statistics 74*(3), 381–393.

Ossa, R. (2011). A "New Trade" Theory of GATT/WTO Negotiations. *Journal of Political Economy 119*(1), 122–152.

Ossa, R. (2014). Trade Wars and Trade Talks with Data. *The American Economic Review 104*(12), 4104–46.

Ossa, R. (2016). Quantitative Models of Commercial Policy. In *Handbook of Commercial Policy*, Volume 1, pp. 207–259. Elsevier.

Pesaran, M. H., R. P. Smith, and T. Akiyama (1998). *Energy demand in Asian developing economies*. Number BOOK. Oxford University Press.

Pierce, J. R. and P. K. Schott (2012). Concording U.S. Harmonized System Codes over Time. *Journal of Official Statistics 28*(1), 53–68.

Ramondo, N., A. Rodríguez-Clare, and M. Saborío-Rodríguez (2016, October). Trade, Domestic Frictions, and Scale Effects. *The American Economic Review 106*(10), 3159–84.

Redding, S. J. and D. E. Weinstein (2016). A Unified Approach to Estimating Demand and Welfare. CEP Discussion Papers dp1445, Centre for Economic Performance, LSE.

Rodrik, D. (2019). Putting global governance in its place.

Sen, A. K. (1962). An aspect of indian agriculture. *Economic Weekly 14*(4-6), 243–246.

Simonovska, I. and M. E. Waugh (2014). The elasticity of trade: Estimates and evidence. *Journal of International Economics 92*(1), 34–50.

Soderbery, A. (2015). Estimating import supply and demand elasticities: Analysis and implications. *Journal of International Economics 96*(1), 1–17.

Spearot, A. (2016). Unpacking the Long-Run Effects of Tariff Shocks: New Structural Implications From Firm Heterogeneity Models. *American Economic Journal: Microeconomics 8*(2), 128–67.

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.

A The Redundancy of Consumption Taxes

Consider any arbitrary combination of taxes that includes (i) a domestic production tax, s_k, (ii) a domestic consumption tax, c_k, (iii) an import tax, t_k, and (iv) an export tax, x_k, in industry k. This complete set of taxes produces the following wedges between producer and consumer prices in both markets:

$$\begin{cases} \tilde{P}_{hh,k} = (1+c_{h,k})\,(1+s_{h,k})\,P_{hh,k} \\ \tilde{P}_{fh,k} = (1+t_{h,k})\,(1+c_{h,k})\,P_{hh,k} \\ \tilde{P}_{hf,k} = (1+x_{h,k})\,(1+s_{h,k})\,P_{hh,k} \\ \tilde{P}_{ff,k} = P_{ff,k} \end{cases}$$

Our claim here is that the same wedges can be replicated without appealing to the consumption tax, c_k. This claim can be established by considering the following set of production, export, and import taxes: (i) $1+s'_{h,k} = (1+c_{h,k})\,(1+s_{h,k})$; (ii) $\left(1+t'_{h,k}\right) = (1+t_{h,k})\,(1+c_{h,k})$; and (iii) $\left(1+x'_{h,k}\right) = (1+x_{h,k})\,/\,(1+c_{h,k})$. It is straightforward to see that these "three" tax instruments, by construction, reproduce the same wedge between producer and consumer prices as those expressed above:

$$\begin{cases} \tilde{P}_{hh,k} = \left(1+s'_{h,k}\right) P_{hh,k} \\ \tilde{P}_{fh,k} = \left(1+t'_{h,k}\right) P_{hh,k} \\ \tilde{P}_{hf,k} = \left(1+x'_{h,k}\right)\left(1+s'_{h,k}\right) P_{hh,k} \\ \tilde{P}_{ff,k} = P_{ff,k} \end{cases} .$$

That being the case, consumption taxes are redundant in the presence of the other three instruments. The same can be said about any other tax instrument. More specifically, the effect of an import tax can be perfectly replicated with a consumption, production, and export tax combination. However, due to product differentiation, if one of the four tax instruments is restricted, the argument falls apart. For instance, if production taxes are restricted, export and import taxes cannot fully replicate the effect of a domestic consumption tax.

B Proof of Lemma 1

Consider the policy-wage combinations $A = (\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{x}; \boldsymbol{w})$ and $A' = (\boldsymbol{s}', \boldsymbol{t}', \boldsymbol{x}'; \boldsymbol{w})$, where ($a$ and $\tilde{a} \in \mathbb{R}_+$)

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

Our goal is to prove that (i) if $A \in \mathbb{F}$ then $A' \in \mathbb{F}$, and $(ii) W_i(A) = W_i(A')$ for all i. To prove, these claims it suffices to establish two intermediate claims, C1 and C2. Claim C1 follows trivially from the Marshallian demand function being homogeneous of degree zero. i.e., $\mathcal{D}_{ji,k}(Y_i, \tilde{\boldsymbol{P}}_i) = \mathcal{D}_{ji,k}(aY_i, a\tilde{\boldsymbol{P}}_i)$:

$$[\text{C1}] \quad \begin{cases} \tilde{\boldsymbol{P}}_h(A') = a\tilde{\boldsymbol{P}}_h(A) \\ \tilde{\boldsymbol{P}}_f(A') = \tilde{\boldsymbol{P}}_h(A) \\ Y_h(A') = aY_h(A) \\ Y_f(A') = Y_f(A) \end{cases} \implies Q_{ji,k}(A') = Q_{ji,k}(A)\ \forall ji, k$$

Claim C2 follows (after some algebra) from Assumption A2 and the balanced trade condition, $\sum_k \left[\tilde{P}_{hf,k} Q_{hf,k}/(1 + t_{hf,k})\right] = \sum_k \left[\tilde{P}_{fh,k} Q_{fh,k}/(1 + t_{fh,k})\right]$:

$$[\text{C2}] \quad Q_{ji,k}(A') = Q_{ji,k}(A)\ \forall ji, k \implies \begin{cases} \tilde{\boldsymbol{P}}_h(A') = a\tilde{\boldsymbol{P}}_h(A) \\ \tilde{\boldsymbol{P}}_f(A') = \tilde{\boldsymbol{P}}_h(A) \\ Y_h(A') = aY_h(A) \\ Y_f(A') = Y_f(A) \end{cases}$$

Together, Claims C1 and C2 indicate that combinations A and A' correspond to the same allocation, $\{Q_{ji,k}\}$, and yield the same welfare levels due to the indirect utility being homogeneous of degree zero. i.e., $V_i(Y_i, \tilde{\boldsymbol{P}}_i) = V(aY_i, a\tilde{\boldsymbol{P}}_i)$:

$$W_i(A') = V_i(Y_i(A'), \tilde{\boldsymbol{P}}_i(A')) = V(Y_i(A), \tilde{\boldsymbol{P}}_i(A)) = W_i(A).$$

C Nested-Eaton and Kortum (2002) Framework

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

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

As in our main model, country j hosts $\overline{M}_{j,k}$ firms indexed by ω, with $\Omega_{j,k}$ denotes the set of all firms serving industry k from country j.[45] Each firm ω supplies good ν to market i at the following *quality-adjusted* price:

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

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

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

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

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

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

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

To determine the share of goods for which firm ω is the most competitive supplier,

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

we can invoke the theorem of "General Extreme Value." Specifically, define $G(\tilde{\boldsymbol{p}}_i)$ as follows

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

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

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

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

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

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

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

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

which is the exact same distribution specified above, then the probability of choosing variety ω is equal to

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

Rearranging the above equation yields the following expression:

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

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

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

which is identical to the nested-CES function specified by A1, with $\xi_k = 1 + \vartheta_k$ and $\epsilon_k = 1 + \theta_k$.

D Firm-Selection under Pareto

In this appendix, we outline the isomorphism between our baseline model and one that admits selection effects. In doing so, we borrow heavily from Kucheryavyy et al. (2016). The two key assumptions here are,

a) the firm-level productivity distribution, $G_{i,k}(z)$, is Pareto with an industry-level shape parameter γ_k, and

b) the fixed "marketing" cost is paid in terms of labor in the destination market.

Following Kucheryavyy et al. (2016), we also assume that cross-industry utility aggregator is Cobb-Douglas with $e_{i,k}$ denoting the constant share of country i's expenditure on industry k. Following Kucheryavyy et al. (2016), we can define the effective supply of production labor in country i as

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

with the labor market clearing condition given by

$$\sum w_i L_{i,k} = w_i \widetilde{L}_i.$$

First, we can appeal to the well-known result that the profit margin in each industry is given by the following expression:

$$\mu_k \equiv \frac{\sum_n P_{in,k} Q_{in,k}}{w_i L_{i,k}} = \frac{\xi_k - 1}{\gamma_k \xi_k}.$$

Second, following the steps in Appendix B.2 of Kucheryavyy et al. (2016), we can express the national-level demand as

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

where $\theta_k \equiv \gamma_k / \left[1+\gamma_k(1-\alpha_k)/(\epsilon_k-1)\right]$, with $\alpha_k \equiv (\epsilon_k-1)/(\xi_k-1)$ as defined in the main text. Abstracting from taxes, the national-level price indexes in the above expression are given by,

$$\tilde{P}_{ji,k} = \rho_{ji,k} w_j M_{j,k}^{-\frac{1}{\gamma_k}} = \rho_{ji,k} w_j Q_{j,k}^{-\frac{1}{1+\gamma_k}},$$

where $\rho_{ji,k}$ is composed of structural parameters that are invariant to policy.[46] Correspondingly, $\tilde{P}_{i,k} = \left(\sum \tilde{P}_{ji,k}^{-\theta_k}\right)^{-1/\theta_k}$ is the industry-level consumer price index that shows up in indirect utility $V_i(.)$. Combining the above expressions, the above model is isomorphic to a version of the baseline model which feature the following elasticity values:

$$\begin{cases} \psi_k = 1/\gamma_k, & \forall k \\ \mu_k = (\xi_k - 1)/\gamma_k \xi_k, & \forall k \\ \varepsilon_{ji,k}^{ji,k} = -1 - \theta_k \left(1 - \lambda_{ji,k}/e_{i,k}\right), & \forall k;\ j = j \\ \varepsilon_{ji,k}^{ji,k} = \theta_k \lambda_{ji,g}/e_{i,k}, & \forall k;\ j \neq j \end{cases}$$

To pin down all the above elasticities, we also need an estimate for θ_k, which is essentially the trade elasticity under firm selection. Doing so, requires estimating $\partial \ln Q_{ji,k} = \partial \ln \left(1+t_{ji,k}\right)$, assuming that tariffs are imposed on cost rather than revenue.

E Proof of Theorem 1.A

The solve for the optimal tax schedule, we take the dual approach. We can simplify the problem by formulating the optimal tax problem as on where government directly chooses optimal consumer prices $\{\tilde{P}^*_{hh,k}, \tilde{P}^*_{fh,k}, \tilde{P}^*_{hf,k}\}$. This choice implicitly pins down the optimal tax schedule

$$\begin{aligned} 1+s_k^* &\equiv \tilde{P}^*_{hh,k}/P_{hh,k} \quad \forall k \\ 1+t_k^* &\equiv \tilde{P}^*_{fh,k}/P_{fh,k} \quad \forall k \\ 1+x_k^* &\equiv \frac{\tilde{P}^*_{hf,k}/P_{hh,k}}{\tilde{P}^*_{hh,k}/P_{hh,k}} \quad \forall k \end{aligned}$$

To cast the problem in this way, note that any feasible combination $(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}; \boldsymbol{w}) \in \mathbb{F}$ corresponds to a feasible price-wage combination $(\tilde{\boldsymbol{P}}; \boldsymbol{w})$, where $\tilde{\boldsymbol{P}} \equiv \{\tilde{P}_{ji,k}\}$. Accordingly, the optimal tax problem is characterized by the following system of first order conditions (F.O.C.s):

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

$$\begin{cases} \nabla_{\tilde{\boldsymbol{P}}_{hh}} W_h(\tilde{\boldsymbol{P}}; \boldsymbol{w}) = 0 \\ \nabla_{\tilde{\boldsymbol{P}}_{fh}} W_h\left(\tilde{\boldsymbol{P}}; \boldsymbol{w}\right) = 0 \\ \nabla_{\tilde{\boldsymbol{P}}_{hf}} W_h\left(\tilde{\boldsymbol{P}}; \boldsymbol{w}\right) = 0 \end{cases},$$

Our objective is to solve this system subject to $(\tilde{\boldsymbol{P}}; \boldsymbol{w}) \sim (\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}; \boldsymbol{w}) \in \mathbb{F}$. Focusing on the feasible set of tax-wage combination means that we can invoke envelope conditions to evaluate the F.O.C.s. We, hereafter, treat labor in Foreign as the numeraire, i.e., $w_f = 1$.

F.O.C. with respect to $\tilde{P}_{hh,k}$

The F.O.C. with respect to $\tilde{P}_{hh,k}$ features terms accounting for direct effects and general equilibrium wage effects:

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}}; w_h)}{\mathrm{d}\tilde{P}_{hh,k}} = \frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial \tilde{P}_{hh,k}} + \frac{\partial V_h(.)}{\partial \tilde{P}_{hh,k}} + \sum_g \frac{\partial W_h}{\partial \tilde{P}_{ff,g}}\frac{\partial \tilde{P}_{ff,g}}{\partial \tilde{P}_{hh,k}} + \left(\frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f}\sum_g \left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial w_h}\right]\right)\frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{hh,k}} = 0.$$

Note that by choice of numeraire, $\tilde{P}_{ff,g}$ is invariant to policy, i.e., $\partial \tilde{P}_{ff,g}/\partial \tilde{P}_{hh,k} = 0$. We can also appeal to Roy's identity, $\frac{\partial V_h/\partial \tilde{P}_{ih,k}}{\partial V_h/\partial Y_h} = -Q_{ih,k}$, to further simplify the above equation as follows:

$$\frac{\mathrm{d}W_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\mathrm{d}\ln \tilde{P}_{hh,k}} = \frac{\partial V_h}{\partial Y_h}\left\{\frac{\partial Y_h}{\partial \ln \tilde{P}_{hh,k}} - \tilde{P}_{hh,k}Q_{hh,k} + \left(\frac{\partial Y_h}{\partial \ln w_h} - \sum_{i=h,f}\sum_g \left[\tilde{P}_{ih,g}Q_{ih,g}\frac{\partial \ln \tilde{P}_{ih,g}}{\partial \ln w_h}\right]\right)\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln \tilde{P}_{hh,k}}\right\} = 0. \tag{17}$$

To determine $\partial Y_h/\partial \ln \tilde{P}_{hh,k}$, note that total income is the sum of factor income, profits, and tax revenue,

$$Y_h\left(\tilde{\boldsymbol{P}}; w_h\right) = w_h L_h + \sum_g \sum_i \left(\frac{\mu_g}{1+\mu_g} P_{hi,g} Q_{hi,g}\right) + \mathcal{R}_h(\tilde{\boldsymbol{P}}; w_h), \tag{18}$$

where

$$\mathcal{R}_h(\tilde{\boldsymbol{P}}; w_h) = \sum_g \left[(\tilde{P}_{hh,g} - P_{hh,g})Q_{hh,g} + (\tilde{P}_{fh,g} - P_{fh,g})Q_{fh,g} + (\tilde{P}_{hf,g} - P_{hf,g})Q_{hf,g}\right]. \tag{19}$$

Taking a basic partial derivative from Equation 18, yields the following

$$
\begin{aligned}
\frac{\partial Y_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial\ln\tilde{P}_{hh,k}} =& \sum_g\left(\frac{\mu_g}{1+\mu_g}P_{hh,g}Q_{hh,g}\left(\frac{\partial\ln Q_{hh,g}}{\partial\ln\tilde{P}_{hh,k}}+\frac{\partial\ln Q_{hh,g}}{\partial\ln Y_h}\frac{\partial\ln Y_h}{\ln\tilde{P}_{hh,k}}\right)\right)\\
&+\tilde{P}_{hh,k}Q_{hh,k}+\sum_g\left((\tilde{P}_{hh,g}-P_{hh,g})Q_{hh,g}\left[\frac{\partial\ln Q_{hh,g}}{\partial\ln\tilde{P}_{hh,k}}+\frac{\partial\ln Q_{hh,g}}{\partial\ln Y_h}\frac{\partial\ln Y_h}{\ln\tilde{P}_{hh,k}}\right]\right)\\
&+\sum_g\left((\tilde{P}_{fh,g}-P_{fh,g})Q_{fh,g}\left[\frac{\partial\ln Q_{fh,g}}{\partial\ln\tilde{P}_{hh,k}}+\frac{\partial\ln Q_{fh,g}}{\partial\ln Y_h}\frac{\partial\ln Y_h}{\ln\tilde{P}_{hh,k}}\right]\right)\\
=& \sum_g\left(\left[\frac{\mu_g}{1+\mu_g}+(\frac{\tilde{P}_{hh,g}}{P_{hh,g}}-1)\right]P_{hh,g}Q_{hh,g}\varepsilon^{hh,k}_{hh,g}\right)+\tilde{P}_{hh,k}Q_{fh,k}+\sum_g\left[(\tilde{P}_{fh,g}-P_{fh,g})Q_{fh,g}\varepsilon^{hh,k}_{fh,g}\right]\\
&+\sum_g\left((\tilde{P}_{fh,g}-P_{fh,g})Q_{fh,g}Q_{fh,g}\eta_{fh,g}+\left[\frac{\mu_g}{1+\mu_g}P_{hh,g}+(\tilde{P}_{hh,g}-P_{hh,g})\right]Q_{hh,g}\eta_{hh,g}\right)\frac{\partial\ln Y_h}{\partial\ln\tilde{P}_{hh,k}}.
\end{aligned}
$$

where $\eta_{fh,g}\equiv\partial\ln Q_{fh,g}/\partial\ln Y_h$ denotes the income elasticity of demand. Similarly, $\mathrm{d}w_h/\mathrm{d}\tilde{P}_{hh,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition,

$$
D_h\left(\tilde{\boldsymbol{P}};w_h\right)\equiv\sum_k\left(P_{fh,k}Q_{fh,k}-\tilde{P}_{hf,k}Q_{hf,k}\right)=0,
$$

which yields the following:

$$
\begin{aligned}
\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln\tilde{P}_{hh,k}} &= \left[\frac{\partial}{\partial\ln\tilde{P}_{hh,k}}\sum_g\left(P_{fh,g}Q_{fh,g}-\tilde{P}_{hf,g}Q_{hf,g}\right)\right]\Bigg/\frac{\partial D_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial\ln w_h}\\
&= \left[\sum_g\left(P_{fh,g}Q_{fh,g}\frac{\partial\ln Q_{fh,g}}{\partial\ln\tilde{P}_{hh,k}}\right)+\sum_g\left(P_{fh,g}Q_{fh,g}\eta_{fh,g}\right)\frac{\partial\ln Y_h}{\partial\ln\tilde{P}_{hh,k}}\right]\Bigg/\frac{\partial D_h(.)}{\partial\ln w_h}\\
&= \left[\sum_g\left(P_{fh,g}Q_{fh,g}\varepsilon^{hh,k}_{fh,g}\right)+\sum_g\left(P_{fh,g}Q_{fh,g}\eta_{fh,g}\right)\frac{\partial Y_h}{\partial\ln\tilde{P}_{hh,k}}\right]\Bigg/\frac{\partial D_h(.)}{\partial\ln w_h}.
\end{aligned}
$$

Plugging the expressions for $\partial Y_h/\partial\ln\tilde{P}_{hh,k}$ and $\mathrm{d}w_h/\mathrm{d}\ln\tilde{P}_{hh,k}$ back into the F.O.C., defining $\bar{\tau}\equiv\left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial w_h}+\sum_i\sum_g\frac{\partial V_h}{\partial\tilde{P}_{ih,g}}\frac{\partial\tilde{P}_{ih,g}}{\partial\ln w_h}\right)/\left(\partial\ln D_h/\partial w_h\right)$, we can write the F.O.C. (Equation 17) as

$$
\begin{aligned}
&\sum_g\left(\left[1-\frac{1}{1+\mu_g}\frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right]\tilde{P}_{hh,g}Q_{hh,g}\varepsilon^{hh,k}_{hh,g}\right)+\sum_g\left[\left(1-(1+\bar{\tau})\frac{P_{fh,g}}{\tilde{P}_{fh,g}}\right)\tilde{P}_{fh,g}Q_{fh,g}\varepsilon^{hh,k}_{fh,g}\right]\\
&+\sum_g\left([1-(1+\bar{\tau})\frac{P_{fh,g}}{\tilde{P}_{fh,g}}]\tilde{P}_{fh,g}Q_{fh,g}\eta_{fh,g}+\left[1-\frac{1}{1+\mu_g}\frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right]\tilde{P}_{hh,g}Q_{hh,g}\eta_{hh,g}\right)\frac{\partial\ln Y_h}{\ln\tilde{P}_{hh,k}}=0
\end{aligned}
$$

Dividing the above equation by Y_h and noting that $\acute{\lambda}_{ji,k} \equiv \tilde{P}_{ji,k}Q_{ji,k}/Y_i$, $1+s_k \equiv \tilde{P}_{hh,k}/P_{hh,k}$, and $1+t_k \equiv \tilde{P}_{fh,k}/P_{fh,k}$, yields the following simplified expression:

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\mu_g}\right] \acute{\lambda}_{hh,g}\varepsilon^{hh,k}_{hh,g} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \acute{\lambda}_{fh,g}\varepsilon^{hh,k}_{fh,g} \right]$$
$$+ \sum_g \left(\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \acute{\lambda}_{fh,g}\eta_{fh,g} + \left[\frac{1+s_g}{1+\mu_g} - 1\right] \acute{\lambda}_{hh,g}\eta_{hh,g} \right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0$$

We can further simplify the above expression, we can rely on the observations that (i) there are multiple optimal tax combinations per Lemma 1, (ii) the term in parenthesis in the second line of the above equation, namely,

$$\Psi\left(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}\right) \equiv \sum_g \left(\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \acute{\lambda}_{fh,g}\eta_{fh,g} + \left[\frac{1+s_g}{1+\mu_g} - 1\right] \acute{\lambda}_{hh,g}\eta_{hh,g} \right),$$

is not industry-specific. Hence, by choice of $\bar{s}$ in Lemma 1, there exists an optimal tax combination for which $\Psi\left(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*\right) = 0$. We can first restrict attention to this particular solution, and then trivially identify the other solutions with an across-the-board shift in all tax instruments. Doing so, the F.O.C. with respect to $\tilde{P}_{hh,k}$ reduces to

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\mu_g}\right] \acute{\lambda}_{hh,g}\varepsilon^{hh,k}_{hh,g} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \acute{\lambda}_{fh,g}\varepsilon^{hh,k}_{fh,g} \right] = 0 \tag{20}$$

F.O.C. with respect to $\tilde{P}_{fh,k}$

The F.O.C. with respect to $\tilde{P}_{fh,k}$ also features terms accounting for direct effects and general equilibrium wage effects:

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}}_h; w_h)}{\mathrm{d}\tilde{P}_{fh,k}} = \frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial \tilde{P}_{fh,k}} + \frac{\partial V_h(.)}{\partial \tilde{P}_{fh,k}} + \sum_g \frac{\partial W_h}{\partial \tilde{P}_{ff,g}}\frac{\partial \tilde{P}_{ff,g}}{\partial \tilde{P}_{fh,k}}$$
$$+ \left(\frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f}\sum_g \left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial w_h} \right] \right) \frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{fh,k}} = 0.$$

Appealing to Roy's identity, $\frac{\partial V_h/\partial \tilde{P}_{ih,k}}{\partial V_h/\partial Y_h} = -Q_{ih,k}$, and given that $\partial \tilde{P}_{ff,g}/\partial \tilde{P}_{fh,k} = 0$, we can simplify the above equation as follows:

$$\frac{\mathrm{d}W_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\mathrm{d}\ln \tilde{P}_{fh,k}} = \frac{\partial V_h}{\partial Y_h} \left\{ \frac{\partial Y_h}{\partial \ln \tilde{P}_{fh,k}} - \tilde{P}_{fh,k}Q_{fh,k} \right.$$
$$\left. + \left(\frac{\partial Y_h}{\partial \ln w_h} - \sum_{i=h,f}\sum_g \left[\tilde{P}_{ih,g}Q_{ih,g}\frac{\partial \ln \tilde{P}_{ih,g}}{\partial \ln w_h} \right] \right) \frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln \tilde{P}_{fh,k}} \right\} = 0. \tag{21}$$

Taking a basic partial derivative from Equation 18, yields

$$\begin{aligned}
\frac{\partial Y_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial\ln\tilde{P}_{fh,k}} =& \sum_g\left(\frac{\mu_g}{1+\mu_g}P_{hh,g}Q_{hh,g}\left(\frac{\partial\ln Q_{hh,g}}{\partial\ln\tilde{P}_{fh,k}}+\frac{\partial\ln Q_{hh,g}}{\partial\ln Y_h}\frac{\partial\ln Y_h}{\ln\tilde{P}_{fh,k}}\right)\right)\\
&+\tilde{P}_{fh,k}Q_{fh,k}+\sum_g\left((\tilde{P}_{hh,g}-P_{hh,g})Q_{hh,g}\left[\frac{\partial\ln Q_{hh,g}}{\partial\ln\tilde{P}_{fh,k}}+\frac{\partial\ln Q_{hh,g}}{\partial\ln Y_h}\frac{\partial\ln Y_h}{\ln\tilde{P}_{fh,k}}\right]\right)\\
&+\sum_g\left((\tilde{P}_{fh,g}-P_{fh,g})Q_{fh,g}\left[\frac{\partial\ln Q_{fh,g}}{\partial\ln\tilde{P}_{fh,k}}+\frac{\partial\ln Q_{fh,g}}{\partial\ln Y_h}\frac{\partial\ln Y_h}{\ln\tilde{P}_{fh,k}}\right]\right)\\
=&\sum_g\left(\left[\frac{\mu_g}{1+\mu_g}+(\frac{\tilde{P}_{hh,g}}{P_{hh,g}}-1)\right]P_{hh,g}Q_{hh,g}\varepsilon^{fh,k}_{hh,g}\right)+\tilde{P}_{fh,k}Q_{fh,k}+\sum_g\left[(\tilde{P}_{fh,g}-P_{fh,g})Q_{fh,g}\varepsilon^{fh,k}_{fh,g}\right]\\
&+\sum_g\left((\tilde{P}_{fh,g}-P_{fh,g})Q_{fh,g}Q_{fh,g}\eta_{fh,g}+\left[\frac{\mu_g}{1+\mu_g}P_{hh,g}+(\tilde{P}_{hh,g}-P_{hh,g})\right]Q_{hh,g}\eta_{hh,g}\right)\frac{\partial\ln Y_h}{\partial\ln\tilde{P}_{fh,k}}.
\end{aligned}$$

where $\eta_{fh,g}\equiv\partial\ln Q_{fh,g}/\partial\ln Y_h$ denotes the income elasticity of demand. Similarly, $\mathrm{d}w_h/\mathrm{d}\tilde{P}_{hh,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition,

$$D_h\left(\tilde{\boldsymbol{P}};w_h\right)\equiv\sum_k\left(P_{fh,k}Q_{fh,k}-\tilde{P}_{hf,k}Q_{hf,k}\right)=0,$$

which yields the following:

$$\begin{aligned}
\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln\tilde{P}_{fh,k}} &= \left[\frac{\partial}{\partial\ln\tilde{P}_{fh,k}}\sum_g\left(P_{fh,g}Q_{fh,g}-\tilde{P}_{hf,g}Q_{hf,g}\right)\right]\Bigg/\frac{\partial D_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial\ln w_h}\\
&=\left[\sum_g\left(P_{fh,g}Q_{fh,g}\frac{\partial\ln Q_{fh,g}}{\partial\ln\tilde{P}_{fh,k}}\right)+\sum_g\left(P_{fh,g}Q_{fh,g}\eta_{fh,g}\right)\frac{\partial\ln Y_h}{\partial\ln\tilde{P}_{fh,k}}\right]\Bigg/\frac{\partial D_h(.)}{\partial\ln w_h}\\
&=\left[\sum_g\left(P_{fh,g}Q_{fh,g}\varepsilon^{fh,k}_{fh,g}\right)+\sum_g\left(P_{fh,g}Q_{fh,g}\eta_{fh,g}\right)\frac{\partial Y_h}{\partial\ln\tilde{P}_{fh,k}}\right]\Bigg/\frac{\partial D_h(.)}{\partial\ln w_h}.
\end{aligned}$$

Plugging the expressions for $\partial Y_h/\partial\ln\tilde{P}_{hh,k}$ and $\mathrm{d}w_h/\mathrm{d}\ln\tilde{P}_{hh,k}$ back into the F.O.C., defining $\bar{\tau}\equiv\left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial w_h}+\sum_i\sum_g\frac{\partial V_h}{\partial\tilde{P}_{ih,g}}\frac{\partial\tilde{P}_{ih,g}}{\partial\ln w_h}\right)/\left(\partial\ln D_h/\partial w_h\right)$, we can write the F.O.C. (Equation 21) as

$$\begin{aligned}
&\sum_g\left(\left[1-\frac{1}{1+\mu_g}\frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right]\tilde{P}_{hh,g}Q_{hh,g}\varepsilon^{fh,k}_{hh,g}\right)+\sum_g\left[\left(1-(1+\bar{\tau})\frac{P_{fh,g}}{\tilde{P}_{fh,g}}\right)\tilde{P}_{fh,g}Q_{fh,g}\varepsilon^{fh,k}_{fh,g}\right]\\
&+\sum_g\left([1-(1+\bar{\tau})\frac{P_{fh,g}}{\tilde{P}_{fh,g}}]\tilde{P}_{fh,g}Q_{fh,g}\eta_{fh,g}+\left[1-\frac{1}{1+\mu_g}\frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right]\tilde{P}_{hh,g}Q_{hh,g}\eta_{hh,g}\right)\frac{\partial\ln Y_h}{\ln\tilde{P}_{hh,k}}=0
\end{aligned}$$

Dividing the above equation by Y_h and noting that $\acute{\lambda}_{ji,k} \equiv \tilde{P}_{ji,k} Q_{ji,k}/Y_i$, $1 + s_k \equiv \tilde{P}_{hh,k}/P_{hh,k}$, and $1 + t_k \equiv \tilde{P}_{fh,k}/P_{fh,k}$, yields the following simplified expression:

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\mu_g} \right] \acute{\lambda}_{hh,g} \varepsilon_{hh,g}^{fh,k} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g} \right) \acute{\lambda}_{fh,g} \varepsilon_{fh,g}^{fh,k} \right] + \Psi(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0$$

As noted earlier, there exists an optimal tax combination for which $\Psi(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*) = 0$. Restricting attention to this particular solution, the F.O.C. with respect to $\tilde{P}_{fh,k}$ reduces to

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\mu_g} \right] \acute{\lambda}_{hh,g} \varepsilon_{hh,g}^{fh,k} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g} \right) \acute{\lambda}_{fh,g} \varepsilon_{fh,g}^{fh,k} \right] = 0 \tag{22}$$

F.O.C. with respect to $\tilde{P}_{hf,k}$

Noting that $\partial \tilde{P}_{ff,g}/\partial \tilde{P}_{fh,k} = 0$, the F.O.C. with respect to $\tilde{P}_{hf,k}$ can be stated as

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}}; w_h)}{\mathrm{d}\tilde{P}_{hf,k}} = \frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial \tilde{P}_{hf,k}} + \left(\frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f} \sum_g \left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}} \frac{\partial \tilde{P}_{ih,g}}{\partial w_h} \right] \right) \frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{hf,k}} = 0$$

As before, we can derive $\partial Y_h / \partial \ln \tilde{P}_{hf,k}$ by taking a derivative from Equation 18, which yields

$$\begin{aligned}
\frac{\partial Y_h(\tilde{\boldsymbol{P}}; w_h)}{\partial \ln \tilde{P}_{hf,k}} = & \sum_g \left(\frac{\mu_g}{1+\mu_g} \left(P_{hf,g} Q_{hf,g} \frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}} + P_{hh,g} Q_{hh,g} \frac{\partial \ln Q_{hh,g}}{\partial \ln Y_h} \frac{\partial \ln Y_h}{\ln \tilde{P}_{hf,k}} \right) \right) \\
& + \tilde{P}_{hf,k} Q_{hf,k} + \sum_g \left((\tilde{P}_{hf,g} - P_{hf,g}) \frac{\partial Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}} \right) \\
& + \sum_g \left((\tilde{P}_{hh,g} - P_{hh,g}) \frac{\partial Q_{hh,g}}{\partial \ln Y_h} + (\tilde{P}_{fh,g} - P_{fh,g}) \frac{\partial Q_{fh,g}}{\partial \ln Y_h} \right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hf,k}} \\
= & \sum_g \left(\left[\frac{\mu_g}{1+\mu_g} + \left(\frac{\tilde{P}_{hf,g}}{P_{hf,g}} - 1 \right) \right] P_{hf,g} Q_{hf,g} \varepsilon_{hf,g}^{hf,k} \right) + \tilde{P}_{hf,k} Q_{hf,k} + \sum_g \left[(\tilde{P}_{hf,g} - P_{hf,g}) Q_{hf,g} \varepsilon_{hf,g}^{hf,k} \right] \\
& + \sum_g \left((\tilde{P}_{fh,g} - P_{fh,g}) Q_{fh,g} Q_{fh,g} \eta_{fh,g} + \left[\frac{\mu_g}{1+\mu_g} P_{hh,g} + (\tilde{P}_{hh,g} - P_{hh,g}) \right] Q_{hh,g} \eta_{hh,g} \right) \frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hf,k}}.
\end{aligned}$$

Similarly, $\mathrm{d}w_h / \mathrm{d}\tilde{P}_{hf,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition, $D_h(\tilde{\boldsymbol{P}}; w_h)$, which yields the following:

$$\begin{aligned}
\frac{\mathrm{d} \ln w_h}{\mathrm{d}\ln \tilde{P}_{hf,k}} &= \left[\frac{\partial}{\partial \ln \tilde{P}_{hf,k}} \sum_g (P_{fh,g} Q_{fh,g} - \tilde{P}_{hf,g} Q_{hf,g}) \right] \Big/ \frac{\partial D_h(\tilde{\boldsymbol{P}}; w_h)}{\partial \ln w_h} \\
&= \left[\sum_g (P_{fh,g} Q_{fh,g} \eta_{fh,g}) \frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hf,k}} - \tilde{P}_{hf,k} Q_{hf,k} - \sum_g \left(\tilde{P}_{hf,g} Q_{hf,g} \frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}} \right) \right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h} \\
&= \left[\sum_g (P_{fh,g} Q_{fh,g} \eta_{fh,g}) \frac{\partial Y_h}{\partial \ln \tilde{P}_{hf,k}} - \tilde{P}_{hf,k} Q_{hf,k} - \sum_g \left(\tilde{P}_{hf,g} Q_{hf,g} \varepsilon_{hf,g}^{hf,k} \right) \right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h}.
\end{aligned}$$

Plugging the expressions for $\partial Y_h/\partial \ln \tilde{P}_{hf,k}$ and $dw_h/d\ln\tilde{P}_{hf,k}$ back into the F.O.C.; adopting our earlier definition $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial \ln w_h} + \sum_i \sum_g \frac{\partial V_h}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial \ln w_h}\right) / \left(\partial D_h/\partial \ln w_h\right)$; and focusing in the solution for which $\Psi\left(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*\right) = 0$, we can write the F.O.C. as

$$\acute{\lambda}_{hf,k} + \sum_g \left[1 - \frac{1}{(1+x_k)(1+\bar{\tau})(1+s_g)(1+\mu_g)}\right] \acute{\lambda}_{hf,g}\varepsilon_{hf,g}^{hf,k} = 0, \tag{23}$$

where in deriving the above expression, we use the fact $\tilde{P}_{hf,k}/P_{hf,k} = (1+x_k)(1+s_k)$, and $\acute{\lambda}_{hf,g} \equiv \tilde{P}_{hf,g}Q_{hf,g}/Y_h$.

Simultaneously Solving the System of F.O.C.

Combining Equations 20, 22, and 23 yields the following system of F.O.C.:

$$\begin{aligned}
\sum_g \left(\left[1 - \frac{1+s_g}{1+\mu_g}\right] \acute{\lambda}_{hh,g}\varepsilon_{hh,g}^{hh,k}\right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \acute{\lambda}_{fh,g}\varepsilon_{fh,g}^{hh,k}\right] &= 0 \quad [\tilde{P}_{hh,k}] \\
\sum_g \left(\left[1 - \frac{1+s_g}{1+\mu_g}\right] \acute{\lambda}_{hh,g}\varepsilon_{hh,g}^{fh,k}\right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \acute{\lambda}_{fh,g}\varepsilon_{fh,g}^{fh,k}\right] &= 0 \quad [\tilde{P}_{fh,k}] \\
\acute{\lambda}_{hf,k} + \sum_g \left[1 - \frac{1}{(1+x_k)(1+\bar{\tau})(1+s_g)(1+\mu_g)}\right] \acute{\lambda}_{hf,g}\varepsilon_{hf,g}^{hf,k} &= 0 \quad [\tilde{P}_{hf,k}]
\end{aligned}$$

Solving the above system implies that $1+s_g = 1+\mu_g$ and $1+t_g = 1+\bar{\tau}$. The term $\bar{\tau}$ is, however, redundant due to multiplicity of the optimal policy up to two uniform tax shifters (Lemma 1).[47] So, altogether, for any $\bar{s}$ and $\bar{t} \in \mathbb{R}_+$, the following formula characterizes an optimal tax combination:

$$\begin{cases}
(1+s_k)(1+\bar{s}) = 1+\mu_k & \forall k \in \mathbb{K} \\
1+t_k = 1+\bar{t} & \forall k \in \mathbb{K}\,, \\
\left[1/(1+x_k)(1+\bar{t})\right]_k = \boldsymbol{\mathcal{E}}_{hf}^{-1}\left(\boldsymbol{I} + \boldsymbol{\mathcal{E}}_{hf}\right)\boldsymbol{1}
\end{cases}$$

where $\boldsymbol{\mathcal{E}}_{hf} \equiv \left[\frac{\lambda_{hf,g}}{\lambda_{hf,k}}\varepsilon_{hf,g}^{hf,k}\right]_{g,k}$ is a $K \times K$ matrix composed of reduced-form demand elasticities and expenditure shares. Under Homothetic preferences, $\frac{\lambda_{hf,g}}{\lambda_{hf,k}}\varepsilon_{hf,g}^{hf,k} = \varepsilon_{hf,k}^{hf,g}$, which implies that $\boldsymbol{\mathcal{E}}_{hf} \equiv \left[\varepsilon_{hf,k}^{hf,g}\right]_{g,k}$.

F Proof of Theorem 1.B

The optimal tax problem under free entry is characterized by the following system of first order conditions (F.O.C.s):

[47] Considering the notation in Lemma 1, by choice of $\bar{s}$ we were able to set $\Psi\left(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*\right) = 0$, and by choice of $\bar{t}$ we are able to set $\bar{\tau} = 0$. These choices identify one of the multiple solutions. The remaining solutions then differ from this particular solution in only the uniform tax-shifters $\bar{s}$ and $\bar{t}$.

$$\begin{cases} \nabla_{\tilde{\boldsymbol{P}}_{hh}} W_h(\tilde{\boldsymbol{P}};\boldsymbol{w}) = 0 \\ \nabla_{\tilde{\boldsymbol{P}}_{fh}} W_h\left(\tilde{\boldsymbol{P}};\boldsymbol{w}\right) = 0 \\ \nabla_{\tilde{\boldsymbol{P}}_{hf}} W_h\left(\tilde{\boldsymbol{P}};\boldsymbol{w}\right) = 0 \end{cases},$$

Our objective is to solve this system subject to $(\tilde{\boldsymbol{P}};\boldsymbol{w}) \sim (\boldsymbol{t},\boldsymbol{x},\boldsymbol{s};\boldsymbol{w}) \in \mathbb{F}$. As noted in the previous appendix, focusing on the feasible set of tax-wage combination means that we can invoke envelope conditions to evaluate the F.O.C.s. We, hereafter, treat labor in Foreign as the numeraire, i.e., $w_f = 1$.

F.O.C. with respect to $\tilde{P}_{hh,k}$

The F.O.C. with respect to $\tilde{P}_{hh,k}$ features terms accounting for direct effects and general equilibrium wage effects:

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}};w_h)}{\mathrm{d}\tilde{P}_{hh,k}} = \frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial \tilde{P}_{hh,k}} + \frac{\partial V_h(.)}{\partial \tilde{P}_{hh,k}} + \sum_g \frac{\partial W_h}{\partial \tilde{P}_{ff,g}}\frac{\partial \tilde{P}_{ff,g}}{\partial \tilde{P}_{hh,k}}$$
$$+ \left(\frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f}\sum_g\left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial w_h}\right]\right)\frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{hh,k}} = 0.$$

where $\tilde{\boldsymbol{P}} \equiv \{\tilde{P}_{ji,k}\}$. The above problem is greatly simplified by the observation that $\partial W_h/\partial \tilde{P}_{ff,g} = 0$. Appealing to Roy's identity that $\frac{\partial V_h/\partial \tilde{P}_{ih,k}}{\partial V_h/\partial Y_h} = -Q_{ih,k}$, we can simplify the above equation as follows:

$$\frac{\mathrm{d}W_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\mathrm{d}\ln \tilde{P}_{hh,k}} = \frac{\partial V_h}{\partial Y_h}\left\{\frac{\partial Y_h}{\partial \ln \tilde{P}_{hh,k}} - \tilde{P}_{hh,k}Q_{hh,k}\right.$$
$$\left. + \left(\frac{\partial Y_h}{\partial \ln w_h} - \sum_{i=h,f}\sum_g\left[\tilde{P}_{ih,g}Q_{ih,g}\frac{\partial \ln \tilde{P}_{ih,g}}{\partial \ln w_h}\right]\right)\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln \tilde{P}_{hh,k}}\right\} = 0. \tag{24}$$

To determine $\partial Y_h/\partial \ln \tilde{P}_{hh,k}$, note that total income is the sum of factor income, profits, and tax revenue,

$$Y_h\left(\tilde{\boldsymbol{P}};w_h\right) = w_h L_h + \mathcal{R}_h(\tilde{\boldsymbol{P}}_h;w_h), \tag{25}$$

where $\mathcal{R}_h(.)$ is given by 19. Taking a basic partial derivative from Equation 25, yields

$$\frac{\partial Y_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial \ln \tilde{P}_{hh,k}} = -\sum_g\sum_i\left(Q_{hi,g}\frac{\partial P_{hi,g}}{\partial \ln Q_{hh,g}}\left[\frac{\partial \ln Q_{hh,g}}{\partial \ln \tilde{P}_{hh,k}} + \frac{\partial \ln Q_{hh,g}}{\partial \ln Y_h}\frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}}\right]\right)$$
$$+\tilde{P}_{hh,k}Q_{hh,k} + \sum_g\left((\tilde{P}_{hh,g} - P_{hh,g})Q_{hh,g}\left[\frac{\partial \ln Q_{hh,g}}{\partial \ln \tilde{P}_{hh,k}} + \frac{\partial \ln Q_{hh,g}}{\partial \ln Y_h}\frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}}\right]\right)$$
$$+\sum_g\left((\tilde{P}_{fh,g} - P_{fh,g}[1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}])Q_{fh,g}\left[\frac{\partial \ln Q_{fh,g}}{\partial \ln \tilde{P}_{hh,k}} + \frac{\partial \ln Q_{fh,g}}{\partial \ln Y_h}\frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}}\right]\right)$$

Given $P_{hi,k} = \bar{\rho}_{hi,k} \left(\sum_\iota \tau_{h\iota,k} Q_{h\iota,k}\right)^{-\frac{\psi_k}{1+\psi_k}}$, we can simplify the first line in the above equation as

$$-\sum_g \sum_i \left(Q_{hi,g} \frac{\partial P_{hi,g}}{\partial \ln Q_{hh,g}} \frac{\partial \ln Q_{hh,g}}{\partial \ln \tilde{P}_{hh,k}} \right) = \sum_g \left(\frac{\psi_g}{1+\psi_g} P_{hh,g} Q_{hh,g} \frac{\partial \ln Q_{hh,g}}{\partial \ln \tilde{P}_{hh,k}} \right)$$

Plugging the above equation back into the expression for $\partial Y_h / \partial \ln \tilde{P}_{hh,k}$, yields

$$\begin{aligned}
\frac{\partial Y_h \left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln \tilde{P}_{hh,k}} =& \sum_g \left(\left[\frac{\psi_g}{1+\psi_g} + \left(\frac{\tilde{P}_{hh,g}}{P_{hh,g}} - 1\right) \right] P_{hh,g} Q_{hh,g} \varepsilon^{hh,k}_{hh,g} \right) + \\
&+ \tilde{P}_{hh,k} Q_{hh,k} + \sum_g \left[(\tilde{P}_{fh,g} - P_{fh,g} [1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}]) Q_{fh,g} \varepsilon^{hh,k}_{fh,g} \right] \\
&+ \sum_g \left((\tilde{P}_{fh,g} - P_{fh,g}) Q_{fh,g} \eta_{fh,g} + \left[\tilde{P}_{hh,g} - \frac{\psi_g}{1+\psi_g} P_{hh,g} \right] Q_{hh,g} \eta_{hh,g} \right) \frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hh,k}}.
\end{aligned}$$

where $\eta_{fh,g} \equiv \partial \ln Q_{fh,g} / \partial \ln Y_h$ denotes the income elasticity of demand. Similarly, $\mathrm{d} w_h / \mathrm{d} \tilde{P}_{hh,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition,

$$D_h \left(\tilde{\boldsymbol{P}}; w_h\right) \equiv \sum_k \left(P_{fh,k} Q_{fh,k} - \tilde{P}_{hf,k} Q_{hf,k} \right) = 0,$$

which yields the following:

$$\begin{aligned}
\frac{\mathrm{d} \ln w_h}{\mathrm{d} \ln \tilde{P}_{hh,k}} &= \left[\frac{\partial}{\partial \ln \tilde{P}_{hh,k}} \sum_g \left(P_{fh,g} Q_{fh,g} - \tilde{P}_{hf,g} Q_{hf,g} \right) \right] \Big/ \frac{\partial D_h \left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln w_h} \\
&= \left[\sum_g \left(P_{fh,g} Q_{fh,g} [1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}] \frac{\partial \ln Q_{fh,g}}{\partial \ln \tilde{P}_{hh,k}} \right) + \sum_g \left(P_{fh,g} Q_{fh,g} \eta_{fh,g} \right) \frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hh,k}} \right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h} \\
&= \left[\sum_g \left(P_{fh,g} Q_{fh,g} [1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}] \varepsilon^{hh,k}_{fh,g} \right) + \sum_g \left(P_{fh,g} Q_{fh,g} \eta_{fh,g} \right) \frac{\partial Y_h}{\partial \ln \tilde{P}_{hh,k}} \right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h}.
\end{aligned}$$

Plugging the expressions for $\partial Y_h / \partial \ln \tilde{P}_{hh,k}$ and $\mathrm{d} w_h / \mathrm{d} \ln \tilde{P}_{hh,k}$ back into the F.O.C., defining $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial Y_h} \frac{\partial Y_h}{\partial w_h} + \sum_i \sum_g \frac{\partial V_h}{\partial \tilde{P}_{ih,g}} \frac{\partial \tilde{P}_{ih,g}}{\partial \ln w_h} \right) / \left(\partial \ln D_h / \partial w_h \right)$, we can write the F.O.C. (Equation 24) as

$$\begin{aligned}
&\sum_g \left(\left[1 - \frac{1}{1+\psi_g} \frac{\tilde{P}_{hh,g}}{P_{hh,g}} \right] \tilde{P}_{hh,g} Q_{hh,g} \varepsilon^{hh,k}_{hh,g} \right) + \sum_g \left[\left(1 - (1+\bar{\tau}) [1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}] \frac{P_{fh,g}}{\tilde{P}_{fh,g}} \right) \tilde{P}_{fh,k} Q_{fh,k} \varepsilon^{hh,k}_{fh,g} \right] \\
+ &\sum_g \left([1 - (1+\bar{\tau}) \frac{P_{fh,g}}{\tilde{P}_{fh,g}}] \tilde{P}_{fh,g} Q_{fh,g} \eta_{fh,g} + \left[1 - \frac{1}{1+\psi_g} \frac{\tilde{P}_{hh,g}}{P_{hh,g}} \right] \tilde{P}_{hh,g} Q_{hh,g} \eta_{hh,g} \right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0
\end{aligned}$$

Dividing the above equation by Y_h and noting that $\hat{\lambda}_{ji,k} \equiv \tilde{P}_{ji,k}Q_{ji,k}/Y_i$, $1+s_k \equiv \tilde{P}_{hh,k}/P_{hh,k}$, $1+t_k \equiv \tilde{P}_{fh,k}/P_{fh,k}$, and

$$\frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}} = \frac{\partial \ln \left(\sum_{\iota} \tau_{f\iota,g} Q_{f\iota,g}\right)^{-\frac{\psi_g}{1+\psi_g}}}{\partial \ln Q_{fh,g}} = -\frac{\psi_g}{1+\psi_g} r_{fh,g}$$

yields the following simplified expression:

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\psi_g}\right] \hat{\lambda}_{hh,g} \varepsilon^{hh,k}_{hh,g} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g} \left[\frac{1+\psi_g r_{ff,g}}{1+\psi_g}\right]\right) \hat{\lambda}_{fh,g} \varepsilon^{hh,k}_{fh,g} \right]$$
$$+ \sum_g \left(\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \hat{\lambda}_{fh,g} \eta_{fh,g} + \left[\frac{1+s_g}{1+\psi_g} - 1\right] \hat{\lambda}_{hh,g} \eta_{hh,g} \right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0$$

We can further simplify the above expression, we can rely on the observations that (i) there are multiple optimal tax combinations per Lemma 1, (ii) the term in parenthesis in the second line of the above equation, namely,

$$\Psi\left(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}\right) \equiv \sum_g \left(\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right) \hat{\lambda}_{fh,g} \eta_{fh,g} + \left[\frac{1+s_g}{1+\psi_g} - 1\right] \hat{\lambda}_{hh,g} \eta_{hh,g} \right),$$

is not industry-specific. Hence, by choice of $\bar{s}$ in Lemma 1, there exists an optimal tax combination for which $\Psi\left(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*\right) = 0$. We can first restrict attention to this particular solution, and then trivially identify the other solutions with an across-the-board shift in all tax instruments. Doing so, the F.O.C. with respect to $\tilde{P}_{hh,k}$ reduces to

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\psi_g}\right] \hat{\lambda}_{hh,g} \varepsilon^{hh,k}_{hh,g} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g} \left[\frac{1+\psi_g r_{ff,g}}{1+\psi_g}\right]\right) \hat{\lambda}_{fh,g} \varepsilon^{hh,k}_{fh,g} \right] = 0 \quad (26)$$

F.O.C. with respect to $\tilde{P}_{fh,k}$

The F.O.C. with respect to $\tilde{P}_{fh,k}$ also features terms accounting for direct effects and general equilibrium wage effects:

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}}; w_h)}{\mathrm{d}\tilde{P}_{fh,k}} = \frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial \tilde{P}_{fh,k}} + \frac{\partial V_h(.)}{\partial \tilde{P}_{fh,k}} + \sum_g \frac{\partial W_h}{\partial \tilde{P}_{ff,g}} \frac{\partial \tilde{P}_{ff,g}}{\partial \tilde{P}_{fh,k}}$$
$$+ \left(\frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f} \sum_g \left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}} \frac{\partial \tilde{P}_{ih,g}}{\partial w_h} \right] \right) \frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{fh,k}} = 0.$$

Noting that $\partial W_h / \partial \tilde{P}_{ff,g} = 0$ and appealing to Roy's identity that $\frac{\partial V_h / \partial \tilde{P}_{ih,k}}{\partial V_h / \partial Y_h} = -Q_{ih,k}$, we can simplify the above equation as follows:

$$\frac{\mathrm{d} W_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\mathrm{d} \ln \tilde{P}_{fh,k}} = \frac{\partial V_h}{\partial Y_h}\left\{\frac{\partial Y_h}{\partial \ln \tilde{P}_{fh,k}} - \tilde{P}_{fh,k} Q_{fh,k} + \left(\frac{\partial Y_h}{\partial \ln w_h} - \sum_{i=h,f}\sum_{g}\left[\tilde{P}_{ih,g} Q_{ih,g} \frac{\partial \ln \tilde{P}_{ih,g}}{\partial \ln w_h}\right]\right)\frac{\mathrm{d} \ln w_h}{\mathrm{d} \ln \tilde{P}_{fh,k}}\right\} = 0. \qquad (27)$$

Taking a basic partial derivative from Equation 25, yields

$$\frac{\partial Y_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln \tilde{P}_{fh,k}} = \sum_g \left(\left[\frac{\psi_g}{1+\psi_g} + (\frac{\tilde{P}_{hh,g}}{P_{hh,g}} - 1)\right] P_{hh,g} Q_{hh,g} \varepsilon^{fh,k}_{hh,g}\right)$$
$$\tilde{P}_{fh,k} Q_{hh,k} + \sum_g \left[(\tilde{P}_{fh,g} - P_{fh,g}[1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}]) Q_{fh,g} \varepsilon^{fh,k}_{fh,g}\right]$$
$$+ \sum_g \left((\tilde{P}_{fh,g} - P_{fh,g}) Q_{fh,g} \eta_{fh,g} + \left[\tilde{P}_{hh,g} - \frac{\psi_g}{1+\psi_g} P_{hh,g}\right] Q_{hh,g} \eta_{hh,g}\right) \frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hh,k}}.$$

where $\eta_{fh,g} \equiv \partial \ln Q_{fh,g} / \partial \ln Y_h$ denotes the income elasticity of demand. Similarly, $\mathrm{d} w_h / \mathrm{d} \tilde{P}_{hh,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition, which yields the following:

$$\frac{\mathrm{d} \ln w_h}{\mathrm{d} \ln \tilde{P}_{fh,k}} = \left[\frac{\partial}{\partial \ln \tilde{P}_{fh,k}} \sum_g \left(P_{fh,g} Q_{fh,g} - \tilde{P}_{hf,g} Q_{hf,g}\right)\right] \Big/ \frac{\partial D_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln w_h}$$
$$= \left[\sum_g \left(P_{fh,g} Q_{fh,g}\left[1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}\right] \varepsilon^{fh,k}_{fh,g}\right) + \sum_g \left(P_{fh,g} Q_{fh,g} \eta_{fh,g}\right) \frac{\partial Y_h}{\partial \ln \tilde{P}_{fh,k}}\right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h}.$$

Plugging the expressions for $\partial Y_h / \partial \ln \tilde{P}_{hh,k}$ and $\mathrm{d} w_h / \mathrm{d}\ln \tilde{P}_{hh,k}$ back into the F.O.C., defining $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial w_h} + \sum_i \sum_g \frac{\partial V_h}{\partial \tilde{P}_{ih,g}} \frac{\partial \tilde{P}_{ih,g}}{\partial \ln w_h}\right) / \left(\partial \ln D_h / \partial w_h\right)$, we can write the F.O.C. (Equation 27) as

$$\sum_g \left(\left[1 - \frac{1}{1+\psi_g}\frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right] \tilde{P}_{hh,g} Q_{hh,g} \varepsilon^{fh,k}_{hh,g}\right) + \sum_g \left[\left(1 - (1+\bar{\tau})\left[1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}\right]\frac{P_{fh,g}}{\tilde{P}_{fh,g}}\right) \tilde{P}_{fh,g} Q_{fh,g} \varepsilon^{fh,k}_{fh,g}\right]$$
$$+ \sum_g \left([1 - (1+\bar{\tau})\frac{P_{fh,g}}{\tilde{P}_{fh,g}}] \tilde{P}_{fh,g} Q_{fh,g} \eta_{fh,g} + \left[1 - \frac{1}{1+\psi_g}\frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right] \tilde{P}_{hh,g} Q_{hh,g} \eta_{hh,g}\right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0$$

Dividing the above equation by Y_h and noting that $\acute{\lambda}_{ji,k} \equiv \tilde{P}_{ji,k} Q_{ji,k} / Y_i$, $1 + s_k \equiv \tilde{P}_{hh,k} / P_{hh,k}$, and $1 + t_k \equiv \tilde{P}_{fh,k} / P_{fh,k}$, yields the following simplified expression:

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\psi_g}\right] \acute{\lambda}_{hh,g} \varepsilon^{fh,k}_{hh,g}\right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\left[\frac{1+\psi_g r_{ff,g}}{1+\psi_g}\right]\right) \acute{\lambda}_{fh,g} \varepsilon^{fh,k}_{fh,g}\right] + \Psi\left(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}\right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0$$

As noted earlier, there exists an optimal tax combination for which $\Psi\left(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*\right) = 0$. Restricting attention to this particular solution, the F.O.C. with respect to $\tilde{P}_{fh,k}$ reduces to

$$\sum_g \left(\left[1 - \frac{1+s_g}{1+\psi_g}\right] \acute{\lambda}_{hh,g} \varepsilon_{hh,g}^{fh,k}\right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\left[\frac{1+\psi_g r_{ff,g}}{1+\psi_g}\right]\right) \acute{\lambda}_{fh,g} \varepsilon_{fh,g}^{fh,k}\right] = 0 \quad (28)$$

F.O.C. with respect to $\tilde{P}_{hf,k}$

Noting that $\partial W_h / \partial \tilde{P}_{ff,g} = 0$, the F.O.C. with respect to $\tilde{P}_{hf,k}$ can be stated as

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}}; w_h)}{\mathrm{d}\tilde{P}_{hf,k}} = \frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial \tilde{P}_{hf,k}} + \left(\frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f} \sum_g \left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}} \frac{\partial \tilde{P}_{ih,g}}{\partial w_h}\right]\right) \frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{hf,k}} = 0$$

As before, we can derive $\partial Y_h / \partial \ln \tilde{P}_{hf,k}$ by taking a derivative from Equation 25, yielding the following

$$\begin{aligned}
\frac{\partial Y_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln \tilde{P}_{hf,k}} = & -\sum_g \sum_i \left[P_{hi,g} Q_{hf,g} \frac{\partial \ln P_{hh,g}}{\partial \ln Q_{hf,g}} \left(\frac{\partial \ln Q_{hh,g}}{\partial \ln \tilde{P}_{hf,k}} + \frac{\partial \ln Q_{hh,g}}{\partial \ln Y_h} \frac{\partial \ln Y_h}{\ln \tilde{P}_{hf,k}}\right)\right] \\
& + \tilde{P}_{hf,k} Q_{hf,k} + \sum_g \left((\tilde{P}_{hf,g} - P_{hf,g}) Q_{hf,g} \frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}}\right) - \sum_g \left(P_{fh,g} Q_{fh,g} \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{ff,g}} \frac{\partial \ln Q_{ff,g}}{\partial \ln \tilde{P}_{hf,k}}\right) \\
& + \sum_g \left((\tilde{P}_{hh,g} - P_{hh,g}) Q_{hh,g} \frac{\partial \ln Q_{hh,g}}{\partial \ln Y_h} + (\tilde{P}_{fh,g} - P_{fh,g}) Q_{fh,g} \frac{\partial \ln Q_{fh,g}}{\partial \ln Y_h}\right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hf,k}}.
\end{aligned}$$

To simplify the last sum in line 2, we can use the fact that $P_{fh,g} Q_{fh,g} \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{ff,g}} = P_{ff,g} Q_{ff,g} \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}$ as well as the following well-known result from consumer theory:

$$\sum_g P_{ff,g} Q_{ff,g} \frac{\partial \ln Q_{ff,g}}{\partial \ln \tilde{P}_{hf,k}} = -\tilde{P}_{hf,k} Q_{hf,k} - \sum_g \tilde{P}_{hf,g} Q_{hf,g} \frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}}.$$

Plugging these equations back into the expression for $\partial Y_h / \partial \ln \tilde{P}_{hf,k}$ yields the following:

$$\begin{aligned}
\frac{\partial Y_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln \tilde{P}_{hf,k}} = & \sum_g \left(\left[\frac{\psi_g}{1+\psi_g} + (\frac{\tilde{P}_{hf,g}}{P_{hf,g}} - 1)\right] P_{hf,g} Q_{hf,g} \varepsilon_{hf,g}^{hf,k}\right) \\
& + \tilde{P}_{hf,k} Q_{hf,k} (1 + \frac{\partial \ln P_{fh,k}}{\partial \ln Q_{fh,k}}) + \sum_g \left[(\tilde{P}_{hf,g}(1 + \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}) - P_{hf,g}) Q_{hf,g} \varepsilon_{hf,g}^{hf,k}\right] \\
& + \sum_g \left((\tilde{P}_{fh,g} - P_{fh,g}) Q_{fh,g} \eta_{fh,g} + \left[\frac{\psi_g}{1+\psi_g} P_{hh,g} + (\tilde{P}_{hh,g} - P_{hh,g})\right] Q_{hh,g} \eta_{hh,g}\right) \frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hf,k}}.
\end{aligned}$$

Similarly, $\mathrm{d}w_h/\mathrm{d}\tilde{P}_{hf,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition, $D_h\left(\tilde{\boldsymbol{P}};w_h\right)$, which yields the following:

$$\begin{aligned}
\frac{\mathrm{d}\ln w_h}{\mathrm{dln}\tilde{P}_{hf,k}} &= \left[\frac{\partial}{\partial \ln \tilde{P}_{hf,k}}\sum_g \left(P_{fh,g}Q_{fh,g} - \tilde{P}_{hf,g}Q_{hf,g}\right)\right] \Big/ \frac{\partial D_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial \ln w_h} \\
&= \left[\sum_g \left(P_{fh,g}Q_{fh,g}\left[\frac{\partial \ln P_{fh,g}}{\partial \ln Q_{ff,g}}\frac{\partial \ln Q_{ff,g}}{\partial \ln \tilde{P}_{hf,k}} + \eta_{fh,g}\frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hf,k}}\right]\right)\right. \\
&\quad \left. - \tilde{P}_{hf,k}Q_{hf,k} - \sum_g \left(\tilde{P}_{hf,g}Q_{hf,g}\frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}}\right)\right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h}.
\end{aligned}$$

Simplifying the term $P_{fh,g}Q_{fh,g}\frac{\partial \ln P_{fh,g}}{\partial \ln Q_{ff,g}}\frac{\partial \ln Q_{ff,g}}{\partial \ln \tilde{P}_{hf,k}}$ along the same lines outlined earlier, we can rewrite the expression for $\mathrm{d}\ln w_h/\mathrm{dln}\tilde{P}_{hf,k}$ as

$$\begin{aligned}
\frac{\mathrm{d}\ln w_h}{\mathrm{dln}\tilde{P}_{hf,k}} = &\left[\sum_g \left(P_{fh,g}Q_{fh,g}\eta_{fh,g}\right)\frac{\partial Y_h}{\partial \ln \tilde{P}_{hf,k}}\right. \\
&\left. - \tilde{P}_{hf,k}Q_{hf,k}\left(1+\frac{\partial \ln P_{fh,k}}{\partial \ln Q_{fh,k}}\right) - \sum_g \left(\tilde{P}_{hf,g}Q_{hf,g}\left(1+\frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}\right)\varepsilon_{hf,g}^{hf,k}\right)\right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h}.
\end{aligned}$$

Plugging the expressions for $\partial Y_h/\partial \ln \tilde{P}_{hf,k}$ and $\mathrm{d}w_h/\mathrm{dln}\tilde{P}_{hf,k}$ back into the F.O.C.; adopting our earlier definition $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial \ln w_h} + \sum_i \sum_g \frac{\partial V_h}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial \ln w_h}\right) / \left(\partial D_h/\partial \ln w_h\right)$; focusing in the solution for which $\Psi\left(\boldsymbol{t}^*,\boldsymbol{x}^*,\boldsymbol{s}^*\right) = 0$; and noting that $1+\frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}} = \frac{1+\psi_g r_{ff,g}}{1+\psi_g}$; we can write the F.O.C. as

$$\left(\frac{1+\psi_k r_{ff,k}}{1+\psi_k}\right)\acute{\lambda}_{hf,k} + \sum_g \left[\frac{1+\psi_g r_{ff,g}}{1+\psi_g} - \frac{1}{(1+x_g)(1+\bar{\tau})(1+s_g)(1+\mu_g)}\right]\acute{\lambda}_{hf,g}\varepsilon_{hf,g}^{hf,k} = 0, \tag{29}$$

where in deriving the above expression, we use the fact $\tilde{P}_{hf,k}/P_{hf,k} = (1+x_k)(1+s_k)$, and $\acute{\lambda}_{hf,g} \equiv \tilde{P}_{hf,g}Q_{hf,g}/Y_h$.

Simultaneously Solving the System of F.O.C.

Combining Equations 26, 28, and 29 yields the following system of F.O.C.:

$$\begin{aligned}
&[\tilde{P}_{hh,k}] \quad \sum_g \left(\left[1-\frac{1+s_g}{1+\mu_g}\right]\acute{\lambda}_{hh,g}\varepsilon_{hh,g}^{hh,k}\right) + \sum_g \left[\left(1-\frac{1+\bar{\tau}}{1+t_k}\right)\acute{\lambda}_{fh,g}\varepsilon_{fh,g}^{hh,k}\right] = 0 \quad [\tilde{P}_{hh,k}] \\
&[\tilde{P}_{fh,k}] \quad \sum_g \left(\left[1-\frac{1+s_g}{1+\mu_g}\right]\acute{\lambda}_{hh,g}\varepsilon_{hh,g}^{fh,k}\right) + \sum_g \left[\left(1-\frac{1+\bar{\tau}}{1+t_k}\left(\frac{1+\psi_g r_{ff,g}}{1+\psi_g}\right)\right)\acute{\lambda}_{fh,g}\varepsilon_{fh,g}^{fh,k}\right] = 0 \\
&[\tilde{P}_{hf,k}] \quad \acute{\lambda}_{hf,k} + \sum_g \left[1-\frac{1}{(1+x_k)(1+\bar{\tau})(1+s_g)(1+\psi_g)}\left(\frac{1+\psi_g}{1+\psi_g r_{ff,g}}\right)\right]\acute{\lambda}_{hf,g}\varepsilon_{hf,g}^{hf,k}\frac{(1+\psi_g)(1+\psi_k r_{ff,k})}{(1+\psi_k)(1+\psi_g r_{ff,g})} = 0
\end{aligned}$$

Solving the above system implies that $1+s_g = 1+\mu_g$ and $1+t_g = 1+\bar{\tau}$. The term $\bar{\tau}$ is, however, redundant due to multiplicity of the optimal policy up to two uniform tax shifters (Lemma 1). So, altogether, for any $\bar{s}$ and $\bar{t} \in \mathbb{R}_+$, the following formula characterizes an optimal tax combination:

$$\begin{cases} (1+s_k)(1+\bar{s}) = 1+\mu_k & \forall k \in \mathbb{K} \\ 1+t_k = (1+\bar{t})\left(\frac{1+\psi_k r_{ff,k}}{1+\psi_k}\right) & \forall k \in \mathbb{K}\,, \\ \left[\left(\frac{1+\psi_k}{1+\psi_k r_{ff,k}}\right) / (1+x_k)(1+\bar{t})\right]_k = \boldsymbol{\mathcal{E}}_{hf}^{-1}\left(\boldsymbol{I}+\boldsymbol{\mathcal{E}}_{hf}\right)\mathbf{1} \end{cases}$$

where $\boldsymbol{\mathcal{E}}_{hf} \equiv \left[\frac{(1+\psi_g)(1+\psi_k r_{ff,k})}{(1+\psi_k)(1+\psi_g r_{ff,g})}\frac{\lambda_{hf,g}}{\lambda_{hf,k}}\varepsilon_{hf,g}^{hf,k}\right]_{g,k}$ is a $K \times K$ matrix composed of reduced-form demand elasticities and expenditure shares. Under Homothetic preferences, $\frac{\lambda_{hf,g}}{\lambda_{hf,k}}\varepsilon_{hf,g}^{hf,k} = \varepsilon_{hf,k}^{hf,g}$, which implies that $\boldsymbol{\mathcal{E}}_{hf} \equiv \left[\frac{(1+\psi_g)(1+\psi_k r_{ff,k})}{(1+\psi_k)(1+\psi_g r_{ff,g})}\varepsilon_{hf,k}^{hf,g}\right]_{g,k}$.

G Proof of Theorem 2

Without domestic taxes the F.O.C. with respect to trade taxes x_k and t_k (Equations 20 and 23) can be stated as follows:

$$\sum_g\left[\left(1-\frac{1}{1+\mu_g}\right)\tilde{P}_{hh,g}Q_{hh,g}\varepsilon_{hh,g}^{fh,k}\right]+\sum_g\left[\left(1-\frac{1+\bar{\tau}}{1+t_g}\right)\tilde{P}_{fh,g}Q_{fh,g}\varepsilon_{fh,g}^{fh,k}\right]=0$$

$$\tilde{P}_{hf,k}Q_{hf,k}+\sum_g\left(\left[1-\frac{1}{(1+\mu_g)(1+x_g)(1+\bar{\tau})}\right]\tilde{P}_{hf,g}Q_{hf,g}\varepsilon_{hf,g}^{hf,k}\right)=0$$

Given that $\acute{\lambda}_{ji,g} = \tilde{P}_{ji,g}Q_{ji,g}/Y_i$ and $\sum_g\sum_j \acute{\lambda}_{jh,g}\varepsilon_{jh,g}^{fh,k} = -\acute{\lambda}_{fh,k}$, the first equation can be simplified as

$$\sum_g\left[\frac{1+\bar{\tau}}{1+t_g}\frac{\acute{\lambda}_{fh,g}}{\acute{\lambda}_{fh,k}}\varepsilon_{fh,g}^{fh,k}\right] = -1-\left(\sum_g\frac{1}{1+\mu_g}\frac{\acute{\lambda}_{hh,g}}{\acute{\lambda}_{fh,k}}\varepsilon_{hh,g}^{fh,k}\right).$$

Writing the above equation in matrix form; combining it withe F.O.C. for the export tax; and accounting for the multiplicity of the trade tax, implies the following optimal tax formula for any $\bar{t} \in \mathbb{R}_+$:

$$\left[\frac{1+t_k^*}{1+\bar{t}}\right]_k = -\boldsymbol{\mathcal{E}}_{fh}^{-1}\left(\boldsymbol{I}+\boldsymbol{\mathcal{E}}_{hh}^{fh}\boldsymbol{m}\right)$$

$$\left[\frac{1}{(1+x_k^*)(1+\bar{t})(1+\mu_k)}\right]_k = \boldsymbol{\mathcal{E}}_{hf}^{-1}\left(\boldsymbol{I}+\boldsymbol{\mathcal{E}}_{hf}\right)\mathbf{1}$$

where $\boldsymbol{\mathcal{E}}_{fh} \equiv \left[\frac{\acute{\lambda}_{fh,g}}{\acute{\lambda}_{fh,k}}\varepsilon_{fh,g}^{fh,k}\right]_{g,k}$, $\boldsymbol{\mathcal{E}}_{hh}^{fh} \equiv \left[\frac{\acute{\lambda}_{hh,g}}{\acute{\lambda}_{hf,k}}\varepsilon_{hh,g}^{fh,k}\right]_{g,k}$, and $\boldsymbol{m} \equiv \left[\frac{1}{1+\mu_k}\right]$. Finally, assuming zero cross-substitutability between industries, i.e., $\varepsilon_{fh,g}^{fh,k} = 0$ and $\varepsilon_{hh,g}^{fh,k}$ if $g \neq k$, the above formula reduces to:

$$1 + t_k^* = \frac{(1+\mu_k)\left[1 + \epsilon_k \lambda_{hh,k}\right]}{1 + \mu_k + \epsilon_k \lambda_{hh,k}}(1+\bar{t});$$
$$1 + x_k^* = \frac{1 + 1/\epsilon_k \lambda_{hh,k}}{1+\mu_k}(1+\bar{t})^{-1}.$$

H Proof of Theorem 3

Following Section E, when $s_k = x_k = 0$, the F.O.C. with respect to t_k (Equation 20) can be stated as

$$\sum_g \left[\frac{1+\bar{\tau}}{1+t_g}\frac{\acute{\lambda}_{fh,g}}{\acute{\lambda}_{fh,k}}\varepsilon_{fh,g}^{fh,kQ}\right] = -1 - \left(\sum_g \frac{1}{1+\mu_g}\frac{\acute{\lambda}_{hh,g}}{\acute{\lambda}_{fh,k}}\varepsilon_{hh,g}^{fh,k}\right).$$

Rearranging the above expression while noting that $\varepsilon_{fh,g}^{fh,k} = -1 - \epsilon_k \lambda_{hh,k}$ if $g = k$ and $\varepsilon_{fh,g}^{fh,k} = 0$ if $g \neq k$ implies the following optimal tax formula:

$$1 + t_k^* = \frac{(1+\mu_k)\left[1 + \epsilon_k \lambda_{hh,k}\right]}{1 + \mu_k + \epsilon_k \lambda_{hh,k}}(1+\bar{\tau}).$$

However, since the use of export taxes is restricted , $\bar{\tau}$ is no longer redundant and should be formally characterized. To so so, we can appeal to the definition of $\bar{\tau}$ as follows:

$$\bar{\tau} = -\frac{\partial W_h/\partial w_h}{\partial D_h/\partial w_h} = -\frac{\partial W_h/\partial w_f}{\partial D_h/\partial w_f} = \frac{\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial w_f} + \sum_i \sum_g \frac{\partial V_h}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial w_f}}{\frac{\partial}{\partial w_h}\sum_k \left[P_{hf,k}Q_{hf,k} - P_{fh,k}Q_{fh,k}\right]}.$$

Given that $\partial V_h(.)/\partial \tilde{P}_{ih,g} = -Q_{ih,g}$, $\partial \ln P_{fi,g}/\partial \ln w_f$ =1, and $Y_h = w_h L_h + \Pi_h$, we can rewrite the above expression as

$$\bar{\tau} = \frac{\sum_g \sum_i \left(\frac{\mu_g}{1+\mu_g}P_{hi,g}Q_{hi,g}\varepsilon_{hi,g}^{fi,g}\right) + \sum_k \left(t_g P_{fh,g}Q_{fh,g}\varepsilon_{fh,g}^{hh,k}\right) - \sum_k \tilde{P}_{fh,k}Q_{fh,k}}{\sum_k \left[P_{fh,k}Q_{fh,k}(1+\varepsilon_{fh,k}) - P_{hf,k}Q_{hf,k}(\varepsilon_{hf,k}^{ff,k} + \gamma_{hf,k})\right]}.$$

To simplify the above expression we can appeal to the F.O.C., which states that

$$\sum (t_k - \bar{\tau})\, P_{fh,k}Q_{fh,k}\varepsilon_{fh,k} = -\sum_g \left[\frac{\mu_g}{1+\mu_g}\tilde{P}_{hh,g}Q_{hh,g}\varepsilon_{hh,g}^{fh,k}\right].$$

Plugging the above expression back into Equation and noting that $\varepsilon_{hi,g}^{fi,g} = \epsilon_k \lambda_{fi,g}$, $\gamma_{hf,k} = 1$, and $\sum_k P_{fh,k}Q_{fh,k} = \sum P_{hf,k}Q_{hf,k}$, yields the following expression for $\bar{\tau}$:

$$\bar{\tau} = \frac{\sum_g \sum_i \left(\frac{\mu_g}{1+\mu_g} P_{hi,g} Q_{hi,g} \epsilon_g \lambda_{fi,g} \right) - \sum_g \left(\frac{\mu_g}{1+\mu_g} \tilde{P}_{hh,g} Q_{hh,g} \epsilon_g \lambda_{fh,g} \right) - \sum_k \left(P_{hf,k} Q_{hf,k} \right)}{- \sum_k P_{hf,k} Q_{hf,k} \epsilon_k \lambda_{ff,k}}.$$

Plugging the above equation back in the optimal tax formula and noting that $\frac{P_{hf,k}Q_{hf,k}}{\sum_g P_{hf,g}Q_{hf,g}} = \frac{\lambda_{hf,k}}{\lambda_{hf}}$ yields the sufficient statistics tax formulas specified under Theorem 3:

$$1 + t_k^* = \frac{(1+\mu_k)\left[1 + \epsilon_k \lambda_{hh,k}\right]}{1 + \mu_k + \epsilon_k \lambda_{hh,k}} \left(\frac{1 + \sum_g \left[\frac{1}{1+\mu_g} \frac{\lambda_{hf,g}}{\lambda_{hf}} \epsilon_g \lambda_{ff,g} \right]}{\sum_g \frac{\lambda_{hf,g}}{\lambda_{hf}} \epsilon_g \lambda_{ff,g}} \right).$$

I Optimal Tax Formula with IO Linkages

We can appeal to the theorem in Kopczuk (2003), whereby one can first eliminate the underlying market inefficiency with the appropriate domestic tax and then solve for optimal taxes as if markets where efficient. The optimal policy will then be the sum of corrective taxes and taxes corresponding to an efficient market. In the present context, we apply this result by solving for optimal taxes while abstracting from profits, i.e., $Y_h(\tilde{\boldsymbol{P}}; w_h) = w_h L_h + \mathcal{R}_h(\tilde{\boldsymbol{P}}_h; w_h)$. After solving for the optimal taxes under efficient markets, we amend them with corrective Pigouvian taxes that eliminate markup heterogeneity.

To streamline the notation, we let $\boldsymbol{\alpha} \equiv \left[\alpha_{ji,k}^{\jmath\iota,g} \right]_{\jmath g, jik}$ denote the full $4K \times 4K$ *"cost-based"* global IO matrix, with element $\alpha_{ji,k}^{\jmath\iota,g}$ corresponding to the share of intermediate input $\jmath\iota, g$ in the total cost of output variety ji, k. Then, by Shepherd's Lemma:

$$\frac{\partial \ln P_{\jmath\iota,g}(\tilde{\boldsymbol{P}}; w_h)}{\partial \ln \tilde{P}_{ji,k}^{\mathcal{I}}} = \frac{\tilde{P}_{ji,k}^{\mathcal{I}} Q_{ji,k}^{\jmath\iota,g}}{P_{\jmath\iota,g} Q_{\jmath\iota,g}}, \quad \forall \jmath, j, i \in \mathbb{C}; \;\; g, k \in \mathbb{K}$$

where $\alpha_{\jmath\iota,g}^{ji,k} \equiv \tilde{P}_{ji,k}^{\mathcal{I}} Q_{ji,k}^{\jmath\iota,g} / P_{\jmath\iota,g} Q_{\jmath\iota,g}$ denotes the share of intermediate input ji, k in the total production cost of output variety $\jmath\iota, g$ ($Q_{ji,k}^{\jmath\iota,g}$ denotes the amount of ji, k-type inputs used in the production of $\jmath\iota, g$).

F.O.C. with respect to $\tilde{P}_{hh,k}$

The F.O.C. with respect to $\tilde{P}_{hh,k}$ features terms accounting for direct effects and general equilibrium wage effects:

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}}; w_h)}{\mathrm{d}\tilde{P}_{hh,k}} = \frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial \tilde{P}_{hh,k}} + \frac{\partial V_h(.)}{\partial \tilde{P}_{hh,k}} + \sum_g \frac{\partial W_h}{\partial \tilde{P}_{ff,g}} \frac{\partial \tilde{P}_{ff,g}}{\partial \tilde{P}_{hh,k}}$$
$$+ \left(\frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f} \sum_g \left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}} \frac{\partial \tilde{P}_{ih,g}}{\partial w_h} \right] \right) \frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{hh,k}} = 0.$$

where $\tilde{\boldsymbol{P}} \equiv \{\tilde{P}_{ji,k}\}$. The above problem is greatly simplified by the observation that $\partial \tilde{P}_{ff,g}/\partial \tilde{P}_{hh,k} = 0$—i.e., holding $\tilde{P}_{fh,k}$, $\tilde{P}_{hf,k}$, and w_h fixed, $\tilde{P}_{hh,k}$ has no effect on $\tilde{P}_{ff,k}$. Appealing to Roy's identity that $\frac{\partial V_h/\partial \tilde{P}_{ih,k}}{\partial V_h/\partial Y_h} = -Q^{\mathcal{C}}_{ih,k}$, we can simplify the above equation as follows:

$$\frac{\mathrm{d}W_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\mathrm{d}\ln \tilde{P}_{hh,k}} = \frac{\partial V_h}{\partial Y_h}\left\{\frac{\partial Y_h}{\partial \ln \tilde{P}_{hh,k}} - \tilde{P}_{hh,k}Q^{\mathcal{C}}_{hh,k}\right.$$
$$\left. + \left(\frac{\partial Y_h}{\partial \ln w_h} - \sum_{i=h,f}\sum_{g}\left[\tilde{P}_{ih,g}Q^{\mathcal{C}}_{ih,g}\frac{\partial \ln \tilde{P}_{ih,g}}{\partial \ln w_h}\right]\right)\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln \tilde{P}_{hh,k}}\right\} = 0. \tag{30}$$

To determine $\partial Y_h/\partial \ln \tilde{P}_{hh,k}$, note that total income is the sum of factor income, profits, and tax revenue,

$$Y_h\left(\tilde{\boldsymbol{P}}; w_h\right) = w_h L_h + \mathcal{R}_h(\tilde{\boldsymbol{P}}_h; w_h), \tag{31}$$

where $\mathcal{R}_h(.)$ is given by 19. Taking a basic partial derivative from Equation 31, yields

$$\frac{\partial Y_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln \tilde{P}_{hh,k}} = -\sum_g\sum_i\left(Q_{hi,g}P_{hi,g}\frac{\partial \ln P_{hi,g}}{\partial \ln \tilde{P}^{\mathcal{I}}_{hh,g}}\right) + \tilde{P}_{hh,k}Q_{hh,k}$$
$$+\sum_g\left((\tilde{P}_{hh,g} - P_{hh,g})Q_{hh,g}\left[\frac{\partial \ln Q_{hh,g}}{\partial \ln \tilde{P}_{hh,k}} + \frac{\partial \ln Q_{hh,g}}{\partial \ln Y_h}\frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}}\right]\right)$$
$$+\sum_g\left((\tilde{P}_{fh,g} - P_{fh,g})Q_{fh,g}\left[\frac{\partial \ln Q_{fh,g}}{\partial \ln \tilde{P}_{hh,k}} + \frac{\partial \ln Q_{fh,g}}{\partial \ln Y_h}\frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}}\right]\right)$$

Noting that by Shepherd's Lemma,

$$\frac{\partial \ln P_{\jmath\iota,g}(\tilde{\boldsymbol{P}}; w_h)}{\partial \ln \tilde{P}^{\mathcal{I}}_{ji,k}} = \frac{\tilde{P}^{\mathcal{I}}_{ji,k}Q^{\jmath\iota,g}_{ji,k}}{P_{\jmath\iota,g}Q_{\jmath\iota,g}} = \alpha^{ji,k}_{\jmath\iota,g}, \quad \forall \jmath, j, i \in \mathbb{C}; \;\; g,k \in \mathbb{K}$$

we can simplify the first line in the above equation as

$$-\sum_g\sum_i\left(Q_{hi,g}P_{hi,g}\frac{\partial \ln P_{hi,g}}{\partial \ln \tilde{P}^{\mathcal{I}}_{hh,g}}\right) = -\sum_g\left(\alpha^{hh,k}_{hi,g}Q_{hi,g}P_{hi,g}\right) = -\tilde{P}_{hh,k}Q^{\mathcal{I}}_{hh,k}$$

Plugging the above equation back into the expression for $\partial Y_h/\partial \ln \tilde{P}_{hh,k}$, yields

$$\frac{\partial Y_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial \ln \tilde{P}_{hh,k}} = \tilde{P}^{\mathcal{C}}_{hh,k}Q_{hh,k} + \sum_g\left[(\tilde{P}_{fh,g} - P_{fh,g})Q_{fh,g}\varepsilon^{hh,k}_{fh,g} + (\tilde{P}_{hh,g} - P_{hh,g})Q_{hh,g}\varepsilon^{hh,k}_{hh,g}\right]$$
$$+\sum_g\left((\tilde{P}_{fh,g} - P_{fh,g})Q_{fh,g}\eta_{fh,g} + \left[\tilde{P}_{hh,g} - P_{hh,g}\right]Q_{hh,g}\eta_{hh,g}\right)\frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hh,k}}.$$

where $\eta_{fh,g} \equiv \partial \ln Q_{fh,g}/\partial \ln Y_h$ denotes the income elasticity of demand. Similarly, $\mathrm{d}w_h/\mathrm{d}\tilde{P}_{hh,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition,

$$D_h\left(\tilde{\boldsymbol{P}}; w_h\right) \equiv \sum_k \left(P_{fh,k}Q_{fh,k} - \tilde{P}_{hf,k}Q_{hf,k}\right) = 0,$$

which yields the following:

$$\begin{aligned}\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln\tilde{P}_{hh,k}} &= \left[\frac{\partial}{\partial\ln\tilde{P}_{hh,k}}\sum_g\left(P_{fh,g}Q_{fh,g} - \tilde{P}_{hf,g}Q_{hf,g}\right)\right] \Big/ \frac{\partial D_h\left(\tilde{\boldsymbol{P}}; w_h\right)}{\partial\ln w_h} \\ &= \left[\sum_g\left(P_{fh,g}Q_{fh,g}\varepsilon_{fh,g}^{hh,k}\right) + \sum_g\left(P_{fh,g}Q_{fh,g}\eta_{fh,g}\right)\frac{\partial Y_h}{\partial\ln\tilde{P}_{hh,k}}\right] \Big/ \frac{\partial D_h(.)}{\partial\ln w_h}.\end{aligned}$$

Plugging the expressions for $\partial Y_h/\partial\ln\tilde{P}_{hh,k}$ and $\mathrm{d}w_h/\mathrm{d}\ln\tilde{P}_{hh,k}$ back into the F.O.C., defining $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial w_h} + \sum_i\sum_g \frac{\partial V_h}{\partial\tilde{P}_{ih,g}}\frac{\partial\tilde{P}_{ih,g}}{\partial\ln w_h}\right) / \left(\partial\ln D_h/\partial w_h\right)$, we can write the F.O.C. (Equation 30) as

$$\begin{aligned}&\sum_g\left(\left[1 - \frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right]\tilde{P}_{hh,g}Q_{hh,g}\varepsilon_{hh,g}^{hh,k}\right) + \sum_g\left[\left(1 - (1+\bar{\tau})\frac{P_{fh,k}}{\tilde{P}_{fh,k}}\right)\tilde{P}_{fh,k}Q_{fh,k}\varepsilon_{fh,g}^{hh,k}\right] \\ &+ \sum_g\left([1 - (1+\bar{\tau})\frac{P_{fh,k}}{\tilde{P}_{fh,k}}]\tilde{P}_{fh,g}Q_{fh,g}\eta_{fh,g} + \left[1 - \frac{\tilde{P}_{hh,g}}{P_{hh,g}}\right]\tilde{P}_{hh,g}Q_{hh,g}\eta_{hh,g}\right)\frac{\partial\ln Y_h}{\ln\tilde{P}_{hh,k}} = 0\end{aligned}$$

yields the following simplified expression:

$$\begin{aligned}&\sum_g\left(\left[1 - (1+s_g)\right]\lambda_{hh,g}\varepsilon_{hh,g}^{hh,k}\right) + \sum_g\left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right)\lambda_{fh,g}\varepsilon_{fh,g}^{hh,k}\right] \\ &+ \sum_g\left(\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right)\lambda_{fh,g}\eta_{fh,g} + \left[(1+s_g) - 1\right]\lambda_{hh,g}\eta_{hh,g}\right)\frac{\partial\ln Y_h}{\ln\tilde{P}_{hh,k}} = 0\end{aligned}$$

We can further simplify the above expression, we can rely on the observations that (i) there are multiple optimal tax combinations per Lemma 1, (ii) the term in parenthesis in the second line of the above equation, namely,

$$\Psi\left(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}\right) \equiv \sum_g\left(\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right)\lambda_{fh,g}\eta_{fh,g} + \left[(1+s_g) - 1\right]\lambda_{hh,g}\eta_{hh,g}\right),$$

is not industry-specific. Hence, by choice of $\bar{s}$ in Lemma 1, there exists an optimal tax combination for which $\Psi\left(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*\right) = 0$. We can first restrict attention to this particular solution, and then trivially identify the other solutions with an across-the-board shift in all tax instruments. Doing so, the F.O.C. with respect to $\tilde{P}_{hh,k}$ reduces to

$$\sum_g\left(\left[1 - (1+s_g)\right]\lambda_{hh,g}\varepsilon_{hh,g}^{hh,k}\right) + \sum_g\left[\left(1 - \frac{1+\bar{\tau}}{1+t_g}\right)\lambda_{fh,g}\varepsilon_{fh,g}^{hh,k}\right] \tag{32}$$

F.O.C. with respect to $\tilde{P}_{fh,k}$

The F.O.C. with respect to $\tilde{P}_{fh,k}$ also features terms accounting for direct effects and general equilibrium wage effects:

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}};w_h)}{\mathrm{d}\tilde{P}_{fh,k}} = \frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial \tilde{P}_{fh,k}} + \frac{\partial V_h(.)}{\partial \tilde{P}_{fh,k}} + \sum_g \frac{\partial W_h}{\partial \tilde{P}_{ff,g}}\frac{\partial \tilde{P}_{ff,g}}{\partial \tilde{P}_{fh,k}}$$
$$+\left(\frac{\partial V_h(.)}{\partial Y_h}\frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f}\sum_g\left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial w_h}\right]\right)\frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{fh,k}} = 0.$$

Noting that $\partial\tilde{P}_{ff,g}/\tilde{P}_{fh,g} = 0$ and appealing to Roy's identity that $\frac{\partial V_h/\partial \tilde{P}_{ih,k}}{\partial V_h/\partial Y_h} = -Q^{\mathcal{C}}_{ih,k}$, we can simplify the above equation as follows:

$$\frac{\mathrm{d}W_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\mathrm{d}\ln\tilde{P}_{fh,k}} = \frac{\partial V_h}{\partial Y_h}\left\{\frac{\partial Y_h}{\partial \ln \tilde{P}_{fh,k}} - \tilde{P}_{fh,k}Q^{\mathcal{C}}_{fh,k}\right.$$
$$\left.+\left(\frac{\partial Y_h}{\partial \ln w_h} - \sum_{i=h,f}\sum_g\left[\tilde{P}_{ih,g}Q^{\mathcal{C}}_{ih,g}\frac{\partial \ln \tilde{P}_{ih,g}}{\partial \ln w_h}\right]\right)\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln\tilde{P}_{fh,k}}\right\} = 0. \qquad (33)$$

Taking a basic partial derivative from Equation 31, yields

$$\frac{\partial Y_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial \ln \tilde{P}_{fh,k}} = \sum_g\left(\left[\frac{\tilde{P}_{hh,g}}{P_{hh,g}} - 1\right]P_{hh,g}Q_{hh,g}\varepsilon^{fh,k}_{hh,g}\right)$$
$$+\tilde{P}_{fh,k}Q^{\mathcal{C}}_{hh,k} + \sum_g\left[(\tilde{P}_{fh,g} - P_{fh,g})Q_{fh,g}\varepsilon^{fh,k}_{fh,g}\right]$$
$$+\sum_g\left((\tilde{P}_{fh,g} - P_{fh,g})Q_{fh,g}\eta_{fh,g} + \left[\tilde{P}_{hh,g} - P_{hh,g}\right]Q_{hh,g}\eta_{hh,g}\right)\frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hh,k}}.$$

where $\eta_{fh,g} \equiv \partial \ln Q_{fh,g}/\partial \ln Y_h$ denotes the income elasticity of demand. Similarly, $\mathrm{d}w_h/\mathrm{d}\tilde{P}_{hh,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition:

$$\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln\tilde{P}_{fh,k}} = \left[\frac{\partial}{\partial \ln \tilde{P}_{fh,k}}\sum_g\left(P_{fh,g}Q_{fh,g} - \tilde{P}_{hf,g}Q_{hf,g}\right)\right]\Big/\frac{\partial D_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial \ln w_h}$$
$$= \left[\sum_g\left(P_{fh,g}Q_{fh,g}\varepsilon^{fh,k}_{fh,g}\right) + \sum_g\left(P_{fh,g}Q_{fh,g}\eta_{fh,g}\right)\frac{\partial Y_h}{\partial \ln \tilde{P}_{fh,k}}\right]\Big/\frac{\partial D_h(.)}{\partial \ln w_h}.$$

Plugging the expressions for $\partial Y_h/\partial \ln \tilde{P}_{hh,k}$ and $\mathrm{d}w_h/\mathrm{d}\ln\tilde{P}_{hh,k}$ back into the F.O.C., defining $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial w_h} + \sum_i\sum_g \frac{\partial V_h}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial \ln w_h}\right) / \left(\partial \ln D_h/\partial w_h\right)$, we can write the F.O.C. (Equation

33) as

$$\sum_g \left(\left[1 - \frac{\tilde{P}_{hh,g}}{P_{hh,g}} \right] \tilde{P}_{hh,g} Q_{hh,g} \varepsilon_{hh,g}^{fh,k} \right) + \sum_g \left[\left(1 - (1+\bar{\tau}) \frac{P_{fh,k}}{\tilde{P}_{fh,k}} \right) \tilde{P}_{fh,k} Q_{fh,k} \varepsilon_{fh,g}^{fh,k} \right]$$
$$+ \sum_g \left([1 - (1+\bar{\tau}) \frac{P_{fh,k}}{\tilde{P}_{fh,k}}] \tilde{P}_{fh,g} Q_{fh,g} \eta_{fh,g} + \left[1 - \frac{\tilde{P}_{hh,g}}{P_{hh,g}} \right] \tilde{P}_{hh,g} Q_{hh,g} \eta_{hh,g} \right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0$$

Dividing the above equation by Y_h and noting that $\acute{\lambda}_{ji,k} \equiv \tilde{P}_{ji,k} Q_{ji,k} / Y_i$, $1 + s_k \equiv \tilde{P}_{hh,k} / P_{hh,k}$, and $1 + t_k \equiv \tilde{P}_{fh,k} / P_{fh,k}$, yields the following simplified expression:

$$\sum_g \left([1 - (1+s_g)] \, \acute{\lambda}_{hh,g} \varepsilon_{hh,g}^{fh,k} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g} \right) \acute{\lambda}_{fh,g} \varepsilon_{fh,g}^{fh,k} \right] + \Psi(\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{s}) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hh,k}} = 0$$

As noted earlier, there exists an optimal tax combination for which $\Psi(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*) = 0$. Restricting attention to this particular solution, the F.O.C. with respect to $\tilde{P}_{fh,k}$ reduces to

$$\sum_g \left([1 - (1+s_g)] \, \acute{\lambda}_{hh,g} \varepsilon_{hh,g}^{fh,k} \right) + \sum_g \left[\left(1 - \frac{1+\bar{\tau}}{1+t_g} \left[\frac{1+\psi_g r_{ff,g}}{1+\psi_g} \right] \right) \acute{\lambda}_{fh,g} \varepsilon_{fh,g}^{fh,k} \right] = 0 \quad (34)$$

F.O.C. with respect to $\tilde{P}_{hf,k}$

Noting that $\partial W_h / \partial \tilde{P}_{ff,g} = 0$, the F.O.C. with respect to $\tilde{P}_{hf,k}$ can be stated as

$$\frac{\mathrm{d}W_h(\tilde{\boldsymbol{P}}; w_h)}{\mathrm{d}\tilde{P}_{hf,k}} = \frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial \tilde{P}_{hf,k}} + \left(\frac{\partial V_h(.)}{\partial Y_h} \frac{\partial Y_h}{\partial w_h} + \sum_{i=h,f} \sum_g \left[\frac{\partial V_h(.)}{\partial \tilde{P}_{ih,g}} \frac{\partial \tilde{P}_{ih,g}}{\partial w_h} \right] \right) \frac{\mathrm{d}w_h}{\mathrm{d}\tilde{P}_{hf,k}} = 0$$

As before, we can derive $\partial Y_h / \partial \ln \tilde{P}_{hf,k}$ by taking a derivative from Equation 31, yielding the following

$$\frac{\partial Y_h(\tilde{\boldsymbol{P}}; w_h)}{\partial \ln \tilde{P}_{hf,k}} = -\sum_g \left[\left(P_{hf,g} Q_{hf,g} \frac{\partial \ln P_{hf,g}}{\partial \ln \tilde{P}_{hf,g}^{\mathcal{I}}} + P_{hh,g} Q_{hh,g} \frac{\partial \ln P_{hh,g}}{\partial \ln \tilde{P}_{hf,g}^{\mathcal{I}}} \right) \right]$$
$$+ \tilde{P}_{hf,k} Q_{hf,k} + \sum_g \left((\tilde{P}_{hf,g} - P_{hf,g}) Q_{hf,g} \frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}} \right) - \sum_g \left(P_{fh,g} Q_{fh,g} \frac{\partial \ln P_{fh,g}}{\partial \ln \tilde{P}_{hf,g}^{\mathcal{I}}} \right)$$
$$+ \sum_g \left((\tilde{P}_{hh,g} - P_{hh,g}) Q_{hh,g} \frac{\partial \ln Q_{hh,g}}{\partial \ln Y_h} + (\tilde{P}_{fh,g} - P_{fh,g}) Q_{fh,g} \frac{\partial \ln Q_{fh,g}}{\partial \ln Y_h} \right) \frac{\partial \ln Y_h}{\ln \tilde{P}_{hf,k}}.$$

To simplify the last sum in line 2, we can use the fact that $P_{fh,g} Q_{fh,g} \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{ff,g}} = P_{ff,g} Q_{ff,g} \frac{\partial \ln P_{fh,g}}{\partial \ln Q_{fh,g}}$ as well as the following well-known result from consumer theory:

$$\sum_g P_{ff,g} Q_{ff,g} \frac{\partial \ln Q_{ff,g}}{\partial \ln \tilde{P}_{hf,k}} = -\tilde{P}_{hf,k} Q_{hf,k} - \sum_g \tilde{P}_{hf,g} Q_{hf,g} \frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}}.$$

Plugging these equations back into the expression for $\partial Y_h / \partial \ln \tilde{P}_{hf,k}$ yields the following:

$$\begin{aligned}\frac{\partial Y_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial \ln \tilde{P}_{hf,k}} =& \tilde{P}_{hf,k}Q_{hf,k} + \sum_g \left[(\tilde{P}_{hf,g} - P_{hf,g})Q_{hf,g}\varepsilon^{hf,k}_{hf,g}\right] + \sum_g \left(P_{fh,g}Q_{fh,g}\alpha^{hf,k}_{fh,g}\right) \\ &+ \sum_g \left((\tilde{P}_{fh,g} - P_{fh,g})Q_{fh,g}\eta_{fh,g} + \left[\tilde{P}_{hh,g} - P_{hh,g}\right]Q_{hh,g}\eta_{hh,g}\right)\frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hf,k}}.\end{aligned}$$

Similarly, $\mathrm{d}w_h/\mathrm{d}\tilde{P}_{hf,k}$ can be calculated by applying the implicit function theorem to the balanced trade condition, $D_h\left(\tilde{\boldsymbol{P}};w_h\right)$, which yields the following:

$$\begin{aligned}\frac{\mathrm{d}\ln w_h}{\mathrm{d}\ln \tilde{P}_{hf,k}} &= \left[\frac{\partial}{\partial \ln \tilde{P}_{hf,k}}\sum_g \left(P_{fh,g}Q_{fh,g} - \tilde{P}_{hf,g}Q_{hf,g}\right)\right] \Big/ \frac{\partial D_h\left(\tilde{\boldsymbol{P}};w_h\right)}{\partial \ln w_h} \\ &= \left[\sum_g \left(P_{fh,g}Q_{fh,g}\left[\alpha^{hf,k}_{fh,g} + \eta_{fh,g}\frac{\partial \ln Y_h}{\partial \ln \tilde{P}_{hf,k}}\right]\right)\right. \\ &\quad \left. - \tilde{P}_{hf,k}Q_{hf,k} - \sum_g \left(\tilde{P}_{hf,g}Q_{hf,g}\frac{\partial \ln Q_{hf,g}}{\partial \ln \tilde{P}_{hf,k}}\right)\right] \Big/ \frac{\partial D_h(.)}{\partial \ln w_h}.\end{aligned}$$

Plugging the expressions for $\partial Y_h / \partial \ln \tilde{P}_{hf,k}$ and $\mathrm{d}w_h/\mathrm{d}\ln\tilde{P}_{hf,k}$ back into the F.O.C.; adopting our earlier definition $\bar{\tau} \equiv \left(\frac{\partial V_h}{\partial Y_h}\frac{\partial Y_h}{\partial \ln w_h} + \sum_i \sum_g \frac{\partial V_h}{\partial \tilde{P}_{ih,g}}\frac{\partial \tilde{P}_{ih,g}}{\partial \ln w_h}\right) / \left(\partial D_h / \partial \ln w_h\right)$; and focusing in the solution for which $\Psi\left(\boldsymbol{t}^*, \boldsymbol{x}^*, \boldsymbol{s}^*\right) = 0$; we can write the F.O.C. as

$$\begin{aligned}\acute{\lambda}_{hf,k} + \sum_g \left[1 - \frac{1}{(1+x_g)(1+\bar{\tau})(1+s_g)(1+\mu_g)}\right]\acute{\lambda}_{hf,g}\varepsilon^{hf,k}_{hf,g} - \frac{\sum_k \tilde{P}_{hf,g}Q_{hf,g}}{\tilde{P}_{hf,k}Q_{hf,k}}\frac{\sum_k P_{fh,g}Q_{fh,gg}\alpha^{hf,k}_{fh,g}}{\sum_k P_{fh,g}Q_{fh,g}} &= \\ \acute{\lambda}_{hf,k} + \sum_g \left[1 - \frac{1}{(1+x_g)(1+\bar{\tau})(1+s_g)(1+\mu_g)}\right]\acute{\lambda}_{hf,g}\varepsilon^{hf,k}_{hf,g} - \frac{\acute{\lambda}_{hf}}{\acute{\lambda}_{hf,k}}\alpha^{hf,k}_{fh} &= 0\end{aligned} \tag{35}$$

where in deriving the above expression we use the fact $\tilde{P}_{hf,k}/P_{hf,k} = (1+x_k)(1+s_k)$, and $\acute{\lambda}_{hf,g} \equiv \tilde{P}_{hf,g}Q_{hf,g}/Y_h$.

Simultaneously Solving the System of F.O.C.

Combining Equations 32, 34, and 35 Solving the above system implies that $1+s_g = 1$ and $1+t_g = 1+\bar{\tau}$. The term $\bar{\tau}$ is, however, redundant due to multiplicity of the optimal policy up to two uniform tax shifters (Lemma 1). So, introducing the Pigouvian domestic tax, the following formula characterizes an optimal tax combination for any

$\bar{s}$ and $\bar{t} \in \mathbb{R}_+$:

$$\begin{cases} \left(1+s_k^*\right)\left(1+\bar{s}\right) = 1+\mu_k & \forall k \in \mathbb{K} \\ 1+t_k = 1+\bar{t} & \forall k \in \mathbb{K} \\ 1+x_k^* = \left(1+\frac{1}{\lambda_{ff,k}\epsilon_k}[1-\frac{\acute{\lambda}_{hf}}{\acute{\lambda}_{hf,k}}\alpha_{fh}^{hf,k}]\right)\left(1+\bar{t}\right)^{-1} \end{cases}.$$

Considering the above theorem, the optimal policy is fully determined by industry-level expenditure and revenue shares, as well as $\{\mu_k\}$, $\{\epsilon_k\}$, and the global IO matrix, $\boldsymbol{\alpha}$. Here, however, the *corrective gains* from policy no longer depend on only $Var(\ln \mu_k)$, but also depend on the centrality-adjusted variance of μ_k's–see Liu (2018) for a more formal discussion. Similarly, the *terms-of-trade gains* from policy depend on a IO-weighted average of the trade elasticities. That is because IO linkages provide the Home country with extraterritorial taxing power—see Beshkar and Lashkaripour (2019).

Note that the optimal non-cooperative policy is still given by Theorem 1.A, even in the presence of IO linkages. Though, the gains from the optimal cooperative policy can be magnified once we account for IO linkages. Moreover, supposing that the elements of the IO matrix, $\boldsymbol{\alpha}$, are invariant to policy, we can easily produce an analog of Proposition 1 to quantify the gains from policy using Theorem 5.

J Optimal Tax Formula with Arbitrarily Many Countries

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

$$\frac{\mathrm{d}W_i(.)}{\mathrm{d}\ln(1+t_{ji,k})} = \frac{\partial W_i(.)}{\partial \ln(1+t_{ji,k})} + \sum_{\jmath \neq j} \frac{\partial W_i(.)}{\partial \ln w_\jmath} \frac{\mathrm{d}\ln w_\jmath}{\mathrm{d}\ln(1+t_{ji,k})}.$$

It is straightforward to show that $\frac{\partial W_i(.)}{\partial \ln w_\jmath} \frac{\mathrm{d}\ln w_\jmath}{\mathrm{d}\ln(1+t_{ji,k})} \propto \lambda_{\jmath i} r_{\jmath i,k} \lambda_{ji,k}$. Based on actual trade data, $\lambda_{\jmath i} r_{\jmath i,k} / \lambda_{ii} r_{ii,k} \approx 0$ for $\jmath \neq i$; so, changes in welfare can be approximated as

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

The same applies to other types of taxes. Wee can, thus, cast the Country i's optimal policy problem as one that maximizes $W_i(\tilde{\boldsymbol{P}}, w_i)$ by choosing $\{\tilde{P}_{ji,k}, \tilde{P}_{ij,k}\}$subject to feasibility. Following the same steps as in Appendix E, the system of F.O.C. corresponding to this problem can be expressed as:

$$\sum_g \left(\left[1 - \frac{1+s_{i,g}}{1+\mu_g}\right] \acute{\lambda}_{hh,g} \varepsilon_{ii,g}^{ji,k} \right) + \sum_g \sum_{j \neq i} \left[\left(1 - \frac{1+\bar{\tau}}{1+t_{ji,g}}\right) \acute{\lambda}_{ji,g} \varepsilon_{ji,g}^{ji,k} \right] = 0 \quad [\tilde{P}_{ji,k}]$$

$$\acute{\lambda}_{ij,k} + \sum_g \left[1 - \frac{1}{(1+x_{ij,k})(1+\bar{\tau})(1+s_{i,g})(1+\mu_g)} \right] \acute{\lambda}_{hf,g} \varepsilon_{hf,g}^{hf,k} = 0 \quad [\tilde{P}_{ij,k}].$$

Noting that in absence of cross-industry demand effects,

$$\varepsilon_{ji,k}^{\jmath i,k} = \begin{cases} -1 - \epsilon_k(1 - \lambda_{ji,k}) & \jmath = j \\ \epsilon_k \lambda_{\jmath i,g} & \jmath \neq j \end{cases},$$

the above system of F.O.C.s implies the following policy schedule for country i

$$\begin{aligned}
1 + s_{i,k}^* &= \frac{1+\bar{s}}{1+\mu_k}, \quad \forall k \\
1 + t_{ji,k}^* &= 1 + \bar{t}, \quad \forall j, k \\
1 + x_{ij,k}^* &= \left(1 + \frac{1}{\epsilon_k(1-\lambda_{ij,k})}\right)(1+\bar{t})^{-1}, \quad \forall j, k.
\end{aligned}$$

K Technical Appendix: Cleaning the data on the identity/name of exporting firms

Utilizing the information on the identity of the foreign exporting firm is a critical part of our empirical exercise. Unfortunately, the names of the exporting firms in our dataset are not standardized. As a result, there are instances when the same firm is recorded differently due to using or not using the abbreviations, capital and lower-case letters, spaces, dots, other special characters, etc. To standardize the names of the exporting firms, we used the following procedure.[48]

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

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

3. We eliminated abbreviation “LLC,” spaces, parentheses, and other special characters (. , ; / @ ‘ } - & “) from the firms names.

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

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

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

7. We also analyzed all observations with the same contact information, but slightly different name spelling. We only focused on the cases in which there are up to three

[48]The corresponding Stata code is in the cleanFirmsNames.do.

Table 6: Summary Statistics of the Colombian Import Data.

	Year						
Statistic	2007	2008	2009	2010	2011	2012	2013
F.O.B. value (billion dollars)	30.77	37.26	31.39	38.41	52.00	55.79	56.92
$\frac{\text{C.I.F. value}}{\text{F.O.B. value}}$	1.08	1.07	1.05	1.06	1.05	1.05	1.05
$\frac{\text{C.I.F. + tax value}}{\text{F.O.B. value}}$	1.28	1.21	1.14	1.19	1.15	1.18	1.15
No. of exporting countries	210	219	213	216	213	221	224
No. of imported varieties	483,286	480,363	457,000	509,524	594,918	633,008	649,561

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

different variants of the firm name. For these cases, we calculated the Levenshtein distance in the names, which is the smallest number of edits required to match one name to another. We treat all export observations as coming from the same firm if the contact phone number (or email) is the same and the Levenshtein distance is four or less.

L Examining the Plausibility of Estimates

In this section, we look beyond policy implications and ask if our estimates can help resolve the *income-size* elasticity puzzle. This puzzle, as noted by Ramondo et al. (2016), concerns the fact that a large class of quantitative trade models, including Krugman (1980), Eaton and Kortum (2001), and Melitz (2003), predict an income-size elasticity (i.e., the elasticity at which real per capita income increases with population size) that is counterfactually high. One straightforward remedy for this counterfactual prediction is introducing domestic trade frictions into the aforementioned models. This treatment, however, is only a partial remedy. As shown by Ramondo et al. (2016), even after controlling for direct measures of internal trade frictions, the predicted income-size elasticity remains counterfactually strong.

To understand the income-size elasticity puzzle, consider a single-industry version of the model presented in Section 2. Such a model implies the following expression relating country i's real income per worker or TFP ($W_i = w_i/P_i$) to its structural efficiency, A_i, population size, L_i, trade-to-GDP ratio, λ_{ii}, and a measure of internal trade frictions, τ_{ii}:

$$W_i = \gamma \, A_i \, L_i^{\psi} \, \lambda_{ii}^{-\frac{1}{\epsilon-1}} \, \tau_{ii}^{-1}. \tag{36}$$

The standard Krugman model assumes extreme love-of-variety (or extreme scale economies), which implies $\psi = 1/(\epsilon - 1)$ and precludes internal trade frictions, which results in $\tau_{ii} = 1$. Given these two assumptions, we can compute the real income per worker predicted by the standard Krugman model and contrast it to actual data for a cross-section of countries.

For this exercise, we use data on the trade-to-GDP ratio, real GDP per worker, and population size for 116 countries from the PENN WORLD TABLES in the year 2011. Given our micro-estimated trade elasticity, $\epsilon - 1$, and plugging $\tau_{ii} = 1$ as well as $\psi = 1/(\epsilon - 1)$ into Equation 36, we can compute the real income per worker predicted by the Krugman model. Figure 3 (top panel) reports these predicted values and contrasts them to factual values. Clearly, there is a sizable discrepancy between the income-size elasticity predicted by the standard Krugman model (0.36, standard error 0.03) and the factual elasticity (-0.04, standard error 0.06). To gain intuition, note that small countries import a higher share of their GDP (i.e., posses a lower λ_{ii}), which partially mitigates their size disadvantage. However, even after accounting for observable levels of trade openness, the scale economies underlying the Krugman model are so strong that they lead to a counterfactually high income-size elasticity.

One solution to the income-size elasticity puzzle is introducing internal trade frictions into the Krugman model (i.e., relaxing the $\tau_{ii} = 1$ assumption). Ramondo et al. (2016) perform this task using direct measures of domestic trade frictions. Their calibration is suggestive of $\tau_{ii} \propto L_i^{0.17}$. Plugging this implicit relationship into Equation 36 and using data on population size and trade openness, we compute the model-predicted real income per worker and contrast it with actual data in Figure 3 (middle panel). Expectedly, accounting for internal frictions shrinks the income-size elasticity. However, as pointed out by Ramondo et al. (2016), the income-size elasticity remains puzzlingly large.

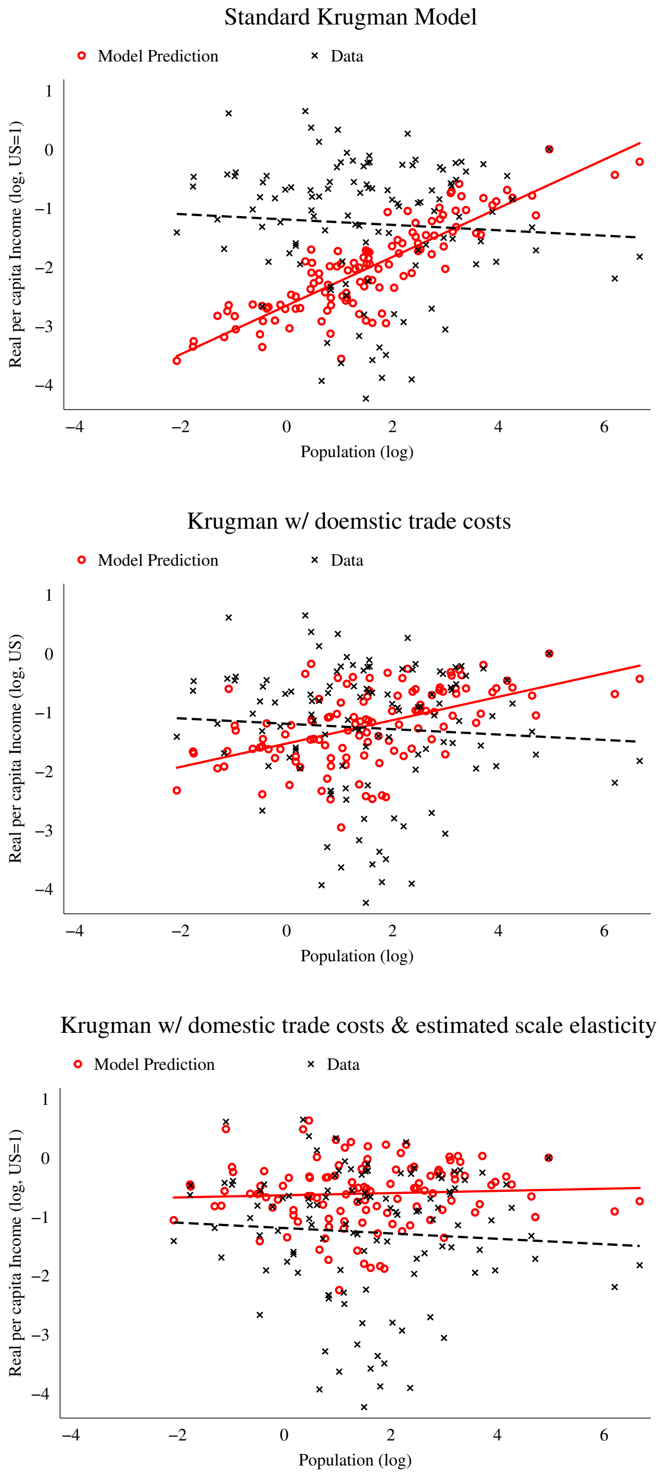
We ask if our micro-estimated scale elasticity can help resolve the remaining income-size elasticity puzzle. To this end, in Equation 36, we set the scale elasticity to $\psi = \alpha/(\epsilon - 1)$ where α is set to 0.65 as implied by our micro-level estimation. Then, using data on population size and trade-to-GDP ratios, we compute the real income per capita predicted by a model that features both domestic trade frictions *and* adjusted scale economies. Figure 3 plots these predicted values, indicating that this adjustment indeed resolves the income-size elasticity puzzle. In particular, the income-size elasticity predicted by the Krugman model with adjusted scale economies is statistically insignificant (0.02, standard error 0.03), aligning very closely with the factual elasticity.

M Mapping Tax Formulas to Data under Free Entry

In this appendix, we present an analog to Proposition 1, but under free entry. To this end, we appeal to Theorem 1.B and adopt the hat-algebra notation, $\hat{x} \equiv x'/x$ to produce the following proposition.

Proposition 4. *Suppose the observed data is generated by a model that exhibits the cross-industry utility aggregator 16. Then, the optimal (first-best) taxes and their effects on wages,*

Figure 3: *The income-size elasticity*



profits, and income can be fully characterized by solving the following system of equations:

$$\begin{cases} x_{hf,k} = \frac{1}{\epsilon_k \hat{\lambda}_{ff,k}\lambda_{ff,k}}; \quad t_{fh,k} = \frac{1+\psi_k \hat{r}_{ff,k} r_{ff,k}}{1+\psi_k}; \quad 1+s_{h,k} = \frac{1}{1+\mu_k} & \forall k \in \mathbb{K} \\ x_{fh,k} = 0, \quad t_{hf,k} = 0; \quad s_k = 0 & \forall k \in \mathbb{K} \\ \hat{\lambda}_{ji,k} = \hat{r}_{ji,k}^{\psi_k} \left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k} \hat{\tilde{P}}_{i,k}^{\epsilon_k} & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \hat{\tilde{P}}_{i,k}^{-\epsilon_k} = \sum_j \left(\hat{r}_{ji,k}^{\psi_k} \left[(1+t_{ji,k})(1+x_{ji,k})(1+s_{j,k})\widehat{w}_j\right]^{-\epsilon_k} \lambda_{ji,k}\right) & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \\ \widehat{w}_i w_i L_i = \sum_k \sum_j \left(\frac{\hat{\lambda}_{ij,k}\lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\right) & \forall i \in \mathbb{C} \\ \widehat{\Pi}_i \Pi_i = \sum_k \sum_j \left(\frac{\mu_k \hat{\lambda}_{ij,k}\lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j}{(1+\mu_k)(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})}\right) & \forall i \in \mathbb{C} \\ \widehat{\mathcal{R}}_i \mathcal{R}_i = \sum_k \sum_j \left(\frac{t_{ji,k}}{1+t_{ji,k}} \hat{\lambda}_{ji,k}\lambda_{ji,k} e_{i,k} \hat{Y}_i Y_i + \frac{(1+x_{ij,k})(1+s_{i,k})-1}{(1+x_{ij,k})(1+s_{i,k})(1+t_{ij,k})} \hat{\lambda}_{ij,k}\lambda_{ij,k} e_{j,k} \hat{Y}_j Y_j\right) & \forall i \in \mathbb{C} \\ \hat{Y}_i Y_i = \ \widehat{w}_i w_i L_i \ + \ \widehat{\Pi}_i \Pi_i \ + \widehat{\mathcal{R}}_i \mathcal{R}_i & \forall i \in \mathbb{C} \\ \hat{r}_{ji,k} r_{ji,k} = \left(\frac{\hat{\lambda}_{ji,k}\lambda_{ji,k} e_{i,k} \hat{Y}_i Y_i}{(1+x_{ji,k})(1+s_{j,k})(1+t_{ji,k})}\right) / \left(\sum_\iota \frac{\hat{\lambda}_{j\iota,k}\lambda_{j\iota,k} e_{\iota,k} \hat{Y}_\iota Y_\iota}{(1+x_{j\iota,k})(1+s_{j,k})(1+t_{j\iota,k})}\right) & \forall j,i \in \mathbb{C},\ k \in \mathbb{K} \end{cases} .$$

The above system solves $\{t_{fh,k}\}$, $\{x_{hf,k}\}$, $\{s_{h,k}\}$, $\{\hat{\lambda}_{ji,k}\}$, $\{\hat{r}_{ff,k}\}$, $\{\hat{\tilde{P}}_{i,k}\}$, $\{\hat{w}_i\}$, and $\{\hat{Y}_i\}$, as a function of (i) industry-level scale elasticities, $\{\psi_k\}$; (ii) industry-level trade elasticities, $\{\epsilon_k\}$; (iii) observed expenditure shares, $\{\lambda_{ji,k}\}$;(iv) observed revenue shares, $\{r_{ji,k}\}$; and (v) total expenditure and wage revenue, $w_i L_i$ and Y_i.

Note that once we solve for the optimal policy, we can immediately calculate the gains from policy as $\hat{Y}_i / \prod_k \hat{\tilde{P}}_{i,k}^{e_{i,k}}$.

N WIOD Industry Details

Table 7: *List of industries in quantitative analysis*

WIOD Sector	Sector's Description	Trade Ealsticity	Scale Ealsticity
1	Agriculture, Hunting, Forestry and Fishing	6.212	0.141
2	Mining and Quarrying	6.212	0.141
3	Food, Beverages and Tobacco	3.333	0.265
4	Textiles and Textile Products Leather and Footwear	3.413	0.207
5	Wood and Products of Wood and Cork	3.329	0.270
6	Pulp, Paper, Paper , Printing and Publishing	2.046	0.397
7	Coke, Refined Petroleum and Nuclear Fuel	0.397	1.758
8	Chemicals and Chemical Products	4.320	0.212
9	Rubber and Plastics	3.599	0.162
10	Other Non-Metallic Mineral	4.561	0.186
11	Basic Metals and Fabricated Metal	2.959	0.189
12	Machinery, Nec	8.682	0.100
13	Electrical and Optical Equipment	1.392	0.453
14	Transport Equipment	2.173	0.133
15	Manufacturing, Nec; Recycling	6.704	0.142
16	Electricity, Gas and Water Supply Construction Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods Hotels and Restaurants Inland Transport Water Transport Air Transport Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies Post and Telecommunications Financial Intermediation Real Estate Activities Renting of M&Eq and Other Business Activities Education Health and Social Work Public Admin and Defence; Compulsory Social Security Other Community, Social and Personal Services Private Households with Employed Persons	100	0