Journal of International Economics 145 (2023) 103824

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Full length articles

A second-best argument for low optimal tariffs on intermediate inputs[☆]

Lorenzo Caliendo[a,b], Robert C. Feenstra[c,b], John Romalis[d,e], Alan M. Taylor[c,b,f,*]

[a] *Yale University, United States of America*
[b] *NBER, United States of America*
[c] *UC Davis, United States of America*
[d] *Macquarie University, Australia*
[e] *ABFER, Singapore*
[f] *CEPR, United Kingdom*

ARTICLE INFO

Dataset link: https://rcfeenstra.github.io/CFRT/

JEL classification:
F12
F13
F17
F61

Keywords:
Trade policy
Monopolistic competition
Gains from trade
Input–output linkages

ABSTRACT

We derive a new formula for the optimal uniform tariff in a small-country, heterogeneous-firm model with roundabout production and a nontraded good. Tariffs are applied on imported intermediate inputs. First-best policy requires that markups on domestic intermediate inputs are offset by subsidies. In a second-best setting where such subsidies are not used, roundabout production and the monopoly distortion in the traded sector create strong incentives to lower the optimal tariff on imported inputs. In a quantitative version of our two-sector small open economy, we find that the optimal tariff is lowered under nearly all parameter values considered, and can be negative.

1. Introduction

The use of tariffs to protect traded goods such as manufactures has a long history. In his famous *Report on Manufactures*, Alexander Hamilton argued for moderate tariffs combined with direct subsidies to promote manufacturing. Opposition to these subsidies came from Thomas Jefferson and James Madison, who favored even higher tariffs, and Madison's administration would put in place the first protectionist tariff in the United States (Irwin, 2004). The administration of President Donald Trump enacted tariffs, often at 25%, to protect several manufacturing industries and against a broad range of products from China. Significantly, the Chinese products were initially selected to minimize the direct impact on consumer prices, leaving American businesses facing the brunt of tariffs on their imported inputs (Fajgelbaum et al., 2020).

Does modern trade theory offer any new answer to this old question of whether to protect the traded sector? To answer this, we investigate a small open economy (SOE) with two sectors – one traded and the other nontraded – and with heterogeneous firms,

☆ For their helpful comments we thank the editor, Andrés Rodríguez-Clare, and the referees, along with Pol Antràs, Kyle Bagwell, John Fernald, Teresa Fort, Robert Staiger, and participants at Dartmouth College and the NBER. We also acknowledge Anna Ignatenko and Ninghui Li for providing excellent research assistance. Romalis acknowledges support from the Australian Research Council. The usual disclaimer applies.

* Corresponding author at: UC Davis, United States of America.
E-mail addresses: lorenzo.caliendo@yale.edu (L. Caliendo), rcfeenstra@ucdavis.edu (R.C. Feenstra), john.romalis@mq.edu.au (J. Romalis), amtaylor@ucdavis.edu (A.M. Taylor).

https://doi.org/10.1016/j.jinteco.2023.103824
Received 25 April 2021; Received in revised form 21 September 2023; Accepted 21 September 2023
Available online 12 October 2023

monopolistic competition and CES preferences (as in Melitz, 2003). We adopt a Pareto distribution for productivity (as in Chaney, 2008) and also roundabout production.[1] As described in Section 2, the differentiated intermediate inputs in each sector are bundled into a finished good that is sold to home consumers and firms in that sector, but not traded, while the differentiated inputs are traded in one sector. A tariff is applied to imports of these differentiated intermediate inputs.

Demidova and Rodríguez-Clare (2009) obtain a formula for the optimal uniform tariff in a SOE with one sector, no roundabout production and heterogeneous firms which we denote by t^{het}. Because there is no roundabout production, we can think of this tariff as applying to imported final good varieties. They argue that this single tariff instrument obtains the first-best by offsetting two distortions: the need to correct for the markup on domestic final goods (by applying a tariff equal to that markup) and the externality present because imported varieties bring surplus that is not taken into account in domestic spending (by slightly lowering the tariff). When there is roundabout production, however, then t^{het} does not correct for the markup on domestic input varieties that is passed-through to the price of the bundled finished good, which is further used as an input to the production of other differentiated inputs. Labor is also used in production, so that markup distorts the use of the finished good relative to labor. In addition, the presence of a second (nontraded) sector creates a further monopoly distortion. The question we address is: what is the optimal tariff on the imported inputs, in the absence of other policy instruments?[2]

In a closed economy, analyzed in Section 3, we show that the distortion created by the markup on differentiated inputs is corrected by applying subsidies to the finished good purchased in both sectors.[3] In the open economy analyzed in Section 4, first best policy requires subsidies on the finished goods in addition to the tariff t^{het}. When subsidies are not used, however, then the second-best policy in an open economy is to lower the import tariff below t^{het}, thereby lowering the price of the finished good in the traded sector. *Our key result shows conditions under which the optimal second-best tariff on imported varieties is below t^{het}, due to the presence of roundabout production and the monopoly distortions in both sectors.*

We obtain the optimal uniform, second-best tariff as a fixed-point of a formula described in Section 5 that has two new terms: a term M that reflects the relative monopoly distortion between the traded and nontraded sectors; and a term R that reflects roundabout production in the traded sector as well as the monopoly distortion there. In Section 6 we consider a quantitative version of our two-sector SOE, where we find that the optimal tariff is lower that t^{het} under nearly all parameter values considered, and can be negative.[4] Further conclusions are in Section 7.

1.1. Related literature

Costinot et al. (2020) analyze optimal tariffs on final differentiated goods with very general tastes and technologies, and they show that optimal tariffs can be lowered (and even made negative) by having a non-Pareto distribution for productivity or linear foreign preferences. They are the first to extend the analysis to *nonuniform* tariffs, and they find that the importing country should use an import subsidy on the least efficient foreign exporters. Haaland and Venables (2016) demonstrate a potential second-best role for reduced trade taxes to offset a monopoly distortion, as does the earlier work by Flam and Helpman (1987). These papers all focus on trade in final goods, while the impact of tariffs on inputs along global supply chains is examined by Antràs and Chor (2021), Beshkar and Lashkaripour (2020), Blanchard et al. (2016) and Grossman and Helpman (2021).

Recently, Antràs et al. (2022) have analyzed "tariff escalation", which means higher (optimal) tariffs on final goods than on intermediate inputs. Their model and ours differ in many of the details: they have two sectors with sequential production, with the strongest results obtained when labor is used in the downstream sector; whereas we have two sectors with roundabout production, and no labor used downstream. Despite these differences, we believe that the underlying distortion is the same and arises from the markup on domestic inputs. As a result, subsidizing domestic inputs (in the first-best) or lowering the tariff on imported inputs (in the second-best) is needed to offset those markups.[5]

Our work is most closely related to Lashkaripour and Lugovsky (2020) and Lashkaripour (2021). The former authors analyze optimal uniform first-best tariffs with multiple sectors and input–output linkages. When considering second-best tariffs, however, they do not incorporate these linkages. Still, we build on their result that the first-best policy in the presence of input–output linkages will be to offset the markups charged by sellers of intermediate inputs by providing a subsidy to those buyers (and we show the

[1] Roundabout production means that the output of a sector is used as an input into the same sector: see Krugman and Venables (1995) and Yi (2010).

[2] We are assuming that the imported differentiated good is not purchased by consumers directly. If so, and if the government could prevent resale between consumers and firms, then it is possible that a different tariff should be applied on the two groups. But this action would not offset the need to charge a *low* tariff on the input varieties that firms purchase – as we shall argue – so as to offset the markup on the domestic varieties, that is passed-through to the price of the bundled, finished good sold to firms

[3] For consumers, the first-best subsidy is in relative terms (see Section 3), since it does not matter if the consumer prices in both sectors are high provided that the tax revenue is redistributed. But for firms purchasing the finished good, the subsidy must *exactly* offset the markup that is passed-through from the input varieties.

[4] In our working paper Caliendo, Feenstra, Romalis and Taylor (CFRT, 2021), we analyze a 186-country, 15-sector quantitative model for 2010 with a general input–output structure. For manufacturing, the one-sector, no roundabout, first-best tariff is 27.3% for our parameter values. We find that the optimal second-best tariff has a median value of only 10% (or 7.5% for countries with above-median shares of manufacturing production), and is negative for five countries: Bhutan, Myanmar, New Caledonia, Hong Kong, and Spain.

[5] There is one important distinction between our models, which arises from the impact of a tariff on intermediate inputs on domestic entry into that sector. Because we have only one traded sector, the import tariff is equivalent to an export tax on that sector (due to Lerner symmetry) and it inhibits entry. In contrast, Antràs et al. (2022) have two traded sectors, so that a tariff on the upstream sector alone is not equivalent to an export tariff on that sector, and it is quite possible that entry increases as in the *firm-delocation* literature. See further discussion in Sections 2.1 and 7.

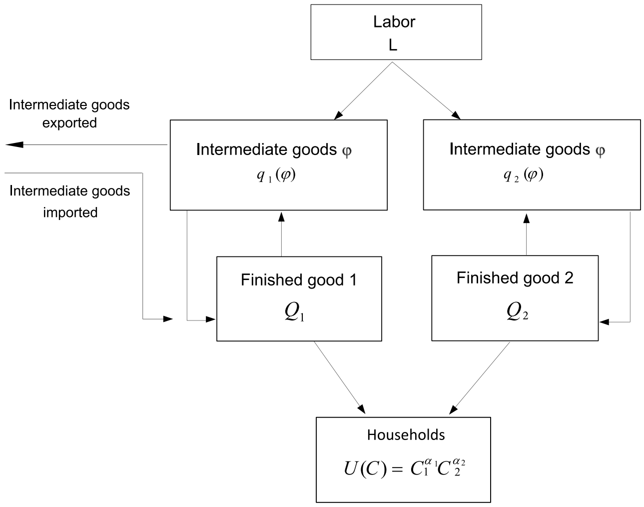

Fig. 1. Schematic production structure.

same result in a closed economy). Our main interest is in the second-best optimal tariff in the absence of the subsidy offsetting the markup.

Lashkaripour (2021) analyzes the second-best use of tariffs in a setting that incorporates input–output linkages. He assumes that there is "duty drawback" on the tariffs applied to imported intermediate inputs, meaning that those duties are forgiven when the imported inputs are used in the production of exported goods. We do not rely on this assumption. Despite differences in the questions that we address (Lashkaripour analyzes Nash-equilibrium tariffs whereas we investigate tariffs for a SOE), there are similarities in our results. Lashkaripour stresses that the welfare impact of tariffs depend on their ability to raise wages in the importing country, and that input–output linkages reduce the calculated optimal tariffs. We similarly show that the wage impact of tariffs is reduced due to roundabout production in the traded sector, which is one reason for the optimal tariff to be lowered. In the presence of markups, Lashkaripour (2021) argues that the second-best tariff should offset those domestic distortions. We likewise find that the monopoly distortion in the traded sector – in conjunction with roundabout production there – is another reason to lower the optimal tariff. Adding the nontraded sector creates a further distortion, and whether this increases or decreases the tariff depends on which sector is more distorted.

2. Two-sector economy with roundabout production

We analyze a two-sector Melitz (2003)–Chaney (2008) model with roundabout production, similar to Arkolakis et al. (2012, section IV) and Costinot and Rodríguez-Clare (2014). We summarize key equations here and Online Appendix A describes the full model with heterogeneous firms, while Appendix B outlines the model with homogeneous firms.

There are two countries, where the home country is a small open economy (SOE) and the foreign country is denoted by an asterisk. As illustrated in Fig. 1, there are two sectors $s = 1, 2$ at home, where sector 1 is traded and sector 2 is nontraded. In both sectors, firms produce differentiated inputs under monopolistic competition, which are costlessly bundled into a *finished good* in CES fashion, with elasticity $\sigma_s > 1$. The finished good is nontraded in both sectors, and it is sold to domestic consumers as a final good and also to domestic firms in the same sector as an intermediate input, used to produce differentiated inputs (e.g., firms produce machinery parts using machines). In sector 1, the imported differentiated inputs are subject to iceberg costs and a tariff, where one plus the *ad valorem* tariff is denoted by t_1.

The finished output in each sector has quantity Q_s, price index P_s, and value $Y_s \equiv P_s Q_s$. With roundabout production, the marginal cost of producing a differentiated input for a firm with productivity $\varphi_s = 1$ in sector s is

$$c_s \equiv w^{(1-\gamma_s)} P_s^{\gamma_s}, \tag{1}$$

where $0 < (1 - \gamma_s) \leq 1$ is the labor share so that γ_s indicates the amount of roundabout production. We refer to (1) as the *input cost index*.

A mass of firms N_s^e incur fixed labor costs of entry f_s^e to enter in each sector. In both the homogeneous and heterogeneous firms models, that mass is endogenously determined from the full-employment conditions for the economy. With homogeneous firms, all firms receive a productivity of unity; with heterogeneous firms, each firm receives a productivity draw from a Pareto distribution, $G_s(\varphi_s) = 1 - \varphi_s^{-\theta_s}$, with $\varphi_s \geq 1$ and $\theta_s > \sigma_s - 1$. As is familiar in the Melitz–Chaney model, firms choose to produce the differentiated input for the domestic market or to export if their productivities exceed some *cutoff* levels, and in each case, the firms then incur additional fixed labor costs.

Consumers have Cobb–Douglas preferences over final goods in the two sectors, with

$$U = C_1^{\alpha_1} C_2^{\alpha_2}, \quad \alpha_1 + \alpha_2 = 1, \quad \alpha_1 > 0 \text{ and } \alpha_2 \geq 0, \tag{2}$$

where α_s is the expenditure share on the sector $s = 1, 2$. Consumer income I includes labor income wL (labor is the only factor of production), plus rebated tariff and tax revenue B, while free entry ensures that expected firm profits equal zero.

Domestic consumer demand for finished goods equals $\alpha_s(wL + B)$ in sector s. The finished good is also sold to firms in the same sector who are producing the differentiated intermediate inputs. To compute those sales, we start with the value of the finished good Y_s, which reflects the value of all intermediates – local and imported – that are bundled together. Let λ_{ds} denote the share of differentiated inputs that are purchased locally, where $\lambda_{d1} \leq 1$ in the traded sector 1 but $\lambda_{d2} = 1$ in the nontraded sector 2. Then $\lambda_{ds} Y_s$ is the value of locally-produced differentiated inputs. We also need to eliminate the markup on those inputs by dividing by $\sigma_s / (\sigma_s - 1)$ to obtain their costs of production, and then we take the share γ_s to obtain the value of finished goods that are sold as an input to firms. We denote the markup-adjusted cost share by

$$\tilde{\gamma}_s \equiv \gamma_s \rho_s \text{ with } \rho_s \equiv \frac{(\sigma_s - 1)}{\sigma_s}. \tag{3}$$

The market clearing condition for the nontraded sector 2 is then $Y_2 = \alpha_2(wL + B) + \tilde{\gamma}_2 Y_2$ where $\lambda_{d2} \equiv 1$. In sector 1, the market clearing condition is more complex, with

$$Y_1 = \alpha_1(wL + B) + \tilde{\gamma}_1 \left(\lambda_{d1} Y_1 + \lambda_{x1} Y_1^* \right). \tag{4}$$

The term $\tilde{\gamma}_1 \lambda_{d1} Y_1$ on the right reflects the sale of the finished good to home firms. The next term, $\tilde{\gamma}_1 \lambda_{x1} Y_1^*$ reflects the home finished good *used in the production of the sector 1 differentiated inputs that are exported* (remember that the finished good is not directly exported). To obtain this term, we start with the *foreign* value of the finished good Y_1^*, and we define λ_{x1} as the home share of intermediate inputs that are bundled together to obtain Y_1^*. Then $\lambda_{x1} Y_1^*$ is the value of home exports of differentiated inputs, and once again we apply the parameter $\tilde{\gamma}_1$ to obtain the finished good that is sold to home firms to create those exports.

The expenditure shares are determined in equilibrium: in a heterogeneous firm model these depend on the optimal choice of cutoff productivities by firms, while in a homogeneous firm model the productivities are exogenously fixed at unity.[6] The cutoff productivities depend on the fixed costs of domestic production and exporting, and we assume that all fixed costs are paid in terms of labor in the source country, with the foreign wage chosen as the numeraire ($w^* \equiv 1$.)

To close the model, we need to use trade balance. As noted above, the term $\lambda_{x1} Y_1^*$ in (4) is the value of home exports of the differentiated inputs. Under balanced trade, this must equal the net-of-tariff value of imports. Letting t_1 denote one plus the *ad valorem* home import tariff in sector 1, the trade balance condition is then

$$\lambda_{x1} Y_1^* = \frac{\lambda_{m1}}{t_1} Y_1, \tag{5}$$

where λ_{m1} is the share of intermediate inputs used in sector 1 that are imported, with $\lambda_{d1} + \lambda_{m1} = 1$.[7]

As described by Demidova and Rodríguez-Clare (2013), the trade balance condition determines the wage w in the SOE, taking the foreign wage w^* as the numeraire. The equilibrium conditions of the SOE assume that changes in the tariff t_1 have a negligible impact on the foreign price index P_1^* and value of output Y_1^*. Fixing the values of P_1^* and Y_1^* means that the location of the foreign demand curve for a home exported variety is itself fixed, though that CES demand curve is not infinitely elastic as in a small-country *competitive* model. This means that trade policy has an impact on the small country's export price and therefore on its terms of trade.[8] We stress that the definition of a small open economy from Demidova and Rodríguez-Clare (2013) allows for a wide range of values for the home expenditure share, $0 < \lambda_{d1} < 1$, and likewise for its import share, $0 < \lambda_{m1} = 1 - \lambda_{d1} < 1$. A *special case* of the small open economy would be to consider $\lambda_{d1} \to 0$, so that the small country is importing nearly all of its intermediate inputs from abroad. We will not make use of this condition except as a limiting example after deriving our main results.

2.1. Response of output and entry to the tariff

Before examining optimal policy, we describe the response of the finished good Y_1 and entry into each sector to the tariff. Using trade balance in (5), we can rewrite market clearing (4) as

$$Y_1 = \alpha_1(wL + B) + \tilde{\gamma}_1 \Lambda_1 Y_1, \quad \text{with } \Lambda_1 \equiv \left(\lambda_{d1} + \frac{\lambda_{m1}}{t_1} \right). \tag{6}$$

The first term on the right (6) is the demand for Y_1 as a final good, whereas the second term is the demand for Y_1 as an intermediate input, where Λ_1 equals the domestic share plus the *duty-free* import share. While this term is unity under either free trade ($t_1 = 1$) or autarky ($t_1 \to +\infty$ so $\lambda_{d1} = 1$ and $\lambda_{m1} = 0$), it has a lower value $\Lambda_1 < 1$ for all finite tariffs $t_1 > 1$.

[6] The full equilibrium conditions are in Online Appendix A, Definition 1 for heterogeneous firms, and Appendix B, Definition 2 for homogeneous firms.

[7] Note that the import share is evaluated using the foreign export prices p_{x1}^* that are inclusive of the iceberg costs of trade, the markup, and the tariff t_1. In other words, we are assuming that the tariff is applied to the c.i.f. value of imports — including the markup. See further discussion in note 34. For simplicity, we assume no foreign tariff.

[8] To use the apt phrase of Bartelme et al. (2019), a small country is "an economy that is large enough to affect the price of its own good relative to goods from other countries, but too small to affect relative prices in the rest of the world".

We can simplify (6) by substituting for tariff revenue $B = \frac{t_1 - 1}{t_1}\lambda_{m1}Y_1$. Using $\lambda_{d1} + \lambda_{m1} = 1$, we can re-express tariff revenue as $B = \left(1 - \Lambda_1\right)Y_1$, and substituting above we obtain

$$Y_1 = \alpha_1[wL + (1 - \Lambda_1)Y_1] + \tilde{\gamma}_1\Lambda_1 Y_1. \tag{7}$$

We see that starting at free trade, a tariff exerts two different forces on the value of the finished good, Y_1. On one hand, it raises tariff revenue $B = \left(1 - \Lambda_1\right)Y_1$ and increases consumer demand. On the other hand, it lowers duty-free imports and therefore lowers exports and Λ_1. Which of these forces dominates depends on the parameters α_1 and $\tilde{\gamma}_1$. We can readily solve for real output Y_1/w from (7) as

$$\frac{Y_1}{w} = \frac{\alpha_1 L}{\alpha_2 + (\alpha_1 - \tilde{\gamma}_1)\Lambda_1}. \tag{8}$$

We see that these two forces just offset each other when $\alpha_1 = \tilde{\gamma}_1$, in which case Y_1/w does not vary with the tariff. When $\alpha_1 > \tilde{\gamma}_1$, then consumer demand dominates and Y_1/w is a ∩-shaped function of the tariff (i.e., the inverse shape of Λ_1). In contrast, when $\alpha_1 < \tilde{\gamma}_1$ then exports dominate and Y_1/w is a ∪-shaped function of the tariff.

Substituting (8) back into $B = \left(1 - \Lambda_1\right)Y_1$, we obtain

$$\frac{B}{w} = \frac{\alpha_1 L\left(1 - \Lambda_1\right)}{1 - \tilde{\gamma}_1 - (\alpha_1 - \tilde{\gamma}_1)(1 - \Lambda_1)} = \frac{\alpha_1 L}{\frac{1 - \tilde{\gamma}_1}{(1 - \Lambda_1)} - (\alpha_1 - \tilde{\gamma}_1)}. \tag{9}$$

Because $\tilde{\gamma}_1 < 1$, from this final equation we see that B/w is monotonically decreasing in Λ_1, so their critical points are at the same tariff which we refer to as the *maximum (real) revenue* tariff. It follows from (8) that Y_1/w also has a critical point at that tariff.

The ambiguity in the shape of Y_1 does not extend to the entry of firms producing differentiated inputs in sector 1. Entry is proportional to the demand for those inputs for home sales, $\lambda_{d1}Y_1$, plus the demand for exports, $\lambda_{x1}Y_1^* = \frac{\lambda_{m1}}{t_1}Y_1$, which sum to $\Lambda_1 Y_1$. Using output in (8), entry is then

$$N_1^e = \frac{(\sigma_1 - 1)\Lambda_1 Y_1}{f_1^e \theta_1 \sigma_1} = \frac{\alpha_1(\sigma_1 - 1)}{f_1^e \theta_1 \sigma_1}\left[\frac{L}{\frac{\alpha_2}{\Lambda_1} + (\alpha_1 - \tilde{\gamma}_1)}\right]. \tag{10}$$

The ∪-shape for Λ_1 means that N_1^e is also ∪ -shaped provided that $\alpha_2 > 0$: it falls as the tariff is increased from free trade, has a minimum at the maximum-revenue tariff, and then rises again to the same value in autarky and free trade. The intuition for this result is Lerner symmetry (Costinot and Werning, 2019): the import tariff acts like an export tax, and starting from free trade the tariff depresses entry into the traded sector and moves resources into the nontraded sector. In particular, entry into the nontraded sector is

$$N_2^e = \frac{\alpha_2(\sigma_2 - 1)}{f_2^e \theta_2 \sigma_2 (1 - \tilde{\gamma}_2)}\left(L + \frac{B}{w}\right),$$

which is a ∩-shaped function of the tariff because of tariff revenue. When $\alpha_1 = 1$ and there is no nontraded sector, however, then entry into the traded sector does not change with the tariff.

While the above equations have used the Pareto parameter θ_s from the heterogeneous firm model, the same conditions are obtained with homogeneous firms (see Appendix B) where we replace θ_s with:

$$\theta_s = \sigma_s^{hom} - 1. \tag{11}$$

Making this substitution in the equations above, the homogeneous firm model has the properties just discussed: in the presence of a nontraded sector, entry into the traded sector is *reduced* for an increase in the tariff starting from free trade, and then returns to its autarky value as the tariff becomes prohibitive. Condition (11) is familiar from Arkolakis et al. (2012), who demonstrate that the heterogeneous and homogeneous firm models are very similar in certain respects that include, as we have just argued, the impact of a tariff on entry.[9] We will see, however, that selection still plays a distinct role on the impact of a tariff with heterogeneous firms, especially in the presence of roundabout production.[10]

These results on entry contrast with the quite different results in the *firm-delocation* literature that combines a monopolistically competitive traded sector with a competitive *traded* outside good (see, e.g., Venables, 1987; Melitz and Ottaviano, 2008; Ossa, 2011), where a tariff *attracts* firms into the country applying it. In those models, the freely-traded outside good produced pins down the relative wage across countries, and a tariff on the monopolistically competitive imports *is not* the same as an export tax on those goods (since Lerner symmetry in this case implies that a uniform import tariff is equivalent to an export tax across *both* sectors). We return to this contrast in our concluding section.

[9] The full isomorphism between the models requires, however, that the fixed costs of exporting use resources of the destination country, as Arkolakis et al. (2012) explain. That is not our assumption, so there will be some differences between the models for this reason: see note 16.

[10] Note that Costinot and Rodríguez-Clare (2014) also find a distinct role for selection in the presence of multiple sectors and roundabout production, as we discuss in note 17.

3. Optimal consumer and producer taxes in a closed economy

Before considering a tariff, we discuss the distortions arising in a closed economy from having monopolistic production of the differentiated inputs, where both sectors $s = 1, 2$ are nontraded. The markup on the differentiated inputs is fully passed-through to the price of the bundled, finished good. That distortion then operates on two margins: consumer purchases of finished goods; and firm purchases of finished goods as inputs, where the higher price on the finished good leads to inefficiently low purchases of the finished good as compared to labor. Rather than correcting the monopoly distortion at its source (i.e., in the price of differentiated inputs), it will be instructive to explore how one could correct it by using taxes/subsidies on purchases of the finished goods on these two margins. So we consider both consumer and producer taxes/subsidies on purchases of the finished goods, where one plus the *ad valorem* rates are denoted by t_s^c and t_s^p, respectively.

With heterogeneous firms, the cutoff productivities chosen in each sector do not depend on these tax/subsidy instruments (see Appendix C). It follows that the optimal policy with homogeneous firms or heterogeneous firms is identical. We consider two solutions to the closed-economy problem: first, choosing both the consumer and producer taxes/subsidies optimally; and second, using only the consumer tax/subsidy while setting $t_s^p \equiv 1$. When both instruments are used, we obtain the solution

$$t_s^p = \rho_s = \left(\frac{\sigma_s - 1}{\sigma_s}\right) < 1 \quad \text{and} \quad \frac{t_1^c}{t_2^c} = \frac{\rho_1}{\rho_2}. \tag{12}$$

The optimal producer subsidies $t_s^p < 1$ exactly counteract the markups on differentiated inputs which would otherwise be fully passed-through to finished goods prices.[11] With these subsidies, firms pay prices for finished goods that reflect their marginal costs. In addition, optimal consumption taxes/subsidies are needed so that, in relative terms, these prices offset the markups implicit in finished goods' prices faced by consumers.

In contrast to this first-best case, consider the second-best policy that involves consumption taxes/subsidies only. In that case, the distortion that arises from having a high price of the finished good as an input (due to the markup on differentiated inputs that is fully passed-through to the finished good price) is not corrected. It is instructive in this case to solve for the price of the finished good. That price index, in the absence of imports, is

$$P_s = \left(N_s^e \int_{\varphi_{ds}}^{\infty} p_{ds}(\varphi)^{1-\sigma_s} g_s(\varphi)\, d\varphi\right)^{\frac{1}{1-\sigma_s}} = \left(N_{ds}\right)^{\frac{-1}{(\sigma_s-1)}} \left(\frac{\sigma_s}{\sigma_s - 1}\right) \frac{c_s}{\bar{\varphi}_{ds}}, \tag{13}$$

where N_s^e is the mass of entering firms and φ_{ds} is the cutoff productivity to remain in the market, while $N_{ds} = N_s^e[1 - G_s(\varphi_{ds})]$ is the mass of surviving firms (equal to domestic product variety) and $\bar{\varphi}_{ds}$ is their average productivity,[12] so that $\left(\frac{\sigma_s}{\sigma_s-1}\right)\frac{c_s}{\bar{\varphi}_{ds}}$ is their average price. We now substitute from the input cost index in (1) to solve for the price index,

$$P_s = \left(N_{ds}\right)^{\frac{-1}{(1-\gamma_s)(\sigma_s-1)}} w \left(\frac{\sigma_s}{\sigma_s - 1}\frac{1}{\bar{\varphi}_{ds}}\right)^{\frac{1}{(1-\gamma_s)}}. \tag{14}$$

Notice that the impact of product variety on reducing the price index has increased from $1/(\sigma_s - 1)$ in (13) to $1/(1-\gamma_s)(\sigma_s - 1)$ in (14). That change carries through to other equations for the closed economy equilibrium, so the economy with roundabout production is effectively acting like a closed economy *without* roundabout but with a lower elasticity of substitution, defined by

$$\tilde{\sigma}_s \equiv 1 + (1 - \gamma_s)(\sigma_s - 1) < \sigma_s. \tag{15}$$

A change in entry in (14) changes the price index with the exponent $1/(\tilde{\sigma}_s - 1)$, so that $\tilde{\sigma}_s$ is the *effective* elasticity of substitution. We find the second-best optimal consumption taxes/subsidies (see Appendix C) are given by

$$\frac{t_1^c}{t_2^c} = \left(\frac{\tilde{\sigma}_1 - 1}{\tilde{\sigma}_1}\right) \Big/ \left(\frac{\tilde{\sigma}_2 - 1}{\tilde{\sigma}_2}\right). \tag{16}$$

Notice that the consumption taxes/subsidies in (16) are similar to those in (12), but are now evaluated using the effective elasticities of substitution: the sector with the lowest effective elasticity must have the lowest tax (i.e., greatest subsidy) to offset the effective monopoly distortion. Even if the elasticities $\sigma_s \equiv \sigma > 1$ are identical then the sector with the stronger roundabout production (higher γ_s) will have the lower effective elasticity in (15) and should be subsidized. The role of the effective elasticities in this second-best case for a closed economy will be useful as we examine tariffs on trade, to which we turn next.

4. First-best uniform tariff in a small open economy

Demidova and Rodríguez-Clare (2009) analyze a SOE with one sector and no roundabout production. They identify two distortions arising from monopolistic competition. The first is the markup charged on the domestic differentiated varieties which

[11] The need for such subsidies in a dynamic monopolistic competition model was noted by Judd (1997, 2002).

[12] This average productivity is defined as in Melitz (2003) and equals $\bar{\varphi}_{ds} = \varphi_{ds}\left(\frac{\theta_s}{\theta_s - \sigma_s + 1}\right)^{\frac{1}{\sigma_s - 1}}$.

can be corrected by subsidizing domestic buyers of those inputs, where one minus the *ad valorem* subsidy is set equal to the inverse of the markup,

$$t_1^d = \rho_1 = \frac{\sigma_1 - 1}{\sigma_1}. \tag{17}$$

Alternatively, the markup on domestic varieties can be offset by using a tariff on imported varieties equal to the markup, $t^{hom} = \frac{1}{\rho_1} = \frac{\sigma_1}{\sigma_1 - 1}$, which offsets the domestic markup in relative terms by introducing the same distortion on import prices. This is the optimal tariff in a one-sector SOE with monopolistic competition and homogeneous firms (Gros, 1987).

With heterogeneous firms, however, Demidova and Rodríguez-Clare (2009) find that there is a second distortion: each new foreign variety brings surplus, which domestic buyers do not take account of in their spending. One way to correct this externality is to use an import subsidy, and they find that one minus the optimal *ad valorem* subsidy is

$$t_1^m = \frac{\theta_1 \rho_1}{(\theta_1 - \rho_1)} < 1, \tag{18}$$

where the inequality follows from $\theta_1 > \sigma_1 - 1$. So the first-best is achieved by using the two subsidies t_1^d, $t_1^m < 1$. Furthermore, they argue that an *equivalent* policy is to multiply the tariff $t^{hom} = \frac{1}{\rho_1}$ by the import subsidy in (18), and then both distortions are corrected by a single instrument, which is the optimal tariff in a one-sector model with heterogeneous firms,

$$t^{het} \equiv t^{hom} \times t_1^m = \frac{\theta_1}{(\theta_1 - \rho_1)} > 1. \tag{19}$$

It is immediate that $t^{het} < t^{hom}$ when evaluated with the same parameter σ_1, since $t_1^m < 1$.[13] But even when we do the more exact comparison across models, then we still find that $t^{het} < t^{hom}$ because $\theta_1 > \sigma_1^{het} - 1$ and so using (11), $\sigma_1^{hom} > \sigma_1^{het}$ and $\rho_1^{hom} > \rho_1^{het}$.[14]

If we add a second sector or roundabout production, then the equivalence of using the policy $t_1^d, t_1^m < 1$ and the optimal tariff $t^{het} > 1$ no longer holds, however. To see this, suppose that we "scale up" t_1^d, t_1^m by dividing by ρ_1 (i.e., multiplying by $\frac{\sigma_1}{\sigma_1 - 1}$), thereby obtaining $t_1^d = 1$ and the import tariff of t^{het}, and then use a subsidy of ρ_1 on the finished good to *offset* this scaling-up. With a single sector and no roundabout production, this subsidy does not make any difference because consumers cannot substitute away from the finished good and firms do not purchase it. But once we add multiple sectors and/or roundabout production, then substitution by consumers and firms means that the subsidy of ρ_1 is needed to avoid the downstream impact of the markup $\frac{\sigma_1}{\sigma_1 - 1}$, as we found in the closed economy. In general, for an open economy with multiple sectors and input–output linkages, Lashkaripour and Lugovsky (2020) argue that such subsidies must be applied in the first-best: in that case, the first-best tariffs for a *small* country are the same with and without input–output linkages.[15] Our interest is in the second-best tariff obtained in the absence of such subsidies, which we turn to next.

5. Second-best uniform tariff in a small open economy

We now add the nontraded sector 2, which can also have roundabout production, and we suppose that the *only* policy instrument available is a uniform import tariff (or subsidy) t_1 with an optimal second-best value t_1^*. The fact that a subsidy on the finished good is not used creates a robust reason for lowering the optimal tariff below t^{het}. A slight reduction of the tariff below its first-best value ordinarily causes only a second-order loss in welfare, but it now brings a first-order gain in welfare because it lowers the price of the finished good purchased by firms. There are two other reasons to have $t_1^* < t^{het}$, which arise from the response of wages and the response of entry to changes in the tariff. We consider each of these in the following sections.

5.1. Response of the home wage to the tariff

A key insight of Demidova and Rodríguez-Clare (2013) is that in the monopolistic competition model, even a SOE experiences an increase in its wage from applying a tariff. That wage increase results in a rise in its export prices, which is analogous to the terms of trade effect of a tariff that occurs in competitive models. Lashkaripour (2021) stresses the importance of this wage elasticity in determining the welfare impact of tariffs changes, and therefore the Nash-equilibrium tariffs in his model.

When solving for the impact of the tariff on wages, we would like to compare the solutions with homogeneous firms and heterogeneous firms, and also understand the impact of roundabout production in either case. We begin by examining the trade balance condition (5) in the homogeneous firm model, where the export share λ_{x1} equals

$$\lambda_{x1} = N_1^e \left(\frac{\sigma_1}{\sigma_1 - 1} \frac{c_1 \tau_{x1}}{P_1^*} \right)^{1-\sigma_1}, \tag{20}$$

[13] The same small-country formula for the optimal tariff as (19) is obtained by Felbermayr et al. (2013), who show that the optimal tariff in a large country is higher.

[14] From (11) and (19) we then have $t^{het} = 1/[1 - (\rho_1^{het}/\theta_1)] < 1/[1 - (\rho_1^{hom}/\theta_1)] = 1/[1 - (1/\sigma_1^{hom})] = t^{hom}$.

[15] See their Section 4(ii) and especially footnote 23, which explains that for a small open economy the equations for the first-best taxes and tariffs are identical with and without input–output linkages.

where N_1^e is the endogenous entry of firms into sector 1, τ_{x1} are iceberg trade costs, and P_1^* is the foreign price index in sector 1 which is exogenous for the SOE. The import share λ_{m1} equals

$$\lambda_{m1} = N_1^{e*}\left(\frac{\sigma_1}{\sigma_1 - 1}\frac{c_1^*\tau_{x1}^* t_1}{P_1}\right)^{1-\sigma_1}, \tag{21}$$

where N_1^{e*} is the entry of foreign firms into sector 1 and c_1^* are their input costs, both of which are exogenous. The home price index P_1 is endogenous, but given its value then an increase in the tariff t_1 reduces the import share, and reduces the tariff-free import share λ_{m1}/t_1 even more. Given Y_1, then to satisfy trade balance this reduction in duty-free imports $\lambda_{m1}Y_1/t_1$ would need to be matched by a reduction in exports $\lambda_{x1}Y_1^*$. That can be achieved by an increase in home wages, which raise the input costs c_1 in (20). This reasoning illustrates the positive terms of trade impact of a tariff in the SOE, but it needs to be sharpened to take into account the endogenous price index P_1 and also roundabout production.

To solve for P_1, we proceed indirectly by focusing on the *domestic* share, which equals

$$\lambda_{d1} = N_1^e\left(\frac{\sigma_1}{\sigma_1 - 1}\frac{c_1}{P_1}\right)^{1-\sigma_1}. \tag{22}$$

Inverting this equation we obtain an expression for the sector 1 price index,

$$P_1 = \left(\frac{\lambda_{d1}}{N_1^e}\right)^{\frac{1}{(\sigma_1-1)}}\frac{\sigma_1 c_1}{(\sigma_1 - 1)}. \tag{23}$$

Replacing the domestic share λ_{d1} by $1-\lambda_{m1}$ in this expression, and also substituting from (1), we solve for the price index as

$$P_1 = \left(\frac{1-\lambda_{m1}}{N_1^e}\right)^{\frac{1}{(1-\gamma_1)(\sigma_1-1)}} w\left(\frac{\sigma_1}{\sigma_1 - 1}\right)^{\frac{1}{(1-\gamma_1)}}. \tag{24}$$

Notice that the impact effect of the tariff on reducing the import share λ_{m1} will increase the price index P_1, and this index is increasingly sensitive to the import share as the extent of roundabout production grows, so that γ_1 rises.

Substituting P_1 back in the input cost index in (1), and totally differentiating, we obtain

$$\hat{c}_1 = \hat{w} - \frac{1}{(\sigma_1 - 1)}\left(\eta_{m1}\hat{\lambda}_{m1} + \frac{\gamma_1\hat{N}_1^e}{(1-\gamma_1)}\right) \quad \text{where} \quad \eta_{m1} \equiv \frac{\gamma_1\lambda_{m1}}{(1-\gamma_1)(1-\lambda_{m1})}. \tag{25}$$

Intuitively, after the impact effect of the tariff on reducing duty-free imports, think of the equilibrium being restored by a rise in the input costs c_1, which reduces exports. In the absence of roundabout production, the rise in c_1 is achieved by an increase in the wage. With roundabout, however, we see from (25) that the fall in the import share itself – by raising the price index in (24) – contributes to restoring equilibrium, so that a *smaller* increase in the wage is needed. The coefficient η_{m1} on $\hat{\lambda}_{m1}$ in (25) is an endogenous variable that depends on the import share, and it is increasing in the amount of roundabout production γ_1. By this argument, the wage impact of the tariff is moderated by the extent of roundabout production, as will be confirmed below. In addition, notice that the induced exit from sector 1 – as we discussed in Section 2.1 – also moderates the increase in the wage needed to obtain a given rise in c_1.

The argument we have just made on how roundabout production reduces the terms of trade impact of the tariff applies with heterogeneous firms, too, in which case selection effects come into play. The above equation for the change in marginal costs (see Appendix A.6) then becomes

$$\hat{c}_1 = \hat{w} - \frac{1}{\theta_1}\left(\eta_{m1}\hat{\lambda}_{m1} + \frac{\gamma_1\hat{N}_1^e}{(1-\gamma_1)} + \frac{\gamma_1(\theta_1 - \sigma_1 + 1)}{(1-\gamma_1)(\sigma_1 - 1)}(\hat{Y}_1 - \hat{w})\right). \tag{26}$$

Using the parameter restriction (11), we see that the fall in the import share has the same impact in (25) and (26), and reduces the increase in the wage needed to restore equilibrium. Induced exit from sector 1 also moderates the increase in the wage. In addition, a third term appears on the right of (26), and that is the change in real output Y_1/w. Recall from our discussion in Section 2.1 that an increase in the tariff from free trade increases (decreases) the real value of output Y_1/w when $\alpha_1 > (<)\ \tilde{\gamma}_1$. The presence of this term in (26) reflects the selection effect of real output on the cutoff productivity for domestic firms. In particular, when roundabout production is strong enough so that $\tilde{\gamma}_1 > \alpha_1$ and $\hat{Y}_1 - \hat{w} < 0$, then this selection effect in the domestic market increases the cutoff, reduces product variety and increases the price index, further moderating the increase in the wage needed to restore equilibrium to the trade balance. This result is our first illustration of how selection due to heterogeneous firms – in conjunction with roundabout production – influences the impact of a tariff.

There is another selection effect that also reduces the terms of trade impact of the tariff with heterogeneous firms, even in the absence of roundabout production. Consider the share of home exporters in the foreign market, which is

$$\lambda_{x1} = \varphi_{x1}^{-\theta_1} N_1^e\left(\frac{\sigma_1}{\sigma_1 - 1}\frac{c_1\tau_{x1}}{\bar{\varphi}_{x1}P_1^*}\right)^{1-\sigma_1}, \tag{27}$$

where $\varphi_{x1} > 1$ is the cutoff productivity for home exporters with average productivity $\bar{\varphi}_{x1}$. The first terms on the right, $\varphi_{x1}^{-\theta_1} N_1^e$, equals the mass of exported varieties and is influenced by the selection of exporters. By solving for the cutoff productivity (see Appendix A.6), we obtain an alternative expression for the export share,

$$\lambda_{x1} = N_1^e \left(\frac{\sigma_1}{\sigma_1 - 1} \frac{c_1 \tau_{x1}}{P_1^*} \right)^{-\theta_1} \left(\frac{\sigma_1 w f_{x1}}{Y_1^*} \right)^{1-\frac{\theta_1}{(\sigma_1-1)}} \left(\frac{\theta_1}{\theta_1 - \sigma_1 + 1} \right). \tag{28}$$

This expression is very similar to the export share in the homogeneous firm model in (20), except for the middle term on the right of (28), involving w/Y_1^*, that reflects the selection of home exporters. The rise in wages from a tariff increases this middle term, which raises the cutoff productivity for home exporters and reduces their export share. This selection effect works in the direction of restoring equilibrium in the trade balance, and therefore reduces the increase in wages needed for equilibrium.[16] This is our second illustration of how selection influences the impact of a tariff.

To summarize, we have argued the impact of the tariff on home wages is reduced when there is roundabout production, and reduced when firms are heterogeneous. To confirm these results, we solve for the marginal impact of the tariff and sector 1 entry on the wage (see Appendix D.1) for either homogeneous or heterogeneous firms as denoted by the superscript z, writing this as

$$\hat{w} = \mathcal{E}_1^z(\gamma_1)\hat{t}_1 + \mathcal{E}_2^z(\gamma_1)\hat{N}_1^e, \quad \text{for } z = hom, het, \tag{29}$$

where $\hat{N}_1^e$ denotes the change in entry into sector 1 and the elasticities $\mathcal{E}_n^z(\gamma_1),\ n = 1, 2$ are the marginal impact of the tariff and entry on the wage that depend on $\gamma_1 \in [0, 1)$ and the market structure $z = hom, het$ (as well as on other parameters and the endogenous import share). With only a single sector, $\alpha_1 = 1$, the tariff has no impact on entry in sector 1 so that $\hat{N}_1^e = 0$. When evaluating at free trade for simplicity, so that $t_1 = 1$, then we can compare the marginal impact of the tariff on wages depending on the amount of roundabout production.

We confirm (see Appendix D.1) that with either homogeneous or heterogeneous firms, an increase in the extent of roundabout production moderates the wage impact of the tariff, a result we state as:

$$\begin{aligned} &\text{For } \alpha_1 = t_1 = 1 \text{ and } \sigma_1 > 2: \ \mathcal{E}_1^z(0) > 0 \text{ and } \mathcal{E}_1^z(\gamma_1) \text{ is declining in } \gamma_1, \text{ with } \mathcal{E}_1^{hom}(\gamma_1) < 0 \\ &\qquad \text{for } \eta_{m1} > \tfrac{\sigma_1}{\sigma_1 - 2} \text{ and } \mathcal{E}_1^{het}(\gamma_1) < 0 \text{ for } \eta_{m1} > \tfrac{\sigma_1}{\sigma_1 - 2 + \lambda_{1m}}. \end{aligned} \tag{30}$$

As expected from our arguments above, the marginal impact of the tariff on the wage is *reduced* by the extent of roundabout production (which in this statement is a parametric increase in γ_1 while holding the import share constant). Surprisingly, we find that $\mathcal{E}_1^z(\gamma_1) < 0$ so the wage *falls* rather than rises with the tariff when the extent of roundabout and the import share – as reflected by η_{m1} – are sufficiently large. This occurs because of the large impact of the reduced import share on the price index P_1 and therefore the input costs in (25) and (26), so that a *fall* in the wage is needed to restore equilibrium. In that case, an import *subsidy* rather than a tariff would be needed to raise the home wage. We will explore in later results whether an import subsidy can be the optimal second-best policy.

We compare across the two market structures using the parameter restriction in (11) (and assuming the same import share under free trade), with no roundabout production for simplicity, to obtain:

$$\text{For } \alpha_1 = t_1 = 1, \sigma_1 > 2 \ \text{ and using (11)}: \quad \mathcal{E}_1^{het}(0) < \mathcal{E}_1^{hom}(0). \tag{31}$$

This result shows the impact of selection in reducing the terms of trade effect in the heterogeneous firm model, and by continuity it continues to hold for a range of positive values for γ_1.

5.2. Entry and welfare

Aside from its reduced impact on the wage, another reason for the tariff to be lower in a second-best setting is through changing the entry of firms. Starting from free trade we found in Section 2.1 that an increase in t_1 from free trade leads to the *exit* of firms from the traded sector 1 and entry into the nontraded sector 2. To solve for the impact of that exit and entry on welfare, we start with indirect utility corresponding to (2), which is (up to a constant): $U = (wL + B)/(P_1^{\alpha_1} P_2^{\alpha_2})$. We totally differentiate utility for a change in the tariff, using the expressions for the price indexes (see Appendix D.2), to obtain

$$\begin{aligned} \hat{U} \ = \ & -\frac{\alpha_1}{\theta_1(1-\gamma_1)}\hat{\lambda}_{d1} + \sum_{s=1,2} \alpha_s \left[1 + \frac{(1-\Gamma_s)}{\theta_s(1-\gamma_s)} \left(\frac{\theta_s}{\sigma_s - 1} - 1 \right) \right] \frac{B}{wL+B}(\hat{B} - \hat{w}) \\ & + \sum_{s=1,2} \alpha_s \left[\frac{(1-\Gamma_s)}{\theta_s(1-\gamma_s)} + \frac{\Gamma_s}{(\sigma_s - 1)(1-\gamma_s)} \right] \hat{N}_s^e, \end{aligned} \tag{32}$$

[16] This extra impact of a tariff due to selection arises from our assumption that the fixed costs of exporting use domestic labor rather than using foreign labor (whose wage is fixed as the numeraire). Likewise, when foreign firms pay their fixed costs of exporting using their own labor, then there is an extra impact of selection on the import share at home, as discussed in Appendix A.6. When we make the alternative assumption that the fixed costs of exporting use labor in the destination country, then these two extra impacts disappear.

where $\Gamma_1 \equiv \tilde{\gamma}_1 \Lambda_1$ denotes the fraction of the sector 1 finished good used as an input in (8), with $\Gamma_2 \equiv \tilde{\gamma}_2$, and $1 - \Gamma_s = \alpha_s (wL + B)/Y_s$ is the fraction used as a final good in each sector. Note that Γ_s is another way to measure the extent of roundabout production in a sector.

The first term in (32) is the change in the domestic share in sector 1 and is familiar from Arkolakis et al. (2012), where it is a sufficient statistic for the welfare change due to a change in iceberg trade costs in a one-sector model with no roundabout. Using a tariff introduces the second term in (32), reflecting the change in real tariff revenue B/w. Most important for our purposes is the third term in (32), which is related to entry. If there is no roundabout production so $\gamma_s = \Gamma_s = 0$, then the third term is simply the weighted sum of $\hat{N}^e_s/\theta_s$ across sectors using the weights $1/\theta_s(1-\gamma_s)$ that appear in the first term. When there is roundabout production, however, then a new mechanism comes into play. The effect of entry in the final term of (32) now depends on Γ_s, the fraction of finished output used as an intermediate input. The coefficient of that term is $1/(\sigma_1 - 1)(1-\gamma_s)$, which *exceeds* $1/\theta_s(1-\gamma_s)$ because $\theta_s > \sigma_1 - 1$. It follows that when the finished output arising from new entry is used more heavily downstream as an *intermediate input to other firms*, rather than just sold to consumers (in which case $\Gamma_s = 0$), then these forward linkages create a magnified effect of entry on welfare.[17]

These results can be contrasted to the case with homogeneous firms. Then using the parameter restriction (11), the weights appearing in the final bracketed term in (32) are both replaced by $1/(\sigma^{hom}_s - 1)(1-\gamma_s)$, so this final term would appear as

$$\sum_{s=1,2} \alpha_s \left[\frac{(1-\Gamma_s)}{(\sigma^{hom}_s - 1)(1-\gamma_s)} + \frac{\Gamma_s}{(\sigma^{hom}_s - 1)(1-\gamma_s)} \right] \hat{N}^e_s = \sum_{s=1,2} \frac{\alpha_s \hat{N}^e_s}{(\sigma^{hom}_s - 1)(1-\gamma_s)}. \tag{33}$$

We see that entry under homogeneous firms has the same impact whether the finished good is used as an intermediate input or a final good, so the share Γ_s no longer appears. Comparing (33) with the final bracketed term of (32), we also see that in both cases the welfare impact of entry depends on $1/(1-\gamma_s)$, so that entry into sectors with more roundabout production (higher γ_s) will have a greater welfare benefit — holding constant the other parameters. But with heterogeneous firms, the finished output arising from new entrants that is used as an *intermediate input* has a *magnified* impact in (32) when $\Gamma_s > 0$, because $\theta_s > \sigma_s - 1$. These results are a third and final illustration of how selection with heterogeneous firms influences the welfare impact of a tariff.[18]

To fully solve for the impact of entry and the tariff on welfare, we focus the remainder of the paper on the heterogeneous firm model: from our arguments above, we are therefore focusing on the case with the greatest potential to lower the second-best tariff. Using the change in the tariff and in wages to compute the change in the expenditure share $\hat{\lambda}_{d1}$ in (32), and also solving for the change in tariff revenue, we obtain the following reduced-form expression for the change in welfare due to selection and entry:[19]

$$\hat{U} = \alpha_1 \left[\mathcal{E}_\varphi \, \hat{\varphi}_{x1} + D(t_1) \hat{N}^e_1 \right], \tag{34}$$

where

$$D(t_1) \equiv \left[\frac{\tilde{\sigma}_1}{(\tilde{\sigma}_1 - 1)} - \frac{\tilde{\sigma}_2}{(\tilde{\sigma}_2 - 1)} \frac{\Lambda_1(1-\tilde{\gamma}_1)}{1 - \tilde{\gamma}_1 \Lambda_1} - \mathcal{E}_d \right]. \tag{35}$$

To interpret (34), the first term appearing on the right in brackets summarizes all the selection effects from the change in the tariff. The second term is the change in sector 1 entry $\hat{N}^e_1$ times $D(t_1)$, which denotes the marginal welfare impact of entry into the traded sector – holding the cutoff productivities constant – relative to the size of that sector (α_1). From (35), the marginal impact of entry equals the terms: $\frac{\tilde{\sigma}_1}{(\tilde{\sigma}_1 - 1)}$, which is the effective distortion in sector 1; minus the effective distortion in sector 2 multiplied by $\frac{\Lambda_1(1-\tilde{\gamma}_1)}{1-\tilde{\gamma}_1\Lambda_1}$ (which is ≤ 1 for $t_1 \ge 1$) that reflects tariff revenue; minus the term $\mathcal{E}_d > 0$ that we interpret as the deadweight loss of the tariff, which is an inefficient instrument to influence entry.[20]

We see from (35) that $D(t_1) > 0$ so that entry into the traded sector leads to a welfare gain – and exit leads to a welfare loss – when that effective distortion there is *sufficiently above* the effective markup in the nontraded sector. We want to allow the effective distortion in the traded sector to be greater or less than that in the nontraded sector, but we do not want the latter distortion to be too high. Accordingly, we will impose an upper-bound on the distortion in the nontraded sector,

$$\frac{\tilde{\sigma}_2}{(\tilde{\sigma}_2 - 1)} < \kappa_0 + \kappa_1 \frac{\tilde{\sigma}_1}{(\tilde{\sigma}_1 - 1)}, \tag{36}$$

where the parameters κ_0, κ_1 will be specified in Theorem 1 below. Our aim is to allow for a wide range of effective distortions in (36).

---

[17] We stress that a weighted sum of the log changes in entry across sectors (using their labor shares as weights) equals zero, as we show in Appendix D.2. So utility can rise only if the beneficial impact of entry in one sector exceeds the cost from reduced entry in the other.

[18] Because this difference between the results with homogeneous and heterogeneous firms arises even when we impose the parameter restriction (11), it shows that the two models are not isomorphic in the presence of roundabout production when entry is changing across sectors, as also found by Costinot and Rodríguez-Clare (2014): compare columns 5 and 6 of their Table 4.3 (p. 232).

[19] Note that the elasticity $\mathcal{E}_\varphi$ incorporates changes in φ_{x1} and all other cutoffs, while $D(t_1)$ incorporates the change in both N^e_1 and N^e_2. In addition, (34) incorporates the change in the wage and in tariff itself, which is inverted so that it is a function of $\hat{\varphi}_{x1}$ and $\hat{N}^e_1$: see Appendix D.4.

[20] All script variables $\mathcal{E}_n$, $n = \varphi, d, a, m$ depend on sector 1 parameters including γ_1 and λ_{d1} and therefore depend on the tariff. They are defined in Appendixes D.4 and D.5.

5.3. Optimal second-best tariff

We can now state a general formula for the optimal second-best tariff t_1^*, as compared to t^{het} (see Appendix E). Specifically, t_1^* is obtained as a fixed point of the equation

$$t_1^* = t^{het} F(t_1^*), \text{ with } F(t_1) \equiv \left[\frac{1-\gamma_1 R(t_1)}{1+\alpha_2 M(t_1)}\right], \tag{37}$$

where $M(t_1)$ captures the impact of the higher monopoly distortion in the traded versus the nontraded sectors, and is defined by

$$M(t_1) \equiv \mathcal{M} \times \left(\mathcal{E}_m - \frac{(t_1-1)}{t_1}\theta_1\right)\frac{D(t_1)}{A(t_1)} \quad \text{with} \quad \mathcal{M} > 0 \quad \text{a constant, } \mathcal{E}_m > 0, \tag{38}$$

and $A(t_1)$ is defined by

$$A(t_1) \equiv \alpha_1 - \tilde{\gamma}_1 + \alpha_2 \mathcal{E}_a \quad \text{with} \quad \mathcal{E}_a > 0, \tag{39}$$

while $R(t_1)$ reflects the impact of roundabout production and is defined by

$$R(t_1) = \mathcal{R} \times \left[\frac{\theta_1 - \rho_1(1-\lambda_{d1})}{\Lambda_1} - \theta_1\rho_1\right], \tag{40}$$

with

$$\mathcal{R} = \left\{\lambda_{d1}\frac{\theta_1}{(\sigma_1-1)}\left(\frac{\theta_1}{\sigma_1-1} - \frac{1}{\sigma_1}\right)\left[(\tilde{\sigma}_1 - 1)\left(1+\frac{\sigma_1}{\Lambda_1}\right)+1\right]\right\}^{-1} > 0. \tag{41}$$

To explain these terms, recall that the distortion term $D(t_1)$ measures the marginal welfare impact of firms moving from the nontraded to the traded sector, and notice that it enters $\alpha_2 M(t_1)$, which appears in the denominator of (37), reflecting the impact of the *relative* monopoly distortion on the optimal tariff. When $\alpha_1 = 1$ so there is only the traded sector, then this term vanishes because there is no impact of the *relative* distortion between traded and nontraded goods. But there is still a monopoly distortion in traded goods alone, where the markup distorts the use of the finished good as an input relative to labor. Notice that $\mathcal{R} > 0$ in (41) is declining in the effective elasticity $(\tilde{\sigma}_1 - 1) = (\sigma_1 - 1)(1-\gamma_1)$, so as that elasticity falls then the term $R(t_1)$ in the numerator of (37) rises, which tends to pull down the optimal tariff. This illustrates a complementary relationship between roundabout production and the monopoly distortion in the traded sector in reducing the optimal tariff. If there was not monopoly distortion, then we would have $\mathcal{R} = 0$ and the presence of roundabout production would not matter for the optimal tariff.[21]

When $\alpha_1 = 1$ and $\gamma_1 = 0$ in (37), then we are back in the one-sector, no-roundabout model and that formula immediately gives $t_1^* = t^{het}$. Outside of that special case, there will be a lower optimal tariff, $t_1^* < t^{het}$, whenever $\alpha_2 M(t_1^*) \geq 0$ and $\gamma_1 R(t_1^*) \geq 0$ with one of these inequalities holding strictly. For example, suppose that $\alpha_1 = 1$ so there is only a traded sector, but $\gamma_1 > 0$ so there is some roundabout production. Then we prove below that $R(t_1^*) > 0$ at the fixed point of (37), so that roundabout production lowers the optimal tariff. Thus, we will find that the optimal tariff is lowered by the monopoly distortion in sector 1, even in the absence of the nontraded sector.

Next, suppose we add the nontraded sector so that $\alpha_2 > 0$, in which case the denominator of $F(t_1^*)$ which is $[1+\alpha_2 M(t_1^*)]$ comes into play. If the relative distortion in the traded sector is positive, $D(t_1^*) > 0$, then provided that the other terms in (38) are positive we will have $M(t_1^*) > 0$, so the denominator *further reduces* the optimal tariff. One of those other terms is $A(t_1)$. Recall that we defined $D(t_1)$ in (35) as the marginal impact of entry into sector 1 relative to the size of that sector (α_1), and we loosely interpret $A(t_1)$ as the effective size of sector 1. As a regularity condition we need to impose $A(t_1) > 0$, which is guaranteed by the sufficient conditions specified in the following main theorem (proved in Appendix E).

Theorem 1.

(a) ***Pure roundabout:*** *If $\alpha_1 = 1$ and $\gamma_1 > 0$, then $R(t_1^*) > 0$ and the optimal tariff is $t_1^* < t^{het}$.*

(b) ***No roundabout:*** *If $\gamma_1 = \gamma_2 = 0$ then (i) $D(t_1^*) > 0$ and the optimal tariff is $t_1^* < t^{het}$ when*

$$\frac{\sigma_2}{(\sigma_2-1)} < \frac{\sigma_1}{(\sigma_1-1)} - \frac{1}{\theta_1}. \tag{42}$$

(ii) If $\frac{\sigma_2}{(\sigma_2-1)} \geq t^{het}\frac{\sigma_1}{(\sigma_1-1)}$, then $D(t_1^*) < 0$ and the optimal tariff is $t_1^* > t^{het}$.

(c) ***Two sectors with roundabout:*** *Assume that $\alpha_2 > 0$ and the following two conditions hold:*

$$\gamma_1 \leq \frac{\frac{\sigma_1}{\rho_1}(\theta_1-\rho_1)(1-\rho_1)}{1+\frac{\sigma_1}{\rho_1}(\theta_1-\rho_1)(1-\rho_1)}, \tag{43}$$

$$\alpha_2 \leq \max\left\{1-\tilde{\gamma}_1, \frac{\frac{\theta_1(1-\rho_1)}{\rho_1} + (1-\gamma_1)\theta_1}{\frac{\theta_1(1-\rho_1)}{\rho_1} + \rho_1\left(1+\frac{\gamma_1}{\sigma_1(1-\gamma_1)}\right)}\right\}. \tag{44}$$

[21] Holding fixed the ratios $\theta_1/(\sigma_1-1)$ in (41), we see that as $\sigma_1 \to +\infty$ then $\mathcal{R} \to 0$, so that roundabout production does not have any impact on the optimal tariff when the differentiated inputs become very strong substitutes and the monopoly distortion in the traded sector vanishes.

Then $A(t_1) > 0$ for $t_1 > t_1'$, where $t_1' < 1$ is an import subsidy. Furthermore, if there is enough roundabout production so that

$$\gamma_1 \geq \frac{\rho_1}{\left[\theta_1(1-\rho_1)+\rho_1^2\right]\left(\theta_1-\rho_1\right)}, \tag{45}$$

and the upper bound in (36) holds as

$$\frac{\tilde{\sigma}_2}{(\tilde{\sigma}_2-1)} < \frac{(t^{het}-\tilde{\gamma}_1)}{(1-\tilde{\gamma}_1)}\frac{\tilde{\sigma}_1}{(\tilde{\sigma}_1-1)}+\kappa_0, \tag{46}$$

where κ_0 is independent of the share λ_{d1},[22] then the optimal tariff is $t_1^* < t^{het}$ with $R(t_1^*) > 0$.

The proof of Theorem 1 does not use the fixed-point formula (37) directly, but rather, uses a slight transformation of it. Taking the difference between the numerator of $F(t_1)$ times t^{het} and the denominator times t_1, we obtain

$$H(t_1) \equiv t^{het}\left[1-\gamma_1 R(t_1)\right] - t_1\left[1+\alpha_2 M(t_1)\right]. \tag{47}$$

The function $H(t_1)$ is a continuous function of the tariff provided that $A(t_1) > 0$ in the interval of tariffs we are interested in, in which case $M(t_1)$ will not have any discontinuities. Our approach for each part of Theorem 1 is to find high and low tariffs at which the sign of $H(t_1)$ switches, and then we apply the intermediate value theorem to obtain a point where $H(t_1^*) = 0$, which by construction is a fixed-point of (37) so that t_1^* is the optimal tariff.

Part (a) of Theorem 1 shows that roundabout production in a one-sector model always lowers the optimal tariff. This result is the simplest demonstration that the tariff t_1^* on intermediate inputs is less than the tariff t^{het} that applies in a model with differentiated final goods. To prove this result, we note that with $\alpha_1 = 1$ then $M(t_1)$ disappears from (47) because there is no monopoly distortion *between* sectors, and we only need to work with the term $R(t_1)$ that incorporates roundabout production and the monopoly distortion *within* sector 1. We first establish (see Appendix E.1) that at $t^{het} > 1$ then $R(t^{het}) > 0$, so that we obtain $H(t^{het}) = -t^{het}\gamma_1 R(t^{het}) < 0$ for $\gamma_1 > 0$. Next, we establish that there is a low enough tariff $t^{R0} < 1$ at which $R(t^{R0}) = 0$, which means that the effect of roundabout production is neutralized.[23] Because $\alpha_1 = 1$ by assumption, it follows from (47) that $H(t^{R0}) = t^{het} - t^{R0} > 0$. It follows from the intermediate value theorem that there exists a tariff t_1^* with $t^{R0} < t_1^* < t^{het}$ at which $H(t_1^*) = 0$. By construction, this optimal tariff is a fixed point of (37) with $t_1^* < t^{het}$.

Part (b) deals with the opposite case where there is no roundabout production. In that case, the term $R(t_1)$ disappears from (47) so we only need to work with the term $M(t_1)$ reflecting the monopoly distortion between sectors. It turns out that $A(t_1) > 0$ is guaranteed in this case, so the sign of M is determined by the sign of D. Condition (42) used in part (b)(i) ensures that the relative distortion in the traded sector sufficiently exceeds that in the nontraded sector so that $D(t_1) > 0$ for $t_1 \in [1, t^{het}]$. It follows that $H(t^{het}) = -t^{het}\alpha_2 M(t^{het}) < 0$. We further show that there exists a sufficiently low tariff $t^{D0} < 1$ at which $D(t^{D0}) = 0$, so the monopoly distortion between sectors is neutralized.[24] In that case, $H(t^{D0}) = t^{het} - t^{D0} > 0$. It follows once again from the intermediate value theorem that there exists an optimal tariff t_1^*, with $t_1^{D0} < t_1^* < t^{het}$.

On the other hand, if the nontraded sector is sufficiently more distorted than the traded sector, with $\frac{\sigma_2}{(\sigma_2-1)} \geq t^{het}\frac{\sigma_1}{(\sigma_1-1)}$, then we have the reverse outcome with $D(t_1^*) < 0$ and $t_1^* > t^{het}$. In this case we find that $D(t^{het}) < 0$ and $M(t^{het}) < 0$, so $H(t^{het}) = -t^{het}\alpha_2 M(t^{het}) > 0$. The negative sign for the monopoly distortion indicates that resources should be shifted away from sector 1. We further show that there is a high enough tariff $t_1'' > t^{het}$ at which $M(t_1'') = 0$, so the monopoly distortion is neutralized.[25] Then we find from (47) with $\gamma_1 = 0$ that $H(t_1'') = t^{het} - t_1'' < 0$. It follows from the intermediate value theorem that there exists an optimal tariff t_1^*, now with $t^{het} < t_1^* < t_1''$. So the general conclusion is that without roundabout production, the tariff on final goods can be greater or less than that found in a one-sector model, depending on the relative monopoly distortion across sectors.

In part (c) we allow for two sectors and roundabout production, and so we need to ensure $A(t_1) > 0$. We establish that $A(t_1) > 0$ for $t_1 > t_1'$, where $t_1' < 1$ is an import subsidy specified in the proof, under the sufficient conditions (43) and (44): the former is an upper-bound on γ_1 and the latter is an upper-bound on α_2 (but also depending on γ_1). There are two further conditions in part (c), and these are used to establish the sign of $H(t_1)$ at two tariff values chosen like in the proof of part (a): namely, t^{R0} and t^{het}, which give the values

$$H(t^{R0}) = (t^{het} - t^{R0}) - t^{R0}\alpha_2 M(t^{R0}), \tag{48}$$

$$H(t^{het}) = -t^{het}\left[\gamma_1 R(t^{het}) + \alpha_2 M(t^{het})\right]. \tag{49}$$

We can establish that $M(t^{R0}) < 0$ provided that conditions (43) and (44) hold so that $A(t^{R0}) > 0$ (see Appendix E.5), and it follows that $H(t^{R0}) > 0$. Then the remaining conditions (45) and (46) in part (c) are used to show that $H(t^{het}) < 0$ in (49), because $\gamma_1 R(t^{het}) > -\alpha_2 M(t^{het})$. We know that $R(t^{het}) > 0$ and we are allowing the relative monopoly distortion to be of either sign, so in the case where $M(t^{het}) < 0$ then we see that (49) requires a sufficient amount of roundabout production, i.e., $\gamma_1 > -\alpha_2 M(t^{het})/R(t^{het}) > 0$.

---

[22] The formula for κ_0 is specified in the Appendix, Lemma 11, and is of either sign.

[23] This result is obtained in Appendix E.1 because for $t_1 < 1$ then $\Lambda_1 > 1$, and so we can prove that the term in brackets in (40) equals zero at a point $t^{R0} < 1$.

[24] This result is obtained in Appendix E.3 because for $t_1 < 1$ then $\frac{\Lambda_1(1-\tilde{\gamma}_1)}{1-\tilde{\gamma}_1\Lambda_1} > 1$, and so under condition (42) we can prove that the terms in (35) sum to zero at a point $t^{D0} < 1$.

[25] This result is obtained in Appendix E.4 because we prove that there exists a high tariff $t_1'' > t^{het}$ at which $\mathcal{E}_m - [(t_1''-1)/t_1'']\theta_1 = 0$ in (38), and therefore $M(t_1'') = 0$.

In that case we can apply the intermediate value theorem one last time to obtain the optimal tariff t_1^* with $t^{R0} < t_1^* < t^{het}$. Condition (46) is an upper-bound on the effective distortion in sector 2 relative to sector 1, and it generalizes condition (42) to now allow for roundabout production. We argue in the next section that the constraints (43)–(45) are satisfied for all countries in our sample, while the upper-bound in (46) is satisfied for most.

We conclude this section by noting that the optimal tariff can be negative, as we earlier suggested following (26) when η_{m1} is sufficiently large. Consider the limiting case as $\lambda_{d1} \to 0$, so that $\eta_{m1} \to +\infty$. For simplicity, let us focus on a one-sector economy, so that $\alpha_1 = 1$. In that case, we can take the limiting value of the fixed-point formula in (37) (see Appendix E.6) to show that

$$\lim_{\lambda_{d1} \to 0} t_1^* = \frac{\theta_1 \rho_1}{(\theta_1 - \rho_1)} < 1, \quad \text{when} \quad \alpha_1 = 1 \text{ and } \gamma_1 > 0. \tag{50}$$

Remarkably, we find that the optimal tariff in this limiting economy exactly equals the subsidy in (18) found by Demidova and Rodríguez-Clare (2009): that subsidy is needed to correct the externality arising in a model with imported differentiated goods, whereby each new foreign variety brings surplus that domestic buyers do not take account of in their spending. Because the share of domestic inputs in the economy is vanishingly small, it appears that the additional tariff of $t^{hom} = 1/\rho_1$ identified by Gros (1987) and used by Demidova and Rodríguez-Clare in conjunction with the subsidy to obtain t^{het} (see Eq. (19)) is not needed anymore, so we are left with just the subsidy as the optimal policy in this limiting case. By continuity, any economy with a sufficiently small domestic share will also have a negative optimal *ad valorem* tariff.

This is not the only example of a negative optimal tariff, however. In Caliendo et al. (2020), we examine the conditions to ensure that the optimal tariff is negative in a model with two symmetric countries, where only one country is applying the tariff. We find that a negative optimal tariff applies in two cases: *Highly Linked Economies* that have high roundabout production (high γ_1) and are very open (low λ_{d1}); and *Remote Economies*, with a small traded sector and with $\lambda_{d1} \to 1$, so that the economy is nearly closed to trade due to high iceberg costs, as may occur for very distant countries. The Highly Linked Economies are very similar to the negative optimal tariff found in (50) (except that (50) holds for all $\gamma_1 > 0$, so it does not require a high amount of roundabout production). The Remote Economies are different, however, and apply at the other extreme of the domestic share when $\lambda_{d1} \to 1$.

6. Second-best tariffs around the world

We now take the model to the data and solve for second-best optimal tariffs. We use data from EORA 26 (Lenzen et al., 2012, 2013) for the year 2010 which contains information for the world economy. Before we compute the optimal tariffs we need to aggregate the data into a two-country two-sector world. We define the Tradable goods sectors as sectors 1 through 12 from the EORA classification and the Nontradable goods sectors as sectors 13 through 25 from EORA classification (sector 26 is re-exports). For each country in EORA we aggregate all variables into these two sectors, and then for each country we aggregate all the others into the rest of the world (RoW). EORA contains information for 189 countries, many of which are small economies. However, as a preliminary step, and in order to determine the reliability of the data, for each country in the sample we compute the total GDP documented in EORA relative to the value documented by the World Bank indicators. Some countries had GDP values in EORA that represented more than 2 times or less than half the GDP value documented by the World Bank. We excluded these countries from the sample. We also excluded countries in the sample for which we do not have 2010 tariff data, which are needed to calibrate the model. As a result, we end up with a list of 164 countries in our sample (and for each country the RoW).[26]

The requirements to take the model to the data for each country are the following: the values of the finished goods produced in each sector Y_s, the domestic expenditure shares λ_{d1}, the labor share in each sector in our model, $1 - \gamma_s$, which more generally should be measured as the share of value added (i.e., payments to labor and capital) in the variable costs of production. We also need information on the elasticities of substitution in each sector σ_s, and the Pareto share parameter θ_s. We first describe how we obtain these variables and then describe how we obtain the elasticities.

When taking the model to the data we need to deal with three issues. First, in order to measure the share of value added in production one cannot take the shares of industry revenue that go to value added directly from the data since the share of value added also includes profits (or "operating surplus"). Second, total intermediate goods includes purchases from your own sector and other sectors, but our model only has purchases of intermediates from your own sector. Third, our model assumes a sector with no trade, but the service industries in EORA have some trade.

For the first issue, we compute the share of intermediate goods in the cost of production for the Tradable and the Nontradable sectors as the intermediate goods purchased from the same sector divided by the sum of the compensation of employees, the consumption of fixed capital and the total intermediate goods purchased. So the latter three terms are used to measure the costs of production (and in particular they exclude "operating surplus", or profits). For the second issue, we include only the intermediate goods purchased from the *same sector* in roundabout production, but we cannot simply ignore the off-diagonal elements of the input–output matrix. By including *all* intermediate purchases in the cost of production, we are essentially attributing the expenditure on goods from other sectors into value added. Note that this is a conservative approach to measuring roundabout production, since it increases the share of value added in production and therefore reduces the share of intermediates. For the final issue, we excluded international trade in services from our calculations, so services is our Nontradable sector.

We measure the value of final goods produced in sector 1, Y_1, as the sum of the total domestic purchases plus total imports. In the case of sector 2 we have that Y_2 is equal to the total domestic purchases. We calculate the domestic expenditure share λ_{d1} as the

[26] See Appendix G, Table 3 for the full set of countries.

Table 1
Elasticities.

Sector(s)	$\frac{\sigma_s \theta_s}{\sigma_s - 1} - 1$	θ_s	σ_s	Global share
Agriculture and Fishing	9.11	8.4	5.8	0.16
Mining and Quarrying	13.53	12.8	8.3	0.09
Manufacturing Sectors	5.55	4.8	3.7	0.75
Total Tradable Sector (above 3 sectors)	–	6.1	4.5	1
Total Nontradable Sector (all services)	–	3.1	2.8	–

Table 2
Distribution of parameters by countries and sectors.

Statistic	Tradable	Nontradable
α_s $(p10)$	0.21	0.60
α_s $(median)$	0.25	0.75
α_s $(p90)$	0.40	0.79
$(1-\gamma_s)$ $(p10)$	0.34	0.75
$(1-\gamma_s)$ $(median)$	0.45	0.84
$(1-\gamma_s)$ $(p90)$	0.57	0.89
$\tilde{\sigma}_s = 1+(1-\gamma_s)(\sigma_s-1)$ $(p10)$	2.20	2.35
$\tilde{\sigma}_s = 1+(1-\gamma_s)(\sigma_s-1)$ $(median)$	2.57	2.51
$\tilde{\sigma}_s = 1+(1-\gamma_s)(\sigma_s-1)$ $(p90)$	3.01	2.60

share of domestic purchases over Y_1. It follows that $\lambda_{m1} = 1-\lambda_{d1}$. Then, given the level of tariffs, we can measure $\Lambda_1 \equiv \lambda_{d1} + (\lambda_{m1}/t)$. Given estimates of σ_s and the definitions in (3), we solve for total value added as

$$wL = \left((1-\gamma_1) + \frac{1}{\sigma_1 - 1}\right)\rho_1 \Lambda_1 Y_1 + \left((1-\gamma_2) + \frac{1}{\sigma_2 - 1}\right)\rho_2 Y_2, \tag{51}$$

and finally, the share of final goods in consumption is obtained using

$$\alpha_1 = \frac{\left(1-\tilde{\gamma}_1 \Lambda_1\right) Y_1}{wL + \left(1-\Lambda_1\right) Y_1}. \tag{52}$$

In order to obtain estimates for the elasticity of substitution and the Pareto parameter we use the estimates from Caliendo and Parro (2015). They show that by triple differencing the gravity equation one can identify the elasticities using tariff policy variation. In the context of our model the elasticity that is estimated is given by $1-\sigma_s\theta_s/(\sigma_s-1)$, and those values are reported in the first column of Table 1. In order to separately identify θ_s and σ_s we rely on estimates from the literature to obtain $\theta_s/(\sigma_s-1)$. The two most cited studies to deal with this issue are Chaney (2008) and Eaton et al. (2011). Chaney finds that $\theta_s/(\sigma_s-1) = 2$ from U.S. sales data, while Eaton et al. (p. 1472) find an initial estimate of $\theta_s/\left(\sigma_s-1\right) = 1.75$ using French data on exporting firms. We rely on the latter estimate and apply it to the first column of Table 1 to obtain values for σ_s of 5.8, 8.3, and 3.7, respectively, for Agriculture and Fishing, Mining and Quarrying, and Manufacturing.[27] Gervais and Jensen (2019) find that services have an elasticity of substitution that is three-quarters the size of the elasticity in manufacturing (though they obtain rather high values for both elasticities using accounting data).[28] We follow them by setting $\sigma_2 = 0.75 \times 3.7 = 2.8$ for services, our Nontradable sector. Finally, we take a weighted average of the elasticities across the Tradable sector using the global shares of output shown in the final column of Table 1, obtaining $\sigma_1 = 4.5$. We therefore have σ_1 for Tradable goods considerably higher than σ_2 for Nontradable services, generating higher markups in the latter sector.

Table 2 reports the shares of industry final consumption, α_s as well as the share of industries revenue that go to value-added, $(1-\gamma_s)$. As expected, the share of expenditure on final goods in the Tradable sector is lower than in the Nontradables sector in our sample. The median share is 25% for Tradables (α_1) and 75% for Nontradables (α_2). We can see that the value added share for Tradables varies across countries from 34% at the 10th percentile to 57% at the 90th. Also reported is the effective elasticity $\tilde{\sigma}_s \equiv 1+(1-\gamma_s)(\sigma_s-1)$ in each sector. We find that the median effective elasticity in Tradables (2.57) is quite close to the median effective elasticity in Nontradables (2.51), so the effective monopoly distortion in the two sectors has similar median but still differs across countries.

To compute the optimal tariffs we then solve numerically the system of equations of the small open economy model using the "hat-algebra" method for large changes. We then verified that the solution coincides with the exact solution to the optimal tariff using our formula $H(t_1^*) = 0$ in (47).[29] Fig. 2 reports the distribution of optimal tariffs for the 164 countries in our sample. The vertical dashed line represents the tariff value of $t^{het} = \frac{\theta_1\rho_1}{(\theta_1-\rho_1)} = 1.146$ or an *ad valorem* value of 14.6%. As we can see, almost all countries in our sample have an optimal tariff that is below t^{het}, and the median *ad valorem* optimal tariff is 11%, with much variation across countries.

[27] These elasticities are slightly revised from our working papers, CFRT (2020, 2021).

[28] This estimate of 0.75 comes from their working paper, Gervais and Jensen (2013).

[29] See Appendix F, where Figure 5 presents a scatter plot between the numerical solution from the hat-algebra and the exact solution, which are closely aligned. Table 3 in the Appendix includes the optimal tariff for each country in our sample along with the parameter values.

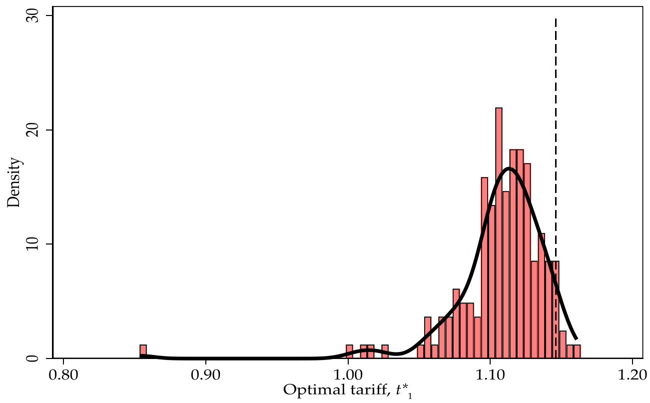

Fig. 2. Distribution of optimal second-best tariffs (exact solution) compared to t^{het} shown by the dashed line.

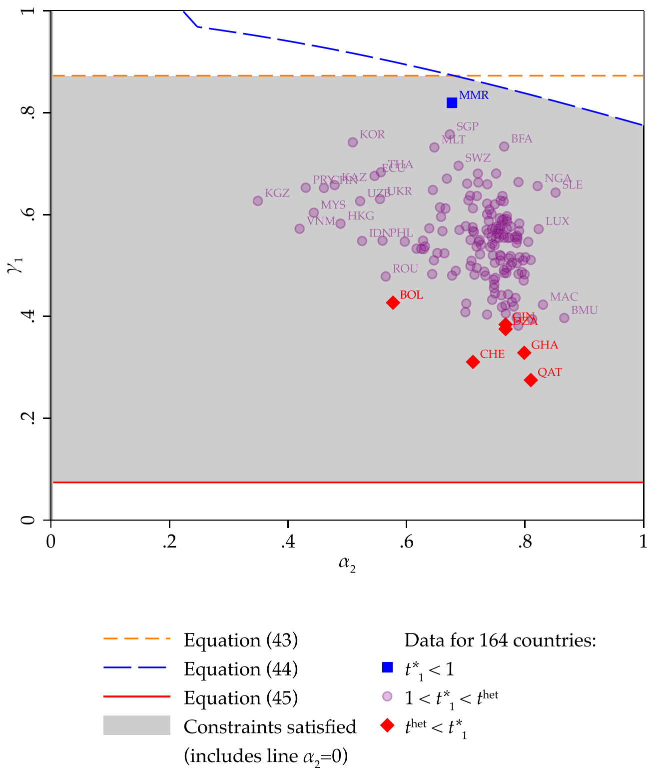

Fig. 3. Parameter restrictions.

The parameters in Table 2 can be used to illustrate how our optimal tariffs from the quantitative model accord with the predictions of Theorem 1. Each scatterplot dot in Fig. 3 corresponds to the values of α_2 and γ_1 for the 164 countries in our sample, and we graph the constraints (43)–(45) from Theorem 1. We see that these constraints are satisfied for all countries in our sample.[30]

The final constraint that should be checked in Theorem 1 concerns the upper-bound on the effective distortion in Nontradables as compared with Tradables, given by (46). This constraint depends on γ_2, so it cannot be graphed here, but rather needs to be

[30] There is one country that is omitted from our sample that lies on the edge of a constraint, and that is Kuwait. However, we found that γ_1 for Kuwait is very sensitive to how we measure value-added: i.e., whether is consists of payments to labor and capital (as followed in this paper), or alternatively, whether it consists of all categories of value-added included in EORA (as followed in our working papers CFRT, 2020, 2021), which in addition to payments to labor and capital also includes operating surplus (i.e., profits), taxes paid, and a miscellaneous category of "mixed income". We did not observe this sensitivity in γ_1 depending on how value-added is measured for other countries, and for this reason, we have excluded Kuwait from our sample.

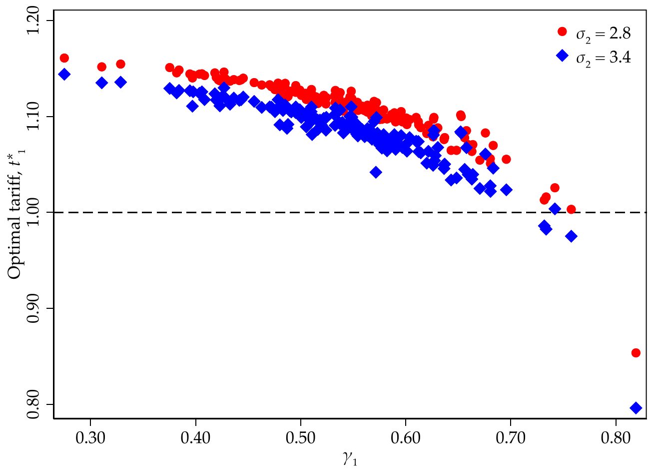

Fig. 4. Optimal tariff t_1^* versus roundabout parameter γ_1.

checked on a country-by-country basis. There are six countries that are highlighted in the lower-portion of Fig. 2 with relatively low values of roundabout production γ_1: these countries all have $t_1^* > t^{het}$ and they violate the constraint in (46).[31] In other words, these six countries have high enough values for the effective distortion in Nontradables that the (modest) amount of roundabout production in the Tradable sector is not enough to lead to $t_1^* < t^{het}$, contrary to what we find for other countries.

There is also one country highlighted at the top of Fig. 3 and that is Myanmar (MMR), which has $t_1^* = 0.85$ so the optimal *ad valorem* tariff is −15%. Myanmar (formerly Burma) is an extremely closed country, and its domestic share evaluated at the optimal tariff is t_1^* is $\lambda_1^* = 0.998$. Just below Myanmar are two other labeled countries that have optimal *ad valorem* tariffs very close to zero, i.e., $t_1^* \in (1, 1.02)$, and are very open: Singapore (SGP, $\lambda_1^* = 0.27$) and Malta (MLT, $\lambda_1^* = 0.48$). Burkina Faso (BFA) is also labeled at the top of Fig. 3 with $t_1^* \in (1, 1.02)$, and it is relatively closed with a domestic share $\lambda_1^* = 0.76$, above the median of 0.70. We will show in a sensitivity analysis below that in an alternative calibration where we modestly increase the value for σ_2 to a still plausible value, which acts to reduce the distortion in the Nontradable sector, then Singapore, Malta, Burkina Faso can also then have negative optimal tariffs.

To gain further insight, we performed a variance decomposition in the spirit of Eaton et al. (2004) to determine the contribution to the variance of the optimal tariff coming from roundabout production in the numerator of (37) versus the relative distortions across sectors in the denominator. Specifically, we write the numerator as $\ln[t^{het}\left(1 - \gamma_1 R(t_1^*)\right)]$ and the denominator as $\ln[1+\alpha_2 M(t_1^*)]$. Using each of these as dependent variables, we run a regression with $\ln t_1^*$ on the right. By construction, the two regression coefficients sum to unity, and they indicate the fraction of the variance in $\ln t_1^*$ explained by the numerator and the denominator of the fixed-point formula. We find that roundabout production explains 47% while the relative distortions across sectors explains 53% of the variation. Thus, in our calibrated model, roundabout production and the relative monopoly distortion are about equally important in explaining the variation in the optimal tariffs.

Recall that in our calibration of elasticities, we have relied on Gervais and Jensen (2013, 2019) who found that σ_2 for services is three-quarters that of σ_1 in manufacturing. That gave us the value $\sigma_2 = 2.8 = 3.7 \times 0.75$ used in our benchmark analysis. Because we also aggregate the Tradable sector over Manufacturing, Mining and Agriculture (see Table 1), we obtain a higher value for $\sigma_1 = 4.5$ in Tradables overall than in Manufacturing, so the elasticity used in Nontradables is considerably lower than that used in Tradables. As an alternative sensitivity check, we make a different assumption: we apply the factor of 0.75 from Gervais and Jensen to the elasticity used in the Tradable sector overall, obtaining the higher value of $\sigma_2 = 3.4 = 4.5 \times 0.75$ for the Nontradable sector.

In Fig. 4 we graph the optimal tariff against γ_1 for our benchmark calibration (with $\sigma_2 = 0.28$) and for this alternative sensitivity check (with $\sigma_2 = 0.34$). In both cases, we see that there is a remarkably strong nonlinear relationship between t_1^* and γ_1. Raising σ_2 lowers all the optimal tariffs. With $\sigma_2 = 0.34$ we find that Myanmar is joined by Burkina Faso, Malta and Singapore in having negative optimal tariffs, with South Korea (KOR) now having an optimal *ad valorem* tariff very close to zero. This set of countries illustrate the theoretical result mentioned at the end of the previous section: negative optimal tariffs are likely to be found for both *Highly Linked* and *Remote* economies, but in all cases we find empirically that having a high value for γ_1 – indicating high roundabout production – is an essential feature.[32] Furthermore, by raising σ_2 we now find that there are *no* countries having a high optimal tariff, with $t_1^* > t^{het}$.

[31] In addition, there are seven other countries – generally appearing in the lower portion of Fig. 3 – that violate (46), which is a sufficient but not necessary conditions to have $t_1^* < t^{het}$. The median value of κ_0 in our sample is −0.184, which is not too different from the value $-1/\theta_1 = -0.164$ appearing in constraint (42) in Theorem 1. But the presence of $\kappa_1 = (t^{het} - \tilde{\gamma}_1)/(1 - \tilde{\gamma}_1)$ in (46), with a median value of 1.26, makes this a notably weaker constraint due to the presence of roundabout production than (42).

[32] As mentioned in note 4, in CFRT (2021), we analyze a 186-country, 15-sector quantitative model for 2010 with a general input–output structure, and we find a negative optimal tariff for five countries, including Myanmar. In Caliendo et al. (2020), we analyze the same quantitative model for 1990, and we find

7. Conclusions

We began by asking whether modern trade theory has anything new to say about arguments for protecting the traded sector. It does, but the answer is nuanced. Gros (1987) showed that the Krugman model of monopolistic competition calls for a positive optimal tariff even for a small country. While we have explained that this tariff equalizes the monopoly markup on the price of domestic goods with the tariff distortion (i.e., one plus the *ad valorem* tariff) on the price of imports, other interpretations are possible. In particular, because of product differentiation in the Krugman model, the foreign demand curve for a home export variety is not infinitely elastic for a small country, but slopes downward. An import tariff – which is equivalent to an export tax by Lerner symmetry – reduces exports and therefore raises the export price, which is a terms of trade gain for the SOE applying the tariff. Even without imperfect competition, the presence of product variety on its own leads to a positive optimal tariff for a small country.[33]

The market structure in the SOE influences the optimal tariff, however. Demidova and Rodríguez-Clare (2009) found that the optimal tariff in a SOE with one sector and heterogeneous firms is lower than that obtained with homogeneous firms, so as to correct an externality in attracting foreign varieties. We have introduced a nontraded sector into the model, with roundabout production in both sectors. We find that there are strong reasons for the optimal tariff to be lower still, though this result is not guaranteed. With roundabout production, the idea of introducing a tariff distortion equal to the domestic markup breaks down: this policy would increase the price of the finished good that is bundled from the imported and domestic varieties, so that firms use too little of this finished good as compared to labor. To offset this distortion in the absence of any other policies, a *lower* value of the tariff is generally needed. The only exception to this rule occurs when the nontraded sector itself has a higher monopoly markup than the traded sector, which argues for a higher tariff to shift resources towards the nontraded good. For the vast majority of countries in our sample, the incentive to lower the tariff (to offset the distortion in the price of the finished good) is greater than the incentive to raise the tariff (when the nontraded sector is more distorted), and we find that the optimal tariff is lowered, and can be negative.

Our results stand in contrast to another literature that to some extent argues in favor of import protection. Specifically, this is the *firm-delocation* literature that combines a monopolistically competitive traded sector with a competitive *traded* outside good (see, e.g., Venables, 1987; Melitz and Ottaviano, 2008, Section 4; Bagwell and Lee, 2020). The traded numeraire good pins down relative wages between countries, so the country applying tariffs is "small" in the sense that its wages do not respond to its tariff. In this literature, encouraging entry into traded goods requires positive import tariffs. Essentially, the ability to attract firms into the home country takes the place of a conventional terms-of-trade motive for tariffs, so that the optimal tariff is positive even though wages are fixed. Of course, with multiple countries pursuing this motive for protection, there is ample scope for trade agreements to reduce the deadweight losses due to the tariffs (Ossa, 2011; Bagwell and Staiger, 2015).

The major differences between this class of models and our own are: (i) roundabout production, so that tariffs are applied on imported intermediate inputs rather than final goods; and (ii) the *nontraded* service sector, which does not fix relative wages between countries. Lerner symmetry holds in the traded sector, so that import tariffs are equivalent to export taxes and inhibit entry into that sector. That logic does not apply when the numeraire good is traded, which gives firm-delocation models a very different flavor: they act like partial equilibrium models because wages are fixed, and perhaps are most appropriate to narrowly targeted tariffs, whereas our results depend on Lerner symmetry, which is a general equilibrium property and depends on having broad tariffs applied to the traded sector. Determining the most appropriate range of applications for each class of models, and therefore the policy implications, is one important area for further research.

A second area for research is to investigate whether the optimal second-best tariff is low in other models beyond those we have investigated here. As we noted in Section 5.1, in the presence of roundabout production the positive impact of an import tariff on the home wage can be reversed: evaluated at free trade, a rise in the tariff can lead to a fall in the home wage, and this is more likely under heterogeneous firms than with homogeneous firms. That negative terms of trade impact is crucial to obtaining an optimal tariff that is negative, and the question is whether this result extends outside the monopolistic competition framework. Consider, for example, the perfect competition Armington and Eaton–Kortum models. In the absence of intermediate inputs, Caliendo and Feenstra (2022) have shown that there is a formula for the optimal tariff that depends critically on the wage impact of the tariff, and this formula holds under monopolistic competition and in these perfect competition models. What has not been investigated is whether the terms of trade impact itself can become negative in these competitive models due to input–output linkages.

Extending this question further, consider the perfect competition model with external economies of scale as analyzed by Bartelme et al. (2019). They have shown that the optimal policy in a small economy is to have production subsidies to internalize the external economies of scale and export taxes to internalize the terms-of-trade externalities. Furthermore, they show that this policy combination continues to hold with intermediate goods and an input–output structure. When production subsidies are not feasible, so that we are in a second-best setting, other questions for research are whether the terms of trade impact of a tariff/tax is reduced due to input–output linkages, and therefore whether the second best tariff/tax is lower than in the first-best.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

---

a negative optimal tariff in ten countries: China, Hong Kong, India, Israel, Vietnam, and five more remote countries. Having a negative optimal tariff suggests that the welfare gains to these countries from unilateral tariff reductions from 1990 were of the first-order.

[33] This point is made by Caliendo and Parro (2022) in the context of a small country in the Eaton–Kortum model.

Data availability and online appendix

The dataset and Online Appendix can be found at https://rcfeenstra.github.io/CFRT/.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jinteco.2023.103824.

References

Antràs, Pol, Chor, Davin, 2021. Global Value Chains. NBER Working Paper 28549.

Antràs, Pol, Fort, Teresa C., Gutiérrez, Agustín, Tintelnot, Felix., 2022. Trade Policy and Global Sourcing: A Rationale for Tariff Escalation. Working paper, https://scholar.harvard.edu/files/antras/files/affgt_draft_01_07_2022.pdf.

Arkolakis, Costas, Costinot, Arnaud, Rodríguez-Clare, Andrés, 2012. New trade models, same old gains? Am. Econ. Rev. 102 (1), 94–130.

Bagwell, Kyle, Lee, Seung Hoon, 2020. Trade policy under monopolistic competition with firm selection. J. Int. Econ. 127, 103379.

Bagwell, Kyle, Staiger, Robert W., 2015. Delocation and trade agreements in imperfectly competitive markets. Res. Econ. 69 (2), 139–156.

Bartelme, Dominick G., Costinot, Arnaud, Donaldson, Dave, Rodríguez-Clare, Andrés, 2019. The Textbook Case for Industrial Policy: Theory Meets Data. NBER Working Paper 26193.

Beshkar, Mostafa, Lashkaripour, Ahmad, 2020. The Cost of Dissolving the WTO: The Role of Global Value Chains. CAEPR Working Paper 2020-005, Indiana University.

Blanchard, Emily J., Bown, Chad P., Johnson, Robert C., 2016. Global Supply Chains and Trade Policy. NBER Working Paper 21883.

Caliendo, Lorenzo, Feenstra, Robert C., 2022. Foundation of the Small Open Economy Model with Product Differentiation. NBER Working Paper 30223.

Caliendo, Lorenzo, Feenstra, Robert C., Romalis, John, Taylor, Alan M., 2020. Tariff Reductions, Entry and Welfare: Theory and Evidence for the Last Two Decades. NBER Working Paper 21768 (revised version).

Caliendo, Lorenzo, Feenstra, Robert C., Romalis, John, Taylor, Alan M., 2021. A Second-Best Argument for Low Optimal Tariffs. NBER Working Paper 28480.

Caliendo, Lorenzo, Parro, Fernando, 2015. Estimates of the trade and welfare effects of NAFTA. Rev. Econom. Stud. 82 (1), 1–44.

Caliendo, Lorenzo, Parro, Fernando, 2022. Trade policy. In: Gopinath, Gita, Helpman, Elhanan, Rogoff, Kenneth (Eds.), Handbook of International Economics, Vol. 5. North-Holland, Amsterdam, pp. 219–295.

Chaney, Thomas, 2008. Distorted gravity: The intensive and extensive margins of international trade. Amer. Econ. Rev. 98 (4), 1707–1721.

Costinot, Arnaud, Rodríguez-Clare, Andrés, 2014. Trade theory with numbers: Quantifying the consequences of globalization. In: Gopinath, Gita, Helpman, Elhanan, Rogoff, Kenneth (Eds.), Handbook of International Economics, Vol. 4. North-Holland, Amsterdam, pp. 197–262.

Costinot, Arnaud, Rodríguez-Clare, Andrés, Werning, Iván, 2020. Micro to macro: Optimal trade policy with firm heterogeneity. Econometrica 88 (6), 2739–2776.

Costinot, Arnaud, Werning, Iván, 2019. Lerner symmetry: A modern treatment. Am. Econ. Rev.: Insights 1 (1), 13–26.

Demidova, Svetlana, Rodríguez-Clare, Andrés, 2009. Trade policy under firm-level heterogeneity in a small economy. J. Int. Econ. 78 (1), 100–112.

Demidova, Svetlana, Rodríguez-Clare, Andrés, 2013. The simple analytics of the Melitz model in a small economy. J. Int. Econ. 90 (2), 266–272.

Eaton, Jonathan, Kortum, Samuel, Kramarz, Francis, 2004. Dissecting trade: Firms, industries, and export destinations. Am. Econ. Rev. 94 (2), 150–154.

Eaton, Jonathan, Kortum, Samuel, Kramarz, Francis, 2011. An anatomy of international trade: Evidence from french firms. Econometrica 79 (5), 1453–1498.

Fajgelbaum, Pablo, Goldberg, Pinelopi, Kennedy, Patrick, Khandelwal, Amit, 2020. The return to protectionism. Q. J. Econ. 135 (1), 1–55.

Felbermayr, Gabriel, Jung, Benjamin, Larch, Mario, 2013. Optimal tariffs, retaliation and the welfare loss from tariff wars in the Melitz model. J. Int. Econ. 89 (1), 13–25.

Flam, Harry, Helpman, Elhanan, 1987. Industrial policy under monopolistic competition. J. Int. Econ. 22 (1–2), 79–102.

Gervais, Antoine, Jensen, J. Bradford, 2013. The Tradability of Services: Geographic Concentration and Trade Costs. NBER Working Paper 19759.

Gervais, Antoine, Jensen, J. Bradford, 2019. The tradability of services: Geographic concentration and trade costs. J. Int. Econ. 118, 331–350.

Gros, Daniel, 1987. A note on the optimal tariff, retaliation and the welfare loss from tariff wars in a framework with intra-industry trade. J. Int. Econ. 23 (3–4), 357–367.

Grossman, Gene, Helpman, Elhanan, 2021. When Tariffs Disturb Global Supply Chains. NBER Working Paper 27722.

Haaland, Jan I., Venables, Anthony J., 2016. Optimal trade policy with monopolistic competition and heterogeneous firms. J. Int. Econ. 102, 85–95.

Irwin, Douglas A., 2004. The aftermath of Hamilton's report on manufactures. J. Econ. Hist. 64 (3), 800–821.

Judd, Kenneth L., 1997. The Optimal Tax on Capital Income Is Negative. NBER Working Paper 6004.

Judd, Kenneth L., 2002. Capital income taxation with imperfect competition. Amer. Econ. Rev. 92 (2), 417–421.

Krugman, Paul, Venables, Anthony J., 1995. Globalization and the inequality of nations. Q. J. Econ. 110 (4), 857–880.

Lashkaripour, Ahmad, 2021. The cost of a global trade war: A sufficient statistics approach. J. Int. Econ. 131, 103419.

Lashkaripour, Ahmad, Lugovsky, Volodymyr, 2020. Profits, Scale Economies, and the Gains from Trade and Industrial Policy. Working Paper, Indiana University, April. https://alashkar.pages.iu.edu/Lashkaripour_Lugovskyy_2020.pdf.

Lenzen, Manfred, Kanemoto, Keiichiro, Moran, Daniel, Geschke, Arne, 2012. Mapping the structure of the world economy. Environ. Sci. Technol. 46 (15), 8374–8381.

Lenzen, Manfred, Moran, Daniel, Kanemoto, Keiichiro, Geschke, Arne, 2013. Building eora: A global multi-regional input-output database at high country and sector resolution. Econ. Syst. Res. 25 (1), 20–49.

Melitz, Marc J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725.

Melitz, Marc J., Ottaviano, Gianmarco I.P., 2008. Market size, trade, and productivity. Rev. Econ. Stud. 75 (1), 295–316.

Ossa, Ralph, 2011. A new trade theory of GATT/WTO negotiations. J. Polit. Econ. 119 (1), 122–152.

Venables, Anthony, 1987. Trade and trade policy with differentiated products: A Chamberlinian-Ricardian model. Econom. J. 97 (387), 700–717.

Yi, Kei-Mu, 2010. Can multi-stage production explain the home bias in trade? Amer. Econ. Rev. 100 (1), 364–393.