Trade Policy and Global Sourcing: An Efficiency Rationale for Tariff Escalation*

Pol Antràs
Harvard University,
CEPR, and NBER

Teresa C. Fort
Tuck School at Dartmouth,
CEPR, and NBER

Agustín Gutiérrez
University of Wisconsin-Madison

Felix Tintelnot
University of Chicago,
CEPR, and NBER

November 16, 2023

Abstract

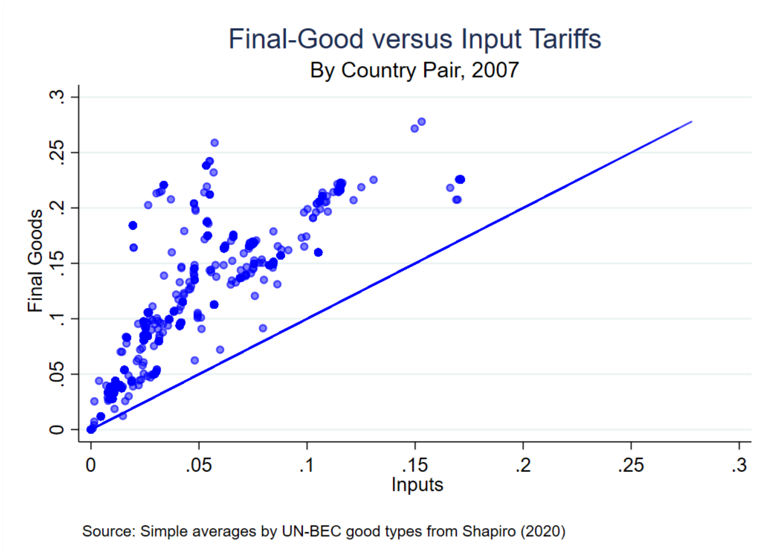
Import tariffs tend to be higher on final goods than inputs, a phenomenon commonly referred to as tariff escalation. Despite its salience, existing trade policy results do not predict that tariff escalation increases social welfare. We show that tariff escalation is often welfare-improving when final-good production occurs under increasing returns to scale. In our model, a country can export inputs directly, or by embodying them into exports of final goods. The latter raises welfare when final-good efficiency is increasing in sector size, and a final-good production subsidy or import tariff are equally effective in exploiting this benefit. When import tariffs are the only policy tools, this force generally dominates other motives for an input tariff, leading to a disproportionately large tariff on final goods. We calibrate our model and show that in such second-best settings, tariff escalation maximizes welfare for empirically relevant parameter values whenever the returns to scale for final goods are at least as large those for inputs. A quantitative evaluation of the US-China trade war demonstrates that any welfare gains are overwhelmingly driven by final-good tariffs.

*We are grateful to Lorenzo Caliendo, Arnaud Costinot, Dave Donaldson, Ahmad Lashkaripour, Steve Redding, Andrés Rodríguez-Clare, Iván Werning, our editor (Ariel Burstein), three anonymous referees, and various seminar audiences for insightful comments, and to Joe Shapiro for sharing data with us. Ignacia Cuevas, Evgenii Fadeev, Reigner Kane, Jack Liang, Nicolo Rizzotti, Nicolas Wesseler, and Shuhan Zou provided excellent research assistance. A previous version of this paper was circulated under the title "Import Tariffs and Global Sourcing."

1 Introduction

Import tariffs tend to be lower on intermediate inputs than on final goods. This pattern has been documented in multiple studies spanning numerous countries across five decades (Travis, 1964; Balassa, 1965; Bown and Crowley, 2016; Shapiro, 2020). It is commonly referred to as 'tariff escalation,' a term that captures the fact that tariffs 'escalate' down the production chain. Figure 1 illustrates the prevalence of tariff escalation across trading partners in 2007: for almost every country-pair, the simple average of final-good tariffs is higher than average input tariffs.

Figure 1: Average Final-Good versus Input Tariffs



Empirical research suggests that these relatively low input tariffs improve downstream firm and worker outcomes. Early papers document significant productivity gains from lower input tariffs (Amiti and Konings, 2007; Goldberg et al., 2010; Topalova and Khandelwal, 2011), while new evidence shows that recent US input-tariff hikes harmed US manufacturing employment (Flaaen and Pierce, 2019) and exports (Handley et al., 2020).[1]

Despite the ubiquity of tariff escalation and evidence that lower input tariffs are beneficial, existing theory does not show that tariff escalation raises welfare. Early neoclassical models analyze input versus final-good tariffs explicitly, but do not find that optimal tariffs are lower for inputs. Modern Ricardian trade models with differentiated goods stress that welfare-maximizing tariffs are *uniform* across sectors; moreover tariff *de*-escalation may maximize welfare in second-best settings without export taxes. Although the classical production efficiency result in Diamond and Mirrlees (1971) indicates that intermediate inputs should not be taxed in a closed economy with constant

[1] Evidence on the costs of raising prices on imported inputs is also documented for anti-dumping duties (Bown et al., 2020; Barattieri and Cacciatore, 2020) and the Bush steel tariffs (Cox, 2021). Relatedly, reductions in US trade policy uncertainty with China on firms' inputs led to relatively higher US export growth (Breinlich et al., 2021).

returns to scale, that result also implies that optimal trade taxes should be uniform when a country has no market power abroad (and in fact implies that free trade is optimal).

The leading explanation for tariff escalation invokes political-economy forces: all firms lobby for protection of their output and counter-lobby against tariffs on their imported inputs (Cadot et al., 2004; Gawande et al., 2012). This counter-lobbying explanation assumes that final-good importers cannot lobby against tariffs on consumer goods, which is increasingly at odds with the well-documented and rising concentration of consumer-good imports by dominant wholesale and retail firms (Basker and Van, 2010; Ganapati, 2018; Smith and Díaz, 2020).[2] More generally, and in line with results in the broader literature, lower tariffs on inputs do not raise welfare in these papers.

In this paper, we document environments in which tariff escalation emerges as a social-welfare maximizing policy. We start with the observation that firm productivity and thus welfare are decreasing in tariffs on traded inputs when final-good production occurs under increasing returns to scale (e.g., as in Antràs et al., 2017; Blaum et al., 2018). Moreover, past work shows that such increasing returns provide an efficiency rationale for final-good tariffs in partial equilibrium (Ossa, 2011). These results do not necessarily imply that tariff escalation is optimal, however, especially if inputs are also produced under increasing returns to scale and thus subject to similar potential efficiency gains. We therefore consider a general-equilibrium framework with an upstream and downstream sector that produce differentiated intermediate and final goods, respectively, both under increasing returns.[3] Inputs are made using only labor, while final goods are produced by combining labor and intermediate inputs.

The key assumption in our vertical model is that inputs are sold only to downstream firms, whereas final goods are all sold to consumers. As a result, our model does not feature misallocation of domestic expenditure across final goods. Instead, the only domestic distortion is the amount of labor allocated to input versus final-good production, and this misallocation depends only on the extent of increasing returns to scale upstream and the downstream labor share. Maximizing welfare in the closed economy thus only requires an upstream production subsidy, and the extent of increasing returns to scale downstream is irrelevant. This finding stands in sharp contrast to results in horizontal models where downstream subsidies are always required to maximize social welfare (Bartelme et al., 2021; Lashkaripour and Lugovskyy, 2023), and illustrates the fact that vertical linkages generate inherently different motives for policy intervention.

We first study optimal tariffs in an open economy with a small Home country and a large Foreign country that trade inputs and final goods. Although Home has market power abroad due to product differentiation, the small-country assumption allows us to hold fixed the prices charged by Foreign, and derive analytic solutions for optimal trade policy following the primal approach (Lucas and Stokey, 1983; Costinot et al., 2015). The Home social planner has the well-known motives to exploit

[2]For example, in September 2018 Walmart responded to the Trump tariffs on consumer goods with a direct letter to US Trade Representative Robert Lighthizer warning of price hikes and consumer harm.

[3]We model both sectors as perfectly competitive with external economies of scale, and show that an isomorphic Krugman-style model with monopolistic competition and firm-level increasing returns to scale generates identical tariff motives.

her country's market power abroad (Gros, 1987), and to shift expenditure towards Home varieties to benefit from increasing returns to scale in production (Venables, 1987; Ossa, 2011).

We analyze how the planner can maximize social welfare using a combination of production subsidies, export taxes, and import tariffs. We first show that any first-best allocation must entail an upstream production subsidy, whereas a downstream production subsidy or a downstream import tariff are interchangeable. The motive for a downstream subsidy or tariff only arises in the open economy, because now Home faces a tradeoff between exporting inputs directly versus using those inputs to make final goods that it then exports. Since in either case Home produces the inputs, that sector's size does not change and its scale elasticity does not affect this decision. However, by using inputs to make final goods it then exports, Home raises the size of the downstream sector and thus its efficiency. As a result, the size of the optimal tariff or subsidy depends on the extent of downstream returns. In line with past work, the planner also exploits Home market power abroad using export taxes and possibly import tariffs. Moreover, if we rule out the downstream subsidy, the first-best necessarily entails a tariff escalation wedge exactly determined by the extent of increasing returns to scale in downstream production.

Export taxes are often illegal and production subsidies face strict World Trade Organization limitations.[4] Since a main goal of this paper is to assess whether real-world tariffs reflect social welfare-maximizing policies, we analyze optimal policies that rely solely on import tariffs. In such a second-best setting, tariffs try to mimic the missing export taxes and upstream production subsidy. Although we cannot derive a simple second-best characterization of optimal tariffs, we find that tariff escalation is decreasing in (i) the relative size of upstream to downstream returns to scale; (ii) the downstream labor share; and (iii) the relative size of Foreign's input versus final-good demand elasticities. The first two conditions imply a greater need for the missing upstream production subsidy, and thus a higher input tariff to replace it. The last condition reflects Home's efforts to exploit its market power abroad without export taxes (as in Beshkar and Lashkaripour, 2020). We also derive conditions that are sufficient, but *not* necessary for our model to predict tariff escalation. For instance, when inputs and final goods have identical domestic scale and Foreign demand elasticities, tariff escalation is necessarily optimal when the share of labor used downstream is not too large.[5]

We evaluate the relative importance of these distinct channels by solving for optimal, second-best tariffs as a function of a wide range of values for these five key parameters. In our grid searches, we find that tariff escalation is essentially always optimal when downstream returns to scale are at least as large as upstream returns. By contrast, tariff *de*-escalation may be optimal when upstream returns are larger than downstream returns, *and* the downstream labor share is quite high. In sum, as long as inputs are sufficiently important in downstream production, tariff escalation maximizes social welfare when import tariffs are the only policy instruments.

We use the calibrated model to analyze optimal tariffs for a large, open economy with endogenous

[4]For example, Article I, Section 9, Clause 5 of the US Constitution bans export taxes.

[5]Tariff escalation is always optimal when final goods are produced under increasing returns, but inputs are made under constant returns to scale.

world prices. The resulting first-best tariff escalation aligns closely with the analytic solutions we derive in a small open economy. In a second-best setting in which tariffs are the only available instruments, the optimal final-good tariff is 29.9 percent, versus an optimal input tariff of only 19.7 percent.[6] Tariff escalation is decreasing in the extent to which the downstream sector uses labor in production, consistent with it mimicking the missing upstream subsidy. As for the small open economy, second-best tariffs are escalated when the returns to scale downstream are at least as large as those upstream; even when upstream returns are larger, tariff escalation remains optimal when the downstream labor share is below 70 percent. A counterfactual analysis of the welfare effects of the US trade war in 2018 to 2019 indicate the welfare benefits from unilateral tariff changes are small and overwhelming driven by final-good tariffs.

Our paper contributes to three strands of literature. First, we add to a growing body of work on optimal trade policy in environments with differentiated final goods. Early papers with single-sector models show that product differentiation provides an incentive for final-good tariffs, even when countries are too small to affect world prices (Gros, 1987; Demidova and Rodríguez-Clare, 2009). Recent studies using multi-sector competitive Ricardian models predict that optimal tariffs are uniform across sectors not only in the first-best allocation, but also in second-best settings when export taxes and production subsidies are not available (Costinot et al., 2015; Lashkaripour and Lugovskyy, 2023). Lashkaripour and Lugovskyy (2023) include an extension with input-output linkages and conclude that increasing returns to scale must be corrected using production subsidies, and that first-best tariffs necessarily remain uniform. Our contribution is to show that in a pure vertical model a downstream tariff or subsidy are interchangeable, and when such a tariff is used instead of the subsidy, first-best tariffs feature a unique tariff escalation wedge determined by the extent of downstream returns to scale. We are also the first to study tariff escalation in a second-best setting and show that it remains optimal for a large range of parameter values.

We also add to work that studies the interaction between increasing returns to scale in domestic production and trade policy. Venables (1987) and Ossa (2011) show that in partial equilibrium with imperfect competition and scale economies, an increase in a final-good import tariff attracts firms to a country, which in turn lowers prices and thus raises welfare. In general equilibrium, agglomeration forces may interact with trade barriers to generate stronger agglomeration effects from changes in trade barriers when production is roundabout (Krugman and Venables, 1995; Puga and Venables, 1999). We contribute to these studies by extending the isomorphism developed in Kucheryavyy et al. (2023), who map a horizontal Krugman-style model to a perfectly competitive economy with external economies of scale, to a vertical setting. Most notably, our framework with distinct input versus final-good sectors allows us to isolate the roles of increasing returns to scale in each sector, and to show that the optimality of tariff escalation is inherently tied to increasing returns to scale in downstream production.[7]

[6] Optimal tariffs tend to be much larger than observed tariff rates in the data in the quantitative trade policy literature (see Costinot and Rodríguez-Clare 2014; Ossa 2014).

[7] While Krugman and Venables (1995) consider multiple equilibria that include full specialization in manufacturing, the general-equilibrium forces in our model imply that input and final-good sectors compete for labor.

Finally, we contribute to a growing literature on trade policy with fragmented production. Early neoclassical models feature trade in inputs and final goods, but do not find that lower input tariffs raise welfare (Ruffin, 1969; Casas, 1973; Das, 1983), while newer work argues that tariff de-escalation is optimal in second-best settings (Beshkar and Lashkaripour, 2020). Recent papers show that global value chains lower optimal tariffs, but these studies are limited to trade in intermediates only with roundabout production (Caliendo et al., 2023), or trade in final goods only but with domestic value added (Blanchard et al., 2021). Amiti (2004) introduces a two-sector, vertical model and performs numerical simulations that suggest tariff de-escalation is optimal. Our contribution is to model input versus final-good sectors separately, to provide closed-form solutions for the first-best welfare allocations that show tariff escalation arises from increasing returns to scale in final-good production, and to demonstrate that second-best tariffs are escalated for empirically relevant parameter values.[8]

The rest of the paper is structured as follows. In Section 2, we introduce an open-economy model with an upstream and downstream sector, both featuring increasing returns to scale. In Section 3, we solve for the first-best policies for a small open economy, building intuition for our results in Section 4. Optimal second-best import tariffs for a small open economy are developed in Section 5, and we provide quantitative results for a large open economy in Section 6. We perform a counterfactual analysis of the welfare impacts of the Trump tariffs in Section 7, and conclude in Section 8.

2 A Model of Global Sourcing with Scale Economies

In this section, we introduce a competitive open-economy model with an upstream (input) sector and a downstream (final-good) sector, both producing differentiated goods under increasing returns to scale.

2.1 Environment

We consider a world economy with two countries (Home and Foreign), indexed by i or j (and sometimes by H and F), each populated by L_i consumers/workers. Labor is employed to produce four goods: a Home final good, a Foreign final good, a Home intermediate input, and a Foreign intermediate input. Preferences in country $i = \{H, F\}$ are given by

$$U\left(Q_{ii}^d, Q_{ji}^d\right) = \left(\left(Q_{ii}^d\right)^{\frac{\sigma-1}{\sigma}} + \left(Q_{ji}^d\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \quad (1)$$

where Q_{ii}^d is i's consumption of its local good, and Q_{ji}^d are imports by i of country j's good. The parameter σ governs the substitutability between Home and Foreign goods.

[8] Other papers study tariff escalation under various market structures (e.g., Spencer and Jones, 1991, 1992; McCorriston and Sheldon, 2011; Hwang et al., 2017); and the effects of input tariffs under incomplete contracts (Ornelas and Turner, 2008; Antràs and Staiger, 2012; Ornelas and Turner, 2012; Grossman and Helpman, 2020) or on demand for downstream tariffs (Erbahar and Zi, 2017).

The final good in each country is produced combining local labor (ℓ_i^d), the Home intermediate input (q_{ii}^u), and the Foreign intermediate input (q_{ji}^u), according to

$$x_i^d = \hat{A}_i^d F^d\left(\ell_i^d, q_{ii}^u, q_{ji}^u\right) = \hat{A}_i^d \left(\ell_i^d\right)^\alpha \left(\left(q_{ii}^u\right)^{\frac{\theta-1}{\theta}} + \left(q_{ji}^u\right)^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}(1-\alpha)}, \tag{2}$$

where $\hat{A}_i^d$ is downstream productivity, α captures the labor intensity of final-good production, and θ governs the substitutability between the Home and the Foreign inputs. Downstream productivity $\hat{A}_i^d$ is in turn given by

$$\hat{A}_i^d = \bar{A}_i^d \left(F^d\left(L_i^d, Q_{ii}^u, Q_{ji}^u\right)\right)^{\gamma^d} = \bar{A}_i^d \left(\left(L_i^d\right)^\alpha \left(\left(Q_{ii}^u\right)^{\frac{\theta-1}{\theta}} + \left(Q_{ji}^u\right)^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta(1-\alpha)}{\theta-1}}\right)^{\gamma^d}, \tag{3}$$

and is thus an endogenous function of (i) country i's aggregate allocation of labor to the downstream sector (L_i^d), (ii) its aggregate use of country i's inputs (Q_{ii}^u), and (iii) its aggregate use of country j's inputs (Q_{ji}^u). The parameter γ^d governs the degree of external economies of scale in the downstream sector, and is often referred to as the scale elasticity of this sector.

The intermediate input in each country is produced using local labor (ℓ_i^u) according to

$$x_i^u = \hat{A}_i^u F_i^u\left(\ell_i^u\right) = \hat{A}_i^u \ell_i^u,$$

where upstream productivity is also endogenous and given by

$$\hat{A}_i^u = \bar{A}_i^u \left(L_i^u\right)^{\gamma^u}. \tag{4}$$

In this expression, L_i^u is country i's aggregate allocation of labor to the upstream sector and γ^u governs the scale elasticity of the upstream sector.

We assume that the above technologies are available to a unit measure competitive fringe of producers in each country and sector. These producers take prices of all goods as given, and do not internalize the effects of their choices on the productivity terms $\hat{A}_i^d$ and $\hat{A}_i^u$. Given symmetry, it should be clear that in equilibrium, $\ell_i^d = L_i^d$, $\ell_i^u = L_i^u$, $q_{ii}^u = Q_{ii}^u$, and $q_{ji}^u = Q_{ji}^u$ for all firms. We thus ignore lower-case variables hereafter (though it is evident that, in characterizing the optimality conditions of the decentralized market equilibrium, such a distinction is important).

There exist natural trade barriers between Home and Foreign, which we capture with iceberg trade costs that apply to final goods $\tau^d > 1$ and to inputs $\tau^u > 1$, respectively. We later introduce additional man-made barriers to international trade in the form of import tariffs and export taxes.

The three resource constraints associated with our economic environment are an aggregate labor market constraint

$$L_i = L_i^u + L_i^d, \tag{5}$$

and two equations equating output produced in each sector to its use (for domestic consumption or for export)

$$\begin{aligned} \hat{A}_i^u L_i^u &= Q_{ii}^u + \tau^u Q_{ij}^u; & (6) \\ \hat{A}_i^d \left(L_i^d\right)^\alpha \left((Q_{ii}^u)^{\frac{\theta-1}{\theta}} + \left(Q_{ji}^u\right)^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta(1-\alpha)}{\theta-1}} &= Q_{ii}^d + \tau^d Q_{ij}^d. & (7) \end{aligned}$$

In addition, we impose the following trade-balance condition

$$P_{ji}^d Q_{ji}^d + P_{ji}^u Q_{ji}^u = P_{ij}^d Q_{ij}^d + P_{ij}^u Q_{ij}^u. \tag{8}$$

In this last condition, the prices P_{ji}^d and P_{ji}^u reflect the import prices paid by importers in i (inclusive of trade costs between j and i), and P_{ij}^d and P_{ij}^u similarly correspond to the cost, insurance, and freight-inclusive (CIF) prices paid by buyers in j per unit consumed (though remember that exporters in i need to produce τ^d and τ^u units, respectively, for one unit to make it to j).

3 First-Best Policies for a Small Open Economy

In this section, we solve for the bilateral first-best policies of one country (Home), for the special case in which it is a small open economy. Although Home is small in the sense that it cannot affect the prices charged by Foreign, Home nevertheless enjoys market power for its goods abroad.

Following the primal approach in Costinot et al. (2015),[9] we first consider an environment in which a social planner directly controls consumption and output decisions, and derive four key conditions characterizing the optimal allocation. We then compare these conditions to those from a decentralized market equilibrium in which the government imposes trade taxes and production subsidies. Finally, we show how the optimal allocation can be implemented through a simple combination of those instruments.

3.1 Optimal Allocation for a Small Open Economy

Home's social planner seeks to maximize welfare in equation (1), subject to the four feasibility constraints (5) to (8). We assume that the planner controls the allocation of labor across sectors (L_H^u, L_H^d), domestic consumption (Q_{HH}^d, Q_{HH}^u), imports (Q_{FH}^d, Q_{FH}^u) and exports (Q_{HF}^d, Q_{HF}^u) of

[9] See Costinot et al. (2020) and Kortum and Weisbach (2021) for other recent applications of this approach.

both final goods and intermediate inputs, and chooses these variables to solve:

$$
\begin{aligned}
\max \quad & U\left(Q^d_{HH}, Q^d_{FH}\right) = \left(\left(Q^d_{HH}\right)^{\frac{\sigma-1}{\sigma}} + \left(Q^d_{FH}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \\
s.t. \quad & L^u_H + L^d_H = L_H \\
& \hat{A}^u_H \left(L^u_H\right) L^u_H = Q^u_{HH} + \tau^u Q^u_{HF} \\
& \hat{A}^d_H \left(L^d_H, Q^u_{HH}, Q^u_{FH}\right) F^d \left(L^d_H, Q^u_{HH}, Q^u_{FH}\right) = Q^d_{HH} + \tau^d Q^d_{HF} \\
& P^d_{FH} Q^d_{FH} + P^u_{FH} Q^u_{FH} = \left(Q^d_{HF}\right)^{\frac{\sigma-1}{\sigma}} P^d_{FF} \left(Q^d_{FF}\right)^{\frac{1}{\sigma}} + \left(Q^u_{HF}\right)^{\frac{\theta-1}{\theta}} P^u_{FF} \left(Q^u_{FF}\right)^{\frac{1}{\theta}},
\end{aligned}
$$

where $\hat{A}^d_H \left(L^d_H, Q^u_{HH}, Q^u_{HF}\right)$ and $\hat{A}^u_H \left(L^u_H\right)$ are given in (3) and (4), respectively. The first three constraints in this program are simply the three resource constraints displayed above in equations (5), (6), and (7).

The fourth constraint requires more explanation. First, notice that this trade balance condition is derived from (8), after substituting $P^d_{HF} = P^d_{FF} \left(Q^d_{HF}/Q^d_{FF}\right)^{-1/\sigma}$ and $P^u_{HF} = P^u_{FF} \left(Q^u_{HF}/Q^u_{FF}\right)^{-1/\theta}$, which correspond to Foreign's inverse demand for Home's final and intermediate-input goods, respectively.[10] Although Home's small-country status implies that it takes the prices charged by Foreign as given, it nevertheless produces *differentiated* final goods and inputs. As a result, the planner seeks to exploit Home's market power abroad, which is by governed by Foreign's elasticities of substitution across inputs (θ) and final goods (σ). Second, although it may seem non-standard to introduce prices in the planner problem, this is precisely where our assumption of Home being a small open economy is useful: Home's policies have *no impact* on Foreign's domestic prices P^d_{FF} and P^u_{FF}, or on the prices P^d_{FH} and P^u_{FH} charged by Foreign exporters. As a result, the Home government treats these prices as parameters when choosing its optimal allocations. Similarly, Home takes as given Foreign's demand for its own goods (Q^d_{FF} and Q^u_{FF}).

Working with the first-order conditions, we characterize the first-best allocations via four key conditions (see Appendix A.1.1). First, on the consumption side, the Home social planner equates the representative consumer's marginal rate of substitution between Home and Foreign final goods with the *social* relative costs of consuming those goods, which delivers

$$
\frac{U_{Q^d_{HH}}\left(Q^d_{HH}, Q^d_{FH}\right)}{U_{Q^d_{FH}}\left(Q^d_{HH}, Q^d_{FH}\right)} = \frac{\frac{\sigma-1}{\sigma} P^d_{HF}/\tau^d}{P^d_{FH}}, \tag{9}
$$

where subindices on the functions denote partial derivatives of those functions with respect to the argument in the subindex.

The left-hand side of equation (9) corresponds to the marginal rate of substitution between domestic and foreign final goods. Since Home is a small open economy, P^d_{FH} is fixed and thus equal to the private cost *and* social cost of consuming foreign goods. By contrast, the social cost of consuming domestic goods is equal to the private opportunity cost of collecting export revenue P^d_{HF}

[10] We derive analogous equations below for the Home economy (see equations (13) and (14) below).

on a fraction $1/\tau^d$ of these goods, multiplied by a wedge $\frac{\sigma-1}{\sigma}$. This wedge between the private and social cost of domestic consumption reflects a standard rationale for terms-of-trade manipulation with market power abroad: for a given level of final-good production, an increase in domestic consumption Q^d_{HH} necessarily reduces exports, which in turn raises Home's export prices and thus improves its terms of trade (see Gros, 1987).

The second key efficiency condition equates the marginal rate of substitution between domestic and foreign inputs in Home's production of final goods to the ratio of social costs of these inputs, or

$$\frac{F^d_{Q^u_{HH}}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right)}{F^d_{Q^u_{FH}}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right)} = \frac{\frac{\theta-1}{\theta} P^u_{HF}/\tau^u}{P^u_{FH}}. \tag{10}$$

As before, the social cost of using Foreign inputs is simply equal to their price, while the social cost of using domestic inputs is lower than their private cost. As for final goods, the Home government again has a terms-of-trade incentive to exploit its market power abroad in inputs, with a wedge between the private and social cost based on Foreign's elasticity of substitution across inputs (θ).

The third efficiency condition is given by

$$\frac{\theta-1}{\theta}\frac{P^u_{HF}}{\tau^u} = \left(1+\gamma^d\right)\hat{A}^d_H F^d_{Q^u_{HH}}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right)\frac{\sigma-1}{\sigma}\frac{P^d_{HF}}{\tau^d}, \tag{11}$$

and equates the benefits of exporting domestic inputs to the benefits of using those inputs to produce final goods that are then exported. This third equation balances Home's motive to exploit its market power in inputs abroad against a comparable market-power motive for final goods *and* the productivity-enhancing effects of a larger domestic final-good sector (the first term $1+\gamma^d$ on the right-hand side). Note that only the extent of increasing returns downstream is relevant here, highlighting its interaction with domestic-input use.

The fourth and final optimality condition is given by

$$\left(1+\gamma^d\right)F^d_{L^d_H}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right) = \left(1+\gamma^d\right)\left(1+\gamma^u\right)\hat{A}^u\left(L^u_H\right)F^d_{Q^u_{HH}}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right), \tag{12}$$

and equates the social value of the marginal product of labor in both sectors in terms of final-good production. The left-hand-side captures the social marginal product of directly allocating labor to produce final goods, while the right-hand-side captures the social marginal product of allocating labor to produce inputs, and then using those inputs to raise final-good production. Equation (12) highlights a key distinction between input and final-good production. While the marginal benefit of increasing the input sector size is magnified by the extent of increasing returns downstream $(1+\gamma^d)$, the marginal benefit of increasing the downstream sector size has no such spillover. As a result, the role of increasing returns downstream cancels in this constraint, and only the extent of increasing returns upstream remains.

3.2 Decentralized Market Equilibrium with Taxes and Subsidies

We now consider the role of upstream and downstream import tariffs (t_i^d and t_i^u), export taxes (v_i^d and v_i^u), and production subsidies (s_i^d and s_i^u) in implementing the first-best allocation (see Appendix A.1.2 for details). We assume that all subsidy proceeds are extracted from households (or tax revenue is rebated to households) in a lump-sum manner.

The first two optimality conditions of the decentralized equilibrium simply equate the marginal rates of substitution in final-good and intermediate-input consumption to the domestic (country i) relative prices faced by the buyers of these goods, or

$$\frac{U_{Q_{HH}^d}\left(Q_{HH}^d, Q_{FH}^d\right)}{U_{Q_{FH}^d}\left(Q_{HH}^d, Q_{FH}^d\right)} = \frac{\left(1-v_H^d\right)}{\left(1+t_H^d\right)}\frac{P_{HF}^d/\tau^d}{P_{FH}^d}; \tag{13}$$

$$\frac{F_{Q_{HH}^u}^d\left(L_H^d, Q_{HH}^u, Q_{FH}^u\right)}{F_{Q_{FH}^u}^d\left(L_H^d, Q_{HH}^u, Q_{FH}^u\right)} = \frac{\left(1-v_H^u\right)}{\left(1+t_H^u\right)}\frac{P_{HF}^u/\tau^u}{P_{FH}^u}. \tag{14}$$

To understand the right-hand-side terms, recall that P_{FH}^d and P_{FH}^u reflect prices (inclusive of trade costs but exclusive of trade taxes) paid by Home importers. Thus, the domestic prices paid by Home buyers of those goods are $(1+t_H^d)P_{FH}^d$ and $(1+t_H^u)P_{FH}^u$, respectively. Similarly, the export prices P_{HF}^d and P_{HF}^u in the trade balance condition (8) correspond to the CIF prices paid by Foreign buyers, so the domestic price collected by Home sellers of those goods are $(1-v_H^d)P_{HF}^d/\tau^d$ and $(1-v_H^u)P_{HF}^u/\tau^u$, respectively. The ratios in the right-hand-side of equations (13) and (14) thus capture the private relative cost of domestic versus foreign goods.

The next optimality condition balances the private benefits of exporting domestic inputs (net of an export tax) against the private benefits of using those domestic inputs to produce an additional (potentially subsidized) amount of the final good that is in turn exported:

$$\left(1-v_H^u\right)\frac{P_{HF}^u}{\tau^u} = \hat{A}_H^d F_{Q_{HH}^u}^d\left(L_H^d, Q_{HH}^u, Q_{FH}^u\right)\frac{\left(1-v_H^d\right)}{\left(1-s_H^d\right)}\frac{P_{HF}^d}{\tau^d}. \tag{15}$$

Notice that this condition, which relates to the planner's optimality condition in equation (11), links the downstream subsidy and the two export taxes. Intuitively, in the open economy, Home faces a tradeoff between exporting its inputs versus using them to produce its own final goods which it then exports. When final-good production occurs under increasing returns, this tradeoff must incorporate the extent to which increasing the size of the downstream sector will in turn raise its productivity. The role of downstream returns is thus relevant because of Home's ability to trade inputs, and its desire to expand the downstream sector to raise its efficiency. The role of upstream returns is irrelevant here, since in either case the upstream sector size is held constant.

The final efficiency condition equates the private marginal product of labor in both sectors in terms of a common good (i.e., the final good)

$$F^d_{L^d_H}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right) = \frac{1}{1-s^u_H}\hat{A}^u\left(L^u_H\right)F^d_{Q^u_{HH}}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right). \tag{16}$$

Notice that this final constraint includes no trade instruments, and the upstream subsidy is absent in all the other conditions. These features indicate that achieving the optimal allocation of labor across sectors requires addressing a fundamental misallocation of the closed economy, whereas the other conditions required to achieve the first-best arise only in an open economy. As a result, we will show that trade instruments cannot be used to satisfy this final constraint, whereas they can ensure that equations (13) to (15) are satisfied.

3.3 First-Best Policies: Upstream and Indeterminate Policies

We now characterize the set of first-best policies by identifying the combinations of taxes and subsidies that equate the equilibrium conditions (13) to (16) with the social efficiency conditions (9) to (12). There are two key differences between these two sets of equations. First, the market equilibrium conditions naturally incorporate the effects of trade taxes and production subsidies on consumer and firm private choices. Second, these decentralized-market equations do *not* incorporate the positive impact of sectoral output on productivity, or the positive effects of curtailing exports on Home's terms of trade.

Inspection of equations (12) and (16) reveals that attaining the first-best allocations requires an upstream production subsidy equal to

$$s^u_H = \frac{\gamma^u}{1+\gamma^u}. \tag{17}$$

Intuitively, the market allocates too little labor upstream because input producers do not internalize the impact of their sector size on their own productivity, which in turn affects downstream productivity. An upstream subsidy equal to $\gamma^u/\left(1+\gamma^u\right)$ increases hiring upstream, raising the sector size and thus its productivity. As a result, and despite pulling labor from downstream firms, the subsidy leads to a lower market price for inputs, greater input production, and thus greater downstream output.

It is clear from equation (17) that the optimal subsidy is only needed when upstream production features increasing returns (i.e., $\gamma_u > 0$). Higher increasing returns upstream raise the efficiency gains from a larger input sector, thus also raising the size of the optimal subsidy. Since the subsidy increases aggregate output by reallocating labor upstream to raise input efficiency, it is also only relevant when the downstream sector uses labor to produce (i.e., $\alpha > 0$).

Although the first-best policies necessarily include an upstream subsidy whenever downstream production uses labor ($\alpha > 0$) and upstream production features increasing returns ($\gamma^u > 0$), the nature and levels of the other instruments are indeterminate. The five remaining instruments (the downstream subsidy s^d_H, import tariffs t^d_H and t^u_H, and export taxes v^d_H and v^u_H) must simply satisfy the following three equations:

$$\frac{1+t_H^d}{1-v_H^d} = \frac{\sigma}{\sigma-1}; \tag{18}$$

$$\frac{1+t_H^u}{1-v_H^u} = \frac{\theta}{\theta-1}; \tag{19}$$

$$\left(1-s_H^d\right)\frac{1-v_H^u}{1-v_H^d} = \frac{\frac{\theta-1}{\theta}}{\frac{\sigma-1}{\sigma}}\frac{1}{(1+\gamma^d)}. \tag{20}$$

Equations (18) and (19) relate to equations (9) and (10), and thus reflect Home's terms-of-trade motives to restrict its exports of final goods and inputs in order to exploit its market power abroad. In line with the well-known application of Lerner Symmetry in single-sector models with differentiated goods, Home might achieve each of these individual objectives using an import tariff, an export tax, or a combination of both. An export tax raises the relative price of Home goods in foreign directly, while an import tariff shifts domestic expenditure towards Home varieties, thus also raising Home's terms of trade.[11]

Although the first two constraints suggest that import tariffs and export taxes are perfect substitutes as policy tools, equation (20) shows that this is not the case in our vertical model. When deciding whether to export inputs directly versus embodied in final goods, Home's social planner must also incorporate the productivity-enhancing effects of a larger final-good sector (i.e., the term $1+\gamma^d$ on the right-hand side). Home's decision to export inputs in these two forms thus introduces a new constraint that links its optimal export taxes in both sectors to the extent of increasing returns in downstream production.

Because there are five policy instruments and only three equations, equations (18) to (20) represent an overdetermined system of equations with a continuum of potential solutions. This indeterminacy implies that out of the five instruments (s_H^d, t_H^d, t_H^u, v_H^d and v_H^u), two can be set to an arbitrary level, and when this arbitrary level is 0, the first-best is achieved with an upstream subsidy and three (or fewer) other instruments. As a result, in the general case in which the upstream sector features increasing returns to scale ($\gamma^u > 0$) and downstream production uses labor and inputs ($\alpha > 0$), the upstream subsidy is the only essential policy instrument to achieve the first-best allocation. We summarize this result in the following:

Proposition 1. When both sectors use labor and the upstream sector features increasing returns to scale, the first-best allocation can only be attained with an upstream subsidy set at $(s^u)^* = \gamma^u/(1+\gamma^u)$. By contrast, a downstream production subsidy is not necessary to achieve the first-best.

Proof. Follows from equations (17)–(20) . □

[11] Although we assume that the elasticities of substitution are the same in Home and Foreign, our analysis does not invoke the fact that Home preferences $U_H\left(Q_{HH}^d, Q_{FH}^d\right)$ are a CES aggregator governed by σ, or that the production function $F^d\left(L_H^d, Q_{HH}^u, Q_{FH}^u\right)$ is a Cobb-Douglas-CES aggregator governed by α and θ. As a result, the parameters σ and θ in the first-best policies relate solely to *Foreign* elasticities of substitution. If these elasticities differ across countries, with $\sigma_H \neq \sigma_F$ and $\theta_H \neq \theta_F$, only σ_F and θ_F affect the wedge between private and socials costs.

The asymmetry in the need for upstream versus downstream production subsidies arises because our vertical model features no domestic (mis)allocation of expenditure across final-good sectors. While the upstream subsidy addresses a closed-economy misallocation of labor across sectors, the downstream subsidy is motivated *only* by the open-economy decision about how to export inputs.[12]

3.4 First-Best Policies: Alternative Implementations

We now consider three alternative implementations of the first-best allocation. Since Proposition 1 indicates that any implementation must include an upstream production subsidy, $s^u_H = \gamma^u / (1+\gamma^u)$, we focus on the remaining instruments.

A first possibility, highlighted by Lashkaripour and Lugovskyy (2023), is to attain the first-best using only production subsidies and export taxes. When import tariffs are zero ($t^d_H = t^u_H = 0$), the above system (18) to (20) implies that $v^d_H = 1/\sigma$, $v^u_H = 1/\theta$, and $s^d_H = \gamma^d / \left(1+\gamma^d\right)$. Under this implementation, the upstream and downstream export taxes exploit Home's market power abroad optimally, while the downstream subsidy raises Home's final-good production in line with the extent of its increasing returns.

Although this implementation appears simple and intuitive – export taxes exploit market power abroad in each sector and production subsidies seem to internalize the presence of increasing returns – the symmetry in taxes and subsidies for inputs and final goods obscures the distinct motives for upstream versus downstream policy intervention in our setting. As summarized in Proposition 1, and in contrast to the results in Lashkaripour and Lugovskyy (2023), the downstream production subsidy is *not* required to maximize social welfare in our model.

We demonstrate the irrelevance of the downstream subsidy with alternative implementations of the first-best that do not use it. Setting $s^d_H = 0$, the first-best can still be achieved again using the upstream subsidy from Proposition 1, an upstream export tax $v^u_H = 1/\theta$ and import tariff $t^u_H = 0$ as above, but now with a downstream export tax $v^d_H = 1 - \left(1+\gamma^d\right)\left(\frac{\sigma-1}{\sigma}\right)$ and tariff $t^d = \gamma^d$. Alternatively, we could also use import tariffs $t^d_H = \frac{1}{\sigma-1}$ and $t^u_H = \frac{\sigma}{\sigma-1}\frac{1}{1+\gamma^d} - 1$, and an upstream export tax $v^u_H = 1 - \frac{\theta-1}{\theta}\frac{\sigma}{\sigma-1}\frac{1}{1+\gamma^d}$ to achieve the first-best. Although the levels of tariffs and export taxes are still indeterminate in these implementations without a downstream subsidy, the ratio of optimal gross import tariffs on final goods and inputs, or the "tariff escalation wedge" is unique and given by

$$\frac{1+t^d_H}{1+t^u_H} = 1 + \gamma^d > 1.$$

We summarize this result in the following:

Proposition 2. The first-best allocation can be achieved with a production subsidy for inputs, and (at least two) trade taxes associated with a tariff escalation wedge equal to $1+\gamma^d$, or $(1+t^d_H)/(1+t^u_H) = 1+\gamma^d > 1$.

[12]In the next section, we show that the upstream production subsidy is not only necessary, but also sufficient to maximize welfare in the closed economy.

Proof. Follows from equations (17)–(20). □

As evident in equations (15) and (20), Home weighs the benefit of exporting inputs to Foreign directly against the benefit of using those inputs to produce and export final goods. Since the benefit of the latter is increasing in the extent of downstream returns to scale, a policy that raises the size of that sector in line with those returns, either via a production subsidy or by redirecting domestic expenditure towards Home's final goods via a downstream tariff, maximizes welfare. Indeed, plugging equations (18) and (19) into (20) we obtain

$$\frac{1+t_H^d}{1+t_H^u}\frac{1}{1-s_H^d} = 1+\gamma^d.$$

This expression demonstrates that tariff escalation can perfectly substitute for a downstream production subsidy in achieving the first-best whenever export taxes are available. The fact that tariff escalation arises in the first-best when the downstream subsidy is zero is particularly useful in understanding second-best policies in which subsidies are not available. When the planner is constrained to using only import tariffs, the same forces we highlight here tend to make tariff escalation a welfare-maximizing policy.

The asymmetry between an input and final-good tariff in substituting for a domestic production subsidy can also be understood from the asymmetry in spillovers across sectors. Although both final-good and input tariffs shift domestic expenditure towards Home varieties, thereby raising the size and efficiency of their respective sectors, the upstream tariff simultaneously affects the marginal rate of substitution between domestic and foreign inputs – see equation (14) – which reduces downstream efficiency. By contrast, a downstream tariff has no negative productivity spillover and instead also raises demand for Home inputs (see equation (13)). When the government does not have access to production subsidies, these forces tend to make downstream import tariffs disproportionately more desirable than input tariffs, thus leading to tariff escalation.

Finally, we describe a third implementation that includes production subsidies in both sectors, but rules out export taxes. Setting $v_H^d = v_H^d = 0$, it is straightforward to conclude from equations (18) to (20) that the first best can be attained whenever $t_H^d = 1/(\sigma-1)$, $t_H^u = 1/(\theta-1)$, $s_H^d = 1-\left(\frac{\theta-1}{\theta}\right)/\left[\frac{\sigma-1}{\sigma}\left(1+\gamma^d\right)\right]$, and of course $s_H^u = \gamma^u/(1+\gamma^u)$. In this implementation, import tariffs exploit Home's market power in export markets optimally, but now the downstream subsidy must offset the impact of the input tariff on Home firms' higher input costs. As a result, the optimal subsidy depends not only on the size of the downstream returns to scale, but also on the size of the upstream tariff. This implementation thus illustrates why the government's use of an input tariff is relatively more costly when downstream production occurs under increasing returns: Home's use of an input tariff to exploit its market power abroad must be offset by a counteracting policy.

This section demonstrates the distinct motives for input versus final-good tariffs in achieving a first-best allocation with vertically linked sectors. In the next section, we consider several simplifications of our model to provide additional intuition for the sources of these differences.

4 Economic Primitives that Drive Tariff Escalation

To highlight the primitive economic forces that lead tariff escalation to maximize social welfare, we now present four simplified versions of our model. Our goal is to build intuition for the mechanisms that drive the second-best allocations we analyze in Sections 5 and 6.

4.1 Closed-Economy Equilibrium and Efficiency

We first characterize the tools necessary to achieve the first-best allocation in the closed economy. This variant of the model provides additional intuition for why the first-best requires a subsidy to upstream but not downstream production. In the closed economy, the marginal rates of substitution between domestic and foreign goods, and the no-arbitrage condition captured by equations (11) and (15) become irrelevant. Optimality is thus dictated *solely* by the labor-market efficiency conditions captured in (12) and (16), which are unchanged in the closed economy. It is clear that an upstream production subsidy equal to $s_H^u = \gamma^u / (1 + \gamma^u)$ is necessary and sufficient to restore efficiency. No other policy instrument, and in particular, no downstream subsidy, is needed to achieve the first-best.

This result is distinct from predictions in models with horizontally differentiated sectors featuring scale economies (Lashkaripour and Lugovskyy, 2023; Bartelme et al., 2021). In those settings, downstream subsidies are required to achieve the first-best, and only their relative size is pinned down by the optimality conditions, while their absolute levels are not. In our model with vertically linked sectors, the optimal upstream subsidy is necessarily equal to $\gamma^u / (1 + \gamma^u)$ and downstream subsidies are irrelevant.[13]

The results in our setting are different from past work since there is no (mis)allocation of domestic expenditure across final-good sectors here. Instead, the planner seeks to balance the efficiency gains from employing workers upstream versus downstream. Although greater employment in either sector will raise that sector's efficiency, only increased upstream efficiency has a follow-on effect on downstream output, as evident in equation (12), where $\left(1 + \gamma^d\right)$ cancels on both sides, but $(1 + \gamma^u)$ does not.

As in the open economy, the size of the optimal subsidy is increasing in the extent of the returns to scale in upstream production. Moreover, if $\gamma^u = 0$, the upstream sector features constant returns to scale and the decentralized equilibrium delivers the socially optimal amount of upstream production. Similarly, a subsidy is only required to maximize social welfare when the downstream sector uses both labor and inputs. If there is no input sector ($\alpha = 1$), or if the downstream sector uses no labor ($\alpha = 0$), there is no scope for the planner to reallocate workers upstream, and thus no role for policy to increase social welfare. Despite the two extremes, the upstream subsidy's potential to raise welfare tends to rise with the downstream labor share (α). This relationship arises because the difference between the market versus social planner upstream-labor allocation is increasing in the downstream labor share.

To demonstrate the role of the downstream labor share in the closed-economy labor misallocation,

[13] Bartelme et al. (2021) obtain a similar result in an extension with vertical links.

we consider the special case in which the downstream technology is Cobb-Douglas, as in equation (2). In this case, the decentralized allocation of labor to the upstream sector satisfies

$$\frac{L_H^u}{L_H} = \frac{1-\alpha}{\alpha\left(1-s_H^u\right)+1-\alpha}.$$

Without government intervention ($s_H^u = 0$), the market would thus allocate a share $1-\alpha$ of labor to the upstream sector, while the social planner would instead set

$$\frac{\left(L_H^u\right)^*}{L_H} = \frac{\left(1-\alpha\right)\left(1+\gamma^u\right)}{\alpha+\left(1-\alpha\right)\left(1+\gamma^u\right)} \geq 1-\alpha. \tag{21}$$

Equation (21) thus shows that the relative size of the planner versus market upstream labor allocation is increasing in the downstream labor share (α).

We will use the facts that the optimal upstream subsidy is increasing in the extent of upstream returns, and that the degree of labor misallocation across sectors is increasing in the downstream labor share to build intuition for the second-best policies we study in Section 5. We summarize our closed-economy results as follows:[14]

Proposition 3. In the closed economy, the social planner can restore efficiency in the market equilibrium by subsidizing upstream production at a rate $\left(s^u\right)^* = \gamma^u/\left(1+\gamma^u\right)$. This subsidy is only needed when the downstream sector uses both inputs and labor to produce ($0<\alpha<1$) and when $\gamma^u > 0$. Downstream subsidies are irrelevant in the closed economy.

Proof. See Appendix B.1. □

4.2 A Special Case: An Isomorphic Krugman Economy

The results above are based on a competitive model without firm-level decisions on entry, exporting, importing, and pricing. However, our results also apply in a setting with monopolistic competition, firm-level returns to scale, and free entry. Extending results in Kucheryavyy et al. (2023) to a vertical setting, in Appendix B.2 we obtain an identical set of first-best allocation conditions and decentralized equilibrium conditions in such a Krugman-style model. There are two subtle aspects of this isomorphism. First, the productivity terms $\bar{A}_i^d$ and $\bar{A}_i^u$ in equations (3) and (4) are a function of the primitive parameters of the 'Krugman' model (including final-good and input elasticities of substitution and the levels of fixed costs). Second, the scale elasticities γ^d and γ^u are no longer free parameters and instead are given by $\gamma^d = 1/\left(\sigma-1\right)$ and $\gamma^u = 1/\left(\theta-1\right)$, respectively.

The Krugman model further highlights the distinctions between the social planner's motives for an input versus final-good tariff. Returning to the second implementation summarized in Proposition

[14] In Appendix B.1, we develop two extensions of our closed-economy model. First, we allow the upstream sector to use the same bundle of inputs Q^u used in the final-good sector, while letting labor intensity upstream (denoted by β) differ from that downstream. Second, we outline a multi-stage extension of the model in which the input bundle used in upstream production aggregates varieties from a yet more upstream sector, which in turns uses inputs from an even more upstream sector, and so on. In both extensions, we show that efficiency again calls for the use of subsidies in all input sectors, but not in the most downstream sector.

2, equations (17) to (20) show that the first-best can be attained with *only* two instruments in addition to the upstream subsidy: a downstream tariff set at $1+t_H^d = 1+\gamma^d = \sigma/\left(\sigma-1\right)$, and an upstream export tax set at $v_H^u = 1/\theta$. As in the general model, tariff escalation depends on the extent of increasing returns to scale in downstream production. Furthermore, no other implementation can achieve the first-best with a lower number of policy instruments.[15]

In subsequent sections, we use this *isomorphic Krugman economy* to highlight the fact that our main results are driven by the existence of increasing returns to scale, regardless of whether these are external or internal to firms.

4.3 The Case with No Scale Economies Upstream

In our model, the degree of scale economies in each sector is governed by the parameters γ^u and γ^d, so when $\gamma^u \to 0$, this economy converges to a competitive economy with no scale economies upstream. Inspecting equations (17) to (20), it is straightforward to derive first-best trade policies in this case. Specifically, the upstream subsidy in equation (17) becomes a redundant instrument and can be set to 0, while the system of equations (18) to (20) is still indeterminate and allows for various implementations of the first-best, as outlined above.

A key aspect of this variant is that the first-best can now be attained using *only* trade taxes, and both production subsidies are redundant policy instruments. Furthermore, and despite the indeterminacy implied by the system (18) to (20), any implementation of the first-best that uses only trade taxes necessarily features a tariff escalation wedge given by

$$\frac{1+t_H^d}{1+t_H^u} = 1+\gamma^d > 1.$$

The simplest implementation is to use export taxes in each sector to exploit Home's market power abroad, a final-good tariff to benefit from increasing returns downstream, and no input tariff whatsoever. Such an implementation is even simpler in the isomorphic Krugman economy, since the downstream scale elasticity must satisfy $1+\gamma^d = \sigma/\left(\sigma-1\right)$. The first-best can then be achieved with only two instruments: a downstream import tariff $t_H^d = \gamma^d = 1/\left(\sigma-1\right)$ and an upstream export tax $v_H^u = 1/\theta$.

The results above hold for any environment in which production subsidies are redundant, so they also apply when the downstream sector uses no labor to produce ($\alpha = 0, \gamma^u > 0$), since in that case all labor is employed upstream in both the decentralized equilibrium and the first-best allocation. In sum, we have derived the following result:

Proposition 4. When $\gamma^u = 0$ or when $\alpha = 0$, the first-best allocation can be achieved with a combination of import and export trade taxes. Although the levels of trade taxes are not uniquely determined, the tariff escalation wedge is necessarily given by $\left(1+t_H^d\right)/\left(1+t_H^u\right) = 1+\gamma^d > 1$. Furthermore, in the isomorphic Krugman economy with internal economies of scale downstream,

[15] By contrast, the implementation suggested by Lashkaripour and Lugovskyy (2023) still features all four tax instruments: two production subsidies and two export taxes.

the first-best can be achieved with just a downstream import tariff t_H^d equal to $\gamma^d = 1/(\sigma - 1)$ and an upstream export tax v_H^u equal to $1/\theta$.

Proof. Follows from equations (17)–(20). □

Why do trade policies that implement the first-best involve higher import tariffs on final goods than inputs? And why can the government tax imports of final goods, but must tax exports of inputs when using the minimum set of instruments to attain the first-best?

The key distinction between trade taxes on final goods versus inputs can be understood as follows. A downstream import tariff or export tax both shift domestic consumers' expenditures towards Home final goods, thereby improving Home's terms of trade (as in Gros, 1987) and raising downstream productivity by expanding the sector's size (as in Ossa, 2011). An upstream import tariff similarly redirects expenditure towards Home inputs, boosting Home's terms of trade and the upstream sector size and efficiency. Although the increased upstream efficiency lowers the market price of Home inputs, it does so only by raising the price of Foreign inputs and distorting the marginal rate of substitution between domestic and foreign inputs (see equation (10)). While a downstream tariff also distorts the marginal rate of substitution between domestic and foreign final goods, a crucial difference is that inputs are sold to firms that produce under increasing returns to scale and can 'relocate' to Foreign, whereas final goods are sold to consumers who cannot move or experience efficiency gains. This asymmetry is evident in equation (11), which captures Home's need to balance the relative prices of its exported inputs and final goods against the potential efficiency gains from expanding downstream production. These differences create a disproportionate incentive for Home to manipulate its terms of trade in the input sector via an export tax, since doing so shifts domestic input expenditure towards Home producers without raising downstream firms' Foreign input costs.

4.4 Case with No Scale Economies Downstream

The underlying economic primitive that leads tariff escalation to maximize social welfare is increasing returns to scale in downstream production. To demonstrate the role of these increasing returns, we analyze an economy in which final goods are produced under constant returns to scale ($\gamma^d \to 0$). The only change to equations (18) to (20) characterizing the first-best policies in this setting is in the term $1 + \gamma^d$ on the right-hand-side of equation (20). Provided that $\gamma^u > 0$, attaining the first-best still requires an upstream subsidy equal to $s_H^u = \gamma^u/(1 + \gamma^u)$. If we rule out the use of a downstream production subsidy, the system (18) to (20) immediately implies that all remaining implementations entail a common import tariff for final goods and inputs, which we summarize as follows:[16]

[16] Of course whenever the planner has more instruments than constraints, she can use those redundant instruments to offset the unnecessary use of other instruments. For example, the planner could maximize welfare with a higher upstream versus downstream tariff (i.e., de-escalation), but would then need an additional downstream subsidy to offset the welfare-reducing impact of tariff de-escalation in this setting.

Proposition 5. In the absence of scale economies downstream ($\gamma^d = 0$), the first-best can be attained with an upstream subsidy equal to $s_H^u = \gamma^u / (1+\gamma^u)$ and a combination of import and export taxes. Although, the levels of trade taxes are not uniquely determined, the tariff escalation wedge $\left(1+t_H^d\right)/(1+t_H^u)$ necessarily equals 1. Furthermore, the first-best can be achieved with just the upstream subsidy, a downstream export tax $v_H^d = 1/\sigma$ and an upstream export tax $v_H^u = 1/\theta$.

Proof. Follows from equations (17) to (20). □

This result shows that the emergence of tariff escalation in our framework is directly tied to the existence of scale economies *in the downstream sector*. In their absence, we obtain a result analogous to that derived by Costinot et al. (2015) and by Beshkar and Lashkaripour (2020), namely that optimal trade policy involves uniform import tariffs across sectors (regardless of their differentiation or whether they are inputs or final goods) and differential export taxes that optimally exploit Home's market power.[17]

In our vertical model, common input and final-good import tariffs maximize social welfare when both sectors produce under constant returns to scale. This result suggests that any welfare-maximizing rationale for tariff escalation is unrelated to the production efficiency result in Diamond and Mirrlees (1971). Those authors demonstrate that under constant returns to scale, the optimal choice of domestic commodity taxes (aimed at raising a given amount of government revenue) should be zero for intermediate inputs. Our planner does not set trade taxes to raise revenue *per se*, but rather to exploit Home's market power abroad. Tariff escalation arises because the Home planner can use a final-good import tariff to exploit efficiency gains from a larger downstream sector. By contrast, an input tariff cannot perfectly replace an upstream subsidy.

5 Second-Best Import Tariffs for a Small Open Economy

A central contribution of this paper is to study the motives for tariff escalation when import tariffs are the only available policy instrument. This analysis is crucial for three reasons. First, because the planner no longer has redundant instruments, the only motives for tariff escalation must arise from primitives of the underlying economy (instead of from unnecessary policy interventions such as a downstream subsidy). Second, this setting is more realistic for many countries and political environments. Export taxes are rarely used in practice and expressly prohibited in some countries, (e.g., see the U.S. Constitution, Article 1, Section 9, Clause 5). Similarly, production subsidies are constrained by the WTO Agreement on Subsidies and Countervailing Measures and may be harder to implement or fund than a tax on foreign goods. Third, this setting is most comparable to a large body of trade theory that rules out production subsidies to focus on trade policy determination (e.g., Grossman and Helpman, 1994; Ossa, 2011, among many others).

[17] When applying the argument above to the Krugman economy, the isomorphism conditions impose $\gamma^d \to 0 = 1/(\sigma-1)$. This suggests that if $\gamma^d \to 0$, we must necessarily have $\sigma \to \infty$, which would imply that the Home economy has no market power in downstream markets. Note, however, that even in such a limiting case, the model without downstream scale economies still fails to generate tariff escalation.

The main conclusion from Proposition 1 is that an upstream production subsidy is necessary to achieve the first-best, while analysis of Proposition 2 indicates that a downstream tariff can perfectly substitute for a downstream subsidy. A natural question is thus whether an upstream tariff will similarly, albeit imperfectly, substitute for the missing upstream subsidy and thus make tariff escalation less likely. To analyze this question, we return to the primal approach developed in Costinot et al. (2015).

The second-best optimal allocation seeks to solve the same problem laid out at the beginning of Section 3.1, though now with two additional constraints. Absent export taxes and production subsidies, the decentralized market optimality conditions (15) and (16) reduce to

$$\hat{A}^d_H F^d_{Q^u_{HH}}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right) = \frac{P^u_{HF}/\tau^u}{P^d_{HF}/\tau^d} = \frac{\left(Q^u_{HF}\right)^{-\frac{1}{\theta}} P^u_{FF}\left(Q^u_{FF}\right)^{\frac{1}{\theta}}/\tau^u}{\left(Q^d_{HF}\right)^{-\frac{1}{\sigma}} P^d_{FF}\left(Q^d_{FF}\right)^{\frac{1}{\sigma}}/\tau^d} \tag{22}$$

and

$$F^d_{L^d_H}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right) = \hat{A}^u\left(L^u_H\right) F^d_{Q^u_{HH}}\left(L^d_H, Q^u_{HH}, Q^u_{FH}\right). \tag{23}$$

Since neither of these of conditions can be *directly* affected by import tariffs, they constitute additional constraints that the second-best allocation must satisfy.

5.1 Second-Best Import Tariffs for a Small Open Economy with No Domestic Distortions

We first study second-best import tariffs in the special case in which the market equilibrium features no misallocation of labor across sectors, and equation (23) can thus be ignored. There is no labor misallocation if it is all employed upstream (i.e., $\alpha = 0, \gamma^u \geq 0$), or when the upstream sector produces under constant returns to scale ($\alpha \geq 0, \gamma^u = 0$).

When there is no domestic labor misallocation, Proposition 4 shows that the first-best can be achieved using export taxes that exploit Home's market power abroad and a uniquely determined tariff escalation wedge that depends solely on the extent of increasing returns to scale downstream. This setting thus illustrates the effect of ruling out export taxes on the optimal degree of tariff escalation.[18]

The emergence of tariff escalation is tightly related to the importance of downstream scale economies. With that in mind, we now assume that the scale elasticity γ^d is large enough to ensure that:

Assumption 1. $1 + \gamma^d > \frac{(\theta-1)/\theta}{(\sigma-1)/\sigma}$.

This assumption is satisfied in two natural scenarios. First, when trade elasticities are relatively similar upstream and downstream ($\theta \simeq \sigma$), Assumption 1 simply imposes the existence of scale economies downstream (or $\gamma^d > 0$). Second, in the isomorphic Krugman economy with internal

[18] Identical results are also obtained when $\alpha > 0$ and $\gamma^u > 0$, if the government has access to the two import tariffs and an upstream subsidy set at $s^u_H = \gamma^u/(1+\gamma^u)$, in line with Proposition 2.

economies of scale, we have $1 + \gamma^d = \sigma/(\sigma - 1)$, and Assumption 1 is automatically satisfied for any input elasticity $\theta > 1$.

We compare the first-order conditions of the planner problem to the decentralized market equilibrium using only import tariffs and in Appendix A.2.2 establish that:

Proposition 6. Under Assumption 1, when $\gamma^u = 0$ or when $\alpha = 0$, the second-best optimal combination of import tariffs involves a tariff escalation wedge larger than the first-best wedge, so $(1 + t_H^d)/(1 + t_H^u) > 1 + \gamma^d > 1$.

Proof. See Appendix A.2.2. □

To understand this result, recall that, in the first-best allocation, Home used the now-missing export taxes to exploit its market power abroad. These market-power motives are easily replicated by import tariffs (see equations (18) and (19)), but our vertical model features an additional constraint linking upstream and downstream production. While an upstream export tax or tariff both boost the size of the domestic downstream sector by redirecting Home's inputs to its final-good producers, an upstream tariff raises final-good producers' costs, which lowers downstream production and thus its efficiency (recall that when $\alpha = 0$ or $\gamma^u = 0$, there is no scope for policy to raise upstream efficiency). The incentive to levy import tariffs is thus tempered by this negative spillover across sectors.

Whenever Assumption 1 holds, i.e., $1 + \gamma^d > \frac{(\theta-1)/\theta}{(\sigma-1)/\sigma}$, the balance of these forces leads to a second-best tariff escalation wedge $(1 + t_H^d)/(1 + t_H^u)$ that is above one, and in fact higher than $1 + \gamma^d$ (the first-best wedge when the downstream subsidy is zero). This condition holds for any $\gamma^d > 0$ whenever $\sigma = \theta$; even when input and final-good demand elasticities are symmetric, the vertical link in equation (20) calls for a relatively high cost to export inputs over final goods, which maps to a higher final-good tariff. We state this corollary as follows:

Corollary 1. For common Foreign demand elasticities for final goods and inputs ($\sigma = \theta$), the second-best optimal combination of import tariffs involves a tariff escalation wedge larger than the first-best wedge, $(1 + t_H^d)/(1 + t_H^u) > 1 + \gamma^d > 1$, for any $\gamma^d > 0$.

When export taxes are not available and there is no domestic labor misallocation, tariff escalation is thus higher in the second-best. This higher degree of escalation arises from the downstream efficiency gain Home experiences from redirecting potential exports of inputs to its own downstream production, and underpins the driving force behind tariff escalation.

Although Assumption 1 is a sufficient condition for second-best tariffs to feature escalation, it is not necessary. Whenever $1 + \gamma^d < \frac{(\theta-1)/\theta}{(\sigma-1)/\sigma}$, tariff escalation will be below $1 + \gamma^d$, but it may remain above one. Tariff de-escalation only arises as an optimal policy when γ^d is sufficiently low and when σ is small relative to θ. Although our discussion above explains why tariff escalation is increasing in downstream returns, it is perhaps surprising that tariff de-escalation arises when the ratio of σ to θ is low. In fact, the standard terms-of-trade rationale would seem to imply that relatively more elastic final-good demand in Foreign should dictate a relatively *lower* downstream tariff. To shed further light on this counterintuitive result, we briefly characterize the second-best import tariffs in the absence of scale effects in either sector.

The Case of No Scale Economies To solve for second-best import tariffs under constant returns to scale, we only need to consider the case when $\gamma^u \to 0$ and $\gamma^d \to 0$ in our competitive economy with external scale economies. In Appendix A.2.2, we prove the following result:

Proposition 7. In the absence of scale economies in either sector (or $\gamma^d = \gamma^u = 0$), the second-best optimal combination of import tariffs features escalation (i.e., $\left(1 + t_H^d\right) / \left(1 + t_H^u\right) > 1$) if and only if $\sigma > \theta$.

Proof. See Appendix A.2.2. □

Under constant returns to scale, Proposition 7 states that tariff escalation is more likely, the *lower* is Home's relative degree of downstream to upstream market power. To understand this result, it is useful to focus first on the case of common demand elasticities. In a competitive Ricardian model, as long as optimal export taxes are common across sectors (i.e., $\sigma = \theta$ in our setting), the first-best can be implemented either via uniform export taxes or uniform import tariffs (Costinot et al., 2015; Beshkar and Lashkaripour, 2020). As a result, second-best import tariffs are sufficient to achieve the first-best allocations, and will necessarily be equal across sectors (regardless of their upstreamness). In such a case, this competitive model with no scale effects delivers *no* tariff escalation.

Moving from this benchmark, when $\sigma \neq \theta$, the first-best can no longer be implemented using only import tariffs. Indeed Beshkar and Lashkaripour (2017) show that in a horizontal economy without vertically linked sectors, second-best import tariffs deliver lower social welfare and remain uniform across sectors. By contrast, when sectors have vertical links, Beshkar and Lashkaripour (2020) find that import tariffs partly mimic the effects of missing export taxes by raising the relative price of downstream exports. If the planner would like to set a higher export tax in one sector relative to another, he can thus adjust the relative size of the second-best import tariff on inputs to achieve (some of) the desired differential terms-of-trade manipulation. When Home has greater market power downstream ($\sigma < \theta$), the desired export tax is higher downstream, which leads the government to impose relatively higher input tariffs (i.e., tariff de-escalation). Conversely, when $\sigma > \theta$, the planner prefers a lower export tax downstream, which he partly achieves by setting a lower tariff on intermediate inputs (i.e., tariff escalation), in turn reducing the relative price of downstream exports.

These results illustrate the fact that the welfare rationale we document for tariff escalation is orthogonal to the production efficiency result in Diamond and Mirrlees (1971). While those authors show that under constant returns to scale, domestic inputs should face lower (or no taxes), we show that second-best import tariffs on inputs may well be higher than those applied to final goods. The systematic desirability of tariff escalation instead arises under increasing returns to scale in downstream production.

5.2 Second-Best Import Tariffs for a Small Open Economy with Domestic Distortions

We now analyze second-best import tariffs in the more general case in which downstream production uses labor and inputs, and both sectors produce under increasing returns to scale ($0 < \alpha < 1$, $\gamma^d > 0, \gamma^u > 0$). In this environment, the domestic labor-market allocation is inefficient, and import tariffs are used not only to replace the missing export taxes, but also to (attempt to) replicate the effects of the missing upstream subsidy.

When both sectors use labor ($\alpha > 0$) and there are scale economies upstream ($\gamma^u > 0$), the upstream sector's size and thus productivity are directly affected by the amount of labor it employs. As a result, the Home government can increase upstream efficiency, and possibly welfare, by shifting domestic expenditure towards Home inputs. When production subsidies are unavailable, these potential efficiency gains generate an additional motive for an upstream tariff. The greater potential efficiency gains from increasing the upstream sector size, the more desirable an input tariff will be.

The size of the downstream labor share (α) generates two opposing forces on the welfare effects of an upstream efficiency gain. On the one hand, the difference between the amount of labor allocated upstream by the planner versus the competitive market is increasing in the final-good sector's labor share (see Section 4.1). On the other hand, as the downstream labor share rises, inputs become less important and thus any efficiency gains in that sector have small effects on welfare. The first force tends to dominate such that a high labor share generates a larger welfare gain from the subsidy (and hence a higher input tariff in the second-best), up to a point when the downstream input usage is so low that any input efficiency gains are too small to matter.

Despite these countervailing forces, we obtain a set of sufficient conditions under which tariff escalation is a feature of second-best import tariffs. First, we again impose the (natural) lower bound on scale economies downstream in Assumption 1. Second, we impose the following upper bound on scale economies upstream γ^u:

Assumption 2. $1 + \gamma^u \leq \theta / (\theta - 1)$.

This condition relates the returns to scale upstream to Home's market power abroad: upstream returns cannot be too high and input demand cannot be too elastic. As for Assumption 1, Assumption 2 is automatically satisfied in the isomorphic Krugman economy with internal economies of scale, where $1 + \gamma^u = \theta / (\theta - 1)$.

With these assumptions in hand, we show in Appendix A.2.3 that:

Proposition 8. Under Assumptions 1 and 2, and provided that

$$\left(1 + \gamma^d\right) \min\left\{1 - \alpha\gamma^u \frac{\theta - 1}{\theta}, (1 - \alpha) \frac{\theta}{\theta - 1}\right\} \geq 1, \tag{24}$$

second-best import tariffs feature tariff escalation, so $\left(1 + t_H^d\right) / \left(1 + t_H^u\right) > 1$.

Proof. See Appendix A.2.3. ∎

Within the bounds on γ^d and γ^u in Assumptions 1 and 2, Proposition 8 indicates that tariff escalation is a feature of second-best import tariffs whenever α or θ are sufficiently low, and whenever γ^d is sufficiently large. As the downstream labor share (α) falls, the difference between the market and planner upstream labor allocations also decreases, so the need for an upstream tariff to mimic the missing upstream subsidy lessens.

While one might expect relatively more market power in inputs (lower θ) to imply a relatively higher input tariff (and thus de-escalation), the same intuition from Proposition 7 applies here. In the absence of scale economies, tariff escalation is increasing in the relative size of final-good to input demand elasticities (σ/θ). An input tariff partly mimics a missing downstream export tax, and the desirability of such a downstream tax is higher when final-good demand is relatively more inelastic. Indeed, when our model features constant returns to scale in both sectors, second-best optimal tariffs are escalated if and only if $\sigma > \theta$, even when $\alpha > 0$ (see Appendix A.2.2). It is thus less surprising that relatively inelastic demand for inputs (low θ) makes it relatively more likely that second-best import tariffs are escalated in the more general case of increasing returns.

It is important to stress that the conditions in Proposition 8 are sufficient but *not necessary* for tariff escalation to be optimal in the second-best. Tariff escalation may emerge in environments with symmetric assumptions on the degree of scale economies and the elasticity of Foreign demand for inputs and final goods. Indeed, whenever $\gamma^d = \gamma^u = \gamma$ and $\sigma = \theta = \varepsilon$, it is straightforward to verify that tariff escalation is necessarily (second-best) optimal provided that $1 + \gamma \leq \varepsilon / (\varepsilon - 1)$ and that

$$(1+\gamma)(1-\alpha)\frac{\varepsilon}{\varepsilon - 1} \geq 1.$$

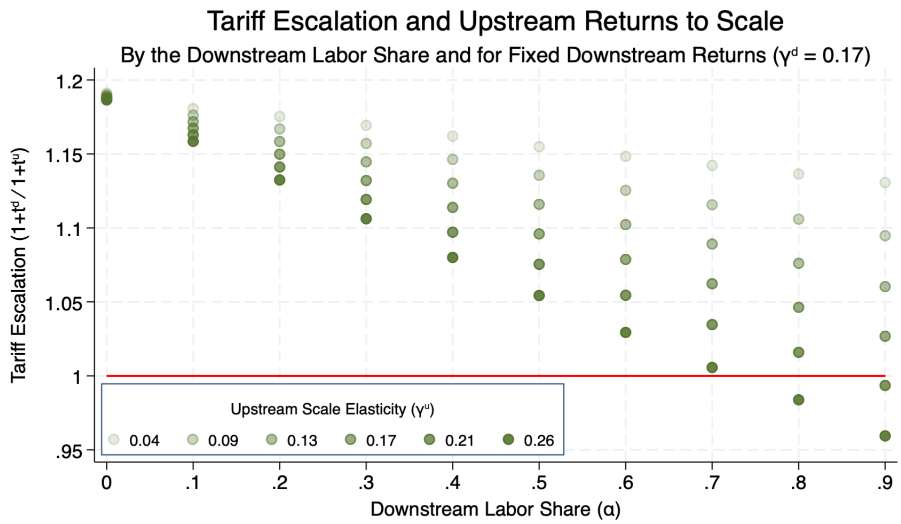
These conditions necessarily hold when α is sufficiently small. Furthermore, these are again sufficient but not necessary conditions for the optimality of tariff escalation.

5.3 Prevalence of Tariff Escalation

To provide a broader characterization of the parameter space under which second-best tariffs are escalated, we compute numerical solutions to the planner problem for a wide range of parameter values. We first fix the input and final-good elasticities to symmetric and standard values in the literature ($\sigma = \theta = 5$) and analyze a wide range of scale elasticities. Using recent estimates from Bartelme et al. (2021) for 15 manufacturing industries, we consider upstream scale elasticities from 0.04 to 0.26, and a downstream labor share (α) ranging from 0 to 0.9. To solve for second-best tariffs, we first use values for import prices, export demand shifters, and productivities in both sectors from the equilibrium solutions in a large open economy described in the next section. Technical details and a full discussion of this analysis are presented in Appendix C.1.

Figure 2 plots optimal, second-best tariff escalation as a function of the upstream scale elasticity and the downstream labor share. We hold downstream returns to scale fixed ($\gamma^d = 0.17$) and use symmetric and constant demand elasticities for inputs and final goods ($\sigma = \theta = 5$). As expected, second-best tariffs are always escalated when downstream returns to scale are at least as large as

Figure 2: Second-Best Tariff Escalation, by the Upstream Scale Elasticity and Downstream Labor Share



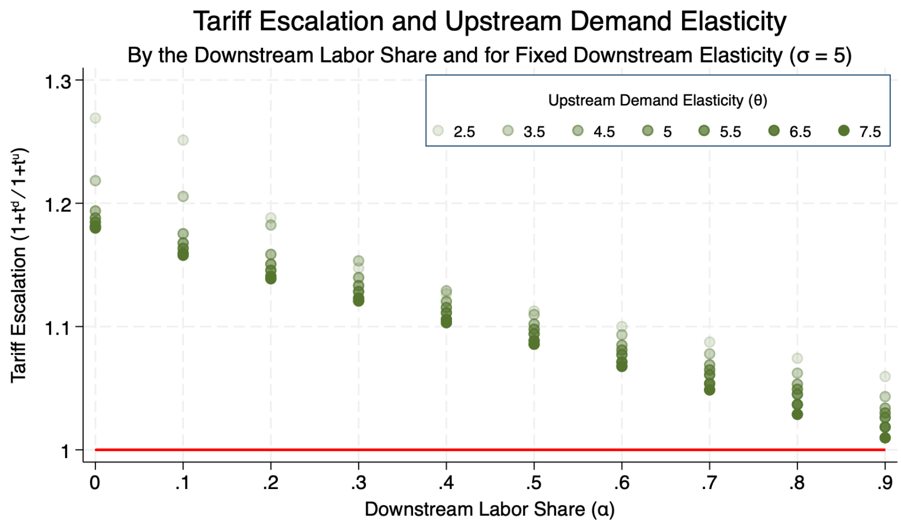
Notes: Figure plots the ratio of optimal downstream to upstream tariffs (i.e., tariff escalation) as a function of the downstream labor share (α) and upstream scale elasticity (γ^u). The downstream scale elasticity (γ^d) is fixed at 0.17, while the demand elasticities (σ and θ) are fixed at 5. Appendix C.1 shows similar patterns for additional values of γ^d.

upstream returns. Tariff escalation is also decreasing in the downstream labor share. This pattern arises in the second-best when the upstream tariff must proxy for the missing upstream subsidy, and the degree of labor misallocation (and thus also the impact of a subsidy) is increasing in that share. These results are all in line with Proposition 8, and demonstrate that even when upstream returns are substantially larger than downstream returns, optimal tariffs are only de-escalated when the downstream labor share is above 70 percent.

We also analyze 360 variants with different values of both sectoral scale elasticities and the downstream labor share. We find that second-best tariffs are escalated in 84 percent of them. The average and median values of tariff escalation are 1.09, with a maximum of 1.29 and minimum of 0.82 (see Appendix Table C.3). In Appendix Figures C.1 and C.2, we show that second-best input tariffs are higher than downstream tariffs only when the upstream scale elasticity is higher than the downstream elasticity and the downstream labor share is high.

We similarly assess second-best tariff escalation for a range of input-demand elasticities (θ) from 2.5 to 7.5. Figure 3 plots optimal tariff escalation as a function of the upstream demand elasticity and the downstream labor share. We hold the downstream elasticity of demand constant ($\sigma = 5$) and the scale elasticities symmetric and fixed ($\gamma^d = \gamma^u = 0.17$). In this setting, optimal tariffs are always escalated, though the extent of escalation is still decreasing in the downstream labor share, as well as in the input-demand elasticity. While it may seem counterintuitive that optimal input tariffs are higher when input demand is more elastic, this result is precisely in line with the prediction from Proposition 7. When export taxes are missing, the upstream tariff serves as an

Figure 3: Second-Best Tariff Escalation, by the Upstream Demand Elasticity and Downstream Labor Share



Notes: Figure plots the ratio of optimal downstream to upstream tariffs (i.e., tariff escalation) as a function of the downstream labor share (α) and upstream demand elasticity (θ). The downstream demand elasticity (σ) is fixed at 5, while the scale elasticities (γ^d and γ^u) are both fixed at 0.17. Appendix C.1 shows similar patterns for additional values of σ.

indirect downstream export tax, which is more appealing when final-good demand is relatively more inelastic (and thus inputs are relatively more elastic).

Finally, we consider the same range of demand elasticities for the downstream sector. There are 490 of these cases. We obtain solutions for 489 of these cases, and find that second-best tariffs are escalated for 99 percent of them. The average and median values of tariff escalation are 1.11, with a maximum of 1.34 and minimum of 0.98 (see Appendix Table C.6). In Appendix Figures C.3 and C.4, we show that second-best input tariffs are higher than downstream tariffs only when the downstream labor share is high and final-good demand is 70 percent more inelastic than input demand.

6 Quantitative Results for a Large Open Economy

In this section, we relax the 'small open economy' assumption by allowing Home and Foreign prices to change, and quantitatively solve for the social-welfare maximizing levels of input and final-good tariffs. We perform the analysis by mapping aggregate data to our two-country model, interpreting the United States as Home, and the rest of the world (RoW) as Foreign. We first analyze policy in the first-best allocation, then study a second-best setting in which only import tariffs are feasible, and finally consider a range of key parameter values. The main message from our SOE analysis - namely that tariff escalation is optimal for a wide range of parameter values - persists in the large open economy.

6.1 Data and Parameters

To solve the model numerically, we must take a stance on a number of parameters and assess whether they deliver equilibrium values for key variables that are consistent with those observed in the data. The main parameters of the model are the elasticities of substitution upstream and downstream (θ and σ), the downstream labor share (α), the scale elasticities upstream and downstream (γ^u and γ^d), iceberg trade costs upstream and downstream (τ^u and τ^d), productivity upstream and downstream in each country (A^u_{US}, A^u_{RoW}, A^d_{US} and A^d_{RoW}), and the Home and Foreign labor endowments (L_{US} and L_{RoW}).

Our quantitative approach constitutes a blend of calibration and estimation. We first discuss various approaches to estimating the key elasticities of substitution and scale elasticities; we then back out the downstream labor share and the labor endowments from readily available public data; and we finally estimate trade costs and the productivity parameters by minimizing the distance between our model and a series of moments from standard sources.

Elasticities of Substitution (θ and σ) We consider three alternative approaches to quantifying the elasticities of substitution across varieties in the upstream and downstream sectors (θ and σ, respectively). We summarize these approaches here and provide additional details in Appendix D.2.

The first approach treats these elasticities as symmetric across sectors. We fix the values of the elasticities of substitution across varieties in each sector to 5 ($\sigma = \theta = 5$), as in Costinot and Rodríguez-Clare (2014). This symmetric case is useful to assess whether real-world tariff escalation might simply be due to differences in input versus final-good demand elasticities.

The second approach is to calibrate demand elasticities from data on markups. Recall that under monopolistic competition and CES preferences, the optimal firm-level markup is equal to $\theta/(\theta-1)$ upstream and $\sigma/(\sigma-1)$ downstream. Using sales and markup data from Baqaee and Farhi (2020) based on publicly listed firms in Compustat, we compute the sales-weighted average markups of firms, that we categorize as upstream or downstream based on their primary sector. This approach leads to estimates of $\theta = 4.43$ for the elasticity of substitution upstream and of $\sigma = 6.44$ for the elasticity of substitution downstream.

The third approach is to estimate demand elasticities using changes in trade flows during the US-China trade war in 2018 to 2019. We follow Amiti et al. (2019) and calculate 12-month changes in US tariffs and imports at the product-country level. Under a CES demand structure, regressing the changes in trade flows on changes in tariffs provides estimates of the trade elasticity. Our preferred specification leads to estimates of $\theta = 2.45$ for the elasticity of substitution upstream, and $\sigma = 3.59$ for the elasticity of substitution downstream. The small magnitude of the trade elasticities is consistent with the findings in Amiti et al. (2020), and might reflect lower responsiveness of trade flows due to uncertainty about the persistence of the tariff changes.

Scale Elasticities (γ^u and γ^d) We use estimates of scale elasticities from Bartelme et al. (2021). These authors provide two-stage least squares estimates of scale elasticities for 15 manufacturing industries, from which we back out scale elasticities for upstream and downstream industries (see Appendix D.2). For the upstream sector we obtain an average scale elasticity of $\gamma^u = 0.166$ and for the downstream sector an average scale elasticity of $\gamma^d = 0.170$.

Downstream Labor Intensity, Trade and Expenditure Shares and Labor Endowments We measure the share of inputs in production, $1 - \alpha = 0.45$, from usage of intermediate inputs by downstream sectors based on the World Input-Output Database (WIOD) (see Appendix D.4 for details). Similarly, we calculate trade and expenditure shares for the upstream (intermediate-input) and downstream (final consumption) sectors based on trade flows in the WIOD, taking into account whether a trade flow is used for final consumption or as an intermediate input.[19] We infer the labor endowment of each country from population data published by CEPII.[20]

Estimation of Productivity Parameters and of Trade Costs Finally, we normalize US productivity in both sectors to one, $A^d_{US} = A^u_{US} = 1$. This leaves us with four parameters to estimate: trade costs in each sector $\{\tau^d, \tau^u\}$, and sectoral productivity in the rest of the world $\{A^d_{RoW}, A^u_{RoW}\}$. To estimate the model, we search for the vector of parameters $\{\tau^d, \tau^u, A^d_{RoW}, A^u_{RoW}\}$ that minimizes the sum of squares of the differences between model-generated and empirical moments, subject to our equilibrium constraints. For this exercise, we set upstream and downstream tariffs to their January 2018 values, as shown in Figure 4 (details in Section 7). Panel B of Table 1 lists the set of moments we target in the estimation. The moments correspond to those that are necessary to solve for the changes in equilibrium outcomes in response to a counterfactual change in tariffs (i.e., the hat algebra approach) and are all retrieved from the WIOD.

Panel A in Table 1 presents the estimated values of the RoW's productivities and iceberg trade costs in each sector obtained under symmetric elasticities upstream and downstream, $\theta = \sigma = 5$. Trade costs appear slightly higher in the downstream sector, but within the range of standard estimates (Anderson and Van Wincoop, 2004). The estimates indicate that the United States is about three times more efficient in final-good production than the rest of the world, and seven times more efficient in terms of input production. Despite only estimating four parameters, the fit of the model is quite good for most moments, except the ratio of total sales in the upstream sector to total expenditure in the downstream sector. This discrepancy is likely due to our assumption that the upstream sector uses only labor in production, which naturally leads to sales of inputs that are significantly lower than sales of final goods. In the data, input production uses both labor *and* intermediate inputs, which leads to a higher ratio of input sales to final-good sales than what our model generates.[21]

[19] We use data for 2014 which is the latest available year in the WIOD.

[20] Specifically, we set $L^{us} = 10 \times \frac{Pop^{us}}{Pop^{us}+Pop^{row}} = 0.45$ and $L^{row} = 10 \times \frac{Pop^{row}}{Pop^{us}+Pop^{row}} = 9.55$.

[21] In Appendix B.1, we extend our closed-economy model by allowing upstream production to use labor and inputs. Caliendo et al. (2023) find that optimal, second-best tariffs in a single-sector model are lower when production is roundabout, suggesting that roundabout production for inputs in our model would lower the optimal input tariff and thus increase the prevalence of escalation via an additional mechanism to the role of downstream returns we study here.

Table 1: Calibrated Parameters and Moments

A. Calibrated Parameters	
Productivity in final-good sector, RoW relative to US, A^d_{row}	0.348
Productivity in input sector, RoW relative to US, A^u_{row}	0.150
Iceberg cost for final goods from US to RoW, τ^d	2.368
Iceberg cost for inputs from US to RoW, τ^u	2.030

B. Estimated Moments	Data	Model
Sales share to US from US in final goods	0.943	0.964
Sales share to RoW from RoW in final goods	0.988	0.986
Sales share to US from US in inputs	0.897	0.889
Sales share to RoW from Row in inputs	0.982	0.978
Expenditure share on US final goods by the US	0.960	0.946
Expenditure share on RoW final goods by the RoW	0.981	0.989
Expenditure share on US inputs by the US	0.905	0.921
Expenditure share on RoW inputs by the RoW	0.980	0.967
Total US input sales over total US final-good expenditure	0.754	0.466
Total RoW input sales over total RoW final-good expenditure	1.250	0.446
Total US final-good sales over total US final-good expenditure	0.996	0.997
Total RoW final-good sales over total RoW final-good expenditure	0.999	0.999
Total expenditure on final goods by the US relative to RoW	0.303	0.285

Sources: World Input Output Database and authors' calculations. Notes: Panel A presents the estimated values of the RoW's productivities and iceberg trade costs in each sector obtained under symmetric demand elasticities upstream and downstream, $\theta = \sigma = 5$, a downstream scale elasticity $\gamma^d = 0.170$, and upstream scale elasticity $\gamma^u = 0.166$. Panel B presents the targeted moments in the estimation. Column 1 presents moments from the data and column 2 presents their estimated counterparts. Note that in the model, total sales upstream to total expenditure downstream cannot be larger than 1, since the upstream sector is pure value added.

6.2 Optimal Tariffs

In this section, we compute counterfactual optimal US tariffs on final goods and inputs, holding RoW tariffs on US goods fixed at their January 2018 levels.

Optimal Import Tariffs under First-Best Policies We begin by considering optimal policy when the Home government has the necessary instruments to achieve the first-best allocation. Proposition 1 indicates that all first-best policies must include an upstream production subsidy. For the remaining instruments, we focus on the implementations in Proposition 2, which rule out the downstream production subsidy and consider import tariffs and export taxes. We normalize the downstream export tax to zero, which leads to an optimal vector of policies given by

$$\left(t^d_H, t^u_H, s^u_H, v^u_H\right) = (0.252, 0.072, 0.143, 0.145)\,.$$

Tariff escalation ($\frac{1+t^d}{1+t^u}=1.168$) is very close to $1+\gamma^d=1.170$, and the optimal domestic production subsidy upstream is essentially indistinguishable from $\gamma^u/\left(1+\gamma^u\right)=0.142$. Accounting for general equilibrium effects that arise in a large open economy delivers first-best policies that are remarkably consistent with the results from Proposition 2.

Optimal Import Tariffs under Second-Best Policies Real-world trade policy rarely features export taxes (they are outlawed by the US Constitution) and production subsidies are often motivated by a host of non-trade-related reasons. We therefore return to second-best policies that rely solely on import tariffs. We again maximize US welfare, assuming that the RoW tariffs on the United States remain fixed at their January 2018 levels. In this case, the vector of optimal import tariffs is given by

$$\left(t_H^d, t_H^u\right) = (0.299, 0.197)\,.$$

Tariff escalation ($\frac{1+t^d}{1+t^u}=1.085$) thus prevails under second-best policies, though to a lesser degree than in first-best policies. This result is in line with an upstream tariff mimicking the missing upstream subsidy by redirecting domestic expenditure towards home inputs, thereby raising upstream and thus also downstream efficiency.

Despite the fact that the optimal downstream tariff t^d is significantly larger than the optimal upstream tariff t^u, our numerical results indicate that the impact of constraining both tariffs to be uniform is quantitatively small. This is not surprising given that the US domestic trade shares in Table 1 are close to 1, thus leading to very modest welfare effects from changes in tariffs (Arkolakis et al., 2012).

6.3 Optimal Tariffs: Robustness

We conclude our analysis of the large open economy by assessing the robustness of tariff escalation to a range of parameter values. We analyze optimal import tariffs for various values of sectoral scale elasticities, using alternative methods to estimate demand elasticities (θ and σ), as well as for different values of the downstream labor share (α). When changing these parameters, we re-calibrate the trade and productivity parameters to provide the best fit of the moments in Panel B of Table 1. As in our baseline analysis, we focus on first-best policies that use an upstream production subsidy and export tax, and import tariffs for both sectors.

Table 2 presents optimal policies for alternative values of the sectoral scale elasticities and downstream labor share. Column 1 imposes symmetric scale elasticities in both sectors ($\gamma^d=\gamma^u=0.17$) and uses the observed downstream labor ($\alpha=0.55$). The optimal upstream subsidy is slightly larger than our baseline analysis, which is due to slightly higher upstream returns to scale considered here. By contrast, tariff escalation is identical to our baseline case, in line with Proposition 2 that predicts it is determined by the extent of downstream returns. Optimal second-best tariffs are still escalated, with a marginally lower value relative to the baseline case, which again reflects the higher degree of upstream returns to scale, and thus a higher upstream tariff to replace the missing subsidy.

Given the key role of relative upstream versus downstream returns to scale, we also consider variants in which the upstream scale elasticity is 25 percent smaller ($\gamma^d = 0.17 > \gamma^u = 0.128$) or 25 percent larger ($\gamma^d = 0.128 < \gamma^u = 0.17$) than the downstream elasticity. In first-best policies and holding the downstream labor share fixed at $\alpha = 0.55$, a lower upstream scale elasticity only affects the upstream subsidy, which necessarily falls. As expected, a lower downstream scale elasticity reduces the degree of tariff escalation, which falls to $\frac{1+t^d}{1+t^u} = 1.126$, which is again quite close to $1 + \gamma^d = 1.128$ predicted by Proposition 2. For both relative sizes of the scale elasticities, we confirm that these patterns are robust for values of the downstream labor share from 0.25 to 0.90.

We similarly analyze second-best import tariffs in Panel B of Table 2. Consistent with the upstream tariff mimicking the missing upstream subsidy, input tariffs are always higher in the second-best, and tariff escalation is lower. This pattern is most pronounced in the last four columns of Table 2, in which upstream returns to scale are greater than downstream returns, and the optimal input subsidy is thus higher. The degree of escalation is also decreasing in the downstream labor share. As explained in Section 5, the difference between the social planner versus market equilibrium upstream labor allocations is increasing in α, such that an input tariff is more important when α is higher. Indeed, the only set of parameter values for which optimal tariffs are less than one is when the downstream labor share is 0.90 and upstream returns are greater than downstream returns.

Table 2: Optimal Tax Policies for Various Scale Elasticities and Labor Share Values

	$\gamma_d = \gamma_u = 0.170$	$\gamma_d = 0.170$ and $\gamma_u = 0.128$				$\gamma_d = 0.128$ and $\gamma_u = 0.170$			
	$\alpha = 0.55$	$\alpha = 0.55$	$\alpha = 0.25$	$\alpha = 0.75$	$\alpha = 0.9$	$\alpha = 0.55$	$\alpha = 0.25$	$\alpha = 0.75$	$\alpha = 0.9$
Panel A: First-Best Policies									
t^d	0.252	0.252	0.255	0.252	0.252	0.252	0.254	0.252	0.252
t^u	0.072	0.072	0.072	0.072	0.071	0.112	0.112	0.112	0.112
ν^u	0.145	0.145	0.147	0.146	0.146	0.113	0.115	0.114	0.115
s^u	0.145	0.114	0.114	0.114	0.114	0.145	0.145	0.145	0.145
$\frac{1+t^d}{1+t^u}$	1.168	1.168	1.170	1.168	1.169	1.126	1.128	1.125	1.126
Panel B: Second-Best Tariffs									
t^d	0.297	0.317	0.329	0.288	0.263	0.268	0.299	0.242	0.240
t^u	0.198	0.192	0.161	0.191	0.186	0.216	0.190	0.228	0.255
$\frac{1+t^d}{1+t^u}$	1.083	1.105	1.145	1.081	1.065	1.043	1.092	1.011	0.988

Notes: Each column presents optimal tariffs, taxes, and subsidies for $\sigma = \theta = 5$ and alternative values of the parameters indicated in the column headers and their corresponding, re-estimated values of τ^d, τ^u, A^d_{row} and A^u_{row}. The set of calibrated parameters that corresponds to each column is displayed in Table C.9 in Appendix C.2. The downstream subsidy is zero, and we normalize the downstream export tax to zero.

We also analyze optimal policies for different values of input and final-good demand elasticities. Table 3 presents the results, now using symmetric upstream and downstream scale elasticities ($\gamma^d = \gamma^u = 0.17$). In the first-best, the optimal upstream production subsidy is 0.14 across all specifications, in line with upstream returns to scale determining the optimal subsidy, as stated in Proposition 2. In the first-best policies, tariff escalation is fairly constant and essentially determined by the extent of downstream returns to scale, as predicted by Proposition 2.

Table 3: Optimal Tax Policies for Various Input and Final-Good Demand Elasticities

	$\alpha = 0.55$					$\theta = 5,\ \sigma = 5$			
	$\theta = 5$ $\sigma = 5$	$\theta = 4.43$ $\sigma = 6.44$	$\theta = 2.45$ $\sigma = 3.59$	$\theta = 2.5$ $\sigma = 4$	$\theta = 5.5$ $\sigma = 4$	$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.75$	$\alpha = 0.9$
First-Best									
t^d	0.252	0.185	0.389	0.336	0.336	0.258	0.255	0.252	0.252
t^u	0.072	0.014	0.191	0.145	0.144	0.073	0.072	0.072	0.072
ν^u	0.145	0.218	0.299	0.317	0.067	0.150	0.147	0.145	0.146
s^u	0.145	0.145	0.145	0.145	0.145	0	0.145	0.145	0.145
$\frac{1+t^d}{1+t^u}$	1.168	1.169	1.66	1.167	1.168	1.173	1.170	1.168	1.169
Second-Best									
t^d	0.297	0.286	0.600	0.557	0.310	0.330	0.322	0.265	0.250
t^u	0.197	0.180	0.444	0.404	0.215	0.126	0.165	0.204	0.217
$\frac{1+t^d}{1+t^u}$	1.083	1.090	1.109	1.109	1.078	1.181	1.135	1.051	1.027

Notes: Each column presents optimal tariffs and taxes for alternative values of the parameters described in Section 6.1 and their corresponding, re-estimated values of τ^d, τ^u, A^d_{row} and A^u_{row}. The set of calibrated parameters that corresponds to each column is displayed in Table C.9 in Appendix C.2. In this table $\gamma_d = \gamma_u = 0.170$. Panels A and B present optimal tariffs and taxes for the cases of policy instruments in Section 6.2. Tariff escalation ($\frac{1+t^d}{1+t^u} > 1$) is a robust feature across all specifications.

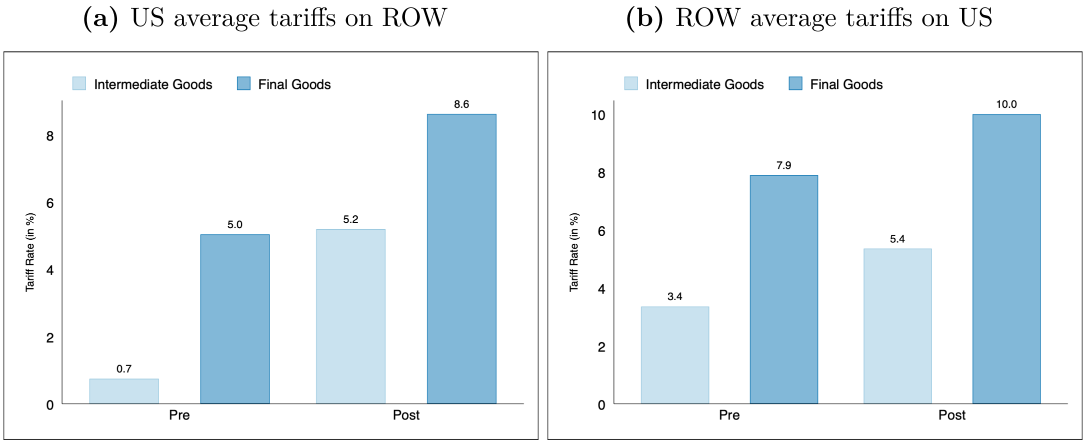
Tariff escalation persists, though is generally lower, in the second-best, as shown in Table 3, Panel B. The only exception is when the downstream sector uses no labor to produce ($\alpha = 0$), in which case there is no domestic labor misallocation and thus no need for the upstream tariff to mimic an upstream subsidy. Indeed, when $\alpha = 0$ the tariffs only need to substitute for the missing export taxes, and tariff escalation is even higher than in the first-best allocation, exactly in line with Proposition 4. More generally, tariff escalation is decreasing in the downstream labor share, highlighting the role of the upstream tariff as a substitute for the missing upstream production subsidy, which is more important when the downstream sector has a higher labor share. From Table 3, we conclude that tariff escalation is a robust feature of second-best trade policy for empirically plausible values of the downstream labor share and demand elasticities.

7 Counterfactuals: Evaluation of the 2018-2019 Tariff Increases

We close the paper by evaluating how the tariff increases during the 2018 to 2019 trade war affected US welfare. Figure 4 illustrates the general pattern of tariff escalation for both US tariffs on other countries and for foreign tariffs applied to the United States, before and after the trade war. Due to their low initial levels, average input tariffs increased the most on a percent basis, though tariff escalation remains a distinct feature of the data.[22] These tariff increases largely arose from the US-China trade war, but also include US tariff changes on washing machines, solar panels, aluminum, and steel, as well as the retaliatory tariffs from the rest of the world.

[22] For details on data construction and for a version of Figure 4 without trade weights, see Appendix D.1.

Figure 4: Tariff Escalation Before and After the US-China Trade War



Notes: Pre: Tariffs in January 2018, Post: Tariffs in December 2019. Tariff data from WTO and USITC. Tariffs are weighted using 2015 imports and exports from the US Census Bureau. Goods are classified as intermediate or final according to the Broad Economic Categories rev. 5 (BEC5) code corresponding to their HS code. BEC5 codes with a third digit of 1 or 2 (intermediate and capital final use in BEC5, respectively) are classified as intermediate goods and BEC5 codes with a third digit of 3 are classified as final consumption goods. When the HS to BEC5 correspondence implies mixed use as intermediate and consumption or capital and consumption, we assign no classification and omit from consideration.

To evaluate the welfare effects of the tariff increases, we calculate US and RoW welfare under five scenarios. First, we use the observed 2017 tariffs and no other domestic instruments. Second, we apply the observed 2019 US tariffs. Third, we evaluate the distinct welfare effects of downstream versus upstream tariffs by changing tariffs in each sector individually. Fourth, we construct a counterfactual tariff that generates the same revenue as the 2019 US tariffs, but only applies to either the downstream or the upstream sector. Finally, we evaluate the welfare implications of optimal second-best tariffs and optimal first-best policies. For each of these scenarios, we first hold RoW tariffs constant at their 2017 values, and then re-calculate welfare using the observed 2019 retaliatory tariff implemented by the RoW. For these calculations, we use our baseline numbers for the demand and scale elasticities ($\theta = \sigma = 5$, $\gamma_u = 0.166$, and $\gamma_d = 0.170$).

Table 4 displays our results, with panels A and B showing the cases with and without retaliation, respectively. Average US import tariffs increased from 0.7 to 5.2 percentage points for intermediate goods, and from 5.2 to 8.6 percentage points for final goods. We find that without tariff retaliation by the RoW, US welfare would have increased by 0.13% from these tariff changes. This gain is consistent with our estimates of the sizable unilateral optimal tariff rates for the US in Section 6 (in the absence of any export taxes or domestic subsidies).

We next evaluate the extent to which this gain was due to larger final-good versus input tariffs. In the third row of Table 4, we consider only the increase in final-good tariffs, while row four considers only input-tariff increases. The comparison reveals that any US welfare gains are driven overwhelmingly by higher final-good tariffs.

The dominant role of final-good tariffs on welfare increases is even starker when considering a counterfactual increase in US final-good tariffs (row 5 of Table 4) that would have (naively) raised the same revenue as the observed tariff increases based on the average tariff rate change and 2017 trade flows. In this case, US welfare – absent any foreign retaliation – would have risen by 0.15%. If instead the tariff increases had been adjusted to apply only to intermediate inputs, US welfare would have increased by only 0.02.

Table 4: Counterfactual Welfare Effects of US-China Trade War

	A. RoW tariff at 2017 level			**B.** RoW tariff at 2019 level		
	U_{US}	U_{RoW}	$\frac{U_{US}}{U_{US,2017}}$	U_{US}	U_{RoW}	$\frac{U_{US}}{U_{US,2017}}$
1. US tariff - 2017 level	0.0725	0.3053				
2. US tariff - 2019 level	0.0726	0.3051	1.0013	0.0726	0.3051	1.0002
3. 2019 US tariff only Downstream	0.0726	0.3051	1.0009	0.0725	0.3052	0.9997
4. 2019 US tariff only Upstream	0.0726	0.3052	1.0004	0.0725	0.3053	0.9993
5. Counterfactual Tariff only Downstream	0.0726	0.3050	1.0015	0.0726	0.3051	1.0004
6. Counterfactual Tariff only Upstream	0.0726	0.3051	1.0002	0.0725	0.3051	0.9991
7. Optimal US Import Tariffs	0.0728	0.3043	1.0037	0.0727	0.3044	1.0025
8. Optimal US Trade & Tax Policies	0.0731	0.3043	1.0075	0.0730	0.3044	1.0063

Notes: Table presents US welfare (U_{US}), RoW welfare (U_{RoW}), and US welfare relative to its initial 2017 level ($\frac{U_{US}}{U_{US,2017}}$). Panel A computes welfare holding the RoW tariffs at their 2017 levels, while Panel B uses the observed 2019 RoW retaliatory tariffs. *1. US tariff - 2017 level* provides baseline welfare values using actual 2017 tariff values; *2. US tariff - 2019 level* uses actual 2019 US tariffs; *3. 2019 US tariff only Downstream* uses 2017 upstream but 2019 downstream tariffs; *4. 2019 US tariff only Upstream* uses 2017 downstream but 2019 upstream tariffs; *5. Counterfactual Tariff only Downstream (6. Upstream)* uses a counterfactual US downstream (upstream) tariff that generates the same revenue as the actual 2019 US tariffs, based on the observed trade flows in 2017. Counterfactual tariffs are $\tilde{t}^d = 0.131$ or $\tilde{t}^u = 0.152$; *7. Optimal US Import Tariffs* uses the second-best optimal import tariffs in response to RoW's trade policy in 2017 (Panel A) or 2019 (Panel B). The optimal policy vector for panel A is $\left(t^d_H, t^u_H\right) = (0.299, 0.197)$ and $\left(t^d_H, t^u_H\right) = (0.299, 0.197)$ for panel B; *8. Optimal US Trade & Tax Policy*: US chooses optimal import tariffs, export tax, and production subsidy, as described in Section 4, in response to RoW's trade policy from 2017 (Panel A) or 2019 (Panel B). The optimal policy vector is $\left(t^d_H, t^u_H, v^u_H, s^u_H\right) = (0.252, 0.071, 0.144, 0.143)$ for panel A and $\left(t^d_H, t^u_H, v^u_H, s^u_H\right) = (0.252, 0.072, 0.145, 0.142)$ for panel B.

In Panel B of Table 4, we repeat these exercises but now take into consideration the observed retaliation by the RoW, which increased its import tariffs on US intermediate inputs from 3.4 to 5.4 percentage points and on US final goods from 7.9 to 10.0 percentage points. In this case, the US welfare gain from the tariff increases shrinks to only 0.02%. Therefore, tariff retaliation by the RoW largely undermines the US welfare gains from higher tariffs, which in turn are overwhelmingly driven by higher final-good (and not input) tariffs. If the US had only placed the tariffs on final goods, while (naively) raising the same revenue (row 5), US welfare would have increased by 0.04%, even accounting for retaliation. If instead those tariffs had only been placed on intermediate inputs (row 6), US welfare would have *declined* by 0.09%.

Rows seven and eight present the potential welfare gains from implementing optimal import tariffs with and without other US policy instruments. The gains from second-best, optimal import tariffs (without production subsidies or export taxes) are displayed in row seven of Table 4. Optimal US tariffs absent any foreign retaliation achieve a welfare gain of 0.37 percent, with a tariff escalation wedge of 1.085. Row eight allows for a full set of instruments that includes both import and export taxes, as well as production subsidies. Optimal trade policy with domestic subsidies and export taxes (and assuming no foreign retaliation) leads to a 0.75 percent increase in welfare, and features close to zero tariffs on inputs and an escalation wedge of 1.17. As in our analysis in the SOE, the tariff escalation wedge is now notably higher since the government uses the input subsidy rather than the tariff to address the domestic labor misallocation. We note that these calculations assume no foreign retaliation (or, in panel B, no retaliation above the observed changes from RoW tariffs from 2017 to 2019).

8 Conclusion

In this paper, we provide an efficiency rationale for the fact that import tariffs on final goods are systematically higher than those on intermediate inputs. This so-called tariff escalation has been widely documented across time and space, but there is little support in the literature for the notion that it constitutes a social welfare-maximizing policy.

We develop a two-sector model with final goods and intermediate inputs, both produced under increasing returns to scale, and show that (i) first-best trade policies are consistent with tariff escalation, and that (ii) second-best import tariffs generally feature tariff escalation. A key result is that tariff escalation is driven by the presence and extent of increasing returns to scale in downstream production. Access to cheaper inputs expands the downstream sector's size and thus raises its efficiency. Relatively higher input tariffs are only optimal in a small set of second-best settings in which an upstream production subsidy cannot be used, and upstream efficiency and thus input prices are sufficiently sensitive to the upstream sector's size.

Although our model generically features domestic distortions related to the existence of scale economies upstream, lower optimal input tariffs are *not* explained by a (second-best) correction of these domestic distortions. Indeed, domestic distortions reduce the desirability of tariff escalation. Instead, input tariffs are less beneficial because they distort the marginal rate of substitution between domestic and foreign inputs. Although final-good tariffs entail a similar distortion, the input distortion is magnified whenever final-good production features increasing returns to scale. By contrast, final goods are purchased by consumers with no potential efficiency gains or ability to relocate abroad. It is thus scale economies downstream, rather than upstream, that shape the optimality of tariff escalation. Investigating the empirical support for relatively higher downstream versus upstream increasing returns in real-world tariffs remains an open and exciting venue for future work.

References

Amiti, Mary, "Are Uniform Tariffs Optimal?," in "Trade Theory, Analytical Models and Development" 2004, p. 47.

_ **and Jozef Konings**, "Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia," *American Economic Review*, 2007, *97* (5), 1611–1638.

_ **, Stephen J. Redding, and David E. Weinstein**, "The Impact of the 2018 Tariffs on Prices and Welfare," *Journal of Economic Perspectives*, 2019, *33* (4), 187–210.

_ **, Stephen J Redding, and David E Weinstein**, "Who's paying for the US tariffs? A longer-term perspective," in "AEA Papers and Proceedings," Vol. 110 2020, pp. 541–46.

Anderson, James E and Eric Van Wincoop, "Trade Costs," *Journal of Economic literature*, 2004, *42* (3), 691–751.

Antràs, Pol and Robert W Staiger, "Offshoring and the role of trade agreements," *American Economic Review*, 2012, *102* (7), 3140–83.

_ **, Teresa Fort, and Felix Tintelnot**, "The Margins of Global Sourcing: Theory and Evidence from US Firms," *American Economic Review*, 2017, *107* (9), 2514–2564.

Arkolakis, Costas, Arnaud Costinot, and Andres Rodríguez-Clare, "New Trade Models, Same Old Gains," *American Economic Review*, 2012, *102* (1), 94–130.

Balassa, Bela, "Tariff protection in industrial countries: an evaluation," *Journal of Political Economy*, 1965, *73* (6), 573–594.

Baqaee, David and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium," *Quarterly Journal of Economics*, February 2020, *135* (1), 105–163.

Barattieri, Alessandro and Matteo Cacciatore, "Self-Harming Trade Policy? Protectionism and Production Networks," 2020. NBER Working Paper No. 27630.

Bartelme, Dominick G, Arnaud Costinot, Dave Donaldson, and Andres Rodríguez-Clare, "The textbook case for industrial policy: Theory meets data," Technical Report, MIT, Mimeo 2021.

Basker, Emek and Pham Hoang Van, "Imports 'R' Us: Retail Chains as Platforms for Developing Country Imports," *American Economic Review*, May 2010, *100* (2), 414–18.

Beshkar, Mostafa and Ahmad Lashkaripour, "Interdependence of Trade Policies in General Equilibrium," *Center for Applied Economics and Policy Research (CAEPR) Working Paper*, 2017, *3*.

_ **and** _ , "The Cost of Dissolving the WTO: The Role of Global Value Chains," Working Paper 2020-005, CAEPR WORKING PAPER SERIES 2020.

Blanchard, Emily J, Chad P Bown, and Robert C Johnson, *Global supply chains and trade policy* 2021.

Blaum, Joaquin, Claire Lelarge, and Michael Peters, "The Gains from Input Trade with Heterogenous Importers," *AEJ: Macro*, 2018. (forthcoming).

Bown, Chad, Paola Conconi, Aksel Erbahar, and Lorenzo Trimarchi, "Trade Protection Along Supply Chains," 2020. mimeo.

Bown, C.P. and M.A. Crowley, "Chapter 1 - The Empirical Landscape of Trade Policy," 2016, *1*, 3 – 108.

Breinlich, Holger, Elsa Leromain, Dennis Novy, and Thomas Sampson, "Import Liberalization as Export Destruction? Evidence from the United States," 2021.

Cadot, Olivier, Jaime De Melo, and Marcelo Olarreaga, "Lobbying, counterlobbying, and the structure of tariff protection in poor and rich countries," *The World Bank Economic Review*, 2004, *18* (3), 345–366.

Caliendo, Lorenzo, Robert C Feenstra, John Romalis, and Alan M Taylor, "A second-best argument for low optimal tariffs on intermediate inputs," *Journal of International Economics*, 2023, *145*, 103824.

Casas, Francisco R, "Optimal effective protection in general equilibrium," *The American Economic Review*, 1973, *63* (4), 714–716.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 197–261.

_ **,** _ **, and Iván Werning**, "Micro to macro: Optimal trade policy with firm heterogeneity," *Econometrica*, 2020, *88* (6), 2739–2776.

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

Cox, Lydia, "The Long-Term Impact of Steel Tariffs on U.S. Manufacturing," Technical Report, Mimeo, Harvard University 2021.

Das, Satya P, "Optimum tariffs on final and intermediate goods," *International Economic Review*, 1983, pp. 493–508.

Demidova, Svetlana and Andres Rodríguez-Clare, "Trade Policy Under Firm-level Heterogeneity in a Small Economy," *Journal of International Economics*, 2009, *78*, 100–12.

Diamond, Peter A and James A Mirrlees, "Optimal taxation and public production I: Production efficiency," *The American economic review*, 1971, *61* (1), 8–27.

Erbahar, Aksel and Yuan Zi, "Cascading trade protection: Evidence from the US," *Journal of International Economics*, 2017, *108*, 274–299.

Fajgelbaum, Pablo D., Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal, "The Return to Protectionism," *Quarterly Journal of Economics*, 2020, *135* (1), 1–55.

Flaaen, Aaron and Justin Pierce, "Disengtangling the Effects of the 2018-2019 Tariffs on a Globally Connected US Manufacturing Sector," Finance and Economics Discussion Series Working Paper 086, Board of Governors of the Federal Reserve System 2019.

Ganapati, Sharat, "The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economies," Working Papers 18-49, Center for Economic Studies, U.S. Census Bureau December 2018.

Gawande, By Kishore, Pravin Krishna, and Marcelo Olarreaga, "Lobbying competition over trade policy," *International Economic Review*, 2012, *53* (1), 115–132.

Goldberg, Pinelopi, Amit Khandelwal, Nina Pavcnik, and Petia Topalova, "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *Quarterly Journal of Economics*, 2010, *125*, 1727–67.

Gros, Daniel, "A Note on the Optimal Tariff, Retaliation, and the Welfare Loss from Tariff Wars in a Framework with Intra-Industry Trade," *Journal of International Economics*, 1987, *23*, 357–67.

Grossman, Gene and Elhanan Helpman, "Protection for Sale," *American Economic Review*, 1994, *84* (4), 833–50.

Grossman, Gene M and Elhanan Helpman, "When Tariffs Disturb Global Supply Chains," Technical Report, National Bureau of Economic Research 2020.

Handley, Kyle, Fariha Kamal, and Ryan Monarch, "Rising Import Tariffs, Falling Export Growth: When Modern Supply Chains Meet Old-Style Protection," Working Paper 26611, NBER 2020.

Hwang, Hong, Chao-Cheng Mai, and Shih-Jye Wu, "Tariff escalation and vertical market structure," *The World Economy*, 2017, *40* (8), 1597–1613.

Kortum, Samuel S and David A Weisbach, "Optimal unilateral carbon policy," 2021.

Krugman, Paul and Anthony J Venables, "Globalization and the Inequality of Nations," *The quarterly journal of economics*, 1995, *110* (4), 857–880.

Kucheryavyy, Konstantin, Gary Lyn, and Andrés Rodríguez-Clare, "Grounded by gravity: A well-behaved trade model with industry-level economies of scale," *American Economic Journal: Macroeconomics*, 2023, *15* (2), 372–412.

Lashkaripour, Ahmad and Volodymyr Lugovskyy, "Profits, Scale Economies, and the Gains from Trade and Industrial Policy," *American Economic Review*, October 2023, *113* (10), 2759–2808.

Li, Minghao, "CARD trade war tariffs database," *Center for Agriculture and Rural Development, Report. https://www. card. iastate. edu/china/trade-war-data*, 2018.

Liu, Ernest, "Industrial policies in production networks," *The Quarterly Journal of Economics*, 2019, *134* (4), 1883–1948.

Lucas, Robert E and Nancy L Stokey, "Optimal fiscal and monetary policy in an economy without capital," *Journal of monetary Economics*, 1983, *12* (1), 55–93.

McCorriston, Steve and Ian Sheldon, "Tariff (De-) Escalation with Successive Oligopoly," *Review of Development Economics*, 2011, *15* (4), 587–600.

Ornelas, Emmanuel and John L. Turner, "Trade Liberalization, Outsourcing, and the Hold-up Problem," *Journal of International Economics*, 2008, *74* (1), 225–241.

_ **and** _ , "Protection and International Sourcing," *The Economic Journal*, 2012, *122* (559), 26–63.

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

_ , "Trade wars and trade talks with data," *American Economic Review*, 2014, *104* (12), 4104–46.

Puga, Diego and Anthony J Venables, "Agglomeration and economic development: Import substitution vs. trade liberalisation," *The Economic Journal*, 1999, *109* (455), 292–311.

Ruffin, Roy J, "Tariffs, intermediate goods, and domestic protection," *The American Economic Review*, 1969, *59* (3), 261–269.

Shapiro, Joseph S, "The Environmental Bias of Trade Policy*," *The Quarterly Journal of Economics*, 12 2020, *136* (2), 831–886.

Smith, Dominic and Sergio Ocampo Díaz, "The Evolution of US Retail Concentration," Technical Report 2020.

Spencer, Barbara J and Ronald W Jones, "Vertical foreclosure and international trade policy," *The Review of Economic Studies*, 1991, *58* (1), 153–170.

_ **and** _ , "Trade and protection in vertically related markets," *Journal of international Economics*, 1992, *32* (1-2), 31–55.

Topalova, Petia and Amit Khandelwal, "Trade Liberalization and Firm Productivity: The Case of India," *The Review of Economics and Statistics*, 2011, *93* (3), 995–1009.

Travis, William Penfield, *The Theory of Trade and Protection* number 121, Harvard University Press, 1964.

Venables, Anthony J., "Trade and Trade Policy with Differentiated Products: A Chamberlinian-Ricardian Model," *The Economic Journal*, 1987, *97* (387), 700–717.